# A DISSERTATION

# FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

# Study on the Patterns of Genetic Variations Using Molecular and Transcriptomic Analysis of the Differentially Expressed Genes of Jeju Native Pig



**Simrinder Singh Sodhi**

**GRADUATE SCHOOL**

**JEJU NATIONAL UNIVERSITY**

**February, 2015**

# Study on the Patterns of Genetic Variations Using Molecular and Transcriptomic Analysis of the Differentially Expressed Genes of Jeju Native Pig

**Simrinder Singh Sodhi**

(Supervised by Professor Dong Kee Jeong)

A dissertation submitted in the partial fulfillment of the requirement for the degree
of
Doctor of Philosophy

2015. 02

This dissertation has been examined and approved by

_____

Professor Sung Jong Oh, Chairman, Department of Animal Biotechnology, Jeju National University



_____

Professor Youn Chul Ryu, Vice Chairman, Department of Animal Biotechnology, Jeju National University

_____

Professor Hee Bal Kim, Department of Agricultural Biotechnology and Research Institute for Agriculture and Life Sciences, Seoul National University

_____

Professor Hak Kyo Lee, Department of Biotechnology, Hankyong National University,

_____

Professor Dong Kee Jeong, Department of Animal Biotechnology, Jeju National University

GRADUATE SCHOOL
JEJU NATIONAL UNIVERSITY

*This work is dedicated to*

*My Father- Dr. Davinder Singh Sodhi*

*and*

*My Jija ji- Late. Mr. Jagtinder Singh Dutt,*

*who dreamt for my present and who actually lived this dream more than me!!*

# ABSTRACT

Pork is a major source of animal protein for humans. The subcutaneous, intermuscular and the intramuscular fat (IMF) are the factors responsible for meat quality. RNA-seq is being rapidly adopted for the profiling of the transcriptomes in different areas of biology, especially in the studies related to gene regulation. The discovery of differentially expressed genes (DEGs) between adult Jeju Native Pig (JNP) and Berkshire breeds of *Sus scrofa*, are of particular interest for the current study. For the better understanding of the gene expression profiles of fat, liver and *longissimus dorsi* muscle, DEGs were identified via RNA-seq. Sequence read numbers were obtained from Ilumina HiSeq2000 and mapped to the pig genome using Tophat2. Total 153 DEGs were identified in fat and among the 71 annotated genes have BLAST matches in the non-redundant database. Similarly 169 and 39 DEGs were identified in fat, liver and muscle of JNP respectively, by comparison with Berkshire breed. Out of all identified genes, 41 genes in liver and 9 genes in muscle have given significant expression. Gene ontology (GO) terms of developmental process and KEGG pathway analysis showed that metabolic, immune response and protein binding were commonly enriched pathways in the three tissues. Further the heat map analysis by ArrayStar has shown the different levels of expression in JNP with respect to the Berkshire breed. The validation through real time PCR and western blotting also confirmed the differential expression of genes in both the breeds. Genes pertaining to metabolic process and inflammatory and immune system are more enriched in Berkshire breed. This comparative transcriptome analysis of two tissues suggests a subset of novel markers genes which expressed differently between the JNP and Berkshire. In our study, biological adhesion, cellular, developmental and multicellular organismal processes in fat were up-regulated in JNP as compare to Berkshire. Multicellular organismal process, developmental process, embryonic morphogenesis and skeletal system development were the most significantly enriched terms in fat of JNP and Berkshire breeds ($p$ = 1.17E-04, 0.044, 3.47E-04 and 4.48E-04 respectively).

The differential expression of genes was also investigated between piglets of Jeju Native Pig (JNP) and Berkshire breeds. Paired end reads of the sequences those passed the quality filters were aligned to *Sus scrofa* genome using Tophat2 (v2.0.2). In the current study, 65% (muscle), 20% (fat) and 54% (liver) genes have shown higher expression in the piglets of JNP than Berkshire. Gene ontology and signaling pathway showed that immune response and lipid metabolisms were commonly enriched pathways in all the three tissues. It was found that the

genes pertaining to body growth and immune system are significantly ($P$<0.01) more expressive in Berkshire piglets. DEGs explored between the piglets of two breeds might influence the identification of the genetic markers for further breed improvement programmes.

Further, the differential expression of a novel liver protein acetyl-CoA acetyltransferase-2 (ACAT2) has been identified in the liver tissue. ACAT2 is onvolved in the beta-oxidation and lipid metabolism. Its comprehensive relative expression, *in silico* non-synonymous single nucleotide polymorphism (nsSNP) analysis, as well as its annotation in terms of metabolic process with another protein from the same family, namely, acetyl-CoA acyltransferase-2 (ACAA2) was performed in *Sus scrofa*. The study was conducted to understand the most important nsSNPs of ACAT2 in terms of their effects on metabolic activities and protein conformation. The two most deleterious mutations at residues 122 (I to V) and 281 (R to H) were found in ACAT2. Validation of expression of genes in the laboratory also supported the idea of differential expression of *ACAT2* and *ACAA2* conceived through the *in silico* analysis. Analysis of the relative expression of *ACAT2* and *ACAA2* in the liver tissue of Jeju native pig showed that the former expressed significantly higher ($P < 0.05$). Overall, the computational prediction supported by wet laboratory analysis suggests that ACAT2 might contribute more to metabolic processes than ACAA2 in swine. Further associations of SNPs in ACAT2 with production traits might guide efforts to improve growth performance in Jeju native pigs.



The demand of pig meat is increasing day by day and during the last decade the quality of pork has become one of the main selection criteria. *MyBPH* has significantly higher transcript level in the *longissimus dorsi* muscle of JNP as compare to Berkshire. The current study targeted the role of the *MyBPH* in the muscle growth of JNP. The systematic *in-silico* annotations of SNPs and *in-vitro* functional analysis of *MyBPH* and other muscle regulatory factors have been performed. Different levels of polymorphism among the SNP sites were confirmed in *MyBPH* by sequence- and structure-based algorithms. More than 80% of total nsSNPs have been predicted as to be highly deleterious in JNP. *In silico* analysis by I-Mutant3 and HOPE tool has indicated that misfiling and intermolecular interaction of the *MyBPH* can significantly ($P$<0.05) affect the structures and functions of the protein. Therefore, such kind of changes due to nsSNPs target the functioning of *MyBPH* and inspite of higher transcript levels, *MyBPH* is still not able to potentiate the myogenesis in JNP. The sharing of specific myogenic features associated with *MyBPH* and MRFs through ToppCluster indicates that *MyBPH* can be a potent candidate gene to potentiate biological processes. The mutations in *MyBPH* can alter the muscle-specific biological and pathophysiological processes involved in muscle growth and metabolism. Our findings reveal *MyBPH* as a candidate gene for the muscle specific biological and pathophysiological processes.

The current study is one of the pioneer transcriptomic studies being conducted in JNP. This study is the first transcriptional analysis for the detection of DEGs from RNA-seq data generated from fat, liver and *longissimus dorsi* muscle tissue samples. This analysis can be used as stepping stone to understand the difference in the genetic mechanisms that might influence the identification of novel transcripts, sequence polymorphisms, isoforms and noncoding RNAs. Our findings give a new horizon for understanding and identification of the candidate genes that are involved in the altered biological functions. Moreover, transcriptome analysis makes it easier to understand the differences between genetic mechanisms of breeds.

# 요 약

돼지고기는 동물성 단백질의 주요 공급원이다. 피하지방, 근육간 지방과 근간지방은 고기의 육질을 결정하는 주요 요소이다. RNA-seq 은 특히 유전자 조절과 연관된 연구를 하는 생물학 분야에서 전사체를 프로파일링하는 방법으로 적용되고 있다. 제주 재래돼지와 버크셔 종 간의 유의발현유전자의 발견은 본 연구에서 가장 흥미로운 결과이다. 지방, 간과 가장 긴 배측근육의 유전자 조절 프로파일 이해를 위해 RNA-seq 을 이용해 유의발현유전자를 확인하였다. Illumina HiSeq2000 을 이용해 시퀀스 데이터를 얻었고, Tpohat2 를 이용해 돼지 지놈을 지도화 하였다. 총 153 개 유의발현유전자가 지방에서 확인되었고, 그 중 71 개 annotated 유전자들이 non-redundant 데이터베이스에서 BLAST matches 했다. 마찬가지로 제주 재래돼지의 지방, 간과 근육 조직에서 169 개와 39 개의 유의발현유전자를 확인하였다. 확인된 유전자 중, 발현에 영향을 미치는 유전자가 간에서 41 개 근육에서 9 개로 확인되었다. 유전자 온톨로지 발달과정과 KEGG pathway 분석은 대사, 면역반응 및 결합 단백질이 일반적으로 세 개의 조직에서 강화된 것으로 나타났다. 또한 ArraryStar 에 의한 heat map 분석은 버크셔 품종과 제주 재래돼지간 서로 다른 발현차이를 보여주었다. Real time PCR 과 western blot 을 통한 검증 실험에서도 두 품종간 유전자 발현 차이가 확인되었다. 신진대사 및 염증, 면역 시스템에 관련된 유전자는 버크셔 품종에서 발현율이 높았다. 두 조직간 전사체 분석에서 제주 재래돼지와 버크셔 간에 발현차이를 보이는 유전자들을 마커로 제안한다. 다세포 유기체 과정, 발달과정, 배아 형태 형성과 골격 시스템 개발은 제주 재래돼지와 버크셔 품종에서 매우 풍부했다($p$ = 1.17E-04, 0.044, 3.47E-04 and 4.48E-04 respectively). 제주 재래돼지와 버크셔 품종의 새끼 돼지간 유전자 발현차이를 조사하였고, paired-end reads 는 Tophat2(v2.0.2)를 이용해 *Sus*

*scrofa* 지놈과의 서열을 정리하였다. 본 연구에서는 버크셔 보다 제주 재래돼지 새끼에서 65%(근육), 20%(지방) 그리고 54%(간) 더 높은 유전자 발현을 보였다. 면역 반응과 지질 대사가 일반적으로 세 개의 조직에서 풍부하다는 것이 유전자 온톨로지 및 신호 전달 pathway 에서 보여졌다. 이것은 몸의 성장과 면역 시스템에 관련된 유전자가 버크셔의 새끼 돼지에서 더 발현되는 것으로 나타났다($P<0.01$). 두 품종의 새끼간 유의 발현유전자의 분석으로 인해 품종 개량 프로그램에 대한 유전자마커 식별에 영향을 미칠 것으로 사료 된다. 또한, 간 조직에서 발현차이를 보이는 새로운 간 단백질 acetyl-CoA acetyltransferase-2 (ACAT2)를 확인하였고, ACAT2 는 베타-산화 및 지질대사에 관여하는 것으로 알려져 있다. 본 연구는 대사활동 및 단백질 입체 구조에 영향을 미치는 ACAT2 의 가장 중요한 nsSNP 를 이해하기 위해 수행하였다. Residues 122(I to V)와 281 (R to H)에서 가장 유해한 2 개의 돌연변이가 ACAT2 에서 발견되었다. 실험실에서 진행한 유전자 발현 확인은 ACAT2 와 ACAA2 의 다른 발현차이를 보이는 *in silico* 분석 결과를 뒷받침할 수 있다. 제주 재래돼지의 간 조직에서 ACAT2 와 ACAA2 의 발현을 비교 분석한 결과, ACAT2 가 더 많이 발현되는 것을 확인하였다($P<0.05$). 전반적으로, 분석을 통한 수치적 추정은 실험적인 분석에 의해서 검증되었으며, ACAT2 는 ACAA2 보다 돼지의 대사과정에 더 기여할 것으로 사료된다. 아울러 ACAT2 의 SNP 와 생산형질과의 관계는 제주 재래돼지의 성장률을 향상시키기 위한 지표가 될 것이다. 돼지고기의 수요가 나날이 증가하고 지난 10 년동안 돼지의 품질이 주요 선택 기준 중 하나가 되고 있다. MyBPH 는 버크셔 종과 비교했을 때 제주 재래돼지의 가장 긴 배측 근육에 매우 높은 전사체 수준을 가지고 있다. 본 연구는 제주 재래돼지의 근육 성장에 있어 MyBPH 의 역할을 알아 보고자 한 것이다. MyBPH 와 근육 조절 요인의 생체 외 기능분석 및 SNPs 의 체계적인 *in-silico* annotation 이 수행되었다. SNP 사이트 중 다형성의 다른 차이는 시퀀스와 구조 기반의 알고리즘에 의해 MyBPH 에서 확인 되었다. 총 nsSNPs 의 80% 이상이 제주 재래돼지에

매우 해로운 것으로 예상되고 있다. I-Mutant3 와 HOPE tool 에 의한 *in-silico* 분석은 MyBPH 의 분자간 상호작용이 단백질의 구조와 기능에 영향을 미칠 수 있음을 나타내고 있다. 따라서 높은 전사물 수준에도 불구하고 MyBPH 의 기능작용을 가지는 nsSNP 의 변화 때문에 MyBPH 는 제주 재래돼지에서 근육 생성을 강화시킬 수 없다. Toppcluster 를 통해 MyBPH, MRFs 와 관련된 특정 근육 조직의 기능 공유는 MyBPH 가 생물학적 과정을 강화시킬 수 있는 강력한 후보 유전자임을 나타낸다. MyBPH 의 돌연변이는 근육의 성장과 대사에 관여하는 근육 고유의 생물학, 병리생리학적 과정을 변경 시킬 수 있다. 우리의 연구결과는 근육 특유의 생물학적, 병리생물학적 과정에 관여하는 후보 유전자로 MyBPH 를 강조했다. 본 연구는 처음으로 제주 재래돼지를 이용한 transcriptomic 의 연구 중 하나이며 지방, 간 및 가장 긴 배측 근육 조직의 샘플을 이용하여 발생된 RNA-seq 데이터로부터 유의발현유전자의 검출을 위한 전사 분석이다. 이 분석은 새로운 전사물, 시퀀스 다형성, 동형단백질 및 비 암호화 RNAs 의 식별에 영향을 줄 수 있는 유전자 메커니즘의 차이를 이해하는 디딤돌로서 사용될 수 있을 것이다. 본 연구결과는 변형된 생물학적 기능에 관여하는 후보 유전자의 이해 및 식별을 위해 새로운 시야를 제시하고 있으며, 또한 전사체의 분석은 품종의 유전학적인 메커니즘 사이의 차이를 이해하기 쉽게 만들 것이다.

# ACKNOWLEDGEMENTS

*At this juncture of my life, it is a pleasant task to express my thanks to all those who contributed in many ways to the success of this study and made it an unforgettable experience for me. Foremost of all, I bow my head in humility to Omnipresent, Omniscient and Omnipotent* ***"Wahe Guru Ji -The Almighty God"*** *for bestowing upon me the wisdom, health and strength to undertake this research task and enabling me to its completion.*

*It is difficult to overstate my gratitude to my advisor,* ***Prof. Dong Kee Jeong,*** *Professor, Department of Animal Biotechnology, Faculty of Biotechnology, Jeju National University, Jeju-si, Jeju-do, Republic of Korea for his unreserved and ever willing support during my PhD study and research. I owe to him for his patience, unflinching encouragement, enthusiasm and immense scientific knowledge. His thoughtful guidance, constructive criticism and meticulous supervision helped me throughout my research and thesis writing.*

*I am grateful to the Directorate of Extension Education and Department of Veterinary and Animal Husbandry Extension Education, Guru Angad Dev Veterinary and Animal Sciences University, Ludhiana for grating me study leave to pursue my Ph.D.*

*I would like to thank members of my defence committee,* ***Prof. Sung Jong Oh,*** *Chairman, Department of Animal Biotechnology, Jeju National University,* ***Prof. Youn Chul Ryu****, Vice- Chairman, Department of Animal Biotechnology, Jeju National University,* ***Prof. Hee Bal Kim,*** *Department of Agricultural Biotechnology and Research Institute for Agriculture and Life Sciences, Seoul National University and* ***Prof. Hak Kyo Lee,*** *Department of Biotechnology, Hankyong National University, Republic of Korea, for their kind counselling, ever willing and untiring help, healthy comments and moral support for the pursuance of this research study.*

*I express my deep sense of gratitude to* ***Prof. Lee Wang Shik, Prof. Kang Min Soo*** *and* ***Prof. Yang Young Hoon,*** *Department of Animal Biotechnology, Jeju National University for their valuable suggestions and constant inspiration while working in the department.*

*Words are compendious to express my sincere thanks to* ***Dr. Simrat Sagar Singh and Dr. (Mrs.) Seema Bedi*** *for their sincere everlasting moral support, constant encouragement and constructive suggestions throughout my carrier.*

*My profound thanks are to my revered friends* ***Dr. Vikas Gawari, Dr. Neelesh Sharma, Dr. Vikas Vohra, Dr. Vijay Pandey, Dr. Navdeep Singh (Honey), Dr. Shahbaz, Dr. Harpreet Dhillon, Dr. Gurpreet Bath, Col. (Dr.) S. P. Singh and Dr. Nitin Thakur*** *for letting me encroach upon their valuable time and providing me with un-ending emotional support during my study. I will be failing in my duty if I do not mention about the unconditional support rendered by* ***Dr. Anil Kumar Khambampati, Dr. Rashid Ahmad, Dr. Mayank Anand Gururani, Ms. Nguyen Thai Quynh Anh*** *and* ***Do Luong Huynh*** *in completion of this research work.*

*My special thanks are due to* ***Dr. Mrinmoy Ghosh, Ms. Kim Nam Eun, Raj Kumar Mogre, Kim Jeong Hyun, Ms. Zhang Jiao Jiao, Satya Chea and Ms. Kang Hyun-Ju*** *for their persistent emotional, moral support and providing me congenial atmosphere in the lab for completion of this work.*

*There is a common saying that behind every success of a man there is a woman, but here I would like to deviate from this saying. I owe my this achievement to the two wonderful ladies of my life;* ***my mother- Mrs. Jasbir Kaur Sodhi*** *and* ***my sister- Mrs. Inderpreet Kaur Dutt****. I am thankful to both of them for their never ending prayers and carving me in such a way that I could achieve my present. I acknowledge the best wishes and blessings of my parents-in-laws;* ***Sh. Harmohinder Singh Keer and Mrs. Amarjit Kaur Keer.***

*The persistent support provided by* ***my brother- Col. (Dr.) Parminder Singh Sodhi*** *and his wife-* ***Dr. Manmeet Kaur Sodhi*** *was the axis of my work. They acted as a trusted comrade and spared me for the work by shouldering my responsibilities. The constant encouragement by my co-brothers-* ***Dr. Tarvinder Pal Singh, Mr. Dalbir Singh*** *and by my co- sisters-* ***Dr. Gurvinder Kaur*** *and* ***Dr. Gurpreet Kaur*** *led me towards successful completion of this research work.*

*One statement is very strongly engraved in my mind which says that- "A wife is like a children's movie; always under- appreciated and without either, life would be incomplete." Therefore, it is a matter of an honour for me to say that I could not have finished my study so smoothly without the support of my wife-* ***Dr. Jaswinder Kaur Sodhi,*** *who stepped back in her professional endeavours so that my dreams could get wings. I would like to thank her for always listening to me, supporting me and encouraging me.*

*I feel it beyond my vocabulary to express my veneration towards my sons* ***Angad Jagtinder Singh Sodhi*** *and* ***Barkat Sodhi*** *for their sacrifice, patience and perseverance in bearing the pangs of separation at the time of their utmost need.*

*It gives me pleasure to acknowledge the love and affection of the angels in my life-* ***GurNoor Singh Dutt, Sant Bux Singh Dutt, Gurmehar Sodhi, Harsidak Sodhi, Pavseerat, Sehaj Bedi, Binal and Arpan Bedi****.*

*In the last but not the least, I am thankful to all those persons who may not be listed here, but without whom the completion of this research work would not have been possible.*

***Jeju, Republic of Korea*** ***(Simrinder Singh Sodhi)***

***Date:***

# Table of Contents

# LIST OF TABLES

# LIST OF FIGURES

# ABBREVATIONS

| | |
|---|---|
| *ACAA2* | - Acetyl-CoA Acyltransferase-2 |
| *ACAT2* | - Acetyl-Coa Acetyltransferase-2 |
| *AIFM3* | - Apoptosis-Inducing Factor Mitochondrion-Associated - 3 |
| *ALDH3A2* | - Aldehyde Dehydrogenase 3 family |
| *ALOX15* | - Arachidonate 15-Lipoxygenase |
| ANOVA | - Analysis Of Variance |
| *APOA4* | - Apolipoprotein- A4 |
| BAC | - Bacterial artificial chromosome |
| bHLH | - basic Helix- Loop- Helix |
| BLASTn | - Basic Local Alignment Search Tool for nucleotide |
| *BMP10* | - Bone Morphogenetic Protein-10 |
| *BMP2* | - Bone Morphogenic Proteins- 2 |
| *BMP4* | - Bone Morphogenic Proteins- 4 |
| CAGE | - Cap Analysis Gene Expression |
| CASAVA -v1.7 | - Consensus Assessment of Sequence and Variation v 1.7 |
| *CCL4* | - Chemokine C-C motif ligand 4 |
| *COL10A1* | - Collagen type 10, Alpha 1 |
| *COL11A2* | - Collagen type 11, alpha 2 |
| *COL21A1* | - Collagen type21-alpha-1 |
| *COL2A1* | - Collagen type 2- alpha 1 |
| *COL9A1* | - Collagen type 9-alpha-1 |
| *COL9A2* | - Collagen type 9-alpha-2 |
| *CRABP2* | - Cellular Retinoic Acid Binding Protein 2 |
| *CRP* | - C - Reactive Protein |
| *CXCL13* | - Chemokine ligand-13 |
| *CYP2E1* | - Cytochrome- P2E1 |
| DAPI | - 4, 6-Diamindino-2-Phenylindole |
| DAVID | - Database for Annotation, Visualization and Integrated Discovery |
| DEGs | - Differentially Expressed Genes |
| DMEM | - Dulbecco's Modified Eagle's Medium |

| | |
|---|---|
| EASE | - Expression Analysis Systematic Explorer |
| *ELOVL6* | - Elongase-6 |
| emPCR | - Emulsion PCR |
| *FASN* | - Fatty acid synthase |
| FBS | - Fetal Bovine Serum |
| *FMO* | - Flavin Containing Mono-Oxygenase |
| *FMO1* | - Flavin Containing Mono-Oxygenase- 1 |
| *FMO2* | - Flavin Containing Mono-Oxygenase- 2 |
| *FMO5* | - Flavin Containing Mono-Oxygenase- 5 |
| GFP | - Green Fluorescent Protein |
| *GFPT1* | - Glutamine--Fructose-6-Phosphate Transaminase 1 |
| GLM | - General Linear Model |
| GO | - Gene ontology |
| HDL | - High density lipids |
| *HMGCS2* | - Hydroxy methyl glutaryl CoA synthetase 2 |
| HMMs | - Hidden Markov Model |
| HOPE | - Have yOur Protein Explained |
| *IL1RAP* | - Interleukin- 1 Receptor Accessory Protein |
| *IL7R* | - Interleukin-7 receptor |
| IMF | - Intramuscular Fat |
| ITRs | - Inverted terminal repeat sequences |
| JNP | - Jeju Native Pig |
| JSSP | - Jeju special self-governing province |
| KEGG | - Kyoto Encyclopedia of Genes and Genomes |
| *LCN2* | - Lipocalin-2 |
| *LGALS3* | - Lactin Galactoside- Binding Soluble 3 |
| MHC | - Myosin Heavy Chain |
| MPSS | - Massively Parallel Signature Sequencing |
| MRFs | - Muscle Regulatory Factors |
| mtDNA | - mitochondrial DNA |
| *MyBPH* | - Myosin binding protein- H |
| *Myf5* | - Myogenic factor-5 |
| *Myf6* | - Myogenic factor-6 |
| *MyoD* | - Myogenic Determination factor |

| | |
|---|---|
| *NASH* | - Non-Alcoholic Steato-Hepatitis |
| NCBI | - National Center for Biotechnology Information |
| NGS | - Next Generation Sequencing |
| NOMAD–Ref | - Normal Mode Analysis, Deformation, and Refinement |
| nsSNP | - non-synonymous Single Nucleotide Polymorphism |
| PBS | - Phosphate Buffer Saline |
| *PCK1* | - Phosphoenolpyruvate carboxykinase-1 |
| PCR | - Polymerase Chain Reaction |
| *PDK4* | - Pyruvate dehydrogenase kinase-4 |
| PE | - Paired-End |
| *PLCH2* | - Phospholipase C, eta 2 |
| PMSF | - Phenyl Methyl Sulfonyl Fluoride |
| *PNPLA3* | - Patatin like Phospholipase domain containing- 3 |
| *POSTN* | - Periostein |
| *PPL* | - Periplakin |
| *PRG4* | - Proteoglycan- 4 |
| PROVEAN | - Protein Variation Effect Analyzer |
| *PSPH* | - Phosphoserine Phosphatase |
| qRT- PCR | - quantitative Real-Time PCR |
| *RDH16* | - Retinol Dehydrogenase -16 |
| RIPA | - Radio Immuno Precipitation Assay |
| RMSD | - Root Mean Square Deviation |
| RNA-seq | - RNA sequencing |
| RPKM | - Reads Per Kilobase Per Million |
| SAGE | - Serial Analysis of Gene Expression |
| SDS | - Sodium Dodecyl Sulfate |
| SIFT | - Sorting Intolerant From Tolerant |
| *SIK1* | - Serine/threonine-protein kinase-1 |
| *SLA* | - Src-like-adaptor |
| *SLA-DRB1* | - Swine Leukocyte Antigen – DRB1 |
| SMRTS | - Single molecule real time sequencing |
| SNPs | - Single Nucleotide Polymorphisms |
| SOLiD | - Sequencing by Oligonucleotide and Detection |
| SRide | - Stabilising Residues |

| | |
|---|---|
| subPSEC | - substitution Position-Specific Evolutionary Conservation |
| *TBX5* | - T- box transcription factor-5 |
| TMA | - Trimethyl Amines |
| TMAO | - Trimethyl Amines oxide |
| TSP | - Transcriptional start points |
| UCSC | - University of California Santa Cruz |

# Chapter 1

# INTRODUCTION

Pork is one of the major sources of animal protein for humans. High impetuous research conducted during last few years has helped in understanding the development, physiology and metabolism of porcine muscle tissue (Liu *et al*., 2012). In 2005, South Korea occupied seventh place in the pork consumption as compared to whole world whereas by 2008 with the increased consumption it jumped up to fifth place. The Pork consumption in Korea has increased over time and this market is expected to keep growing in the future. The preparation process for the Free Trade Agreement between Korea and the United States has played vital role in the increase of pork consumption and production rates. Currently the per capita availability of pork in South Korea is 29.3kg (Oh and Whitley, 2011). Pork produced by China, Japan and South Korea comprise the major share of the global pork export market.

Over the years pig breeding policies have enabled the improvement of European pig breeds with the sole purpose of providing meat to fulfill the human requirements. Till 20$^{th}$ century, pig breed improvement programmes were aimed at to increase the fatness since the fat has been considered as one of the important source of energy for the manual workers of that era (Whittemore, 2006). The cross-breeding of indigenous breeds with the European breeds throughout the 19th century led to increased fatness, decreased age and body size at maturity (Kijas and Andersson, 2001). Pressure created by human demands has acted as the vital force for the domestication and breed improvement policies. A gap between the demand and supply of pig meat also made breeds to think for the alternate breeding strategies. Therefore, over the last few decades, pig breeding is mainly aimed upon the improvement of economic traits like growth, fertility and pork production. However, consumers' preference has also driven the breeders to select for the reduction in fat content (Whittemore, 2006) and increased palatability traits of pork.

Jeju Island (Republic of South Korea), one of the 7 natural wonders of the world has very unique weather conditions. Due to its idiosyncratic environmental conditions, this island has very unique flora and fauna. Native livestock species of Jeju-Do are very well suitable or adjusted to local environmental conditions. Currently, two distinct kinds of pig populations are available in South Korea i.e. indigenous Asian types (e.g. Jeju Native Pig, JNP) in village traditional farms

and commercial European and North American pig breeds. It is estimated that JNP was introduced in 12$^{th}$ century from - China. Amongst the indigenous livestock reared for meat purpose JNP occupies the top rank of the hierarchy in Korea. Meat from JNP is tenderer, juicy and has higher marbling score than western breeds (Jin *et al*., 2001; Cho *et al*., 2011). These qualities of JNP meat bring it on the preference list of consumers especially in China, Japan and Korea. Pork from JNP is regarded as of high quality and is the costliest pork in Korea (Kim *et al*., 2013). Although the above enlisted traits make pork produced by JNP to be of superior quality, but still the major snags faced by JNP are lesser growth rate, lower feed efficiency and small litter size (Kim *et al*., 2009). The lesser growth rate and smaller litter size create a wide gap between the demand and supply of the pork. Inspite of being superior in palatability traits yet pork produced by JNP is not able to meet the market demands.

Most of the commercially viable pig breeds, such as Berkshire, Yorkshire/Large White, Landrace, Duroc and Hampshire have been under intense selection for lean growth. Selection with an aim for improved lean growth in pig led to an increase in muscle growth and a reduction in fat deposition. Twentieth century has been considered as a linear growth period in the pig meat market. During 1930s to 1940s, extensive hybridization took place between JNP and Berkshire breed. Moreover, the black haired Berkshire breed is expected to intervene in paternal linages of JNP and is preferred for meat production (Kim *et al*., 2011; Hur *et al*., 2013). Moreover, the pork from the crossbred animals (JNP and Berkshire) has been also very popular among the Asians especially Korean, Chinese and Japanese. Meat quality characteristics of JNP and Berkshire are very much comparative, but the body growth rate is higher in Berkshire. Due to better meat quality, pork from the crossbreds of JNP and Berkshire occupies unique share in meat industry. Therefore, keeping above mentioned aspects under consideration, these two breeds are frequently chosen for the breed improvement programmes with the aim to improve the meat quality and quantity of JNP (Yu *et al*., 2013).

Skeletal muscles comprise 20 to 50% of the total body mass in pigs and are one of the main tissues which are responsible for meat production in pigs (Wu *et al*., 2013). Moreover, muscularity and growth rate are among the prime economic traits of meat animals. Clear evidences have established that muscle fiber characteristics e.g. fiber size, number, fiber area and density of fibers are in close vicinity with the traits related to meat quality (Ryu and Kim, 2006). It is also reported that composition of fiber type also effect energy metabolism in pre- slaughter and post slaughter metabolism in carcass (Ryu *et al.*, 2008). Prospective for the growth are interlinked with the number of the secondary myotubes of muscles. Further, it has been observed that the postnatal growth and regeneration of skeletal muscles are associated with the presence

and activities of the satellite cells (Molik *et al*., 2011). Muscle regulatory factors (MRFs), which are member of basic helix- loop- helix (bHLH) family control formation of muscles fibers - myogenesis (Buckingham *et al*., 2003). Moreover, bHLH transcription factors are reported to regulate myogenesis and transcription of muscle specific genes (te Pas *et al*., 2007). Although the relationship between carcass weight, muscle fiber characteristics and pork quality traits is not fully understood, but still muscle mass and the weight of the pigs are related to fiber composition (Ryu and Kim, 2006). Meager reports concerning the muscle fiber characteristics of JNP and Berkshire crossbreds are on-record. Therefore, the comparative meat quality characteristics of JNP and Berkshire along with the role of MRFs in muscle development during different developmental stages of growth have acted as the base for the current study. As it is an established fact that muscle tissue plays a major role in the economics of meat production, but the understanding of variation in expression of gene during muscle development in younger age is in its primitive stage.

The skeletal muscles, adipose tissue and liver are the principal organs involved in the regulation of lipid metabolism (Corominas *et al.*, 2013). Fat and variation in fatty acid contents have significant role in the meat quality. Proportion of fat in meat is assessed as marbling score and it is an important factor influencing the meat quality (Sheng *et al.*, 2014). Verbeke *et al.* (1999) reported that intramuscular fat content is one of the important traits which influence the meat tenderness, juiciness and taste. Similarly, different melting points of fatty acids effect the firmness or softness of the fat in meat and thereby affect meat quality, especially the subcutaneous, intermuscular (carcass fats) and the intramuscular fat (IMF) content (marbling) (Wood *et al*., 2004). Therefore, fat tissue is also a one of the major factor responsible for meat quality.



At the same time, liver is one of the most important organs for metabolism and the partitioning of nutrients. It is involved in the transformation of dietary nutrients into fuel and their export via the blood. To meet the changing demands of extra-hepatic tissues for nutrients, liver displays remarkable metabolic flexibility (Baik *et al*., 2009). Liver can also be a tissue of choice to study the immunity and detection of genes related to immune system which will confer an advantage in selection of breeding stock. Since, the weaker immune system in piglets will lead to the development of a herd with poor performance. Worldwide health challenges of the parent stock of the pigs are reported to adversely affect the performance, which causes huge economic loses to the pork industry. Therefore, identification of candidate genes of economic importance in piglets could also be useful for the quick determination of the association of a specific genetic

variant with their phenotype. Therefore, gaining knowledge of the porcine immune system will undoubtedly benefit many aspects of porcine disease research.

During the previous years, genes and different markers between pig breeds have been identified using microarray, qRT-PCR and microsatellite analyses (Kim *et al*., 2005; Kim *et al*., 2008; Moon *et al*., 2009). Recently, Next Generation Sequencing (NGS) has been established as the preferred approach for characterization and quantification of the entire genome. For the serial and cap analysis of gene expression, NGS is more effective than the previous technologies (Mortazavi *et al*., 2008). Developments in RNA-seq technology enable more comprehensive investigation of the transcriptome than microarrays and are preferred for gene expression studies (Mortazavi *et al*., 2008). It allows genome-wide analysis of transcription at single nucleotide resolution, including identification of alternative splicing events and post-transcriptional RNA editing events. The Illumina HiSeq, Ion Torrent and SOLiD systems are the most frequently used platforms of NGS for RNA-seq (Cloonan *et al*., 2008; Eveland *et al*., 2008; Marioni *et al*., 2008). Dynamic ability of RNA –seq to discriminate homologous regions and identification of regions of expression without the prior assumption, makes it as the future of transcriptome research (Shendure and Ji, 2008). Till now transcriptome analysis using RNA seq has been applied in different pig breeds except JNP (Chen *et al*., 2011; Petkov *et al*., 2011; Rustemeyer *et al*., 2011; Jung *et al*., 2012; Looft 2013; Prather, 2013; Samborski, 2013). Moreover, very less comparative studies among the different pig breeds have been conducted for biological processes. The statistical analysis is also critical in transcriptomic studies using RNA-seq; specifically for the normalization of quantitative measurements of expression (Wilhelm and Landry, 2009; Li *et al*., 2010; Robinson and Oshlack, 2010; Hong *et al*., 2012) and detection of DEGs (Hardcastl and Kelly, 2010; Robinson *et al*., 2010; Wang *et al*., 2010; Tarazona *et al*., 2011; Trapnell *et al*., 2012).



The current study is first of its kind, which has been devoted to perform the transcriptome analysis for the divergence in gene expression in adult animals and piglets of JNP. Taken collectively, our integrated RNA-Seq approach aims to reveal differential expression profiles of the genes which may be useful for successful marker-assisted selection for improved growth rate and muscle mass in JNP.

Therefore, this study was planned with the following objectives:

1. To identify differentially expressed key regulatory genes from JNP and Berkshire which can be used as breed specific markers for the discrimination of breeds.

2. Identification and analysis of biomarker genes which can be useful to monitor growth performance, meat quality and immunity in adult and piglets of both the breeds.
3. To identify a novel marker which can be responsible for differential growth pattern between JNP and Berkshire, using stem cells/plasmid mediated system.

# Chapter 2

# REVIEW OF LITERTURE

Pig domestication has played a vital role in the rise of agriculture in the world and the domestic pigs are one of the major sources of animal-based protein for humans. During the process of domestication, the current domesticated pigs exhibit remarkable morphological differences than their wild ancestors. These phenotypic differences have acted as base to consider these two forms as distinct species (Darwin, 1868). Moreover, interaction between the genotype and environment has played significant role for the development of *Sus domesticus.*

## 2.1. History of the domestication of pig and subsequent introgression:

Pig an even-toed ungulate belongs to the genus *Sus* with Suidae family. It has been suggested by an extensive zoo-archeological record that the first domestication of pigs as domestic pig (*Sus domesticus*) from wild boars (*Sus scrofa*) started since 9000 years ago in the East (Epstein and Bichard, 1984). The wild boar is widespread in Eurasia and Northwest Africa. It is proposed that at least 16 different subspecies of wild boar existed (Ruvinsky and Rothschild, 1998). There are marked morphological differences between the wild and domestic pigs. Till now, it is not fully established fact that whether the domestic pigs have single or multiple origin. The two major forms of domestic pigs i.e. a European (*Sus scrofa*) and an Asian form (*Sus indicus*) were recognized by Darwin (1868). The former was assumed to originate from the European wild boar, while the wild ancestors of the Asian form are unknown. These two forms have been considered as distinct species on the basis of profound phenotypic differences. It's a well-documented fact that Asian pigs were involved in the crossbreeding programmes to improve European pig breeds during the 18th and early 19th centuries (Darwin, 1868; Jones, 1998) but upto what level Asian pigs have contributed genetically to different European pig breeds is still unknown. The use of microsatellite markers has estimated the divergence between European breeds and the Chinese Meishan breed to be approximately in the range of 2000 years (Paszek *et al.*, 1998). Few of the mitochondrial DNA (mtDNA) studies have indicated genetic differences between European and Asian pigs but no exact starting time of divergence has been provided (Watanabe *et al.*, 1986; Okumura *et al.*, 1996). Giuffra *et al.* (2000) reported that an independent

domestication of pig at second stage occurred in China. 16 different subspecies of the wild boar and more than 730 domestic breeds of pig have been listed all over the world and two third breeds are available in China and Europe (Rischkowsky *et al*., 2007). Some of the earlier studies have marked the different locations of pig domestication across Eurasia (Larson *et al*., 2005).

### 2.2. Quality and grading of pork:

Increasing demand for lean meat has directed breeders to obtain modern meat type of fattener (Stewart, 1991). To improve the efficiency of pork production, intense selection is targeted at improvement of increased daily gain and carcass leanness. It has resulted in improved growth rate, better feed conversion ratio as well as lean meat content and loin eye area along with decreased back fat thickness and carcass fat content (Tribout, 2010). The characteristics related to meat quality play an important role for the consumer's acceptance of pork and these are also affected by the selection procedure (Schwab, 2006). Quality of meat is a complex parameter which depends on the interaction between genotype of animal, environmental conditions, pre-slaughter handling and slaughtering procedure (Rosenvold and Andersen, 2003). A large segment of genes associated with both structural and metabolic features of muscles are involved in the physiological processes taking place in muscle and affect the quality of meat. The differences in the transcriptome expression profiles of selected and non-selected breeds on the basis of differences in their muscle meat quality could be helpful to understand the biological processes underlying the development of meat quality (Damon *et al*., 2012).

Interaction between muscle fibers, perimortem energy metabolism and different environmental factors determine the postmortem transformation of muscles to meat. Dynamic nature of skeletal muscles makes it difficult to categorize them into distinct units. The metabolic response during the pre-slaughter stage is an important factor for post-mortem changes taking place in the various fiber types which affect the meat quality. Selection for leaner pigs and for a higher proportion of large muscle fibers, especially of type IIB, can result in poor capillarisation, consequently an insufficient delivery of oxygen and elimination of end products which later may lead to a reduced pork quality. The literature indicates possibilities to include muscle fiber characteristics in breeding schemes for improved meat quality, while preserving optimal production traits. In order to use muscle fiber traits in a beneficial way for future breeding programmes, further investigations are required for the better understanding of the physiological mechanisms.

### 2.3. Breed characteristics and importance of JNP

Jeju Province (officially the Jeju Special Autonomous Province) is one of the nine provinces of South Korea. The province is located on and co-terminus with the nation's biggest island of Jeju which is also known as Jeju-do. Jeju Island is a volcanic island, predominated by Halla-san (Halla Mountain): which was a volcano of 1,950 meter (6,400 ft) height and still is the highest mountain in South Korea. The Jeju-do measures approximately 73 kilometres (45 mi) across, east to west, and 41 kilometres (25 mi) from north to south. The unique geographical location of Jeju Island put it among the World's 7 natural wonders of the world. Due to its geographical configuration, Jeju Island has its unique flora and fauna. Bovine goods naturally and undoubtedly had to take a back seat on Jeju because landscape of the island is generally not suitable for grazing cows. Moreover, use of cattle in the farming was also not so popular because field tilling wasn't a requirement for tangerine, which is Jeju's main crop. Therefore, mostly beef has been imported and is expensive. Therefore, the self-sufficient Jeju natives wisely turned to a lower-maintenance form of meaty substance and the role of the pig as a source of animal protein gained importance.



The Jeju Native Pig (JNP, *Sus scrofa coreanus*) is a typical pig breed in Korea. The JNP have black glossy hairs. It has big eyes on the caved face along with long straight nose and straight upright ears (Figure 2.1). Overall body structure of JNP is compact with round shoulders and a narrow rear back (Hur *et al*., 2013). It is an indigenous pig that only lives in Jeju. JNP those are available throughout the country seem to have originated from the northern part of China in the Goguryeo Empire (Sodhi *et al*., 2014). It seems that these small pigs were transported to Jeju and settled here as its native species.

They are highly resistant to diseases and quickly adjustable to environmental changes (Cho *et al*., 2011). Among the traditional foods of Jeju, pork from JNP occupies the most important place among the island's specialties. Jeju's black pork is unconditionally as significant and rooted as a staple of Jeju Island. At the same time pork from JNP is very popular among Koreans and is more preferred than the pork from its counterparts because of high tenderness, juiciness and marbling quality (Jin *et al*., 2001; Cho *et al*., 2011). Pork from JNP is exceptional in quality and taste, therefore JNP is usually bred for table use. Moreover, consumers like black coat color than white coat color in pigs, which makes black pork from JNP to be the most delicious variety throughout the Korea.

**Figure 2.1 Adult Jeju Native Pig**



Normally lesser growth rate and lighter carcass weight are the major disadvantages faced by the JNP than commercial pig breeds (Hur *et al*., 2013). Apart from the growth rate and carcass weight, the other drawbacks which are encountered by JNP are lower feed efficiency with smaller litter size (5~8). The lesser productivity and poor genetic traits have led to a drastic decrease in the population of JNP during the early 20$^{th}$ century (Kim *et al*., 2002). Because of its uniqueness and taste characteristics, JNP pork has been regarded as the most expensive and high-quality pork in Korea. Keeping its importance under consideration, the genetic breeding programs for development of black pigs are highly recommended. During the earlier years extensive crossbreeding of JNP with commercially viable foreign breeds' at large scale has been practiced to improve its economic traits (Kim *et al*., 2002). High redness and intramuscular fat content make the meat from JNP to be preferred pork by Korean consumers as compared with commercial pig breeds (Kim *et al*., 2002). As a consequence of the popularity of JNP, many have been crossbred with the commercial foreign breeds for upgrading their economic traits, such as body weight, fat traits, fecundity, and so forth (Kim *et al*., 2002). However, there is still a relatively poor understanding of quantitative and qualitative traits in JNP.

## 2.4. Breed characteristics and importance of Berkshire

Berkshire pigs produce true ‘heritage’ pork. It has the oldest recorded swine pedigree history in the U.S., starting in 1875 and since then its niche pork market is expanding worldwide. Unique taste of the pork from purebred Berkshire pigs gives consumers a superior eating experience. Pork from the Berkshire is dark colored, tastier, tenderer and it has higher marbling score than crossbred merchandised pork (Li *et al*., 2014). During the past several years the Berkshire is contributing significantly to meet the demands of the swine industry (Figure 2.2).


**Figure 2.2** Adult Berkshire pig

Berkshires are very hardy and are adaptable to both outdoor and indoor production systems. The Berkshire sows are very docile and have an excellent disposition. As far as Berkshire boars are concerned, they have high libido and superior quality of semen. An average litter size for purebred Berkshire sows ranges from 2 to 16 piglets. Under optimal health, and managemental programs an average of 9.0 alive piglets and 7.0 to 8.0 weaned piglets per litter is

an attainable target for the purebred Berkshire sow. Earlier studies on meat quality of pork has demonstrated that purebred Berkshires have advantages over most commodity based pork.

The Berkshire breed has superior meat quality features (Suzuki, 2003). Usually, in the retail market the purebred Berkshire is on an average 50% costlier than that of a usual fattening pig. The meat quality features of the Berkshire breed have been reported by the National Pork Producers Council (NPPC; 1995) and Kawaida (1993) in detail. Pork from Berkshire has thin muscle fibers and an excellent water holding capacity. Several studies have enlisted distinct differences in meat quality parameters between genetically different pure breeds and crossbred pigs (Suzuki *et al.*, 1997; Suzuki *et al.*, 2001). Further, it has been observed that the eating quality of meat is strongly affected by the fatty acid composition of adipose tissue (Cameron and Enser, 1991). Similarly, the melting point of fat is reported to influence the texture of pork (Irie *et al.*, 1985).

Berkshire occupies superior place as far as flavor is concerned as compared to rest of the modern pig breeds (Wood *et al*., 2004). Moreover, it has been observed that an exhaustive crossbreeding has been practiced between Berkshire and JNP during 1930s to 1940s. The JNP has been reported to be extensively hybridized with Berkshire breed. Moreover, the black haired Berkshire breed is reported to be involved in the paternal linages of JNP (Kim *et al*., 2011). Since this era pork from the crossbred animals (JNP and Berkshire) has been very popular among the Koreans.



In the case of the JNP, inbreeding to conserve a pure breed has been initiated at the Institute for Livestock Promotion of the Jeju special self-governing province (JSSP). The effort was initiated in 1986, by selecting five JNP seed stocks (4 females and 1 male) on the basis of their phenotype (Annual Research Plan 2009, Institute for Livestock Promotion of JSSP, http://www.jeju.go.kr/contents/). However, pork from the pure bred JNP is very tastier, juicer and has high marbling score but at the same time poor growth performance in purebred JNP creates a gap between the supply and demand. Therefore to access the comparative performance of the pure breed JNP, Berkshire breed was selected for the current study. The present study attempts to identify marker genes determining the meat quality, body growth rate and immune system of the Berkshire breed and JNP.

### 2.5. Description of the technologies for the Expression correlation analysis:

The genomic sequences of a various organisms including that of humans are being elucidated one after another. The genomes of eukaryotic organisms are long, massive and contain an enormous number of genes. It is thus believed that the majority of biological processes found in a variety of organisms can be explained by the quantity of gene products. Functions the gene are ascertained by its protein and the amount of protein produced is directly dependent on the amount of mRNA that encodes it. Therefore, it can be said that cellular functions can be understood by measuring the species and respective numbers of mRNAs at a point of time. Earlier few techniques to identify genes whose expression has been influenced by a variety of internal or external factors were in vogue. These techniques were classical differential colony hybridization of cDNA clones (Yamamoto *et al*., 2001), subtractive hybridization (Kavathas *et al*., 1984; Hubank and Schatz, 1994) and a more recent differential display method (Liang and Pardee, 1992; Welsh *et al*., 1992). The body mapping project was the unique (Okubo *et al*., 1992). Under this project, a rough pattern of gene expression and identification of mRNAs of highly abundant class were performed by sequencing approximately 1000 cDNAs. Hybridization-based analyses for example DNA microarray, immobilized cDNAs (Schena *et al*., 1995) or oligonucleotides (Lockhart *et al*., 1996) were used to examine the expression patterns of a relatively large number of genes. This method could only examine the expressed sequences that were identified. Quantitative realtime PCR (qRT-PCR) (Higuchi *et al*., 1993; Heid *et al*., 1996; Wittwer *et al*., 1997) can provide a valuable standard for validation because of its accuracy and wide dynamic range.



Further, high throughput sequencing techniques using NGS technologies are used for RNA-Seq. Through such techniques, following the preparation of library, the templates can amplify through emulsion PCR (em PCR) and solid phase amplification. The next generation sequencing (NGS) suggests that RNA-Seq is aplomb to supplant microarray-based approaches for transcriptome analysis.

### *2.5.1. Tag based sequencing*

The techniques with high-throughput gene expression like microarrays (Schena *et al*., 1995), massively parallel signature sequencing (MPSS) (Brenner *et al*., 2000), serial analysis of gene expression (SAGE) (Velculescu *et al*., 1995) and Cap Analysis Gene Expression (CAGE) are known as Tag based sequence methods. These give us a snapshot of the RNA concentrations in the cell at a particular time in a specific experimental condition. Tag based methods were

developed to overcome limitations in the original Sanger sequence methods. The tag based sequencing methods have been individually described as follows:

### *2.5.2. Serial Analysis of Gene Expression (SAGE) technology*

SAGE is an experimental technique which was designed to gain a direct and quantitative measure of gene expression. Generations of library of short sequence tags, the quantitative and concurrent analysis of a huge number of transcripts are efficiently performed through SAGE. It works on the principle that the number of times each tag is detected is correlated to the expression level for the corresponding transcript. Basically SAGE works on two principles (Velculescu *et al*., 1995):

- To uniquely identify a transcript, a small nucleotide sequence tag (9 to 10 base pairs) has sufficient information, provided it is isolated from a defined position within the transcript.
- Concatenation of short sequence tags permits the efficient investigation of transcripts in a serial manner by the sequencing of multiple tags within a single clone.



In addition to providing quantitative information on the abundance of known transcripts, SAGE could also be used to identify novel expressed genes from genome-wide expression analysis. SAGE technology permits comprehensive, digital and genome-wide expression analysis. The expression profiles generated through SAGE are sensitive enough to detect low-abundance transcripts and are cable in identifying the novel genes without prior sequence knowledge. Since, the SAGE procedure has been developed and introduced as a tool for the study of gene expression, a variety of biological phenomena has been analyzed. SAGE analysis has been applied to the study the expression profiles derived from cancer and normal tissue of interest. SAGE analyses have also been conducted to study the immunological phenomena (Chen *et al*., 1998), human monocytes and their differentiated descendants, macrophages and dendritic cells (Hashimoto *et al*., 1999).

SAGE methodology has been used since 1995. In spite of its overwhelming potential its use is limited which quietly indicates its intrinsic difficulty in preparing tag libraries. The major disadvantages of the SAGE procedure are the need of relatively high amount of mRNA and difficulty to construct tag libraries (Datson *et al*., 1999). The earlier disadvantage deals with sequencing error, while the later one is about the making of valid tag to gene assignments. Another issue with the SAGE approach is how to further analyze the unknown tags?

### *2.5.3. Massively Parallel Signature Sequencing (MPSS) technology:*

MPSS is one of the latest tools available for conducting in-depth expression profiling. It's an open-ended platform and the gene expression levels are calculated on the basis of the number of exclusive mRNA molecules formed per gene. It captures data by counting virtually all mRNA molecules in a tissue or cell sample. Additionally, MPSS datasets are supplemental in nature, datasets from diverse analyses but with the similar starting mRNA sample can be combined. MPSS has the potential to capture virtually all genes present within the sample and not just those that have been placed on the microarray. It is most relevant to non-human species whose genomes have not been sequenced, it also applies to genes on the human genome that have not been identified and annotated.

Template sequences in MPSS are finalized by detecting adaptor ligations and a signature. The signature is obtained by monitoring a series of ligations on the surface of a micro-bead in a fixed position of a flow cell. As compared to SAGE, MPSS has few notable advantages. The major advantage of MPSS sequencing method is that it has separate cleavage site from its recognition site on the basis of characteristic number of nucleotides (Brenner *et al*., 2000). It can handle millions of templates at a time in just a few reactions without any separate isolation, processing of templates or complex robotic systems. Another significant advantage it has is the bigger library size as compared with SAGE. An MPSS library commonly can control one million signature tags. It means that MPSS can control 20 times more than SAGE library. The immense susceptibility and complete gene expression strongly support MPSS. However, being an improvised technique MPSS still shares few disadvantages with SAGE. The loss of certain transcripts due to lack of the restriction site of the enzyme and uncertainty in the tagged annotation are the common disadvantages shared by both of the applications.



### *2.5.4. Cap Analysis Gene Expression (CAGE) technology:*

CAGE was introduced in 2003, as a method to determine transcription start sites on a genome-wide scale. CAGE determines the transcription start sites by separating and sequencing short sequence tags originating from the 5′ end of RNA transcripts. Mapping of the tags back to the reference genome aids in the identification of the transcription start sites from which the transcripts originated. CAGE tags those were mapped upstream of known genes may be derived from the corresponding full-length mRNAs. CAGE aids in the high-throughput gene expression

analysis and the profiling of transcriptional start points (TSP), including promoter usage analysis (Shiraki *et al*., 2003).

Recently, a new generation of sequencers surpassing at high-throughput sequencing of short tags has facilitated deep CAGE tag sequencing, generating upward of a million tags from a single experimental condition. To conquer the complications of the partial cDNA sequence and augmented full length cDNAs cap-trapping has been used (Carninci *et al*. 2005). The tag counts in such tryouts are significantly much larger and allow an accurate estimate of the cellular concentration of the RNA molecule corresponding to each CAGE tag. In deep CAGE experimentations, clustering of transcription start site is also based on the similarity in expression profiles. Thus deep CAGE helps to characterize the individual promoters from which genes are transcribed on the basis of the promoter activities.

Many of the ambiguous CAGE tags map to a large number of genome locations, suggesting that they originate from repeat regions of the genome. This is consistent with an appreciable fraction of the human transcriptome arising from the repetitive sequences, which may play a role in the transcriptional regulation of gene expression. Although the Cap tag technology has few the advantages over SAGE and MPSS technology yet it also has certain drawbacks. As isoforms are indistinguishable from one another, therefore CAGE may not be able to pinpoint the transcripts that are lowly abundant.



### *2.5.5. Next generation sequencing (NGS) through RNA sequencing (RNA-seq) based technology*

The development of NGS platforms enforces the increased requirements on statistical approaches and bioinformatic tools for the investigation and the execution of the larger data developed by these technologies. NGS has led to an exponential growth in analytical approach during the past decade, primarily vested with the general interest of the researchers in functionally annotating the human genome. Methodologically improved technologies for transcriptome analysis are available and are in vast use since the development of microarray technology and the complete sequencing. Earlier, mRNA expressions were measured by microarray techniques or real-time PCR techniques. Both of the techniques lacked in an exquisite sensitivity and were quite expensive. Therefore, these could not be used for a genome-wide survey of gene expression (Mardis, 2008). However, in contrast to these NGS offers high throughput profiling of gene expression, annotation of genome or discovery of non-coding RNA more rapidly and less expensively. RNAseq efficiently uses deep-sequencing technologies. In

recent years, RNAseq has emerged as the major quantitative transcriptome profiling system (Wang *et al*., 2010). It provides much more decisive measurement of levels of transcripts and their isoforms than other methods.

NGS is very efficient and it can analyze millions of short DNA fragments ranging from 25–450 base pairs during a single sequencing run depending on the type of NGS platform. Although, the reads are much shorter than those created by Sanger sequencing but still NGS has a higher throughput and creates data sets with up to 50 gigabases per run (Nowrousian, 2010). The RNA-seq a new generation sequence technology is acknowledged by the researchers worldwide and its sequential working pattern has been explained in figure 2.3. It has also certain disadvantages. Mostly RNAseq needs large quantity of RNA for short reads that may map too many fields of the genome. It also faces confrontations for the storage of large amount of data. The high cost of RNAseq and the computational complexities are also encountered during the data analysis (Hitzemann, *et al*., 2013).

The comparative study on the hybridization methods and sequencing methods for quantification of transcriptomes has been enlisted in the table: 2.1.

## 2.6. Platforms used for NGS technology



The more demand for lesser-cost sequencing has driven the development of high-throughput sequencing or NGS technologies that parallelize the sequencing process, producing thousands or millions of sequences repeatedly. High-throughput sequencing technologies are intended to lower the cost of DNA sequencing which is possible with standard dye-terminator methods. Following platforms are most frequently used in the NGS:

### *2.6.1. Roche 454 GS FLX Titanium Genome Sequencer*

The GenomeSequencer instrument was introduced in 2005 as the first next-generation system on the market by 454 Life Sciences. It works on the basis of detection of pyrophosphate (Nyren and Lundin, 1985) and a system based on this principle for DNA sequencing was reported by Hyman (1988). In this sequencing system, DNA fragments are ligated to beads by means of specific adapters.

The major limitation with the 454 sequencing platform is that base calling cannot properly interpret long stretches (>6) of the same nucleotide (homopolymer DNA segments).

Therefore because of this reason homopolymer segments are decumbent to base insertions and deletions during base calling whereas substitution errors are rarely encountered in Roche/454 sequence reads (Alkan *et al*., 2009). Currently, the GS FLX Titanium series allows generation of more than 1,000,000 single reads per run with an average read length of 400 bases.

### *2.6.2. Sequencing by Oligonucleotide Ligation and Detection (SOLiD)*

SOLiD is a system by Applied Bioscience. DNA fragments are ligase-modified with adapters, coupled to micro-particles and applied to an emPCR system (Harismendy, 2009). The ABI's SOLiD sequencer is an extensively used sequencing platform and has its origins in the system described by Shendure *et al.* (2005). The ABI SOLiD uses the similar sequencing process as that by the Solexa work flow with minor differences. In SOLiD, each base is investigated twice: first as the right nucleotide and then as the left one of a pair, as the sequencing machine moves along the read. This procedure helps in deriving each subsequent base if we know the previous one. Moreover, if one of the colors in a read is wrongly identified then it will change all of the subsequent letters in the translation. Although it may seem to create problems in read sequencing, still it can be more usefull during the read alignment to a reference genome. The raw 'per-color' error rate is around 2-4%.

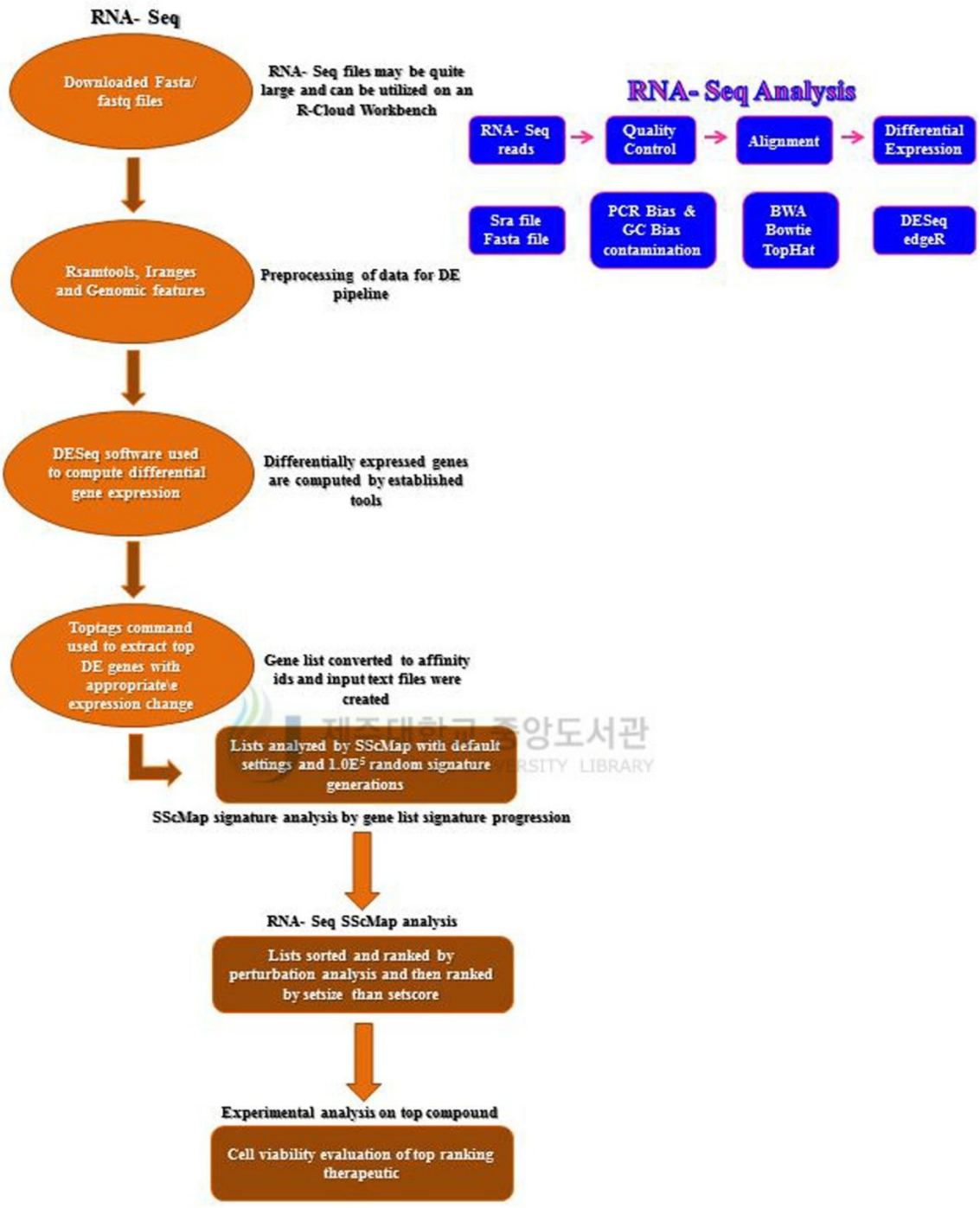


**Figure 2.3** Schematic representation of RNA sequencing analysis**.**

**Table 2.1** The comparative study on the hybridization and sequencing methods for quantification of transcriptome

| Methods | Process | Advantages | Disadvantages |
|---|---|---|---|
| **Microarray** | Non tiling array | • Helps to identify less abundant transcripts.<br>• Relatively inexpensive. | • Do not read the entire genome.<br>• Limited dynamic range of detection due to background and signal saturation,<br>• Require complex normalization for comparing expression levels across the experiments. |
| | Tiling array | • The high throughput through entire genome can help to identify low abundance transcripts, new gene and exons.<br>• It helps in the analysis of alternative splicing, RNA binding protein transcript target identification, polymorphism analysis methylome analysis. | • Limited dynamic range of detection due to lack of techniques for organizing and analyzing the large amount of data<br>• Background and saturation of signals also hinder its range of detection.<br>• Limited number of the probe features for identification of entire genome.<br>• Large numbers of RNA are required for this process. |
| **Tag based sequencing** | SAGE | • Allows the quantitative and simultaneous analysis of a large number of transcripts in any particular cell or tissue, without prior knowledge of the genes. | • Short tags cannot be mapped to the reference genome.<br>• Only a portion of the transcript is analyzed. |

*Continued……*

| | | |
|---|---|---|
| | • The length of the tag is 14 nucleotides.<br>• Rapid and cost-saving sequencing. | • It cannot generate complete datasets for building relational databases. |
| MPSS | • It produces data in a digital format.<br>• The temporal/spatial expression profiles for all genes can be deduced by comparing results from multiple cell types electronically.<br>• The length of the tag is 17 nucleotides.<br>• Produce efficiently a very large datasets of signature sequences. | • Use on sophisticated instrumentation limits its general use.<br>• Express in rare cell types or in response to specific stimuli. |
| CAGE | • Provides digital counts of transcript abundance.<br>• Maps transcriptional start points.<br>• Relies on a cap-trapper system to capture full-length RNAs,<br>• Due to their short size, sequencing of CAGE tags is more efficient for detection of transcription start sites than sequencing full-length cDNAs.<br>• Allows a direct comparison of the expression values of different genes measured in a single deep CAGE experiment. | • Isoforms are indistinguishable from one another.<br>• Ambiguous tags map to multiple genome locations with equal scores. |

*Continued……*

| | | |
|---|---|---|
| **RNA-Seq** | • Without any designed probe this technology provides the expression values.<br>• It can reveal the location of the transcriptional boundaries.<br>• Due to high accuracy and reproducibility it does not need the sophisticated normalization methods.<br>• It can be used for polymorphism analysis. | • Faces difficulty in the development of efficient methods to process, retrieve and store the large amount of data.<br>• Library preparation can bias the outcome of sequencing.<br>• It requires large amount of RNA. |

### *2.6.3. Single Molecule Real Time Sequencing (SMRTS)*

SMRTS was reported by Jett *et al.* (1989). The Heliscope sequencer was the first instrument that was used for sequencing with SMRTS. It works on the principle of cyclic interrogation of a dense array of sequencing parameters with the uniqueness that no clonal multiplication is needed. A very sensitive system to detect fluorescence is used for the investigation of single DNA molecules via sequencing by synthesis.

The main category of error is deletions whereas occurrence of substitution error is comparatively less (0.01–1.0% with one pass) in SMRTS (Magi *et al*., 2010). There has been comparatively small work for developing informatics solutions for the data and this is a very upcoming arena for future algorithm development (Pushkarev and Neff, 2009).

### *2.6.4. Illumina Genome Analyzer*

The Genome analyzer can be used both for traditional DNA sequencing and transcriptome analysis. Interference between DNA polymerase and the used fluorochromes restrict the length of Solexa read which leads to the decreased enzyme activity. It is the most extensively available HTS tool. The bridge PCR is used to generate the amplified sequencing features (Adessi *et al*., 2000) and after confining in the array, all the molecules are sequenced in parallel. While sequencing each and every nucleotide is observed through imaging techniques which are transformed into base calls. These sequencers have capacity to sequence reads up to 100bp. In this plateform the average error rate ranges from 1.0 to 1.5%, but in case of high accuracy levels it reaches up to 0.1%. The most recent Illumina Genome Analyzer IIe has capacity to produce up to 200 million, 100 bp paired-ends reads per run for a total of 20 Gb of data with a throughput of around 2 Gb per day. Comparative analysis of the most commonly used four platforms for NGS technologies has been enlisted in table- 2.2.



## 2.7. Genome wide analysis of tissues

Till now, expression profiling of genes related to the biological pathways existing in different tissues and developmental stages of JNP have not been studied. That's why, through current study, an aim was set to investigate the expression profiles of those genes and most bounteous biological

pathways which caused the variations between JNP and Berkshire. However, the earlier studies have shown that a much larger portion of the transcriptome instability was found in the tissue difference rather than the breed (Zhang *et al*., 2004; Shyamsundar *et al*., 2005; Ferraz *et al*., 2008). Therefore, the current study was planned to understand the capability of DEGs between the two breeds and to know the important molecular mechanisms which are correlated with the biological processes. To know how different tissues achieve specificity is an elemental question in the tissue ontogenesis and its evolution. Genes show differential expressions in different tissues. Few genes have higher expression with respect to their counterparts in a particular tissue and have lesser expression or not expressed at all in other tissues. Such kinds of genes are commonly known as tissue-selective genes. It is a well-established fact that genes are responsible for tissue specific functions and thus can act as the biomarkers for specific biological processes. Many tissue-specific genes have very important role in the pathogenesis of diseases (Lage *et al*., 2008). Many of the genes co-related with the special biological functions and characteristics have the preferred expression in the particular tissues (Goh *et al*., 2007). Moreover, identification of tissue-specific genes also plays an important role in the biomedical research. The demonstration of tissue-specific genes could help scientists to highlight the molecular mechanisms of tissue development and may provide valuable credentials for identification of candidate biomarkers. Therefore, our study may give a new vision towards an understanding and identification of the candidate genes that are involved in the alteration of the biological functions.

**Table 2.2** Comparative study of the widely applied four platforms used in next generation sequencing technologies

| Application | Sequencing methods | Template preparation | Advantages | Disadvantages |
|---|---|---|---|---|
| **Roche/ 454 GS FLX Titanium** | Pseudo Sanger | emulsion PCR (emPCR) | • Fast run time.<br>• Longest short reads among the entire NGS platform. | • The high error rates in homopolymer repeats.<br>• Costly affair. |
| **ABI's SOLiD** | Cleavable probe sequencing by ligation | emPCR | • emPCR can achieve high data densities.<br>• Two base encoding provides inherent error correction and makes this technology a highly accurate. | • Underrepresentation of AT rich and GC reach region.<br>• The DNA library preparation procedures are time consuming.<br>• The emPCR can be cumbersome and challenging. |
| **Helicos BioSciences Heliscope** | One color cyclic reversible termination | Single molecular | • It can directly sequence single DNA molecules | High error rate |



*Continued……*

| | | | | |
|---|---|---|---|---|
| **Illumina/Hiseq 2000** | Four color cyclic reversible termination | Solid phase | • Currently the most widely used platform.<br>• Most adaptable easiest technology to use for sequencing platform | • Substitutions are the most common error type.<br>• Underrepresentation of high rich AT and GC regions. |

## 2.8. Elucidation of fat tissue for genetic basis of gross performance of genes

An increase in the demand of pork has been observed during the last decade and pork has been considered as one of the economical source of animal protein for humans. The increased demand and changes in choice of the human have led to the new trends for meat industry. During the last ten years quality of pork has marched forward than the quantity. The quality of pork and intra muscular fat content are affected by several factors. Intramuscular fat content has been reported to significantly influence the meat tenderness, juiciness and taste (Verbeke *et al*., 1999). Further, it has been also reported that the quality of pork is significantly affected by its fat and fatty acid contents. Further, it has been noticed that the different melting points and variation in fatty acid contents affect the firmness or softness of the fat in pork. Such kinds of variations affect the quality of meat, especially the subcutaneous, carcass fats and the marbling (Wood *et al*., 2004). Therefore, it has been observed that the fat tissue plays the major role for meat quality.

Therefore, it was thought that the analysis of fat tissue for differential gene expression would be useful in the understanding of the biological processes. The large number of genes of fat tissue would have high significance for the sets of genes conferring the phenotypes. The differential expressions of genes have been reported in the previous research for lipid biosynthesis, hormone metabolism, cell-cell signaling, fat-soluble vitamin in the fat tissue and limb development (Rehfeldt and Kuhn, 2006). Further it has been reported earlier that the genes related to body growth, development of skeletal systems and cartilage have significant expression in the fat tissue (Kim *et al*., 2009).

Further the, transcriptome analysis of fat tissue could provide the significant ways of disclosing differences in gene expression since the regulation of activity of genes occurs primarily at the transcription level (Venkatesh *et al*., 2005). These annotations of DEGs would be beneficial for the selection of genes to be used as the selection markers. Growth and cell/tissue differentiation are affected by the lipids whereas their metabolites function as transcriptional regulators of gene expression. The genes involved in ketone, amine and lipid metabolism have significant expression in the fat tissue e.g. amino acid sequence of *PDK4* acknowledged a G/A mutation in its intron 9 and analysis for the association authenticated that it is significantly associated with intramuscular fat and muscle water content of the meat (Lan *et al*., 2009). These reports highlighted the intramuscular fat content in specific pig breeds. Similarly, *FMO3* is reported to be related with the fishy off flavor and pork pH (Glenn *et al*., 2007). It affects the

metabolism of trimethyl amines to trimethyl amine oxide which induces the production of oxidized fishy odor in pork (Zhou and Shephard, 2006).

The interactions among the gene expression, metabolites, nutrients and modifications of protein are important for the organization of cell growth with extracellular and intracellular conditions (Yuan *et al*., 2013). Analysis of adipose tissues by the earlier studies has interpreted the significant enrichment of GO categories associated to immune response, stress, lipid metabolism and carboxylic acid metabolic processes (Pashaj *et al*., 2013). Further, a study on the three Chinese pig strains involving co-transfection with *DRA* and *DRB* genes confirmed that DR molecules could stimulate human T- cells (Chen *et al*., 2012). Moreover, Src-like-adaptor (*SLA*) identified the pig fat as an important tool for the study of immune responses, disease resistance and production traits (Ho *et al*., 2006). At the same time, proteoglycan- 4 (*PRG4*) is reported to have an important role in transformation of the functional superficial zone cartilage, full thickness cartilage with zonal organization and cellular phenotype (Coates *et al*., 2012).

Phosphorylation has an important role in the governance of protein binding and coordination of pathways by accommodating the nature and strength of protein- protein interactions. Therefore, the present study targeted the identification of the DEGs existing in fat; those are responsible for protein binding and phosphorylation. In the earlier studies *PPL* has been reported as a cytoskeletal protein. It provides structural integrity within the skin and is reported to aid in the maintenance of water impermeability and resiliency of the skin (Seo *et al*., 2012). Another gene T- box transcription factor-5 (*TBX5*) is reported to be involved in dorsal eye patterning and limb regeneration (Jassar *et al*., 2013). Further, interleukin- 1 receptor accessory protein (*IL1RAP*) is reported to express significantly in the porcine endometrium and play an important role during peri- implantation period of pregnancy (Sousounis *et al*., 2013). Such findings were used as base to find out the genetic control of placental efficiency and enhance the perception of placental evolution in different breeds (Zhou *et al*., 2009).

It has been noticed that oxidation of proteins led into the biochemical changes in meat which expedites to change in color of meat, offensive flavor and rancidity in meat (Bekhit, 2013); thereby directly affect the consumer's choice. Similarly, it has been observed that the loss of functional mutations in the genes from flavin containing monooxygenase (*FMO*) family is strongly linked with fishy off flavor (Glenn *et al*., 2007). Therefore, wavering expression of *FMO* gene in different pig breeds may mark their role in the quality and shelf life of meat. Therefore, these genes can be influential markers for the studies related to meat quality of pork. Therefore,

the differential expressions of the genes in the fat those are involved in different biological processes have been addressed while outlining the current study.

### 2.9. Elucidation of Muscle and Liver tissues for genetic basis of gross performance of genes

Liver is the main site for the fatty acid synthesis (Corino *et al*., 2002) and it has a significant role in body growth (Mourot *et al*., 1995). Liver is considered as the tissue of choice to analyze the growth, quality of pork, metabolism related processes and immunity. Similarly, composition of muscle, cross-section of fiber and the capillary density of a particular muscle are the important factors which affect various peri-mortal and post-mortal biochemical changes affecting the meat quality and body growth (Malek *et al*., 2001). Therefore, along with the fat in the present study, the *longissimus dorsi* muscle and liver tissues to identify DEGs responsible for growth and meat quality in the JNP were chosen. The genes expressing in the muscle tissue have been ascertained to show extensive relationship enrichments with the genes of liver in single organism biosynthesis and small molecule biosynthesis process.

Studies conducted by earlier researchers have earmarked Acetyl-CoA Acyltransferase-2 *(ACAA2),* Apolipoprotein- A4 *(APOA4),* Cytochrome- P2E1 *(CYP2E1)* and Hydroxy methyl glutaryl CoA synthetase 2 *(HMGCS2),* genes which are linked to growth hormones, hepatic lipid and high density lipids (HDL) (Daniels, 2009; Uddin *et al*., 2011; Graugnard *et al*., 2012; Morlein *et al*., 2012). *CYP2E1* is reported to be linked with the metabolism of endogenous compounds including fatty acids and ketone bodies (Novak and Woodcroft, 2000). Moreover, the interplay due to metabolites, nutrients, protein modifications and gene expressions are involved in the coordination of cell growth with extracellular and intracellular conditions (Yuan *et al*., 2013). Similarly, Lan *et al*. (2009) reported that Pyruvate dehydrogenase kinase-4 (*PDK4),* which expresses in skeletal muscles, is linked with meat quality. A little earlier studies reported that Periostein *(POSTN)* and myosin binding protein- H (*MyBPH),* control fat deposition (Chen *et al*., 2011) and muscle growth (Bilek *et al*., 2008) respectively.

*MyBPH* has been earmarked as a significant gene by the few earlier studies. *MyBPH* is related to myosin in muscle. Actin and myosin are the main protein of Myofibrils (Zhang, 2009). Myosin exists in the thick filament and type of muscle fibers is affected by it. The type of muscle fiber is known to be responsible for water content, color and eating quality of pork. Further, one another gene collagen type 2- alpha 1 (*COL2A1)* is related to the skeletal system morphogenesis and growth of the cartilage (Lefebvre and Smits, 2005; Kim *et al*., 2007). Recently, a study

conducted by Pashaj *et al*. (2013) reported main enhancement of GO categories related to lipid metabolism, carboxylic acid metabolic processes, immune response and stress in liver.

Further, Adler *et al*. (2013) reported that the phenotypes of growth performance and immune response are significantly dependent upon the functional transcript changes throughout in the immune response. Moreover, the selection for high immune response is reported to be associated with the enhanced weight gain (Mallard *et al*., 1998; Clapperton *et al*., 2006). Lateral to it phosphorylation is stated to play a vital role in regulation of protein binding. It relates the pathways by adjusting the nature and strength of protein- protein interactions. Therefore, in the current study we proposed to identify the genes from the samples of Berkshire and JNP, those are liable for protein binding and phosphorylation.

*FMO3* is known to play a significant role in human hepatic metabolism. *FMO5* is stated to be associated with fishy off flavor and pork pH (Glenn *et al*., 2007). Findings on the knockout mice also revealed the vital role of *FMO5* in regulating the endogenous fat and cholesterol metabolism. The genome wide analysis in mouse stated that phosphoenolpyruvate carboxykinase-1 (*PCK1*) encodes the rate- limiting enzyme in hepatic gluconeogenesis (Shin *et al*., 2012).

Based on the earlier findings, in this study, we aimed to investigate most abundant biological pathways and the expression profiles of those genes which caused the variations between JNP and Berkshire. From the leading available resources, the current study is the pioneer transcriptional analysis for the detection of DEGs from RNA-seq data generated from fat, *longissimus dorsi* muscle and liver tissue samples. Transcriptome analysis by RNAseq has been already conducted in different breeds of pig except JNP (Chen *et al*., 2011; Petkov *et al*., 2011; Rustemeyer *et al*., 2011; Jung *et al*., 2012; Looft, 2013; Prather, 2013; Samborski *et al*., 2013). The basic procedure followed during transcriptomic analysis by RNA-seq has been represented in the figure 2.4. Thus, our work proposes a new direction towards an understanding and identification of the candidate genes that are involved in the alteration of the biological functions. The annotation findings by our study may act as the selection marker for the genes related to the body growth.



## 2.10. *In-vitro* model to illustrate the muscle development:

### *2.10.1. Selection of C2C12 cell line*

Mouse myoblast C2C12 cells are the most commonly used cell line to study differentiation of muscles. The thigh muscles of mice have been the source for C2C12 cell line

(Yaffe and Saxel, 1977). They develop as myoblasts and have powerfull expression of myoblast determination protein (MyoD). C2C12 cells are cultured in the mitogen-rich culture medium with 10% to 20% fetal bovine serum (FBS) which is known as growth medium. While growing in the growth media they do not potentiate downstream genes which are linked with the muscle differentiation. When in differentiation media, C2C12 successfully differentiate into skeletal muscle cells and potentiate the expression of muscle genes. The differentiation of C2C12 cells can be hurriedly persuaded to produce contractile myotube forming muscle proteins by simply replacing the growth media with differentiation media at 50%-70% confluence. To study the *in-vitro* model for growth, differentiation, migration and death, C2C12 myoblast cells are repeatedly used on the basis of their following advantages:

i. They are commercially available.
ii. C2C12 cells grow at very fast rate. After seeding in a normal culture plate, they achieve approximately 90% confluence level in two to three days.
iii. Their ease to differentiate into multi-nucleated myofibres.
iv. C2C12 cells can be successfully used for immunostainings, transient transfections with expression vectors or siRNA oligos and stable transfections with bacterial artificial chromosome (BAC) constructs are simple to perform in these cells.



### *2.10.2. Selection of PiggyBac vector system*

The genetic amendments by PiggyBac promote scientists/researchers to:

• reshape the genomes of varied livestock species alongwith a simple transfection
• handle highly conclusive and cost-effective non-viral gene delivery
• make adjustments by reversing the genome with footprint-free transposon removal

DNA transposon which is known as “sleeping beauty” is considered as the avant-grade DNA transposons to be used in mammalian cells (Izsvak *et al*., 2010). PiggyBac has been treated as a propitious alternative to sleeping beauty. The PiggyBac DNA transposon technology has been practiced in different research field e.g. regenerative medicine, cell line engineering, gene therapy and particularly for the establishment of animal models. PiggyBac, a moving genetic element works on the principle of "cut and paste" mechanism. PiggyBac conveniently transpose between vectors and chromosomes. During transposition, PiggyBac conveniently with ease identifies transposon-specific inverted terminal repeat sequences (ITRs). It conveniently merges the contents into TTAA chromosomal sites. PiggyBac vector system efficiently mobilizes the gene of interest into the target genome.

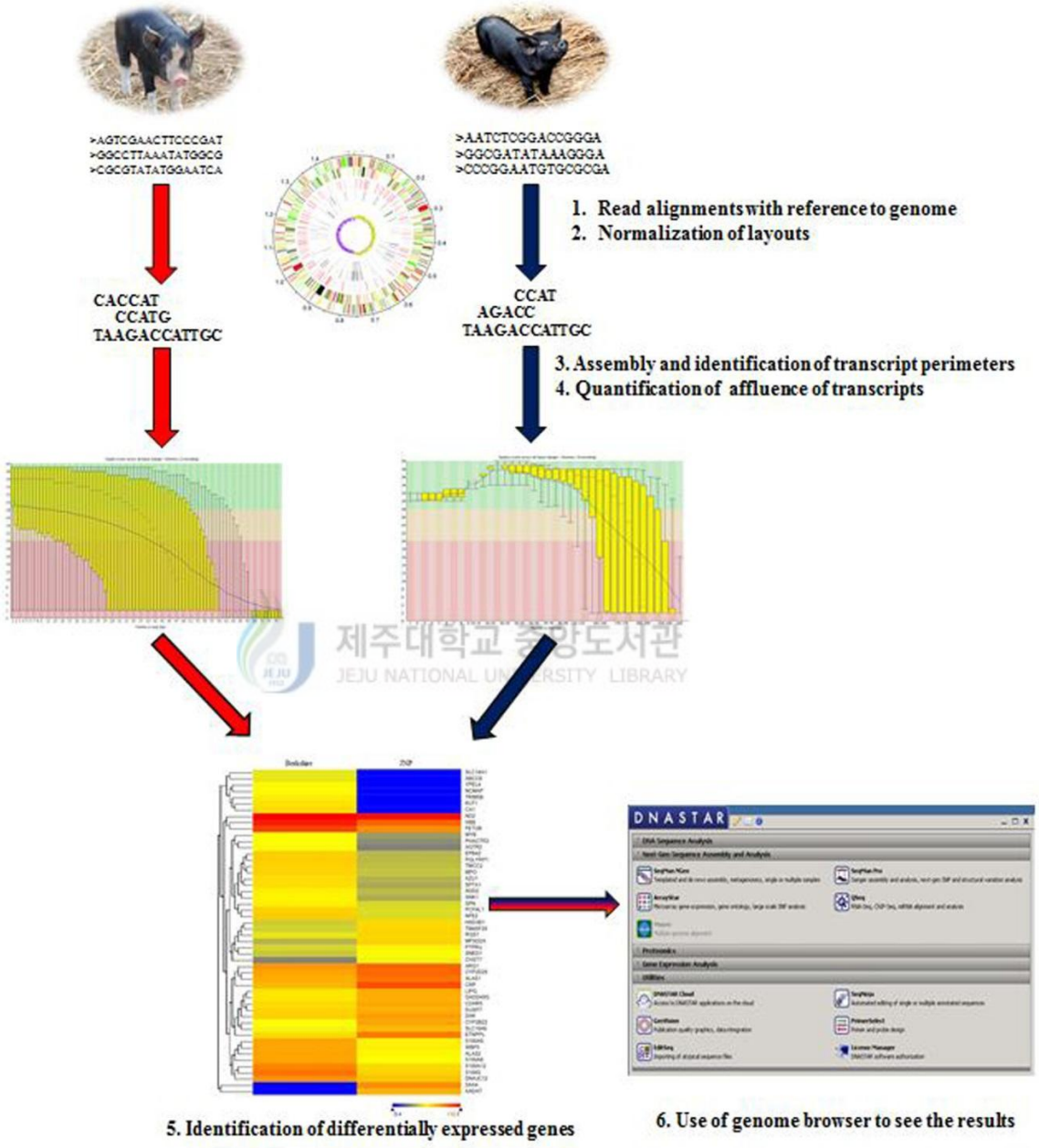


**Figure 2.4** Flow chart of the procedure followed during transcriptomic analysis by RNA-seq .

The unique features of PiggyBac transposons are that there is no cargo limit and it is also reversible **(**Li *et al*., 2011**).** Ding *et al*. (2005) demonstrated its transposition along with cargos without any loss of its efficiency. PiggyBac is extremely functional in assorted cell types, including mammalian cells (Schibli *et al*., 2002; Li *et al*., 2013). PiggyBac comfortably intercedes the *in-vivo* long-term enunciation in mammalian cells (Schibli *et al*., 2002). Genomes having an inserted PiggyBac vector can be transiently re-transfected with the PB tranposase expression vector. Without leaving any marks, the PiggyBac transosase can remove the transposons from the genome.

So, in the present study an attempt has been made to identify the differential expression of genes using molecular and transcriptomic analysis methods.

# Chapter 3

## Comparative transcriptomic analysis to identify differentially expressed genes in fat tissue of adult Berkshire and Jeju Native Pig using RNA-seq

### 3.1. Abstract

Pork is a major source of animal protein for humans. The subcutaneous, intermuscular and the intramuscular fat (IMF) are the factors responsible for meat quality. RNA-sequencing (RNA-seq) is rapidly adopted for the profiling of the transcriptomes in the studies related to gene regulation. The discovery of differentially expressed genes (DEGs) between adult Jeju Native Pig (JNP) and Berkshire breeds are of particular interest for the current study. RNA-seq was used to investigate the transcriptome profiling in the fat tissue. Sequence read numbers were obtained from Ilumina HiSeq2000 and mapped to the pig genome using Tophat2. Total 153 DEGs were identified and among the 71 annotated genes have BLAST matches in the non- redundant database. Metabolic, immune response and protein binding are enriched pathways in the fat tissue. In our study, biological adhesion, cellular, developmental and multicellular organismal processes in fat were up-regulated in JNP as compare to Berkshire. Multicellular organismal process, developmental process, embryonic morphogenesis and skeletal system development were the most significantly enriched terms in fat of JNP and Berkshire breeds ($p$ = 1.17E-04, 0.044, 3.47E-04 and 4.48E-04 respectively). Collagen type X, alpha 1 (*COL10A1*), Collagen type XI, alpha 2 (*COL11A2*), *PDK4* and Patatin like phospholipase domain containing- 3 (*PNPLA3*) genes are responsible for skeletal system morphogenesis and body growth were down regulated in JNP. This study is the first transcriptional analysis for the detection of DEGs from RNA-seq data generated from fat tissue sample. This analysis can be used as stepping stone to understand the difference in the genetic mechanisms that might influence the identification of novel transcripts, sequence polymorphisms, isoforms and noncoding RNAs.

***Keywords*:** RNA-seq, Jeju Native Pig (JNP), Berkshire, Fat, Differentially Expressed Genes (DEGs)

### 3.2. Introduction

Pig meat is an integral and major source of animal protein for humans. The change in consumer's choice has resulted into the new scenarios for world meat market. During the current years quality of pork has attracted much more importance than the quantity and the aspect of lean meat is also a topic of concern. Previously, it has been reported that the quality of pork has declined over the years (Wood, 2001). It is a well reported fact that the quality of pork and its relation with intra muscular fat is influenced by several factors. Verbeke (1999) reported that intramuscular fat content is one of the important traits which influence the meat tenderness, juiciness and taste. Fat and variation in fatty acid contents affects meat quality. Due to their different melting points, variation in fatty acid contents effects the firmness or softness of the fat in meat and thereby affect meat quality, especially the subcutaneous, intermuscular (carcass fats) and the intramuscular fat (IMF)content (marbling) (Wood *et al*., 2004). Therefore, fat tissue is a one of the major factor responsible for meat quality. Jeju Island has distinct flora and fauna. Historically it is reported that JNP was imported in 12th century from main China. Among the meat producing animals, JNP has adapted very well to the unique environmental niche of Jeju Island. Koreans have preferred taste for the meat from JNP as compared to other commercial breeds because of its higher tenderness, juiciness and marbling quality (Cho *et al*., 2011). Because of these qualities meat from JNP is always in high demand but the lower feed efficiency and smaller litter size (5~8) are major drawbacks of JNP which create large gap between demand and supply.



Next Generation Sequencing (NGS) is the preferred approach for characterization and quantification of the entire genome. For the serial and cap analysis of gene expression, NGS is more effective than the previous technologies (Mortazavi, 2008). The RNA-seq allows genome-wide analysis of transcription at single nucleotide resolution, including identification of alternative splicing events and post-transcriptional RNA editing events. The Illumina HiSeq, Ion Torrent and SOLiD systems are the platforms of NGS, which are most frequently used for RNA-seq (Cloonan *et al.,* 2008; Eveland *et al.,* 2008; Marioni *et al.,* 2008). Dynamic ability of RNA-seq to discriminate homologous regions and identification of regions of expression without the prior *assumption*, makes it as the future of transcriptome research (Shendure and Ji, 2008).

Genes and different markers between pig breeds have been identified using microarray, qRT-PCR and microsatellite analyses (Kim *et al.,* 2005; Kim *et al.,* 2008; Moon *et al.,* 2009). Developments in RNA-seq technology enable more comprehensive investigation of the transcriptome than microarrays and are preferred for gene expression studies (Mortazavi, 2008).

Transcriptome analysis using RNA seq has been applied in different pig breeds except JNP (Chen *et al.*, 2011; Petkov *et al.*, 2011; Rustemeyer *et al.*, 2011; Jung *et al.*, 2012; Looft, 2013; Prather, 2013; Samborski *et al.*, 2013) and very less comparative studies among the different pig breeds have been conducted for biological processes. The statistical analysis is also critical in transcriptomic studies using RNA-seq; specifically for the normalization of quantitative measurements of expression (Wilhelm and Landry, 2009; Li *et al.*, 2010; Robinson and Oshlack, 2010; Hong *et al.*, 2012) and detection of DEGs (Hardcastle and Kelly, 2010; Robinson *et al.*, 2010; Wang *et al.*, 2010; Tarazona *et al.*, 2011; Trapnell *et al.*, 2012). Moreover because of extensive crossbreeding between Berkshire and JNP during 1930s to 1940s, the JNP has been widely hybridized with Berkshire breed (Kim *et al.*, 2011). Therefore, keeping all these above mentioned points in mind the current study was planned on the fat tissue harvested from adult JNP and Berkshire pigs to perform transcriptome analysis using RNA-seq. To best of our knowledge, this study is the first statistical analysis for the detection of DEGs from RNA-seq data generated from fat tissue sample.

## 3.3. Materials and methods



### *3.3.1. Animals and sample preparation*

For the current study, 5 female piglets each from JNP and Berkshire breeds were used. They were housed in similar environmental and nutritional conditions. Animals were slaughtered at 5 months of age, according to the standard protocols of Jeju National University. The fat samples were collected immediately after the slaughter at the 5th lumber vertebra level. The samples were stored immediately in dry ice and later were stored at -80℃ until used for RNA extraction. The research proposal and the relevant experimental procedures were approved by the institutional review board of the Department of Animal Biotechnology, Jeju National University.

### *3.3.2. Extraction and analysis of quality of RNA*

RNA was isolated from 100 mg of the fragmented frozen fat tissue samples from 5 animals each from JNP and Berkshire breeds. TRIzol™ (Invitrogen, USA) reagent was used for the isolation of RNA. Tissue samples were homogenized in 1.5 ml of TRIzol reagent and chloroform, which were subsequently precipitated by using isopropanol (Junsei Chemical Co.

Ltd., Japan). Isolated RNA samples were stored at -80℃. To purify RNA from genomic DNA contamination, 25 μg of RNA from each sample was treated with the RNase-free DNase set (QIAGEN, Hilden, Germany) and it was purified with the RNeasy mini kit according to the user guidelines (QIAGEN, Hilden, Germany). A Bioanalyzer 2100 with RNA 6000 Nano Labchips was used to assess the quality and quantity of RNA by automated capillary gel electrophoresis by following user guidelines (Agilent Technologies Ireland, Dublin, Ireland). 28S/18S ratios for the RNA samples ranged from 1.8 to 2.0 and integrity of RNA had ranged from 8.0 to 10.0.

### *3.3.3. Preparation and sequencing of mRNA-seq ilumina library*

0.1-4.0 μg of total RNA has been taken from the each sample for the construction and sequencing of the total RNA-seq library. Five samples per breed have been used for NGS sequencing. The TruSeq RNA Sample Pre Kit was used according to the manufacturer's guidelines. Agilent Technologies Human UHR total RNA has been used as a positive control sample. The library was constructed according to a standard protocol provided by Ilumina, Inc. Libraries with different indexes were pooled together and were sequenced in one lane using an Ilumina HiSeq 2000 high-throughput sequencing instrument with 100 paired-end (PE) sequencing.



### *3.3.4. Read filtering and merging*

The low quality PE reads were removed before further analysis by Consensus Assessment of Sequence and Variation v 1.7 (CASAVA-v1.7) with default parameters. All quality-filtered PE reads were aligned to Sus scrofa genome (Sscrofa10.2) from the University of California Santa Cruz (UCSC) using Tophat2 (v2.0.2) and reads were counted using HTseq (v 0.5.3p3). The remaining reads were later merged on the basis of the overlap of PE reads. To analyze, correlation between adult JNP and Berkshire, the corrplot (Friendly, 2002) and DEGseq (Hardcastle and Kelly, 2010) R packages have been used. Fisher's exact test model was used to statistically identify DEGs (Bloom et al., 2009). Fat samples from JNP and Berkshire were compared by DEGseq for detection of DEGs. Significant DEGs were selected at FDR <0.01.

### *3.3.5. Functional annotation of DEGs and the pathway analysis of the DEGs*

Bovine Ensembl gene IDs were converted to official gene symbols after cross-matching with the gene IDs and official gene symbols of human Ensembl. The official gene symbols of

human homologs for bovine genes were used in functional clustering and enrichment analyses using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Dennis *et al*., 2003). The representations of functional groups in fat relative to the whole genome were investigated using the Expression Analysis Systematic Explorer (EASE) tool (Hosack *et al*., 2003) of DAVID. EASE has been used to measure the enrichment of gene ontology (GO) terms (Alterovitz and Ramoni, 2010). To identify enriched GO terms, functionally clustered genes were filtered according to EASE values <0.1.

The primary object of ToppCluster is the identification of biological themes in data sets involving numerous sets of genes, regulatory networks and systems biology-based dissection of biological states. The shared and specific features associated with any number of genes involved in biological processes have been visualized. ToppCluster with a 0.05 cutoff P-value was selected and no false discovery correction method was applied.

Further, pathway analysis for the DEGs was carried out by KEGG (Kyoto Encyclopedia of Genes and Genomes) tool. KEGG was used to understand the functional meanings of higher-order and utilities of the cell or the organism from their genome information. It is an integrated resource with associated software consisting of three types of databases for genomic, chemical and network information.



***3.3.6. Statistical analysis of the levels of differentially expressed genes***

The raw data was analyzed by adopting general linear model (GLM). DEGs with significant difference in their expression at $P$<0.05 and FDR< 0.001 were selected. Further, analysis was designed to understand the statistical value of the expression levels among the genes between the Berkshire and JNP. Data were analyzed using the ArrayStar V4.1 (DNASTAR, Madison, WI) software for generation of the heat maps of selected genes of interest. The level of gene expression for each gene was measured using reads per kilobase per million (RPKM) (Wagner *et al.*, 2012). It has been calculated by the following equation:

$$\mathrm{RPKM}_g = \frac{r_g \times 10^9}{fl_g \times R} \qquad \text{Eq. 1}$$

Where, $r_g$ represent the number of the reads mapped for the particular gene region, $fl_g$ is feature length and R is the total number of reads obtained from the sequencing run of the respective sample.

## 3.4 Results and discussion

### *3.4.1. Quality analysis of RNA-sequence reads*

Transcriptome data was generated using Ilumina HiSeq2000 to investigate DEGs in fat tissue of both breeds. Highly reliable sequence reads of fat from both the breeds were acquired through RNA-seq. A few sequence reads (<5.0E-0.5%) did not pass the filtration process with respect to their quality. On an average, 90.7% sequence reads from both of the breeds passed the quality control step (Table 3.1) and were mapped successfully to the *Sus scrofa* genome using TopHat (v2.0.2). DEGs between both breeds, with a focus on skeletal system morphogenesis, body growth and collagen fibril organization, metabolic process, ion and protein binding, phosphorylation, inflammatory and innate immunity, oxidation-reduction process and metal ion binding were identified in the present study. GO of all of DEGs along with up-regulated DEGs in the fat tissue of JNP and Berkshire breeds (Figure 3.1 a, b, c) have been summarized. These DEGs will be helpful while planning for the further transcriptomic studies in pig.



**Table 3.1** RNA-seq reads and mapping rate of fat tissue from Berkshire and Jeju Native Pig

| Breeds | No. of | | |
|---|---|---|---|
| | total reads | reads after QC | accepted Hit |
| JNP | 41413088 | 41411459 | 37879239 (91.5%) |
| Berkshire | 40310004 | 40308452 | 36254763(89.9%) |

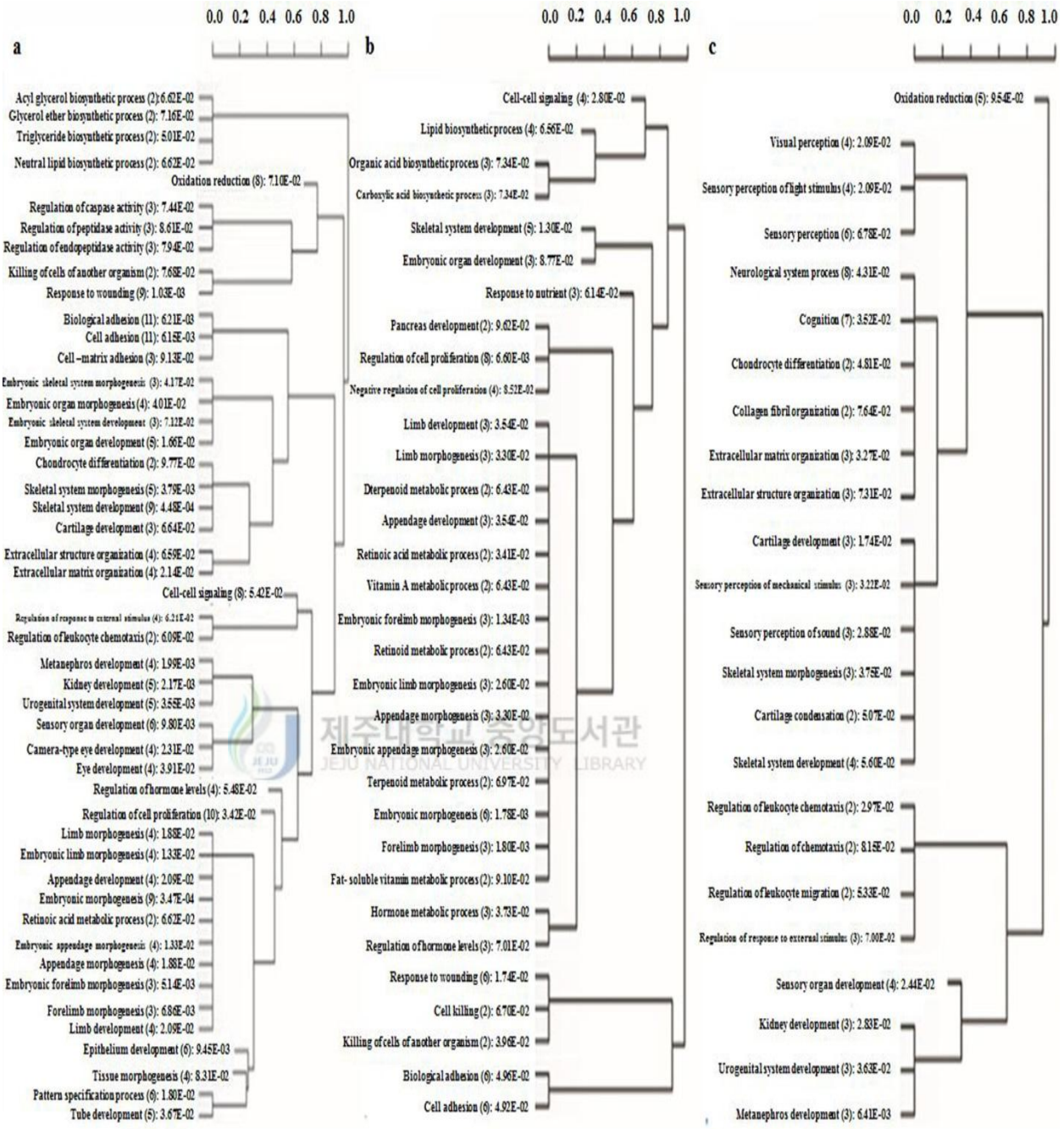


**Figure 3.1** Enriched biological process (a) GO of DEGs expressed in fat tissue. Go of up-regulated DEGs in fat tissue of, (b) JNP (c) Berkshire

### *3.4.2. Identification and analysis of DEGs*

The analysis of fat tissue for differential gene expression was conducted on the basis of raw data. The GLM was used to identify the genes with a significant difference in their expression in the fat tissue of the two breeds. DEGs for the current study were selected on the basis of the *P* value of <0.05 (Table 3.2). 153 DEGs were identified and out of them 96 DEGs were annotated in the study (Figure 3.2) between JNP and Berkshire. Out of the all annotated genes, 71 genes have BLAST matches in the non- redundant database. The rest of the genes either matched unnamed or hypothetical proteins or had no BLAST matches in the non-redundant database. Open source online visualization of these data is available at http://biopopdb.snu.ac.kr/PIG_DEG/ and the matrix prepared in the web can be used to find out the DEGs among the two breeds at different age groups.

### *3.4.3. Analysis of the features of multiple gene lists by clustering and annotation of the pathways for the DEGs of fat*

Results of ToppCluster can be viewed by the graphical network. The network between the genes under study is displayed through the 'abstracted' option excluding the genes from the network and 'gene level' option helps to generate the entire network (Kaimal *et al.*, 2010). To provide a dissected gene-level view, we chose the GO of the biological process which is shared by the genes expressing in the fat (Figure 3.3). From the abstracted network analysis, the distinct functional separation was observed. Notably, the extensive gene list of fat tissue has shown high significance for the sets of genes conferring the phenotypes. Our result is in line with an earlier study and may be quite useful especially when an investigator wants to further explore only some of the enriched terms in the output (Kaimal *et al.*, 2010).

The GO pathways were analyzed in details through the KEGG and have been listed in Table 3.3. It is found that most of the DEGs are involved in metabolism, immune responses and protein binding. A significant up-regulation of genes has been observed for cell-cell signaling, lipid biosynthesis, limb development, hormone metabolism and fat-soluble vitamin in JNP fat (Rehfeldt and Kuhn, 2006; Kim *et al.*, 2009). In contrast to it, development of cartilage and skeletal system is up-regulated in the fat tissue of Berkshire pigs and is associated with body growth. This investigation provides an opportunity to elucidate the efficiency of DEGs between two breeds and helps to understand regarding the key molecular mechanisms associated with the biological process.

**Table 3.2** List of DEGs in the fat tissue of Berkshire and JNP (FDR<0.001)

| Gene | Function | *P*-value | FDR |
|---|---|---|---|
| COL11A2 | Skeletal system morphogenesis, collagen fibril organization | 1.88E-11 | 2.78E-08 |
| COL11A1 | collagen fibril organization, skeletal system morphogenesis | 9.00E-22 | 3.98E-18 |
| PLCH2 | lipid metabolic process | 1.10E-10 | 1.22E-07 |
| PDK4 | regulation of cellular ketone metabolic process, reactive oxygen species metabolic process | 5.99E-12 | 1.06E-08 |
| MCTP2 | protein binding | 1.87E-08 | 1.14E-05 |
| WT1 | tissue development, germ cell development | 7.69E-05 | 0.014322 |
| OTOR | cartilage condensation | 0.000224 | 0.031019 |
| IGLC | protein binding | 4.63E-13 | 9.10E-10 |
| MT1A | metal ion binding | 0.000101 | 0.016842 |
| XPNPEP2 | Proteolysis | 0.000134 | 0.020574 |
| RDH16 | metabolic process, oxidation-reduction process | 3.34E-05 | 0.007209 |
| WDR93 | protein binding | 3.60E-06 | 0.00111 |
| PCK1 | lipid metabolic process, glucose metabolic process | 1.91E-10 | 1.78E-07 |
| SLC25A34 | transport | 9.34E-10 | 7.51E-07 |
| CCL4 | inflammatory response | 0.00022 | 0.030691 |
| IL1RAP | protein binding | 3.32E-07 | 0.000157 |
| KCNK5 | ion transport | 0.000124 | 0.019609 |
| AOC2 | oxidation-reduction process, amine metabolic process | 0.000241 | 0.032486 |
| GREM1 | collagen fibril organization | 0.000142 | 0.021426 |
| LCN2 | innate immune response | 9.63E-07 | 0.000341 |
| RGS3 | termination of G-protein coupled receptor signaling pathway | 7.19E-07 | 0.000276 |
| CXCL2 | immune response | 7.54E-07 | 0.000284 |
| CR2 | (complement receptor mediated signaling pathway) | 6.10E-05 | 0.01227 |
| AIFM3 | oxidation-reduction process | 0.000108 | 0.017685 |
| SIK1 | regulation of cell differentiation | 0.000111 | 0.018018 |
| ACE2 | proteolysis | 7.00E-05 | 0.01318 |
| ADAMTS4 | proteolysis | 4.24E-06 | 0.001249 |
| IGKC | protein binding | 0.000192 | 0.026983 |
| ALDH1L1 | metabolic process, oxidation-reduction process | 1.27E-05 | 0.003173 |

*Continued……*

| | | | |
|---|---|---|---|
| NOTCH4 | Notch signaling pathway, cell differentiation | 1.27E-05 | 0.003173 |
| SLC2A3 | glucose transport | 2.32E-05 | 0.005392 |
| ITGA8 | smooth muscle tissue development | 0.000119 | 0.018964 |
| ACER2 | ceramide metabolic process | 6.91E-05 | 0.013145 |
| FMO1 | oxidation-reduction process | 0.000372 | 0.045734 |
| FMO2 | oxidation-reduction process, organic acid metabolic process | 0.000189 | 0.026983 |
| SLA-DRB | immune response | 0.000387 | 0.046409 |
| ADAMTS1 | proteolysis | 0.000246 | 0.033 |
| SLA-5 | immune response | 0.000377 | 0.045949 |
| NR4A1 | skeletal muscle cell differentiation | 0.000388 | 0.046409 |
| GFPT2 | metabolic process carbohydrate metabolic process | 0.000404 | 0.04791 |
| DDR2 | biomineral tissue development | 0.000339 | 0.04256 |
| APOD | lipid metabolic process glucose metabolic process | 0.000344 | 0.042891 |
| PLAU | proteolysis | 0.000261 | 0.034398 |
| C4A | immunoglobulin mediated immune response | 0.000411 | 0.04791 |
| EPDR1 | calcium ion binding | 0.000303 | 0.038855 |
| LGALS3 | protein binding | 9.87E-05 | 0.016751 |
| PNPLA3 | metabolic process, lipid metabolic process | 0.000386 | 0.046409 |
| PSPH | metabolic process | 6.36E-05 | 0.012639 |
| ACAA | metabolic process | 0.000165 | 0.024078 |
| CD209 | carbohydrate binding | 3.64E-06 | 0.00111 |
| CDON | striated muscle cell differentiation | 3.61E-06 | 0.00111 |
| FABP3 | lipid binding | 1.23E-05 | 0.003162 |
| SULT1B1 | metabolic process | 5.97E-05 | 0.012134 |
| PPL | protein binding | 5.74E-06 | 0.001639 |
| BMP10 | growth | 3.74E-07 | 0.00017 |
| CRABP2 | retinoic acid metabolic process | 0.000172 | 0.024991 |
| TBX5 | protein binding | 9.70E-05 | 0.016751 |
| CD9 | immune response | 2.67E-09 | 1.82E-06 |
| MTMR7 | dephosphorylation | 5.35E-05 | 0.011131 |
| PRG4 | immune response | 3.09E-10 | 2.61E-07 |
| S100A12 | calcium ion binding | 4.05E-09 | 2.65E-06 |



*Continued……*

| | | | |
|---|---|---|---|
| GABRE | ion transport | 0.00041 | 0.04791 |
| ANKEF1 | protein binding | 1.29E-05 | 0.003176 |
| TIMD4 | protein binding | 8.83E-06 | 0.002366 |
| ALOX15 | oxidation-reduction process | 0.000134 | 0.020574 |
| LIPG | lipid metabolic process | 2.46E-09 | 1.74E-06 |
| PKD2L1 | ion transport | 0.000429 | 0.049614 |
| WISP2 | regulation of cell growth | 7.05E-08 | 3.90E-05 |
| KRT8 | apoptotic process | 1.52E-09 | 1.17E-06 |
| KRT18 | apoptotic process | 9.94E-05 | 0.016751 |
| F2RL1 | immune response | 0.000316 | 0.040164 |

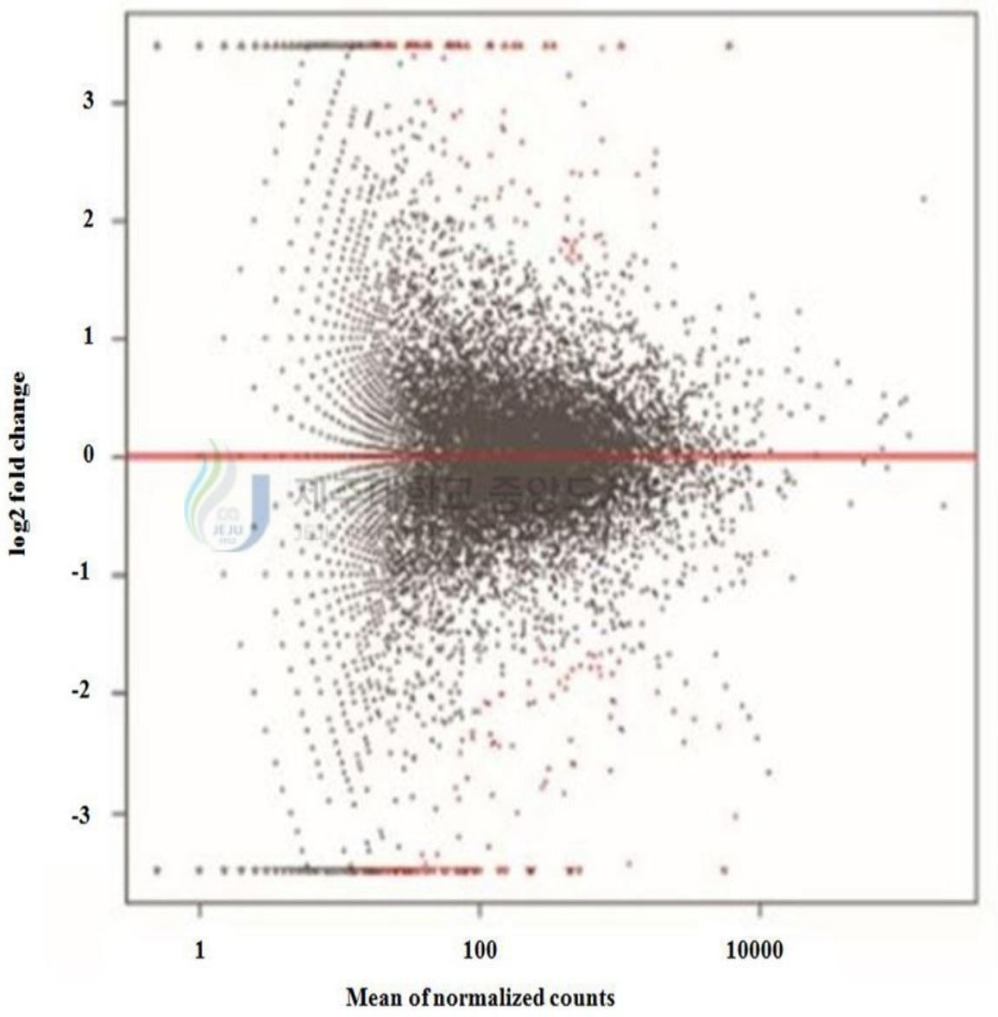


**Figure 3.2** MA plot between JNP and Berkshire

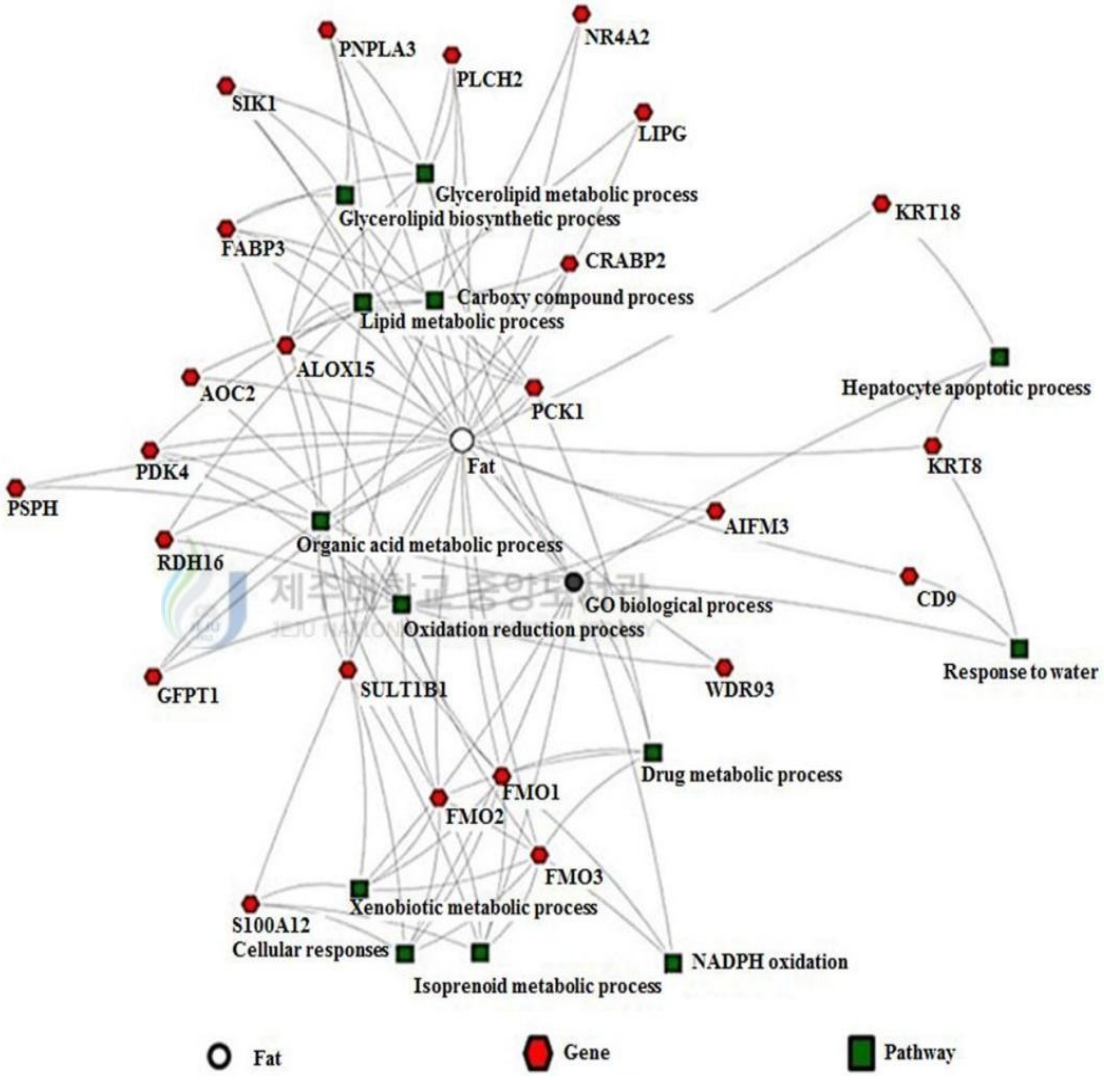


**Figure 3.3** Dissected gene level view sharing GO of biological factors and an abstracted network showing biological pathway enriched with the cluster of genes specific for fat

**Table 3.3** Pathway analysis of the differentially expressed genes in the fat tissue of adult Berkshire and Jeju Native Pig

| Gene name | Pathway |
|---|---|
| RDH16, AOC2, PSPH, PCK1, LIPG, ALOX15 | Metabolic Pathway |
| PCK1, PPL | TCA cycle |
| IL1RAP, CCL4 | Cytokine-Cytokine receptor interaction |
| KCNK-5, XPNPEP2 | Protein Biosynthesis |
| SLA-DRB1 | Antigen-processing and presentation |
| NOTCH4 | Notch signaling pathway |
| CCL4, PLAU | NF kappa biosynthesis |

### *3.4.4. Functional annotation of GO and expression level of DEGs*

The details of the biological pathways and the GO terms of the potentially associated genes were analyzed. The GO analysis revealing significant association with the DEGs and the biological processes of the different genes has been elaborated in the table 3.4. Cellular processes, biological adhesion and developmental processes along with multicellular organismal processes in fat are up-regulated in JNP as compare to Berkshire. Multicellular organismal process, developmental process, embryonic morphogenesis and skeletal system development are the most significantly enriched terms in fat tissues of JNP and Berkshire breeds ($p$ = 1.17E-04, $p$ = 0.044, $p$ = 3.47E-04 and $p$ = 4.48E-04, respectively) (Table 3.2). The GO of cellular components for specific DEGs is related to the extracellular region in fat. Further, the molecular function GO of specific DEGs is reported to be related with cofactor binding in fat.

The levels of expression of DEGs among the JNP and Berkshire have been visualized by the heat maps (Figure 3.4). It is distinctly observed that the genes have significant differential expressions between the two breeds (Table 3.5) and influence genetic drift. Therefore, transcriptome analysis could provide the common way of discovering differences in gene expression because regulation of gene activity occurs primarily at the transcription level (Venkatesh *et al.*, 2005).

### *3.4.5. Development of skeletal system and body growth*

We have found 4 genes related to meat quality and body growth. Most of these genes were up-regulated in Berkshire pigs. Out of all the identified DEGs, collagen type 10- alpha 1 (*COL10A1)* and collagen type 11- alpha 2 (*COL11A2)* are the genes, related to the skeletal system morphogenesis and growth of cartilage (Lefebvre and Smits, 2005; Kim *et al.*, 2007). These two genes are down-regulated in JNP (Table: 3.5). This may provide evidence that in Berkshire, body growth is more efficient than JNP. Pyruvate dehydrogenase kinase -4 (*PDK4),* found in fat is also associated with meat quality (Lan *et al.*, 2009) and was up-regulated in Berkshire. Patatin-like phospholipase domain containing 3 *(PNPLA3),* involved in metabolic process is reported to control fat deposition (Chen *et al.*, 2011) and was down regulated (1.21) in JNP (Table: 3.5). Therefore, the annotation findings of the current investigation can be used as the selection marker and might be helpful in better understanding of genetic mechanisms associated with the constraints of poor body growth in JNP.

### *3.4.6. Metabolic process*

The lipid and their metabolites function as transcriptional regulators of gene expression and fundamentally affect physiological processes such as growth and cell/tissue differentiation. The genes like *PDK4* (ketone metabolism), retinol dehydrogenase -16 *(RDH16)* (amine metabolism), phospholipase C, eta 2 *(PLCH2)* and patatin- like phospholipases-3 (*PNPLA3)* (lipid metabolism) involved in the metabolic processes have significantly high expression in the Berkshire breed, whereas in JNP, phosphoserine phosphatase *(PSPH),* flavin containing mono-oxygenase 3 (*FMO3),* cellular retinoic acid binding protein 2 *(CRABP2)* and lipase, endothelial *(LIPG)* genes have significantly high expression (Supplementary Table S2). The deduced amino acid sequence of *PDK4* revealed a G/A mutation in intron 9 and association analysis showed that it is significantly associated with intramuscular fat and muscle water content (Lan *et al.*, 2009). These findings indicate towards the high intramuscular fat content in Berkshire. *FMO3* has been reported to be associated with fishy off flavor and pork pH (Glenn *et al.*, 2007). It has been reported that it affects the metabolism of trimethyl amines (TMA) to TMA oxide (TMAO) which leads to the production of oxidized fishy odor (Zhou and Shephard, 2006) in pork.

*PNPLA3* is reported to be mainly expressed in porcine adipose and liver tissue to exhibit dual functions (Zhou and Shephard, 2006; Glenn *et al.*, 2007; Chamoun *et al.*, 2013). In the present study, bone morphogenetic protein-10 (*BMP10)* and *FABP3* (lipid) genes have shown the least and equal level of expression pattern in both the breeds. Although the expression level of

**Table 3.4** Gene ontology analyses related to the biological process of the differentially expressed genes in fat tissue of adult Berkshire and Jeju Native Pig

| Gene name | Go term |
|---|---|
| COL10A1, COL11A2 | skeletal system morphogenesis, collagen fibril organization |
| PDK4 (ketone), RDH16, AOC2 (amine), FMO3 (organic acid), PSPH, SULT1B1, CRABP2 (retinoic), PLCH2(lipid), PNPLA3 (lipid), FABP3 (lipid), LIPG | metabolic process |
| SIK1, NOTCH4, WISP1 | cell differentiation |
| WDR93. IL1RAP, LGALS3, PPL, TBX5, ANKEF1 | protein binding |
| KCNK5, S100A12 , PKD2L1 ($Ca^{2+}$ transport) | ion binding |
| AIFM3, ALDH3L3, FMO1, FMO2, ALOX15, | oxidation-reduction process |
| CCL4, LCN2, CXCL2, SLA-DRB1, SLA-5, PRG4, | Inflammatory and immune response |
| XPNPEP2, ADAMTS4, ADAMTS1, PLAU | proteolysis |
| ITGA8, NR4A2, | smooth muscle tissue development and skeletal muscle cell differentiation |
| KRT8, KRT18 | apoptotic process |
| PCK1, GFPT1, CD9 | glucose metabolic process |

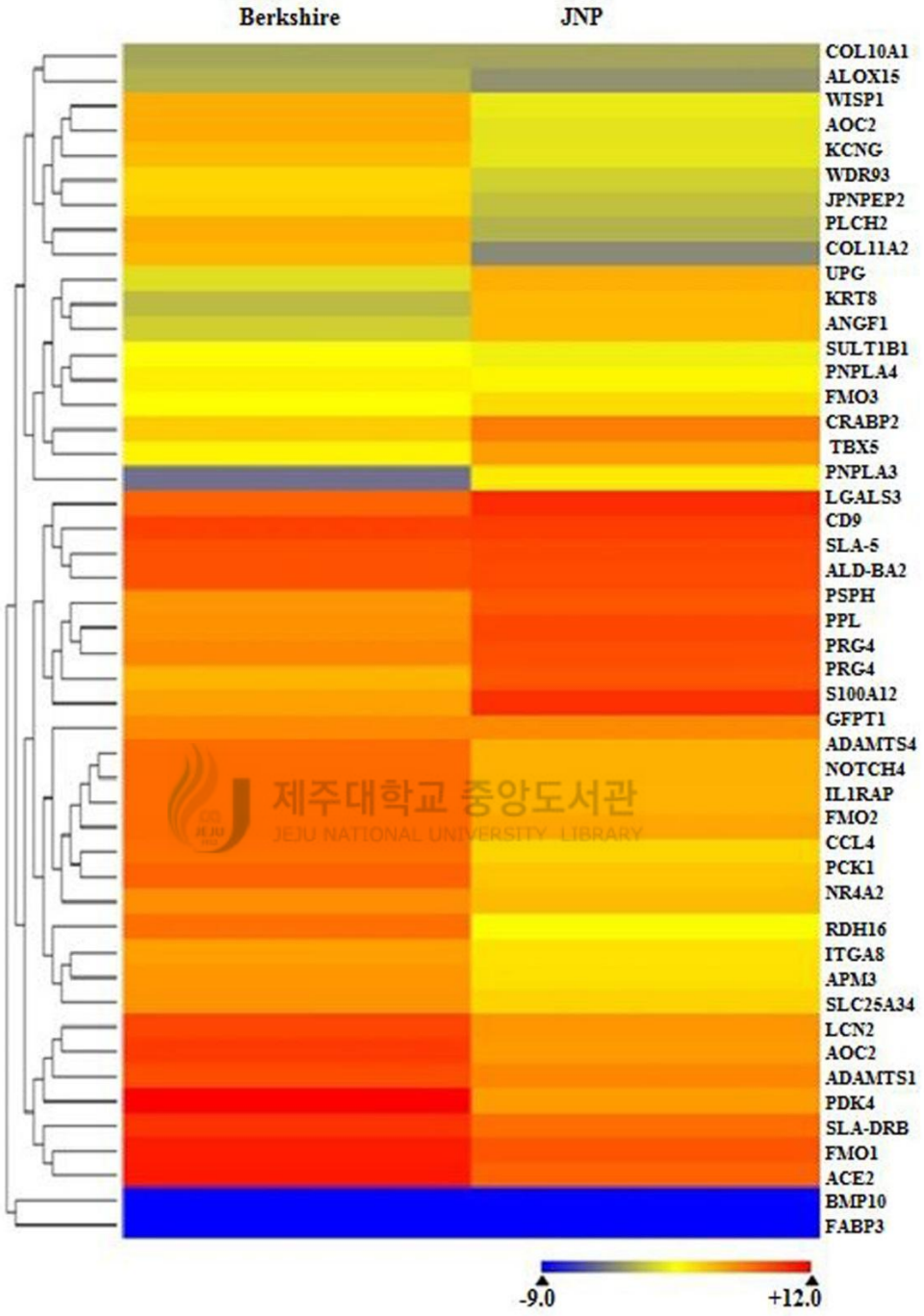



**Figure 3.4** Heatmap illustrating the level of the differentially expressed genes in fat tissue of adult JNP and Berkshire. The red blocks represent the over expressed genes and the blue blocks represented the lowest level expression of the genes

**Table 3.5** The expression of the genes and the log fold changes in fat tissue of adult Berkshire and JNP.

| Gene | Berkshire adult fat tissue* | JNP adult fat tissue* | Fold change+ |
|---|---|---|---|
| COL11A2 | 2.92 | -3.62 | 93.49 down |
| COL10A1 | -2.66 | -2.73 | 1.04 down |
| PDK4 | 9.21 | 3.85 | 41.11 down |
| RDH16 | 5.34 | 0.35 | 31.78 down |
| AOC2 | 3.31 | -0.46 | 14.03 down |
| ACE2 | 8.43 | 5.85 | 5.99 down |
| FMO3 | 0.41 | 1.64 | 2.33 up |
| PSPH | 4.07 | 6.2 | 4.53 up |
| CRABP2 | 2.18 | 4.90 | 6.59 up |
| SIK1 | 4.13 | 1.92 | 4.60 down |
| NOTCH4 | 5.53 | 3.12 | 5.31 down |
| BMP10 | -8.40 | -8.40 | none |
| WISP1 | 3.21 | -0.26 | 11.22 down |
| PLCH2 | 3.20 | -2.20 | 42.59 down |
| PNPLA3 | 0.98 | 0.69 | 1.21 down |
| FABP3 | -8.40 | -8.40 | none |
| LIPG | -0.68 | 7.17 | 14.49 up |
| WDR93 | 1.79 | -1.19 | 7.91 down |
| IL1RAP | 5.53 | 3.05 | 5.56 down |
| LGALS3 | 5.81 | 7.69 | 3.67 up |
| PPL | 4.25 | 6.82 | 5.92 |
| TBX5 | 0.77 | 3.77 | 8.01 up |
| ANKEF1 | -1.23 | 2.84 | 16.83 up |
| KCNK5 | 2.74 | -0.41 | 8.95 down |
| S100A12 | 3.65 | 7.56 | 15.15 up |
| PKD2L1 | -4.47 | 1.14 | 49.34 up |
| AIFM3 | 4.06 | 1.53 | 5.77 down |
| ALDH3A2 | 6.44 | 6.66 | 1.16 up |
| FMO1 | 8.29 | 6.28 | 4.03 down |

*Continued……*

| | | | |
|---|---|---|---|
| FMO2 | 5.60 | 3.50 | 4.30 down |
| ALOX15 | -2.28 | -2.35 | 2.09 down |
| CCL4 | 5.35 | 1.86 | 11.20 down |
| LCN2 | 6.86 | 4.06 | 6.95 down |
| CXCL2 | 7.21 | 3.88 | 10.07 down |
| SLA-DRB | 7.48 | 5.49 | 3.97 down |
| SLA-5 | 6.39 | 6.76 | 1.29 up |
| PRG4 | 3.00 | 6.33 | 10.07 up |
| XPNPEP2 | 2.02 | -1.72 | 13.44 down |
| ADAMTS4 | 5.51 | 3.10 | 5.31 down |
| ADAMTS1 | 6.60 | 4.58 | 4.03 down |
| PLAU | 4.58 | 6.53 | 3.86 up |
| ITGA8 | 3.71 | 1.42 | 4.86 down |
| NR4A2 | 4.32 | 2.78 | 2.90 down |
| KRT18 | -1.88 | 2.84 | 26.41 up |
| PCK1 | 5.83 | 2.41 | 10.71 down |
| GFPT1 | 4.50 | 4.45 | 1.03 down |
| CD9 | 7.02 | 7.11 | 1.06 up |

*FABP3* is quiet less, but the expression of *FABP3* is reported to have positive correlation with intramuscular fat content. Our findings are in line with the earlier study (Serao *et al.*, 2011) which has confirmed that *FABP3* is important for fat deposition in the porcine *L. dorsi* muscle. It has been reported that mutation in *FABP3* can be a useful marker in marker assisted selection which is aimed to improve intramuscular fat, with the condition that gene are segregating and presence of association is noticeable in the population (Switonski *et al.*, 2010).

Phosphoenolpyruvate carboxykinase 1 *(PCK1),* glutamine--fructose-6-phosphate transaminase 1 *(GFPT1)* and *CD9* involved in the glucose metabolism, are down regulated (10.71 and 1.03 fold respectively) whereas *CD9* (1.06 fold) is up regulated in JNP (Supplementary Table S2). These are reported to be associated with gluconeogenesis and in cattle hepatocyte cultures increase in *PCK1* mRNA levels by long chain fatty acids in a dose - dependent manner (Akbar *et al.*, 2013). The genome wide analysis in mouse has reported that it encodes the rate- limiting enzyme in hepatic gluconeogenesis (Shin *et al.*, 2012). It has been found that in contrast to JNP, maximum number of genes related to the metabolic process, are more expressive in the Berkshire. As the interactions among nutrients, metabolites, gene expression and protein modification are important for the coordination of cell growth with extracellular and intracellular conditions (Yuan *et al.*, 2013). Therefore, it can't be discerned from the present data whether the expressions of the genes are directly related to more growth in Berkshire than JNP.



### *3.4.7. Inflammatory and innate immune response*

Analysis of adipose tissues has revealed significant enrichment of GO categories related to immune response, stress, lipid metabolism and carboxylic acid metabolic processes (Pashaj *et al.*, 2013). In the current study, six genes i.e. Chemokine C-C motif ligand 4 (*CCL4*), lipocalin-2 (*LCN2*), *CXCL2,* swine leukocyte antigen – DRB1 (*SLA-DRB1*), *SLA5* and proteoglycan-4 (*PRG4*) are reported to be involved with the immune response (Table 3.4). It is quite interesting that out of total six DEGs, four genes i.e., *CCL4, LCN2, CXCL2, SLA-DRB1* are more expressive in Berkshire. All these genes contribute for the strong immune system of the breed. *CCL4* shows elevated mRNA levels with leuteolytic capacity (Luo *et al.*, 2011). *SLA-DRB1* having high expression in the Berkshire is responsible for antigen processing and presentation of exogenous peptide antigen vis-a-vis MHC class I (Lunney, 2007). A functional test in three Chinese pig strains also indicated that co-transfection with DRA and DRB genes, provided facts that DR molecules could stimulate human T- cells (Chen *et al.*, 2012). Moreover, *SLA* characterized pigs are an important tool for the study of immune responses, disease resistance and production traits

(Ho *et al.*, 2006). Similarly *CXCL2* is reported to induce mobilization of neutrophils from bone marrow (Akha *et al.*, 2013). *LCN2* - a member of the lipocalin superfamily is reported to play an important role in the osteoblast differentiation, regulation of cellular oncogenesis, apoptosis and inflammation (Costa *et al.*, 2013). Soto *et al.* (2008) reported that, different *SLA* haplotypes were associated with the variation in the immunity response and disease. Higher expression of these genes in Berkshire is indicative of better immune system as compare to JNP.

*PRG4* is only gene which expressed more in JNP as compare to Berkshire breed whereas *SLA5* has shown equal expression in both of the breeds (Table 3.5). *PRG4* has the vital role in regenerating functional superficial zone cartilage tissue and full thickness cartilage with zonal organization and cellular phenotype (Coates *et al.*, 2012). These findings suggest that it can be used a marker gene for the immune system related studies in JNP.

### *3.4.8. Protein binding and phosphorylation*

Phosphorylation plays a vital role in regulation of protein binding and coordination of pathways by modulating the nature and strength of protein- protein interactions. In the present study we identified the DEGs those are responsible for protein binding and phosphorylation. Our study reported that *WDR93* and Interleukin-1 receptor accessory protein *(IL1RAP)* have high expression in Berkshire breed whereas Lactin galactoside- binding soluble 3 *(LGALS3)*, Periplakin *(PPL)* and T-box5 (*TBX5)* have a significantly higher expression in JNP (Table 3.5).



*PPL* is a cytoskeletal protein and provide structural integrity within the skin and it helps in maintaining water impermeability and resiliency of the skin (Seo *et al.*, 2012). At the same time *TBX5* which has expressed significantly high in JNP is reported to be involved in dorsal eye patterning and limb regeneration (Jassar *et al*., 2013). It has been observed that *IL1RAP* expresses abundantly in the porcine endometrium and play an important role during peri- implantation period of pregnancy (Sousounis *et al.*, 2013). The lesser expression of *IL1RAP* in JNP may be one of the reasons of lesser litter size. Such findings help to elucidate the genetic control of placental efficiency and improve the understanding of placental development (Zhou *et al.*, 2009).

### *3.4.9. Oxidation reduction process*

Since, oxidation results into the biochemical changes in meat which leads to change in color of meat. It also causes undesirable flavor and rancidity in meat (Bekhit *et al.*, 2013); thereby have impact on consumer's choice. In the course of investigation, pertaining to the genes

responsible for the oxidation and reduction process in the two breeds, quite interesting results have been indicated. It is found that among the annotated genes, five genes which are related to the oxidation and reduction process have differently expressed in both of the breeds. Out of these, three genes such as apoptosis-inducing factor mitochondrion-associated - 3 (*AIFM3),* flavin containing monooxygenase 1 and 2 *(FMO1* and *FMO2)* expressed significantly more in the Berkshire. The expression level of the aldehyde dehydrogenase 3 family *(ALDH3A2)* was found to be equal and arachidonate 15-lipoxygenase *(ALOX15)* have expressed negatively in both the breeds (Table 3.5). It has been reported that loss of function mutations in the genes from *FMO* family is associated with fishy off flavor (Glenn *et al.*, 2007). In our study, lesser expression of *FMO1* and *FMO2* genes in JNP may indicate regarding their role in the better meat quality and long shelf life of meat from JNP. Therefore, these genes can be potent markers for the studies related to meat quality of pork.

### 3.5. Conclusion

To our knowledge till date, expression profiles of genes in different types of tissues and developmental stages related to the biological pathways in JNP have not been studied. Therefore, in this study, we aimed to investigate most abundant biological pathways and the expression profiles of those genes which caused the variations between JNP and Berkshire. To best of the available resources, this study is the first transcriptional analysis for the detection of DEGs from RNA-seq data generated from fat tissue sample. It is a well-known and proven fact that fat and fatty acid content affect meat quality. At the same time it is also observed that different melting points of fatty acids affect the tenderness of meat. Transcriptome analysis by RNAseq has been already conducted in different breeds of pig except JNP (Chen *et al*., 2011; Petkov *et al*., 2011; Rustemeyer *et al*., 2011; Jung *et al*., 2012; Looft, 2013; Prather, 2013; Samborski *et al*., 2013). Therefore, keeping this goal in mind, the current study was planned to analyze the differential expression of genes in fat tissue from JNP and Berkshire by using RNAseq. On the basis of number of DEGs, our findings confirmed the transcriptome activities of fat with biological functions. Notably, the genes up-regulated in JNP are directly or indirectly related to meat quality and body growth. Thus, our results offer new insight towards an understanding and identification of the candidate genes that are involved in the alteration of the biological functions. The annotation findings of our study may act as the selection marker for the

genes related to the body growth and might be helpful in better understanding of genetic mechanisms associated with the constraints of poor body growth in JNP.

Therefore, it can be concluded here that the present investigation would add up in the process of improvement in the pig genome annotations. The detail understanding of genetic and functional variants such as novel transcripts, sequence polymorphisms, isoforms and non-coding RNAs are the part of supported outcomes of the current study. Further, in order to investigate the genetic profiles of fat tissue transcriptome analysis can make the path easier to understand the difference of the genetic mechanisms in JNP with respect to Berkshire and enable the identification of specialized biological functions and regulatory genes which would be used as the selection of marker for further breed improvement programmes. Further extensive *in-vitro and in-vivo* studies can be planned for the regulated expression of genes enlisted by the current study.

# Chapter 4

# Comparative transcriptomic analysis by RNA-seq to discern differential expression of genes in liver and muscle tissues of adult Berkshire and Jeju Native Pig

## 4.1. Abstract

RNA-seq is being rapidly adopted for the profiling of the transcriptomes in different areas of biology, especially in the studies related to gene regulation. The discovery of differentially expressed genes (DEGs) between adult animals of Jeju Native Pig (JNP) and Berkshire breeds of *Sus scrofa*, are of particular interest for the current study. For the better understanding of the gene expression profiles of liver and *longissimus dorsi* muscle, DEGs were identified via RNA-seq. Sequence reads were obtained from Ilumina HiSeq2000 and mapped to the pig reference genome (Sscrofa10.2) using Tophat2. We identified 169 and 39 DEGs in liver and muscle of JNP respectively, by comparison with Berkshire breed. Out of all identified genes, 41 genes in liver and 9 genes in muscle have given significant expression. Gene ontology (GO) terms of developmental process and KEGG pathway analysis showed that metabolic, immune response and protein binding were commonly enriched pathways in the two tissues. Further the heat map analysis by ArrayStar has shown the different levels of expression in JNP with respect to the Berkshire breed. The validation through real time PCR and western blotting also confirmed the differential expression of genes in both the breeds. Genes pertaining to metabolic process and inflammatory and immune system are more enriched in Berkshire breed. This comparative transcriptome analysis of two tissues suggests a subset of novel markers genes which expressed differently between the JNP and Berkshire.



**Key words**: Differentially expressed genes (DEGs), liver, *longissimus dorsi* muscle, RNA-seq, Gene Ontology (GO)

## 4.2. Introduction

*Sus scrofa* is one of the most important domestic animals used for meat production. It is an economical source of animal proteins and thereby the demand of pig meat is increasing day by day. Moreover, pigs are one of the major animal species used in the translational research (Swindle *et al*., 2012). In the tropical Asian and Southern European continents, pigs also play a fundamental role in the social and economic status of meat industry. Jeju Island (Republic of South Korea) has very unique weather conditions and due to such climatic conditions, this island has very unique native species of plants and animals. Among the various meat species of animals, Jeju Native Pig (JNP) is the one which is known for its good meat quality. Pork from JNP is preferred than other commercial breeds in Japan, China and Korea (Cho *et al*., 2011). Palatability characteristics i.e. high tenderness, juiciness and marbling quality are the parameters which make JNP meat more preferred than its contemporaries. At the same time lesser prolificacy and lower feed efficiency breaks the continuity between supply and demand of JNP meat.

Berkshire pork is renowned for its palatability characteristics (Ryu *et al*., 2008). As far as eating quality is concerned, pork loins from Berkshire pigs are tenderer than Landrace and Yorkshire pigs (Lee *et al*., 2012). Moreover, earlier breeding strategies have reported that Asian native pigs are closely related to Berkshire breeds (Kim *et al*., 2002). It is reported that liver is the major site of fatty acid synthesis (Corino *et al*., 2002) and plays a significant role in body growth (Mourot *et al*., 1995). Therefore, liver is the tissue of choice to study the metabolic processes, immunity, growth and meat quality as well. Factors influencing many peri-mortal and post-mortal biochemical processes in meat quality and body growth are muscle composition, area of fiber and the capillary density of specific muscles (Malek *et al*., 2001). Thus, in the present study, we selected the *longissimus dorsi* muscle tissue and liver to identify DEGs responsible for growth and meat quality in the JNP in comparison to Berkshire.

qRT-PCR, microarray and microsatellite techniques are in vogue to identify genes along with their different markers between the pig breeds (Kim *et al*., 2005; Kim *et al*., 2008; Moon *et al*., 2009). To characterize and quantify the entire genome, next generation sequencing (NGS) is one of the preferred approaches. NGS analyzes expression of genes more effectively by the serial and cap analysis gene expression (CAGE) technique by producing a snapshot of 5’ ends of capped transcripts to locate exact transcription start sites in the genome (Mortazavi *et al*., 2008; Takahashi *et al*., 2012). RNA-seq furnishes a far more exact computation of transcript levels and their isoforms than any other techniques (Wang *et al*., 2009). RNA-seq also lay a path for the

detection of DEGs which have low expression levels (Hardcastle and Kelly, 2010; Robinson *et al*., 2010; Wang *et al*., 2010; Tarazona *et al*., 2011; Trapnell *et al*., 2012).

The statistical interpretation of data obtained by RNA-seq is important in transcriptomic studies especially for the normalization of quantitative measurements of expression (Wilhelm and Landry, 2009; Li *et al*., 2010; Robinson and Oshlack, 2010; Hong *et al*., 2012). Such kind of interpretation aids in the identification of new genes along with splice variants of known ones. It aids in the comparison of gene and transcript expressions under two or more conditions (Trapnell *et al*., 2012). RNA-seq has emerged as the future of transcriptome research because of its potential to distinguish homologous regions, the identification of regions of expression without making any prior assumption (Shendure and Ji, 2008) and its ability to provide more comprehensive approach to investigate the transcriptomes than microarrays (Mortazavi *et al*., 2008).

Many researchers have used RNA-seq technology for transcriptome analysis in the commercial pig breeds but no such kind of study has been conducted in JNP (Chen *et al*., 2011; Petkov *et al*., 2011; Rustemeyer *et al*., 2011; Jung *et al*., 2012; Looft, 2013; Prather, 2013; Samborski *et al*., 2013). It is worth mentioning that very less comparative studies among the different pig breeds have been conducted for the biological processes. Therefore, in this study, we sought to obtain DEGs showing differential expression in liver and muscle of JNP and Berkshire by RNA-seq analysis. These DEGs from each breed will be useful to develop breed specific markers for breed discrimination. Moreover, these biomarkers will be useful to monitor growth performance and meat quality.



## 4.3. Materials and methods

### *4.3.1. Selection of animals and collection of tissue samples*

In the current study, five adult (5 months old) female animals each from JNP and Berkshire breeds were selected. Animals were reared under same environmental and nutritional conditions. The commercial feed (Seoul Feed, Jeju- Si, South Korea) and water were offered *ad-libitum* to all the pigs. Pigs were housed in pens with concrete flooring, a nipple bowl drinker and feeder. The experiment was approved by the animal ethics committee of the Faculty of Biotechnology, Jeju National University, Jeju and animals were slaughtered according to the standard protocols of Jeju National University. The *longissimus dorsi* muscle between the 12$^{th}$

and 13th rib space and the median hepatic lobes of liver from each animal were collected immediately after the slaughter. Samples were stored immediately in dry ice and later were stored at -80℃ in the lab until used for RNA extraction.

### *4.3.2. RNA extraction and analysis of its quality*

The fragmented frozen tissue samples (120 mg) (liver and *longissimus dorsi* muscle) in five replicates were used for the RNA isolation from JNP and Berkshire (adult) pigs. To extract RNA from the tissues, TRIzol™ (Invitrogen, USA) reagent was used. 2.0 ml of TRIzol and chloroform reagents were added to tissue samples during the homogenization step. The homogenized tissues were subsequently precipitated by using isopropanol (Junsei Chemical Co. Ltd., Japan). In the final step, precipitated pellet was washed with 1 ml of 75% ethanol. Isolated RNA samples were immediately stored at -80℃. From each isolated RNA samples, 25 μg of RNA was treated with the RNase-free DNase set (QIAGEN, Hilden, Germany) to get rid of genomic DNA contamination, and it was purified with the RNeasy mini kit according to the user guidelines (QIAGEN, Hilden, Germany). The quality and quantity of RNA was assessed by the Bioanalyzer 2100 with RNA 6000 Nano Labchips and automated capillary gel electrophoresis following user guidelines (Agilent Technologies Ireland, Dublin, Ireland). 28S/18S ratios for the RNA samples ranged from 1.8 to 2.0 and integrity of RNA had ranged from 8.0 to 10.0.



### *4.3.3. Extraction of protein from tissues*

The finely minced liver and muscle tissues (500mg each) were used for protein extraction. 1X phenyl methyl sulfonyl fluoride (PMSF) and 1.5 ml Radio Immuno Precipitation Assay (RIPA) buffer (T & I, South Korea) were added to the tissue samples for homogenization. The homogenized samples were incubated in ice for 30 minutes followed by centrifugation at 13,000rpm for 25 minutes at 4℃. The supernatant containing soluble protein was then collected. The manufacturer's guidelines were followed for measurement of the concentration of protein with the Pierce™ BCA protein assay kit (Thermo Scientific, USA) in a Bio-Rad Micro-plate Reader (Model-680). The protein samples were stored at -80 ℃ until their further use.

### *4.3.4. Construction and sequencing of RNA-seq ilumina library*

For the construction and sequencing of the RNA-seq library, 0.1-4.0 µg of RNA in five replicates from both the tissues has been used. For the construction of library the TruSeq RNA Sample Pre Kit was used by following the manufacturer's guidelines. Human UHR total RNA from agilent technologies was used as a positive control sample. The standard protocol of Ilumina, Inc was used to construct the library. Then libraries were pooled on the basis of their indexes and an Ilumina HiSeq 2000 high-throughput sequencing instrument with 100 paired-ends (PE) sequencing was used to sequence in one lane.

### *4.3.5. Filtering and merging of reads*

Consensus Assessment of Sequence and Variation v 1.7 (CASAVA -v1.7) was used with its default parameters to remove the low quality paired-end reads. Tophat2 (v2.0.2) was used to align PE reads to Sus scrofa genome (Sscrofa10.2) from the University of California Santa Cruz (UCSC) after quality filtration. HTseq (v 0.5.3p3) was used to count and merge reads on the basis of the overlapping of PE reads. The corrplot (Friendly, 2002) and DEGseq (Wang et al., 2010) R packages were used to analyze, correlation between adult animals from both of the breeds. DEGs were statistically identified by Fisher's exact test model (Bloom et al., 2009). Samples from both the breeds were compared with the samples of their respective class (i.e. JNP muscle vs Berkshire muscle and JNP liver vs Berkshire liver) by DEGseq for detection of DEGs. Significant DEGs were selected at FDR <0.05.



### *4.3.6. Functional annotation and the pathway analysis of DEGs*

Official gene symbols of DEGs expressed in the current study were allotted after cross-matching gene IDs and official gene symbols from bovine Ensembl and the human. The official gene symbols from human Ensembl which were homologs for bovine genes were used in functional clustering and enrichment analyses using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Dennis et al., 2003). The representations of functional groups in liver and longissimus dorsi muscle tissues of JNP and Berkshire, relative to the whole genome were investigated using the Expression Analysis Systematic Explorer (EASE) tool (Hosack et al., 2003) of DAVID. Enrichment of gene ontology (GO) terms was measured by EASE (Alterovitz and Ramoni, 2010). To identify enriched GO terms, functionally clustered genes were filtered according to EASE values < 0.1.

The primary objective of ToppCluster (Kaimal et al., 2010) is to locate biological themes in data sets involving numerous sets of genes, comparative enrichment clustering of genes and network based dissection of biological systems. Here, an intuitive and efficient analysis has been presented. We have visualized the shared and specific features associated with any number of genes involved in biological processes. ToppCluster with a P-value cutoff 0.05 was selected and no false discovery correction method was applied.

Further, pathway analysis for the DEGs was also carried out by KEGG (Kyoto Encyclopedia of Genes and Genomes) tool. KEGG aided to understand the functional meanings of higher-order and utilities of the cell or the organism from their genome information. It is an integrated resource with associated software consisting of three types of databases for genomic, chemical and network information.

### *4.3.7. quantitative Real-time PCR*

Primers used in quantitative real-time PCR were designed using the online Primer-3 program (Rozen and Skaletsky, 2000) and the primers that were used to conduct PCR along with their detailed information are listed in table 4.1. To evaluate the relative expressions of the differentially expressed genes quantitatively in Berkshire and JNP, real-time qRT-PCR was conducted using an Applied Biosystems, Step-One Real Time PCR system. The dye EvaGreen (Biotium, USA) was used to determine the quantity of transcript of target genes present in each sample. Each individual sample was quantified in triplicate under the following amplification conditions: 95 ℃ for 10 min initially, and then 40 cycles of 95 ℃ for 15 sec and 60 ℃ for 1 min. Standard curve methods were used to define the efficiency of real-time PCR. The efficiency of amplification of the target gene was compared with that of the endogenous *GAPDH* control transcript (Wang *et al*., 2003). Quantification of mRNA levels was performed using the comparative $C_T$ method. The results are reported as the relative expression normalised using the level of the transcript of the endogenous reference (Van Poucke *et al*., 2001; Erkens *et al*., 2006).

### *4.3.8. Western blotting analysis*

Sixty micrograms of protein extract was diluted as 1:1 with 2X loading buffer (4% sodium dodecyl sulfate (SDS), 20% glycerol, 0.004% bromophenol blue, 25% 0.5 M Tris, and 5% β-mercaptoethanol). Protein extracts were denatured by boiling for 5 min before loading onto a 12% SDS-PAGE gel. Proteins were transferred to nitrocellulose membrane after electrophoresis.

The membranes were blocked for two hours at room temperature and incubated with primary and secondary antibodies specific for the respective proteins of the DEGs and β-actin, as listed in table 4.2.

The washed membranes were analysed to observe specific chemiluminescent signals using a Luminescent Image Analyzer (LAS-4000 mini) instrument. The results are depicted as the relative band intensities normalised relative to the band intensities of *β-actin* bands for each gene using Image J software (National Institute of Health, Bethesda, Maryland, USA). The means were compared between JNPs and bovines.

### *4.3.9. Statistical analysis to access the expression levels of differentially expressed genes*

General linear model (GLM) was used to analyze the raw data. DEGs which have significant difference in their expression at $p<0.05$ and $FDR<0.05$ were selected. Further, analysis was designed to understand the statistical value of the expression levels among the genes between the Berkshire and JNP. Data were analyzed using the ArrayStar v4.1 (DNASTAR, Madison, WI) software for generation of the heat maps of selected genes of interest. The level of gene expression for each gene was measured using reads per kilobase per million for the particular gene region (RPKMg) (Wagner *et al*., 2012) and was calculated as follows:



$$\mathrm{RPKM}_g = \frac{r_g \times 10^9}{fl_g \times R} \qquad \text{Eq. 1}$$

Where, $r_g$ represent the number of the reads mapped for the particular gene region, $fl_g$ is feature length and R is the total number of reads obtained from the sequencing run of the respective sample.

The data was analyzed by the analysis of variance (ANOVA) for the relative quantitative expressions of the genes by real-time qRT-PCR and western blotting. The significant differences between the mean expressions of different genes at $P<0.05$ were analyzed by Tukey's b- test. The values have been expressed as mean±SEM.

**Table 4.1** List of the sequences of primers used during the quantitative real time PCR to analyze the differential expression of genes in the liver and *longissimus dorsi* muscles of Berkshire and JNP.

| Gene name | Primer sequences | Product size | Annealing temperature ($T_a$) | Gene bank ID |
|---|---|---|---|---|
| *FMO5* | F- AGGGAGGATGACATTGATGC<br>R- GGAGCTCTGAAATGGGCATA | 199 | 61℃ | XM_005653075 |
| *APOA4* | F- GATGAAGAAGCAGGCAGAGG<br>R- TGTTGAAGGTCTCCCCGTAG | 215 | 61℃ | NM_214388 |
| *SOCS2* | F- ATTGGTTTTGTGGCTGGAAG<br>R- AGCACAGAATCGGAACATCC | 188 | 61℃ | NC_010447 |
| *ACAT2* | F- ATCACCAAGGAGCGAATCC<br>R- CCTCTTCTGCTTGTCCCAAC | 245 | 58.5 ℃ | NM_001243427 |
| *ACAA2* | F- TAGGCTCTGTGGCTCTGGTT<br>R- GTAATTGCCATCGGGATTTG | 225 | 61℃ | NM_001167638 |
| *PDK4* | F- CCTCATTCCTCCACCAAGAA<br>R- GAGAAATGCTCGACCTCTCG | 244 | 61℃ | NC_010451 |
| *POSTN* | F- CAATTAGGCTTGGCTTCTGC<br>R- GAAGACCCTGAGCTGTTTGC | 213 | 61℃ | NC_010453 |
| *ARID5B* | F- GTGGTCCAGTCCTTCGACAT<br>R- GTCATCCCTCGCAATCAGTT | 234 | 61℃ | XM_003359209 |
| *COL2A1* | F- TCAAGGATTTCAAGGCAACC<br>R- TGTGACCTTTGACACCAGGA | 209 | 61℃ | NC_010447 |
| *MyBPH* | F AGTGCAGAAGGCAGACAAA<br>R AAGACCCGGAAGGAGTAAGA | 117 | 62℃ | NC_010451 |
| *β-actin* | F GACATCCGCAAGGACCTCTA<br>R ACACGGAGTACTTGCGCTCT | 157 | 60℃ | XM_003124280 |

**Table 4.2** List of primary and secondary antibodies used in western blotting to analyze the differential expression of proteins in the liver and *longissimus dorsi* muscles of Berkshire and JNP.

| Protein | Primary Antibody* | Secondary Antibody* |
|---|---|---|
| FMO5 | Goat polyclonal 1:50 | Donkey anti-goat, 1:1000 |
| APOA4 | Goat polyclonal 1:50 | Donkey anti-goat, 1:1000 |
| SOCS2 | Rabbit polyclonal 1:50 | Goat anti-rabbit, 1:1000 |
| ACAT2 | Goat polyclonal 1:50 | Donkey anti-goat, 1:1,000 |
| PDK4 | Goat polyclonal 1:50 | Donkey anti-goat, 1:1,000 |
| POSTN | Goat polyclonal 1:50 | Donkey anti-goat, 1:1,000 |
| ARID5B | Goat polyclonal 1:50 | Donkey anti-goat, 1:1,000 |
| COL2A1 | Rabbit polyclonal 1:50 | Goat anti-rabbit, 1:1000 |
| MyBPH | Goat polyclonal 1:50 | Donkey anti- goat, 1:1000 |
| β-actin | Mouse monoclonal 1:1000 | Goat anti-mouse, 1:1000 |



*All antibodies were from Santa Cruz.

## 4.4. Results and discussion

### *4.4.1. Identification and analysis of differentially expressed genes (DEGs)*

Analysis and understanding of RNA-seq data clearly depends on the ultimate goal of the experiment to perform differential gene expression. Identification of alternatively spliced and/or novel isoforms is wildly different to a pipeline (Corney, 2013). Ilumina HiSeq2000 was used to investigate DEGs in the liver and muscle tissue of both breeds. Transcriptome data were generated and RNA-seq reads of two tissues were acquired for adult animals of both the breeds. The data so generated consisted of highly reliable sequence reads. Few of the sequence reads (<5.0E-0.5%) did not pass the quality filtering. It has been observed that on an average, 97.6% (liver) and 90.8% (muscle) sequence reads passed the quality control (Table 4.3). Then these sequences were mapped successfully to the pig genome *Sus scrofa* using TopHat (v2.0.2). Further, these reads can be subjected to annotate with the putative transcripts, quantification of the number of reads per transcript and statistical comparison of transcript abundance across samples or treatments (Jones *et al*., 2012; Corney, 2013).

The genes with a significant difference in their expression were identified by the general linear model (GLM). GO of DEGs and up-regulated DEGs in both the tissues of JNP and Berkshire pigs (Figure 4.1, 4.2 (a) and (b)) have been summarized. Strong correlations with respect to genes between JNP and Berkshire have been observed for both of the tissues. DEGs for the current study were selected on the basis of the *P* value of <0.05 and FDR <0.05. 169 and 39 DEGs respectively were identified in liver and muscle tissues. Out of them 99 and 22 DEGs respectively, were annotated in the study (Figure 4.3) between JNP and Berkshire. Among the annotated genes, 41(liver) and 9 (muscle) functional genes have been identified (Table 4.4 and 4.5). The rest of the genes either matched unnamed or hypothetical proteins or had no BLAST matches in the non- redundant database. This data along with its matrix form can be visualized through an open source online visualization at http://biopopdb.snu.ac.kr/PIG_DEG/. Using NGS platforms, the transcriptional profiling has given a new window to researchers workings on the cell development and diseases (Corney, 2013).

**Table 4.3** RNA-seq reads and mapping rate of liver and *longissimus dorsi* muscle from Berkshire and Jeju Native Pig.

| Breed | No. of | Tissue | |
|---|---|---|---|
| | | **Liver** | **Muscle** |
| JNP | total reads | 45255914 | 35563538 |
| | reads after QC | 45254078 | 35562180 |
| | accepted Hit | 44164596 (97.6%) | 31807200(89.4%) |
| Berkshire | total reads | 37618038 | 37310958 |
| | reads after QC | 37616584 | 37309517 |
| | accepted Hit | 36731445(97.6) | 34331014 (92.0%) |

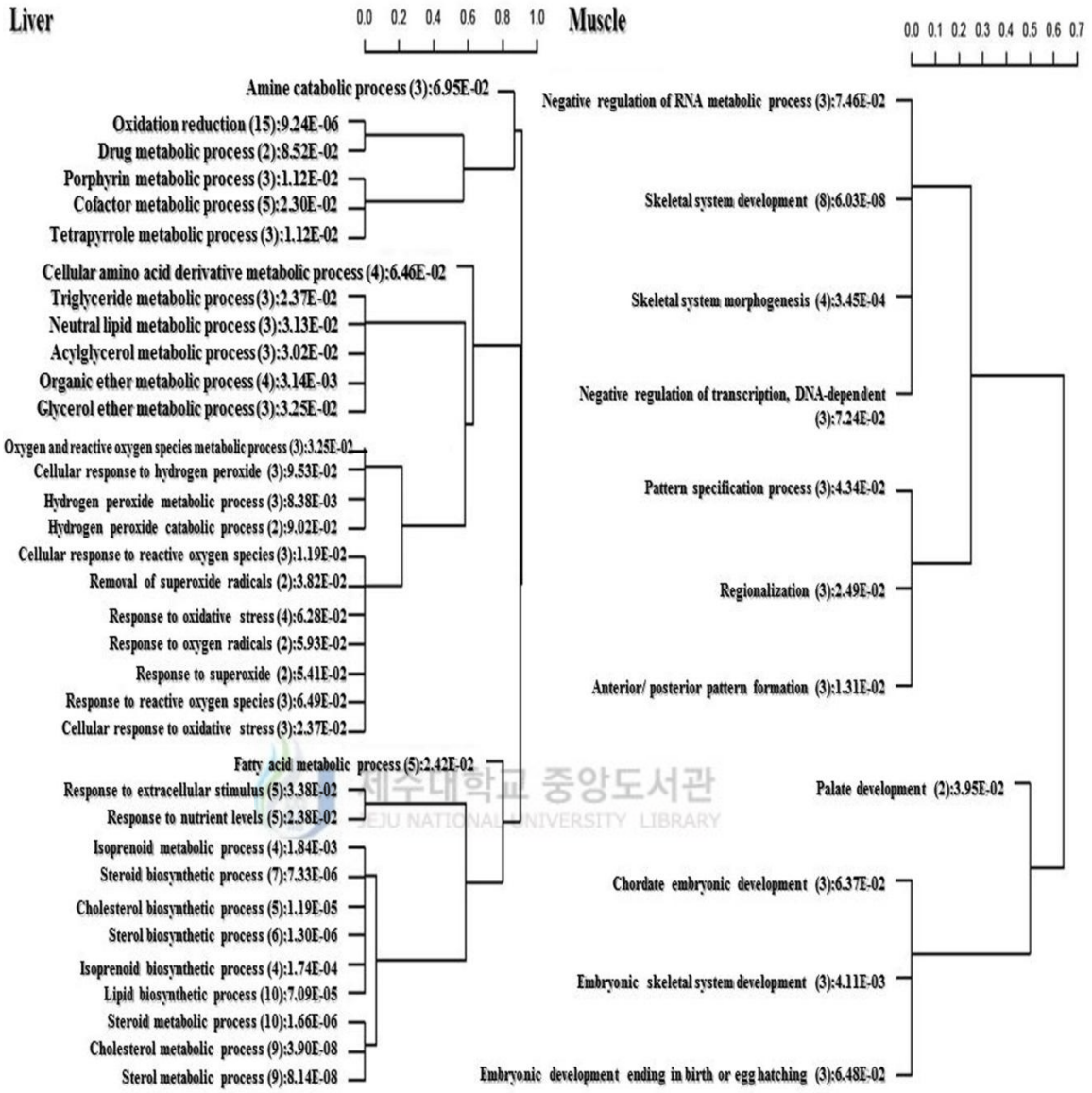




**Figure 4.1.** Consensus phylogenic tree with GO terms of enriched biological processes of differentially expressed genes expressed in liver and *longissimus dorsi* muscle of JNP and Berkshire. The phylogenic tree has been reconstructed by GLM with 100 bootstrap replicates. The scale bar represents expected number of transcripts across tissue in each branch. The genes from the groups with P- values ranging from 9.53E-02 to 1.12E-02 in Liver and 1.31E-02 to 7.46E-02 in muscle have been included in the analysis.

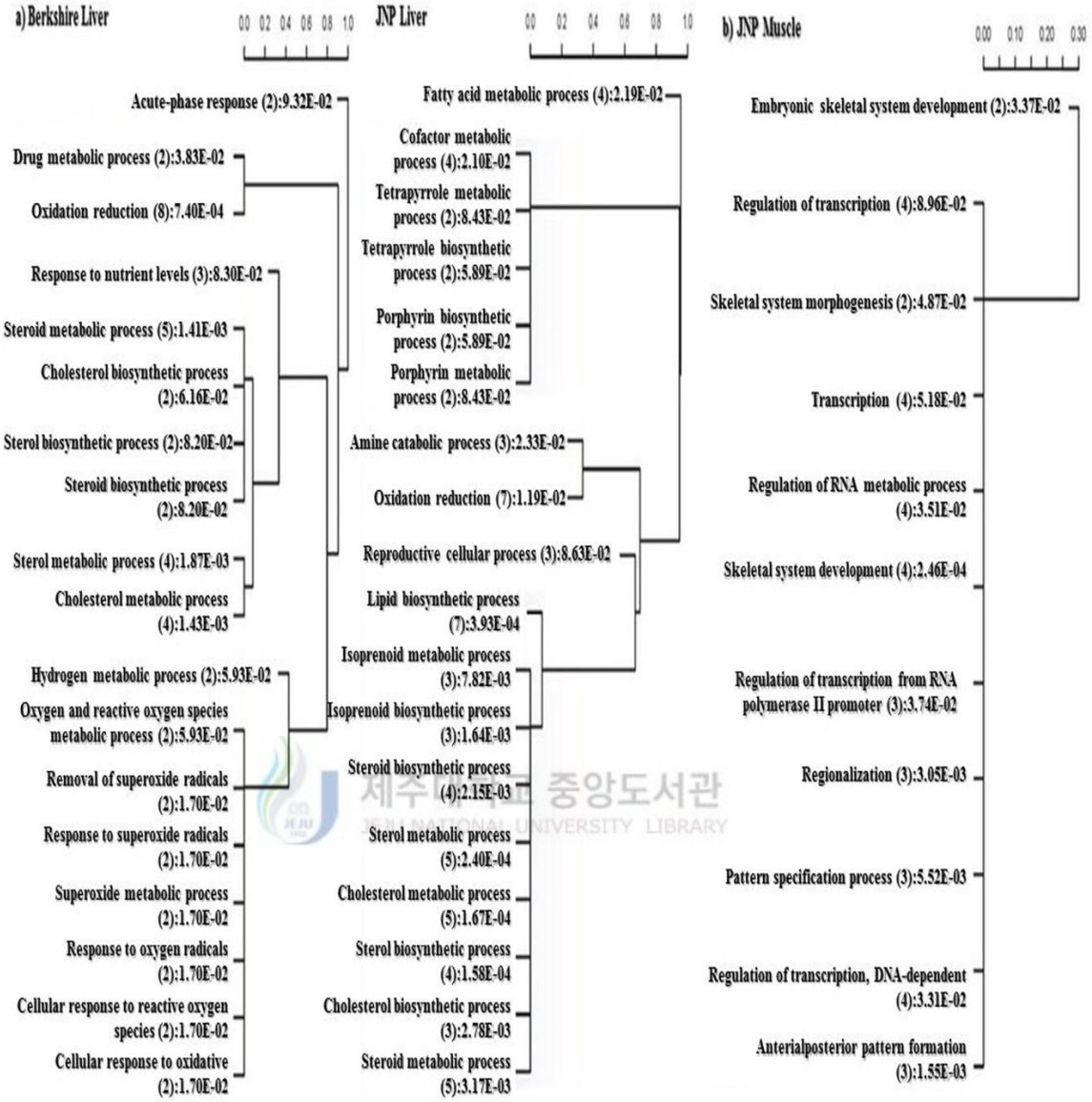




**Figure 4.2**. Phylogenies reconstructed by GLM with 100 bootstrap replicates of GO terms of up-regulated biological processes of (a) liver and (b) *longissimus dorsi* muscle specific DEGs from JNP and Berkshire. The scale bar represents expected number of transcripts across tissue in each branch.

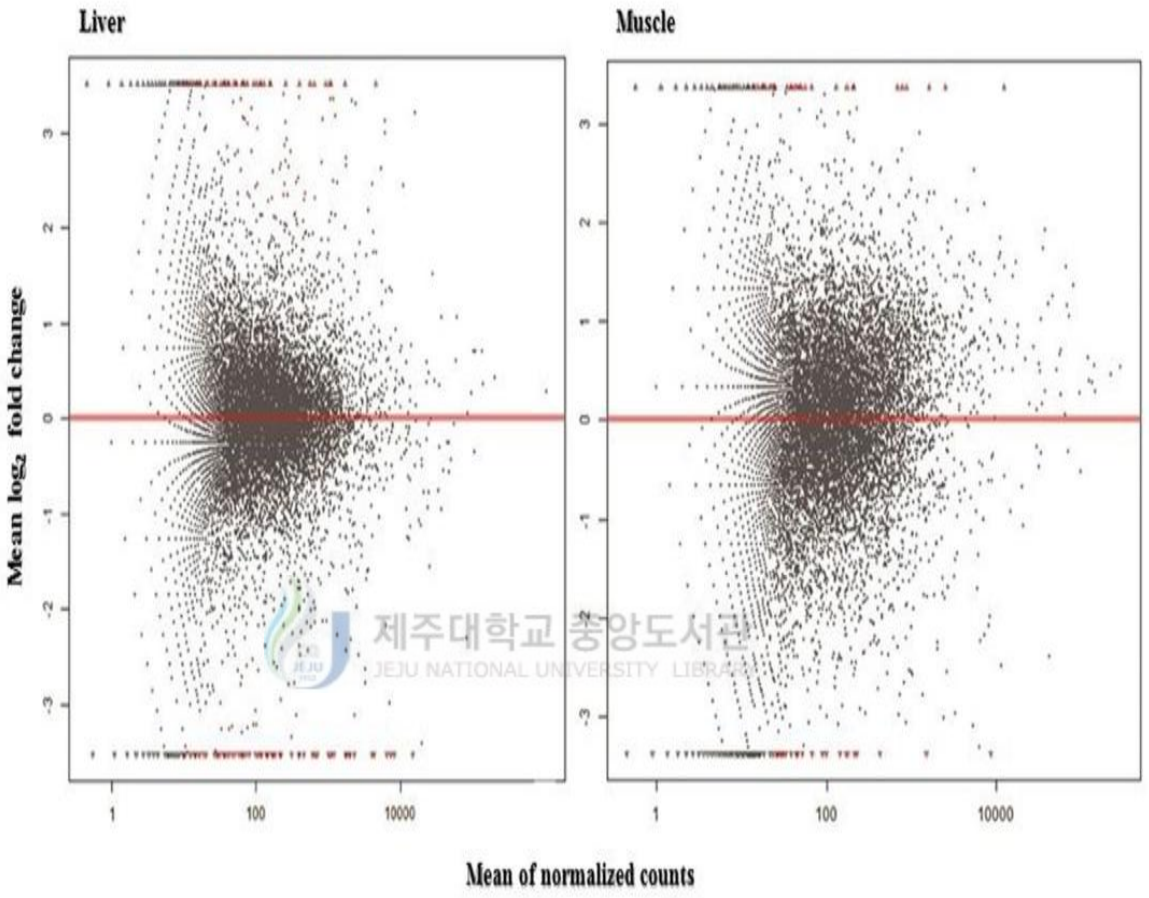



**Figure 4.3**. The MA-plot shows the $log_2$ fold changes in the JNP over the mean of normalized counts of Berkshire in the liver and *longissimus dorsi* muscle. The x-axis represents the average expression of genes over the tissue sample and the y-axis represents the $log_2$ fold change between the JNP (treatment) and Berkshire (control).

**Table 4.4** List of DEGs in the liver tissue of JNP and Berkshire (FDR<0.05).

| Gene | Function | *P*-value | FDR |
|---|---|---|---|
| *CYP2E1* | oxidation-reduction process, steroid metabolic process | 2.11E-08 | 8.21E-06 |
| *FMO5* | oxidation-reduction process | 9.18E-09 | 4.24E-06 |
| *ACAA2* | metabolic process | 7.07E-11 | 4.64E-08 |
| *APOA4* | cholesterol metabolic process, innate immune response | 4.23E-10 | 2.33E-07 |
| | metabolic process | 1.14E-11 | 9.75E-09 |
| *MT1A* | metal ion binding | 5.87E-11 | 4.01E-08 |
| *GALP* | regulation of appetite | 1.91E-07 | 5.94E-05 |
| *IGLV* | protein binding | 1.19E-09 | 6.33E-07 |
| *HSD17B13* | metabolic process, oxidation-reduction process | 1.49E-05 | 0.002761 |
| *SLC13A5* | sodium ion transport | 2.29E-11 | 1.78E-08 |
| *CCBL1* | biosynthetic process | 4.11E-10 | 2.33E-07 |
| *AOX1* | aldehyde oxidase 1 | 3.90E-06 | 0.000889 |
| *CYP1A2* | steroid catabolic process | 2.94E-08 | 1.12E-05 |
| *ACMSD* | metabolic process | 1.16E-07 | 3.97E-05 |
| *LEPR* | cholesterol metabolic process | 1.96E-08 | 7.84E-06 |
| *MIOX* | oxidation-reduction process | 1.52E-08 | 6.49E-06 |
| *MCOLN2* | mucolipin 2 | 4.29E-07 | 0.000122 |
| *CHL1* | cell adhesion molecule | 2.54E-07 | 7.36E-05 |
| *CACNA2D1* | calcium ion transport | 9.93E-08 | 3.53E-05 |
| *GPT2 2* | oxoglutarate metabolic process | 1.49E-07 | 4.82E-05 |
| *SLC25A47* | transport | 5.29E-07 | 0.000144 |
| *TRIB3* | positive regulation of protein binding | 2.27E-05 | 0.003962 |
| *PTPRZ1* | protein dephosphorylation | 0.00022 | 0.025899 |
| *FMO6P* | oxidation-reduction process | 2.11E-05 | 0.003726 |
| *IGLV* | protein binding | 0.000282 | 0.03235 |
| *SLC16A5* | organic anion transport | 6.34E-06 | 0.001321 |
| *MYCL* | regulation of transcription | 8.59E-06 | 0.001727 |
| *NNMT* | methylation | 0.000253 | 0.029572 |
| *ACAT2* | metabolic process | 4.17E-06 | 0.000925 |

*Continued……*

| | | | |
|---|---|---|---|
| *SOCS2* | regulation of growth | 1.58E-05 | 0.002912 |
| *HAVCR2* | protein binding | 0.000212 | 0.025292 |
| *PIM1* | cell proliferation | 2.91E-05 | 0.004669 |
| *SLC25A25* | transport | 2.90E-05 | 0.004669 |
| *SLC38A2* | membrane | 9.44E-05 | 0.012697 |
| *IGLV-7* | protein binding | 0.000267 | 0.030878 |
| *SAA4* | acute-phase response | 0.000168 | 0.020985 |
| *TUFT1* | protein binding | 0.000302 | 0.033093 |
| *RORC* | cell differentiation | 0.000487 | 0.049286 |
| *HMGCS2* | lipid metabolic process | 6.52E-21 | 1.11E-16 |
| *HSD17B2* | oxidation-reduction process | 4.77E-17 | 1.02E-13 |
| *LUZP2* | leucine zipper protein 2 | 8.16E-09 | 3.87E-06 |
| *TGM6* | peptide cross-linking | 2.58E-05 | 0.004315 |



**Table 4.5** List of DEGs in the *longissimus dorsi* muscle tissue of JNP and Berkshire (FDR<0.05).

| Gene | Function | *P*-value | FDR |
|---|---|---|---|
| *PDK4* | protein phosphorylation | 1.67E-08 | 3.51E-05 |
| *PADI2* | cellular protein modification process | 1.19E-09 | 4.99E-06 |
| *ATP1A4* | ATP biosynthetic process | 1.23E-08 | 2.95E-05 |
| *POSTN* | growth | 1.40E-06 | 0.001626 |
| *IGSF10* | protein binding | 5.66E-05 | 0.028004 |
| *HPGD* | oxidation-reduction process | 2.57E-05 | 0.015429 |
| *ARID5B* | muscle organ morphogenesis | 2.75E-05 | 0.015909 |
| *SGK2* | protein phosphorylation | 1.32E-05 | 0.00854 |
| *COL2A1* | skeletal system development | 2.46E-05 | 0.015277 |

### *4.4.2. Clustering of multiple gene list and annotation of the pathways*

Output from the ToppCluster was viewed by the graphical network. For the current study, we selected the biological processes related to the metabolic process, immunity, muscle cell differentiation, organ morphogenesis and oxidation and reduction processes. To provide a dissected gene-level view, we chose the gene ontology of the biological process which is shared by the genes expressed in the liver and muscle genes (Figure 4.4). The clear functional separation was observed through the abstract network analysis. High significance for the sets of genes that bestow the phenotypes is observable through the extensive gene lists of the tissues. The genes from the muscle tissue show extensive relationship enrichments with the genes of liver in single organism biosynthesis and small molecule biosynthesis process. Through ToppCluster we were able to select the terms of interest to be included in the network. This may be very useful, especially when researchers working on it want to further explore only some specific enriched terms in the output (Kaimal *et al*., 2010).

The data was further subjected for the analysis of the GO pathway in details through the KEGG. The results from KEGG revealed that, the differentially expressed genes were mainly involved in the metabolic pathway, TCA cycle, cytokine-cytokine receptor interaction, folate biosynthesis, protein biosynthesis, antigen-processing and presentation, Notch signaling pathway and NF kappa biosynthesis. The detailed GO of biological process terms (Figure 4.4) revealed significant up-regulation of genes for cell-cell signaling lipid biosynthesis, limb development, hormone metabolism and skeletal muscle system. These have been associated with meat quality and body growth in JNP (Kouba *et al*., 2003; Rehfeldt and Kuhn, 2006; Kim *et al*., 2009).



### *4.4.3. Functional explanations of GO and expression levels of DEGs*

The expounded analysis of the biological pathways and the GO terms of the potentially associated genes was conducted. The results revealed that DEGs are significantly associated with the skeletal system morphogenesis and collagen fibril organization, metabolic process, ion binding and innate immunity, oxidation-reduction process, smooth muscle tissue development and skeletal muscle cell differentiation, protein binding, and metal ion binding. The GO analysis corresponding to the biological process of the different genes is given in the table 4.6. Simultaneous up regulation of GO, biological and metabolic processes in liver with multicellular organismal processes were analyzed by functional annotation. The GOs for cellular components and molecular function of DEGs in two tissues of JNP and Berkshire pigs (Table 4.4 and 4.5)

have been summarized. The GO of cellular components for specific DEGs is related to the endoplasmic reticulum and proteinaceous extracellular matrix in liver and muscle, respectively. Further, the molecular function GO of specific DEGs was related to cofactor binding in liver tissues. Cholesterol and sterol metabolic processes were the most significantly enriched groups in liver tissue ($P$ = 4.23-E10 and $P$ = 7.07E-11, respectively) (Table 4.4). Development and morphogenesis of skeletal system were the most significantly enriched groups in muscle tissue ($P$ = 2.46E-05 and $P$ = 2.75E-05, respectively) (Table 4.5). In addition, the biological process GO of specific up-regulated DEGs in two tissues of JNP and Berkshire pigs have been extensively analyzed (Figure 4.4). Metabolic processes were the most significantly enriched term in liver of both breeds ($P$ = 0.001and 0.03 in JNP and Berkshire respectively) whereas developmental processes were the most significantly enriched term in muscle from JNP ($P$ = 0.02).

The levels of expression of DEGs among the JNP and Berkshire were visualized by the heat maps in the current study (Figure 4.5). It is the distinct, clear cut observation that the genes have significantly ($P$<0.05) differential expressions between the two breeds. The significant differential expression levels of the genes in two breeds have been also analyzed through qRT-PCR and western blotting. The qualitative and quantitative differential expression of the genes have indicated the significant findings (Table 4.7 and 4.8); Figure 4.6 and Figure 4.7).



Till now, expression profiling of genes related to the biological pathways existing in different tissues and developmental stages of JNP have not been studied. That's why, through current study, an aim to investigate most profuse biological pathways and the expression profiles of those genes which caused the variations between JNP and Berkshire was set. However, the previous research has shown that a much larger fraction of the transcriptome variability was found in the tissue difference rather than the breed (Zhang *et al*., 2004; Shyamsundar *et al*., 2005; Ferraz *et al*., 2008). Therefore, our results give a new vision towards an understanding and identification of the candidate genes that are involved in the alteration of the biological functions.

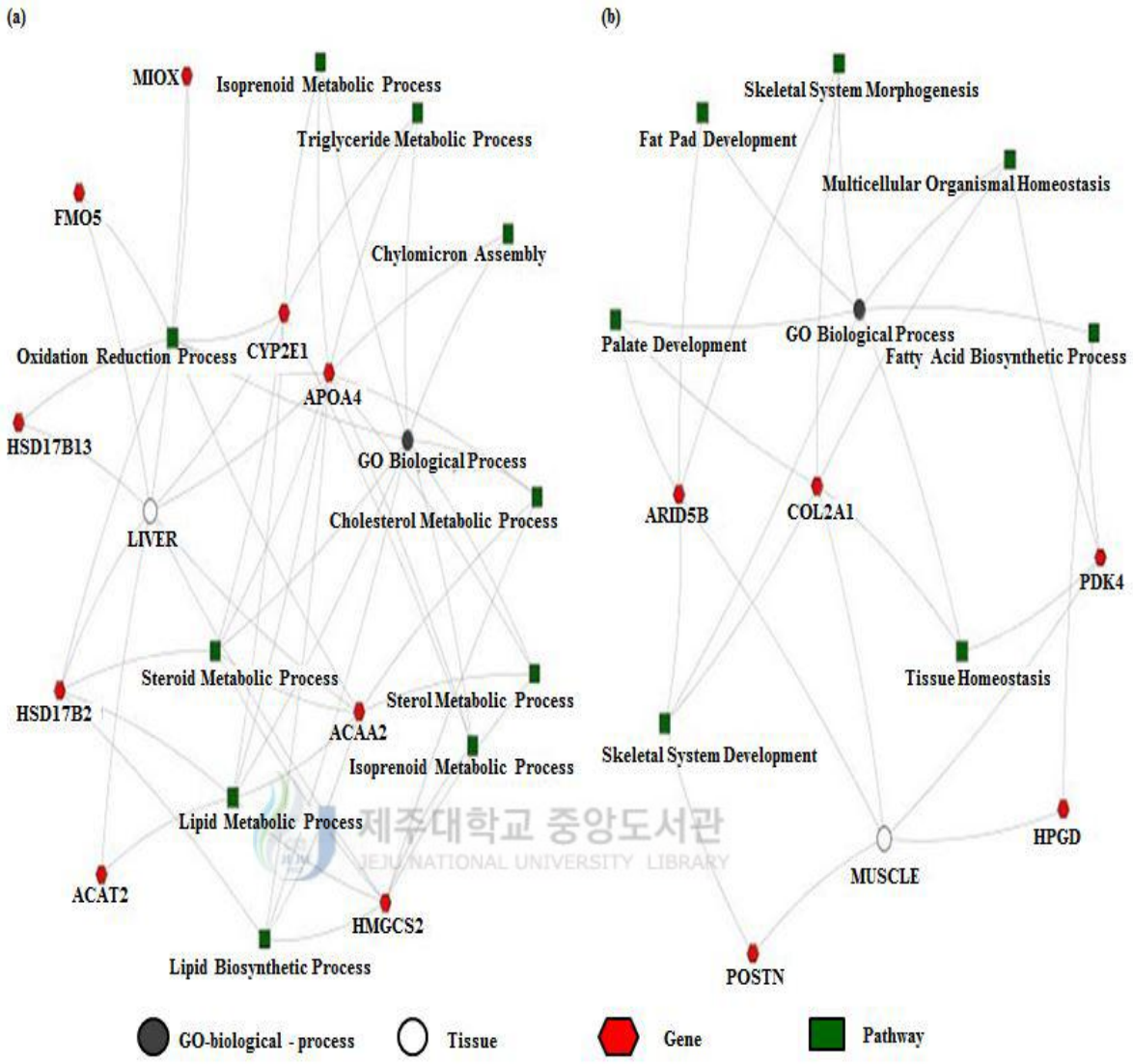


**Figure 4.4**. Dissected gene level view sharing GO of biological process shared by the differentially expressed genes in liver and *longissimus dorsi* muscles of JNP and Berkshire. An abstracted network is showing biological pathways enriched with the cluster of differentially expressed genes specific for (a) liver and (b) *longissimus dorsi* muscle.

**Table 4.6** Gene ontology analysis related to the biological process of the differentially expressed genes in liver and *longissimus* dorsi muscle tissue of Berkshire and JNP.

| Gene name | Go term |
|---|---|
| **(a) Liver** | |
| CYP2E1, FMO5, HSD17B13, MIOX, HSD17B2 | Oxidation-reduction process |
| ACAA2, ACAT2, CACNA2D1, HMGCS2 | Metabolic process |
| APOA4 | Innate immune response |
| MT1A, SLC13A5 | Metal ion binding |
| SOCS2 | Regulation of growth |
| **(b) *Longissimus dorsi* muscle** | |
| PDK4 | Protein phosphorylation |
| POSTN, MyBPH | Growth |
| IGSF10 | Protein binding |
| HPGD | Oxidation-reduction process |
| ARID5B | Muscle organ morphogenesis |
| COL2A1 | Skeletal system development |

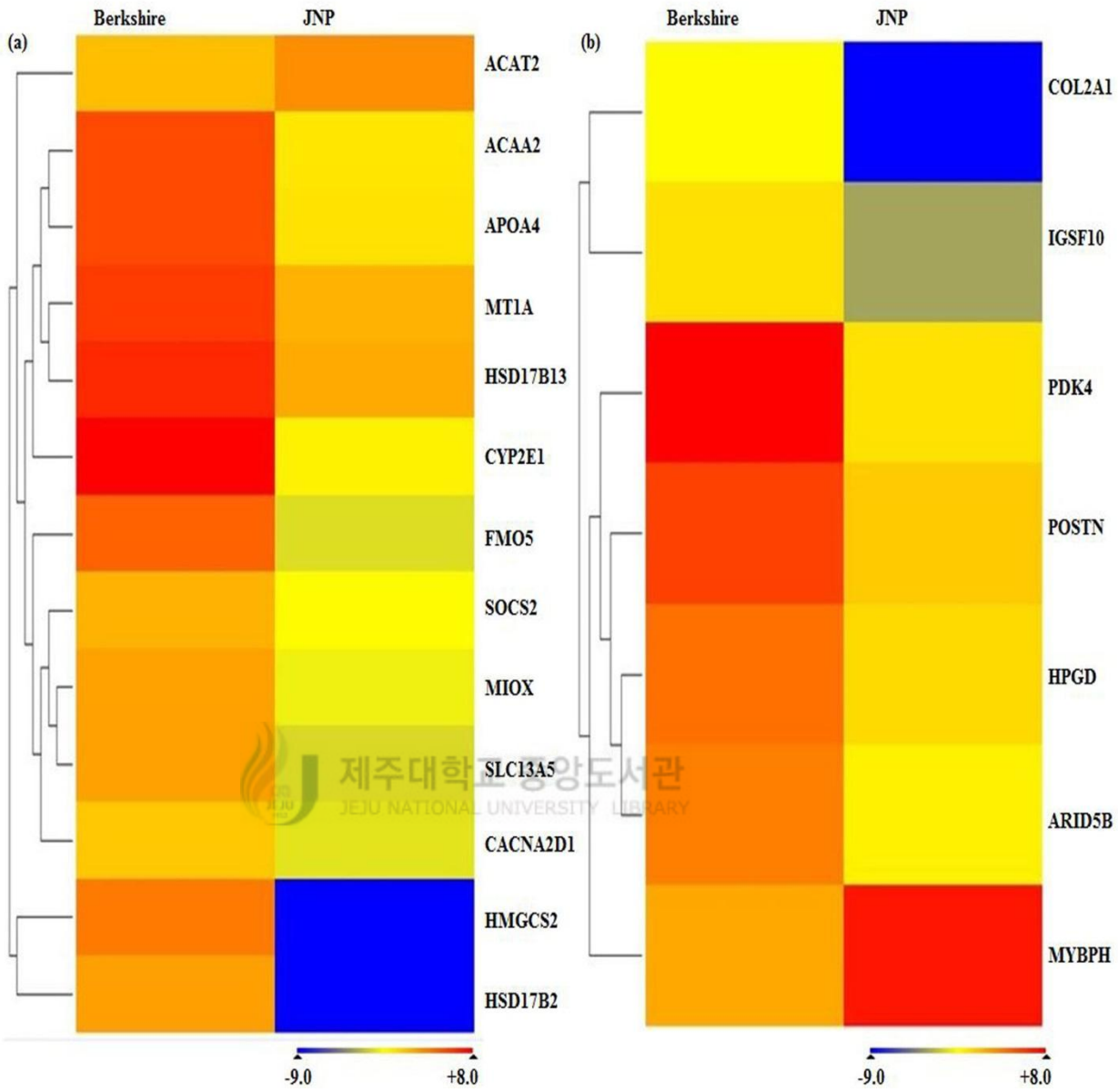




**Figure 4.5**. Heat maps illustrating the level of the differentially expressed genes in (a) liver and (b) *longissimus dorsi* muscle tissue of adult JNP and Berkshire. The red blocks represent the over expressed genes and the blue blocks represented the lowest expression level of the genes.

**Table 4.7** The expression of the genes and the log fold changes in liver tissue of adult Berkshire and JNP.

| Gene | Berkshire adult liver tissue* | JNP adult liver tissue* | Fold change+ |
|---|---|---|---|
| *CYP2E1* | 11.81 | 2.08 | 846.66 down |
| *FMO5* | 7.85 | 0.19 | 202.19 down |
| *MIOX* | 5.31 | 0.97 | 20.27 down |
| *HSD17B2* | 5.42 | -8.73 | 18288.61 down |
| *ACAA2* | 8.86 | 2.50 | 81.59 down |
| *HSD17B13* | 10.16 | 4.99 | 35.85 down |
| *CACNA2D1* | 3.78 | 0.46 | 10.02 down |
| *HMGCS2* | 6.87 | -8.73 | 49920.91 down |
| *APOA4* | 8.83 | 2.72 | 69.00 down |
| *MT1A* | 9.45 | 4.63 | 28.166 down |
| *SLC13A5* | 5.33 | 0.05 | 38.82 down |
| *SOCS2* | 4.66 | 1.72 | 7.65 down |
| *ACAT2* | 4.14 | 6.11 | 3.91 up |



* The level of the expression unit is measured as $\log_2$ total RPKM.

+ The fold change has been calculated for JNP with respect to Berkshire as a reference.

**Table 4.8** The expression of the genes and the log fold changes in *longissimus dorsi* muscle tissue of adult Berkshire and JNP.

| Gene | Berkshire adult muscle tissue* | JNP adult muscle tissue* | Fold change+ |
|---|---|---|---|
| *PDK4* | 9.87 | 1.53 | 324.00 down |
| *POSTN* | 7.47 | 2.39 | 33.85 down |
| *MYBPH* | 3.64 | 9.12 | 44.53 up |
| *IGSF10* | 1.58 | -2.86 | 21.83 down |
| *HPGD* | 5.75 | 1.84 | 14.95 down |
| *ARID5B* | 5.21 | 0.99 | 18.62 down |
| *COL2A1* | 0.63 | -8.94 | 765.95 down |

* The level of the expression unit is measured as $\log_2$ total RPKM.

+ The fold change has been calculated for JNP with respect to Berkshire as a reference.

### *4.4.4. Elucidation of the genes bounded with the metabolic process*



The lipid and their metabolites act as transcriptional regulators of gene expression and fundamentally affect physiological processes such as growth and cell or tissue differentiation. Extracellular nutrients and intracellular metabolite concentrations regulate cell growth. Acetyl-CoA Acyltransferase-2 *(ACAA2),* Apolipoprotein- A4 *(APOA4),* Cytochrome- P2E1 *(CYP2E1)* and Hydroxy methyl glutaryl CoA synthetase 2 *(HMGCS2),* genes are related to hepatic lipid, high density lipids (HDL) and growth hormone (Daniels, 2009; Uddin *et al*., 2011; Graugnard *et al*., 2012; Morlein *et al*., 2012). These genes are down-regulated in the liver of JNP (Table 4.7), except Acetyl-CoA Acetyltransferase-2 (*ACAT2)* which is significantly ($P$<0.05) up-regulated in JNP (Figure 4.6 and Figure 4.7). Moreover, *CACNA2D1* and *HMGCS2* genes in liver tissue are also found to have higher levels in Berkshire breed (Table 4.7). It has been found that in contrast to JNP, maximum number of genes related to the metabolic process, are expressed in the Berkshire. The interactions due to nutrients, metabolites, gene expressions and protein modifications are involved in the coordination of cell growth with extracellular and intracellular conditions (Yuan *et al*., 2013).

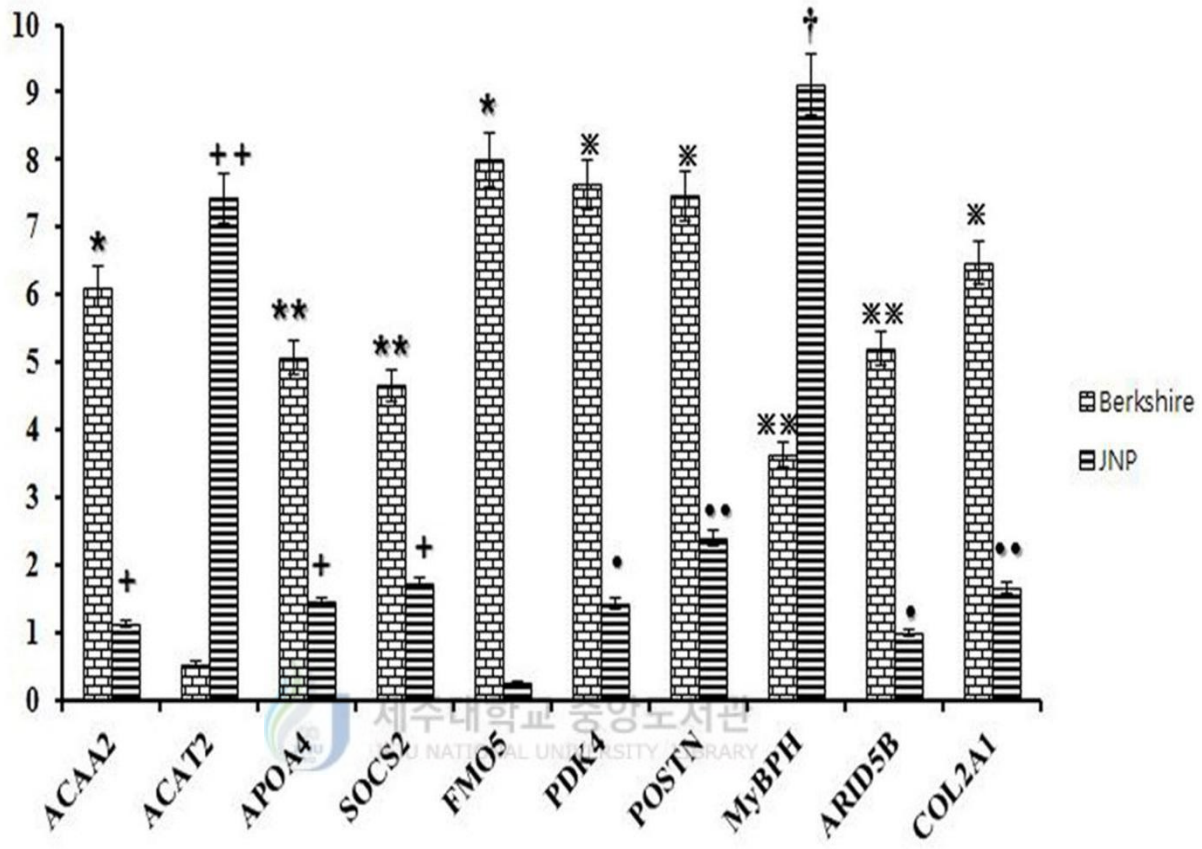


**Figure 4.6**. Relative transcript levels (RQ) of DEGs expressed in the liver and *longissimus dorsi* muscle of JNP and Berkshire. Bars with different superscripts show significant differences in the transcript levels between the two breeds ($P < 0.05$). Values are expressed as mean ± SEM.

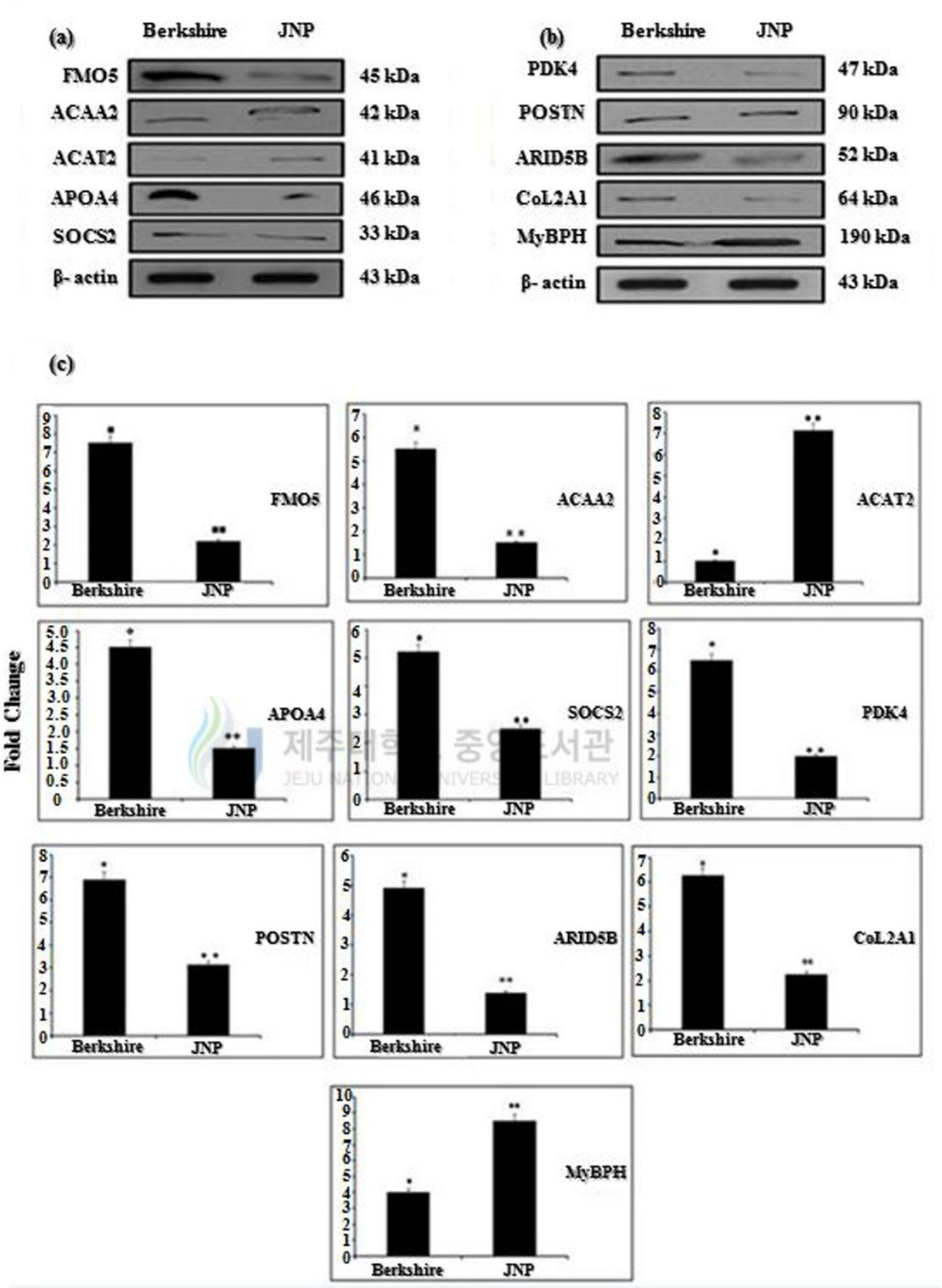


**Figure 4.7**. Relative differential blot expression analysis of proteins expressed in the (a) liver and (b) *longissimus dorsi* muscle of JNP and Berkshire.

Therefore, it can't be discerned from the present data, whether the expressions of the genes are directly related to more growth in Berkshire than JNP.

### *4.4.5. Exposition of genes involved in skeletal system morphogenesis and body growth*

We have found 11 genes related to meat quality and body growth. Most of these genes were up-regulated in Berkshire pigs. Pyruvate dehydrogenase kinase-4 (*PDK4),* found in skeletal muscles is associated with meat quality (Lan *et al*., 2009) and is up-regulated in Berkshire. Periostein *(POSTN)* and myosin binding protein- H (*MyBPH),* controlled fat deposition (Chen *et al*., 2011) and muscle growth (Bilek *et al*., 2008) respectively. *POSTN* and *MyBPH* responsible for body growth have shown contrast results. *POSTN* is significantly down regulated whereas *MyBPH* is significantly ($P$<0.05) up-regulated in JNP (Table 4.7 and 4.8); Figure 4.6 and Figure 4.7). *MyBPH* is related to myosin in muscle. Myofibrils contain actin and myosin as the predominant protein in thin and thick filaments (Zhang, 2009). Myosin is the major protein that comprises the thick filament and affects the muscle fiber type i.e. myosin heavy chains. The muscle fiber type is responsible for water content, color and eating quality of pork. In the current study, it has been found that *MyBPH* is up-regulated (9.12 fold) in JNP which may have affected the meat quality of JNP. It has also been reported that collagen type 2- alpha 1 (*COL2A1)* gene which is related to the skeletal system morphogenesis and growth of cartilage (Lefebvre and Smits, 2005; Kim *et al*., 2007), is down-regulated in JNP (Table 4.8; Figure 4.6 and Figure 4.7). Even the validation of *COL2A1* by quantitative real time PCR and western blotting (Figure 4.6 and Figure 4.7) has also complemented the findings of *in-silico* analysis. This may provide evidence that in Berkshire body growth is more efficient than JNP.



### *4.4.6. Portrayal of genes involved in inflammatory and innate immune response*

In the current study, *APOA4* from liver tissue is reported to be involved with the immune response has higher levels (Table 4.7) in Berkshire breed. This gene contributes to the stronger immune system of the breed. Findings of GO analysis acted as base for further validation of the results. The findings of quantitative real- time RT-PCR and protein band expression through western blotting have also complemented the earlier findings with the conformation about the up-regulation of *APOA4* in Berkshire liver tissue. Pashaj *et al*. (2013) reported that GO analysis of liver revealed significant enrichment of GO categories related to immune response, stress, lipid metabolism and carboxylic acid metabolic processes.

Divergent phenotypes of both growth performance and immune response are significantly dependent upon the functional transcript changes throughout in the immune response (Adler *et al*., 2013). The direct relation between the responsiveness and performance in the pig breeds has been very well reported. It is observed that selection for high immune response is associated with the enhanced weight gain (Mallard *et al*., 1998; Clapperton *et al*., 2006). Our current findings suggest that the gene/s those are differently expressed in the JNP and Berkshire in terms of the immune response can possibly be used as the selection marker for lean growth associates with higher numbers of lymphocytes and monocytes in JNP.

### *4.4.7. Depiction of protein binding and phosphorylation related genes*

Phosphorylation plays a vital role in regulation of protein binding. It coordinates the pathways by modulating the nature and strength of protein- protein interactions. In the present study we identify the genes from the sample of two breeds those are responsible for protein binding and phosphorylation. It was found that *PDK4* and *IGSF10* genes (Table 4.6) have been significantly up- regulated in Berkshire (Table 4.8). Validation of expression profiles through qRT-PCR and western blotting also confirmed significantly ($P$<0.05) higher expression of *PDK4* gene and protein respectively in muscle tissue of Berkshire (Figure 4.6 and 4.7).



*PDK4*, a mitochondrial enzyme, is involved in aerobic oxidation of carbohydrate fuels i.e. glucose and fatty acid oxidation (Majer *et al*., 1998; Boulatnikov and Popov, 2003). A complete cDNA cloning and sequencing of *PDK4* has revealed a G/A mutation in intron 9, which is reported to have a significant association with intramuscular fat and water content (Lan *et al*., 2009). These findings indicate towards the high intramuscular fat content in Berkshire.

### *4.4.8. Interpretation of genes involved in oxidation reduction process*

In the course of investigation, pertaining to the genes responsible for the oxidation and reduction process in the two tissue samples from the both breeds, quite interesting results have been indicated. It was found that six genes related to the oxidation and reduction process have differently expressed in both of the breeds. Out of those genes, five genes such as *CYP2E1, FMO5, MIOX, HSD17B2* and *HPGD* expressed significantly more in the Berkshire (Table 4.7 and 4.8; Figure 4.6 and 4.7).

*CYP2E1* is involved in the metabolism of endogenous compounds including fatty acids and ketone bodies (Novak and Woodcroft, 2000). Moreover, *CYP2E1* may have an important role

in the pathogenesis of high-fat mediated non-alcoholic steatohepatitis (NASH) (Abdelmegeed *et al*., 2012). *FMO3* is known to play a significant role in human hepatic metabolism. *FMO5* a member of *FMO* group of genes is reported to be associated with fishy off flavor and pork pH (Glenn *et al*., 2007). Studies on the knockout mice also revealed the vital role of *FMO5* in regulating the endogenous fat and cholesterol metabolism.

### 4.5. Conclusion

JNP has unique flavor with an unmatched juiciness and tenderness for pork. Unlike usual pork or the white meat pork, Berkshire and JNP pork are distinct in their traits with respect to palatability. Moreover, close relation of Berkshire to Asian native breeds made it breed of choice for the current comparative study with JNP (Kim *et al*., 2002). The study revealed transcriptomic analysis of liver and muscle tissues of adult JNP and Berkshire pigs by using RNA-seq. This comparative transcriptome analysis showed that most genes involved in metabolism; immune response and protein binding were up-regulated and significantly enriched in liver tissue. Hence, in order to investigate the genetic profiles of liver and muscle tissues transcriptome analysis can make the way easier to understand the difference of the genetic mechanisms at JNP in respect of Berkshire. Transcriptome analysis enables the identification of specialized biological functions and regulatory gene/s which would be used as the selection of marker for further breed improvement programmes. Further extensive *in vitro* and *in vivo* studies can be planned for the regulated expression of genes enlisted by the current study.

# Chapter 5

# An approach to identify SNPs in the gene encoding acetyl-CoA acetyltransferase-2 (*ACAT-2*) and their proposed role in metabolic processes in pig

## 5.1. Abstract

The novel liver protein acetyl-CoA acetyltransferase-2 (ACAT2) is involved in the beta-oxidation and lipid metabolism. Its comprehensive relative expression, *in silico* non-synonymous single nucleotide polymorphism (nsSNP) analysis, as well as its annotation in terms of metabolic process with another protein from the same family, namely, acetyl-CoA acyltransferase-2 (ACAA2) was performed in *Sus scrofa*. This investigation was conducted to understand the most important nsSNPs of ACAT2 in terms of their effects on metabolic activities and protein conformation. The two most deleterious mutations at residues 122 (I to V) and 281 (R to H) were found in ACAT2. Validation of expression of genes in the laboratory also supported the idea of differential expression of *ACAT2* and *ACAA2* conceived through the *in silico* analysis. Analysis of the relative expression of *ACAT2* and *ACAA2* in the liver tissue of Jeju native pig showed that the former expressed significantly higher ($P<0.05$). Overall, the computational prediction supported by wet laboratory analysis suggests that ACAT2 might contribute more to metabolic processes than ACAA2 in swine. Further associations of SNPs in ACAT2 with production traits might guide efforts to improve growth performance in Jeju native pigs.



***Keywords*:** acetyl-CoA acyltransferase-2 (*ACAA2*), acetyl-CoA acetyltransferase-2 (*ACAT2*), liver, bovine, Jeju native pig (JNP)

## 5.2. Introduction

The liver protein acetyl-CoA acyltransferase-2 (ACAA2), contributes to metabolism in two ways: it catalyses the last step of beta oxidation of mitochondrial fatty acids and is responsible for lipid catabolism. Apart from ACAA2, the ACAT2 protein is also expressed in the liver. The latter has two isoforms, ACAT1 and ACAT2, with distinct metabolic functions, intracellular localisations and membrane topologies in different mammals (Anderson *et al.*, 1998; Joyce *et al.*, 2000). ACAT2 belongs to the same protein family as ACAA2 and both are present on chromosome 1 in pig.

Liver is one of the most important organs for metabolism and the partitioning of nutrients. It is involved in the transformation of dietary nutrients into fuel and their export via the blood. To meet the changing demands of extra-hepatic tissues for nutrients, liver displays remarkable metabolic flexibility (Baik *et al.*, 2009). The ACAT2 has a unique tissue distribution, being predominantly expressed in the liver and intestine (Parini *et al.*, 2004). It plays an important role in lipid and cholesterol metabolism in human beings and bovines. The high variability of ACAT2 was observed among the human liver samples and suggested that it is a regulated enzyme that is responsible for more than 50% of the ACAT activity in the majority of human population (Parini *et al.*, 2004). Another study, conducted using broiler chickens, investigated the expression of *ACAT2* in liver and revealed that this gene has been subjected to selection to either promote or suppress the accumulation of abdominal fat (Wang *et al.*, 2005). A correlation between *ACAT2* and cholesterol has also been reported in fish (Ushio and Nagasaka, 2013) and cholesterol has further been associated with increased lean mass in human beings (Riechman *et al.*, 2007). Finally, ACAT2 controls enzymes involved in lipogenesis or lipolysis in beef cattle (Jiang *et al.*, 2009).

Moreover, analyses of single-nucleotide polymorphisms (SNPs) provide important information about genetic linkages and are useful in fine-mapping of the regions of candidate genes. SNPs not only have attracted substantial attention in terms of human and vertebrate haplotypes associated with traits of interest, but they have also increased our understanding of the genetic basis of phenotypic diversity within and between populations (Guryev *et al.*, 2004; Wong *et al.*, 2004) Owing to their abundance in the genome, thousands of potentially informative SNP markers can be identified for the development of high-density SNP maps (Zimdahl *et al.*, 2004). SNPs detected in the bovine ACAA2 are reported to be associated with gain in daily weight and loin muscle area ($P<0.05$) significantly (Li, 2008). Further, SNPs in the ovine *ACAA2* gene are reported to be associated with milk production traits (Orford *et al.*, 2012). Moreover, SNPs

associated with production traits were discovered in the *ACAT2* gene of Holstein cattle (Cochran *et al.*, 2013).

Given the importance of the SNPs of *ACAT2* gene in metabolic processes, our study comprised two aspects. First, we conducted comprehensive *in silico* non-synonymous single nucleotide polymorphism (nsSNP) analysis and predicted the possible effects of these nsSNPs on protein structure and function. In the second stage, we explored the level of involvement of the *ACAT2* gene relative to that of *ACAA2* gene in metabolic processes in *Sus scrofa*.

## 5.3. Materials and methods

### *5.3.1. Database retrieval of information on ACAA2 and ACAT2 genes*

Data on the *ACAA2* and *ACAT2* genes from *Sus scrofa* and *Bos taurus* were collected from Entrez Gene of the National Centre for Biotechnology Information (NCBI). The SNP ids and types of SNPs related information were obtained from the NCBI dbSNP (http://www.ncbi.nlm.nih.gov/snp/) and SWISSProt databases (http://expasy.org/) for computational analysis. The online tool Treefam (http://www.treefam.org/) was used to analyze the phylogenic relationships between the *ACAA2* and *ACAT2* genes. A bootstrap threshold of 50% was used for the inclusion of organisms (bovine) in the same clade as Sus *scrofa.* It was followed by multiple sequence alignment to compare the sequences of the *ACAA2* and *ACAT2* from both bovine and swine.

### *5.3.2. Annotation of Gene Ontology (GO) in terms of functional classes*

GO analysis has become a commonly used approach for functional studies of large-scale genomic or transcriptomic data. In the current study, GO analysis has been used to analyse the biological processes associated with the functioning of the *ACAA2* and *ACAT2* genes in *S scrofa* and *B taurus*. CateGOrizer (previously known as "GO terms classifications Counter"; www.animalgenome.org/bioinfo/tools/countgo/) (Hu *et al.*, 2008); was used as a common language tool to annotate *ACAA2* and *ACAT2* in *S scrofa* and *B taurus*.

### *5.3.3. Validation of the expression levels of ACAA2 and ACAT2 in Sus scrofa and Bos taurus liver tissues*

#### *5.3.3.1. Ethics Statement and collection of tissue samples*

Pure bred adult animals from Jeju native pigs (JNPs) at an average body weight of 84.7±3.5 kg were used for the study. The study was carried out in strict accordance with the recommendations of the animal bioethics committee of Jeju National University, Jeju-Si, Jeju-Do, Republic of Korea. Bioethics committee specifically approved this study vide permit number: 2013-0009. Animals were reared under same environmental and nutritional conditions. The commercial feed (Seoul Feed, Jeju- Si, South Korea) and water were offered *ad- libitum* to all the pigs. Pigs were housed in pens with concrete flooring, a nipple bowl drinker, feeder and at an ambient temperature maintained at 25±1℃. *Bos taurus* liver samples were collected from the Jeju joint livestock products market slaughter house, Jeju-Si, Republic of Korea. Five female animals of each species were used for the study. The slaughtering method entailed exsanguination following electric stunning. The liver tissues were dissected immediately after slaughter and samples were frozen in liquid nitrogen. The tissue samples were subsequently stored in the laboratory at -80 ℃ until the isolation of RNA and protein.

#### *5.3.3.2. RNA and protein isolation from liver tissue*

RNA was isolated from 100 mg of the fragmented frozen liver tissue samples of JNPs and bovines. *TRIzol*™ reagent (Invitrogen, USA) was used to isolate RNA. Tissue samples were homogenised in 1.5 ml of TRIzol reagent and chloroform and the RNA extracted was subsequently precipitated using isopropanol (Junsei Chemical Co. Ltd., Japan). Isolated RNA samples were stored at -80 ℃. Genomic DNA contamination was removed by treating 25 μg of RNA from each sample with the RNase-free DNase set (QIAGEN, Hilden, Germany) and purification with the RNeasy mini kit, in accordance with the user guidelines (QIAGEN, Hilden, Germany). A Bioanalyzer 2100 with RNA 6000 Nano Labchips was used to assess the quality and quantity of RNA by automated capillary gel electrophoresis, in accordance with the user guidelines (Agilent Technologies, Dublin, Ireland). 28S/18S ratios for the RNA samples ranged from 1.8 to 2.0. Superscript III enzyme (Invitrogen, USA) was then used to synthesise first-strand cDNA.

Radio Immuno Precipitation Assay (RIPA) buffer was used to extract protein from the homogenised liver tissue. The manufacturer's guidelines were followed for measurement of the concentration of protein with the Pierce™ BCA protein assay kit (Thermo Scientific, USA) in a Bio-Rad Micro-plate Reader (Model-680). All the RNA and protein samples were stored at -80 ℃ until use.

#### *5.3.3.3. RT-PCR and quantitative real-time PCR*

Primers used in RT-PCR and quantitative real-time PCR were designed using the online Primer-3 program (Rozen and Skaletsky, 2000) and the primers that were used to conduct PCR along with their detailed information are listed in Table 5.1. An Eppendorf Mastercycler gradient instrument was used to conduct RT-PCR under the following conditions: an initial 5 min at 94 ℃; 35 cycles of 30 seconds at 94 ℃, 30 seconds at the $T_a$ temperature indicated in Table 5.1 and 1 min at 72 ℃; followed by a single extension period of 5 min at 72 ℃. The results of RT-PCR are presented as the relative expression normalised, using *GAPDH* transcript as an endogenous reference.

To evaluate the relative expressions of *ACAA2* and *ACAT2* genes quantitatively in *Sus scrofa* and *B taurus*, real-time qRT-PCR was conducted using an Applied Biosystems, Step-One Real Time PCR system. The dye EvaGreen (Biotium, USA) was used to determine the quantity of transcript of target genes present in each sample. Each individual sample was quantified in triplicate under the following amplification conditions: 95 ℃ for 10 min initially and then 40 cycles of 95 ℃ for 15 sec and 60 ℃ for 1 min. Standard curve methods were used to define the efficiency of real-time PCR. The efficiency of amplification of the target gene was compared with that of the endogenous *GAPDH* control transcript (Wang *et al.,* 2003). Quantification of mRNA levels was performed using the comparative $C_T$ method. The results are reported as the relative expression normalised using the level of the transcript of the endogenous reference (Van *et al.,* 2001; Erkens *et al.,* 2006).

#### *5.3.3.4. Western blotting analysis*

Sixty micrograms of protein extract was diluted 1:1 with 2X loading buffer (4% sodium dodecyl sulfate (SDS), 20% glycerol, 0.004% bromophenol blue, 25% 0.5 M Tris and 5% β-mercaptoethanol). Protein extracts were denatured by boiling for 5 min before loading onto a 12%

SDS-PAGE gel. Proteins were transferred to nitrocellulose membrane after electrophoresis. The membrane was blocked for two hours at room temperature and incubated with primary and secondary antibodies specific for ACAA2, ACAT2, and β-actin, as listed in Table 5.2.

The washed membranes were analysed to observe specific chemiluminescent signals using a Luminescent Image Analyzer (LAS-4000 mini) instrument. The results are depicted as the relative band intensities normalised relative to the band intensities of β-actin bands for each gene using Image J software (National Institute of Health, Bethesda, Maryland, USA). The means were compared between JNPs and bovines.

### *5.3.4. Functional consequences of the coding of nsSNPs of ACAT2 of Sus scrofa using the sorting* **intolerant from** t*olerant (SIFT) tool*

Studies on SNPs of ACAT2 (*Sus scrofa*) were performed to identify amino acid substitutions which effect metabolic activities and protein conformations. SIFT (available at http://www.blocks.fhcrc.org/sift/SIFT.html) uses sequence homology to predict whether an amino acid substitution affects protein function; if it does, it can be prioritised for further study. In this study, by using SIFT, we submitted SNP queries, selecting those with an intolerance score below a certain threshold for further study. This threshold was fixed at 0.005 or less. SIFT scores were classified as indicating that the SNP would be damaging (0.00–0.05), potentially damaging (0.051–0.10), borderline (0.101–0.20), or tolerable (0.201–1.00) (Hussain *et al.*, 2012).



### *5.3.5. Functional annotation of nsSNPs of ACAT2 (Sus scrofa) using protein variation effect analyzer (PROVEAN)*

PROVEAN is an online software tool that predicts the possible impact of amino acid substitutions on the structural and functional properties of proteins. The amino acid sequences of ACAT2 (*Sus scrofa*) associated with identified SNPs through SIFT were submitted as queries. A delta alignment scoring system was used, where the scores of each supporting sequence were averaged within and across clusters to generate the final PROVEAN score. This computational tool helps to predict the functional effect of variation in amino acids by selecting the common deleterious SNPs through both SIFT and PROVEAN. A protein variant is said to be “deleterious” if the final score is below a certain threshold (default is -2.5) or “neutral” if the score is above this threshold.

**Table 5.1** List of primer sequences.

| Gene name | Primer sequences | Product size | Annealing temperature ($T_a$) | Genebank ID |
|---|---|---|---|---|
| Pig | F- TAGGCTCTGTGGCTCTGGTT | 225 | 61℃ | NM_001167638.1 |
| *ACAA2* | R- GTAATTGCCATCGGGATTTG | | | |
| Bovine | F- GTGTTCATCGTCGCTGCTAA | 235 | 63.1 ℃ | NM_001035342.2 |
| *ACAA2* | R- CTGGGGTCTCTTTTGGGATT | | | |
| Pig | F- ATCACCAAGGAGCGAATCC | 245 | 58.5 ℃ | NM_001243427.1 |
| *ACAT2* | R- CCTCTTCTGCTTGTCCCAAC | | | |
| Bovine | F- CGGATTCTTGTCACCCTGTT | 223 | 61 ℃ | NM_001075549.1 |
| *ACAT2* | R- TTTCCACCCTCACTTTGGTC | | | |
| Pig | F- AGAAGGTGGTGAAGCAGG | 170 | 61 ℃ | NM_001206359 |
| *GAPDH* | R- GTCGTACCAGGAAATGAGC | | | |
| Bovine | F- CCACCCAGAAGACTGTGGAT | 127 | 61 ℃ | NM_001034034.1 |
| *GAPDH* | R- TTGAGCTCAGGGATGACCTT | | | |



**Table 5.2** List of primary and secondary antibodies used in western blotting.

| Protein | Primary antibody* | Secondary antibody* |
|---|---|---|
| ACAA2 | Mouse monoclonal 1:50 (catalogue no. sc-100847) | Rabbit anti-mouse, 1:1,000, (catalogue no. sc-358922) |
| ACAT2 | Goat polyclonal 1:50 (catalogue no. sc-30279) | Donkey anti-goat, 1:1,000 (catalogue no. sc-2020) |
| β- actin | Mouse polyclonal 1:500 (catalogue no. sc-2025) | Rabbit anti-mouse, 1:1,000 (catalogue no. sc-358922) |

*****All antibodies were from Santa Cruz.

### *5.3.6. Protein stability prediction using I-Mutant*

I-Mutant (2.0) is a support vector machine based tool (http://folding.biofold.org/cgi-bin/i-mutant2.0.cgi). In this tool the FASTA format of protein sequence of ACAT2 was used as an input to predict the effect of a mutation on its stability.

### *5.3.7. Homologous modelling and prediction of root mean square deviation (RMSD)*

Protein structures were modelled to compare the structural stability of the native and mutant proteins. In the present study, the three-dimensional structure of ACAT2 from *Sus scrofa* was generated by subjecting the amino acid sequence in the FASTA format to analysis using the SWISS-MODEL expasy and ESyPred3D tools. The obtained structural model that was identical between these two tools was selected and subjected to structural validation as described below.

### *5.3.8. Validation of the model*



The constructed homologous protein model was further subjected to evaluation of its internal consistency and reliability. A Psi/Phi Ramachandran plot was generated by PROCHECK analysis. The packing quality of the refined structure was investigated using the PROCHECK Quality Control Check. Furthermore, NOMAD–Ref (Normal Mode Analysis, Deformation, and Refinement) was applied for both the native and mutant models to refine the structures to those with lowest energy. The energy-minimised structures were then used for further analysis for the RMSD calculation.

### *5.3.9. Identification of the stabilising residues using SRide and the effect of mutation on the protein using Project Hope*

The stabilising residues (SRide) tool was used to predict stabilising residues within native and mutant ACAT2 proteins of *Sus scrofa*. A residue is marked as stabilising if it has a high conservation score, high long-range order, high surrounding hydrophobicity, and belongs to a stabilisation centre. The homologous structures of the native and mutant proteins were uploaded as queries. The Project Hope (Have yOur Protein Explained) tool combines the information to give analyses of the effect of a mutation on the protein structure. The protein sequence was used

as the input for selection of the mutant variants. The output was given in the form of the structural variation between mutant and wild-type residues.

### *5.3.10. Data analysis*

The results from the expression analysis by real-time qRT-PCR and western blotting are expressed as the mean ± SEM. The significant differences between the mean expressions of the genes at $P<0.05$ were analyzed by Tukey's b- test.

## 5.4. Results

### *5.4.1. Functional annotation of ACAA2 and ACAT2 using GO*

To investigate the functional similarity in the metabolic processes, genes from same protein family i.e., *ACAA2* and *ACAT2* in *Sus scrofa* and *Bos taurus* respectively were selected for the current study. The phylogenetic trees of *ACAA2* and *ACAT2* were constructed using Treefam. The Treefam showed that in their respective phylogenetic trees the *ACAA2* and *ACAT2* of *B taurus* and *S scrofa*, respectively are present in same clade.

From the GO annotation, both genes participate in metabolic processes. The CateGOrizer analysis of the *ACAA2* gene from the two species showed fractional differences with respect to their GO categories. On the basis of GO categories it was found that 35.7% and 18.18% of GO functions of *ACAA2* gene in *Bos taurus* and *Sus scrofa*, respectively, are engaged in the regulation of metabolic processes (Table 5.3). However, the GO function related to lipid metabolism was absent in *ACAA2* of *S scrofa*. Surprisingly, the experimental evidence supported by GO annotations indicated that *ACAT2* from *Sus scrofa* has functional similarity to *ACAA2* and *ACAT2* from *Bos taurus* in terms of involvement of both genes in metabolic processes. Furthermore, the CateGOrizer analysis indicated that equal fractions (75%) of *ACAT2* are involved in the metabolic processes in the two species (Table 5.3). In addition to functional similarities, expression profiling was also conducted to compare the level of expression within and between the species.

### *5.4.2. Validation of mRNA and protein expressions of ACAA2 and ACAT2 in the liver tissue of Sus scrofa and Bos taurus*

Both RT-PCR and real-time qRT-PCR were performed to clarify the qualitative and quantitative expressions of the genes under study. Specifically, relative mRNA expression levels of *ACAA2* and *ACAT2* were determined after their normalisation relative to the transcript levels of an endogenous reference, *GAPDH*. Here, *ACAA2* exhibited significantly ($P$<0.05) higher expression in bovine liver than was observed in JNP (Figure 5.1) although opposite was true for *ACAT2*. It has also been observed that the levels of expression of *ACAT2* are not significantly different in the two species.

The results of the protein expression for the candidate proteins under study were obtained by western blot and are presented as relative band intensities. These intensities of the candidate proteins have been normalised to the area of intensities of β-actin bands. The results from western blot analysis are in line with the results from the relative quantification of mRNA expression levels. Whereas ACAA2 protein was expressed at a significantly ($P$<0.05) higher level in bovine liver than in JNP (Figure 5.1), but vice versa expression was observed for ACAT2.



### *5.4.3. Retrieval of SNPs and prediction of deleterious SNPs*

The present study was carried out to investigate the nsSNPs in the *ACAT2* gene of *Sus scrofa*. Five missense variants, two splice variants and 14 synonymous SNPs were identified (Xuelian, 2007). For further analysis, only the five nsSNPs were selected. The consequences of an amino acid substitution on protein function were estimated using the SIFT tool. It was found that the two nsSNPs, rs321479200 and rs342236888 of ACAT2 were categorised as borderline (0.101-0.20) and tolerable (0.201-1.00) on the basis of their index score i.e. 0.15 and 0.35 respectively (Table 5.4). In terms of the nucleotide variations, out of nsSNPs, one nsSNP involved a T>C change, another involved a T>G change and three nsSNPs involved C>T changes.

To predict any deleterious effects, the five nsSNPs were furthermore submitted to PROVEAN. On the basis of their scores (-3.587 and -4.012, respectively), the nsSNPs, rs342236888 and rs321479200 of ACAT2 were predicted to be deleterious (Table 5.4). Analysis of the structural stability associated with these mutations was carried out using I-Mutant; their destabilising effects were evaluated by subtracting the unfolded Gibbs free energy value of the native protein from that of the mutated one.

### *5.4.4. Quality analysis of the stability of the homologous model*

The 3D structures of the native and mutant proteins were modelled using the SWISS- MODEL expasy and ESyPred3D tools. Both tools gave the same 3D structure using the template PDB ID 1wl 4 of chain A (Figure 5.2). The validation by PROCHECK verified that 90% of the residues are in the most favoured region in the Ramachandran plot (Figure 5.3), which indicates the good quality of the model. The QMEAN-Z score (-0.51) for the query model of *ACAT2* indicated the high level of accuracy of the constructed model (Figure 5.4). In the current study, the total free energy and RMSD values of a nsSNPs i.e. rs342236888 of ACAT2 were found to be –17.531 and 0.82 A°, respectively, whereas those of rs321479200 of ACAT2 were reported to be –16.837 and 2.63 A°. This value of 2.63 A° suggests that rs321479200 is a highly deleterious nsSNP.

The stabilising residues in the native and mutant protein models were identified using the SRide tool; variations in these stabilising residues are shown in Table 5.5. The 13 stabilising residues, those in the native protein were compared with their counterparts in the mutants. The stability of the native ACAT2 protein was compared with its mutated homologous structures (Figure 5.5 and 5.6) showing the causative mutations i.e. I122V and R281H. It was found that nsSNP-rs321479200 (R281H) was associated with more residues having a stabilising effect, which may be important in terms of structure and function.



### *5.4.5. Effect of mutation on protein function*

The possible effects of mutations at specific residues on protein function were also evaluated. Interestingly, both the native residues, namely, isoleucine and arginine, of ACAT2 protein of *Sus scrofa* are part of an interpro domain named thiolase (IPR002155). This domain was classified into the following GO categories: 0008152 (metabolic process), 0016747 (transferring acyl groups other than amino-acyl groups) and 0016740 (transferase activity). These same residues, isoleucine and arginine, are also part of an interpro domain called thiolase-like, subgroup (IPR016038), with the GO categories of 0003824 (catalytic activity) and 0008152 (metabolic process). These residues are also part of a third interpro domain named thiolase, N-terminal (IPR020616), which plays a role in metabolic processes (GO: 0008152), transferase

**Table 5.3** Gene ontology analysis suggests the biological processes associated with functioning of the *ACAA2* and *ACAT2* genes in *Sus scrofa* and *Bos taurus.*

| GO Class ID* | Definitions | Counts** | | Fractions*** | |
|---|---|---|---|---|---|
| | | ***Sus Scrofa*** | ***Bos taurus*** | ***Sus Scrofa*** | ***Bos taurus*** |
| ***ACAA2*** | | | | | |
| GO:0006915 | apoptosis | 4 | 4 | 36.36% | 28.57% |
| GO:0008152 | metabolism | 2 | 3 | 18.18% | 21.43% |
| GO:0006629 | lipid metabolism | NA | 2 | NA | 14.29% |
| GO:0016265 | death | 2 | 2 | 18.18% | 14.29% |
| GO:0005739 | mitochondrion | 2 | 2 | 18.18% | 14.29% |
| GO:0042981 | regulation of apoptosis | 1 | 1 | 9.09% | 7.14% |
| ***ACAT2*** | | | | | |
| GO:0008152 | metabolism | 2 | 2 | 50.00% | 50.00% |
| GO:0005739 | mitochondrion | 1 | NA | 25.00% | NA |
| GO:0006629 | lipid metabolism | 1 | 1 | 25.00% | 25.00% |
| GO:0009058 | biosynthesis | NA | 1 | NA | 25.00% |



***** **Go terms/ class IDs:** GO terms into each ancestral term are independent of one another. The counted results may be selectively used and percentages recalculated as the selected terms represent a well-covered spectrum of the scope which helps to avoid redundancy and may make best sense for the data set.

****** **Counts :** Number of similar GOs with the current GO

******* **Fractions** **:** Share of biological process shown with respect to GO functions

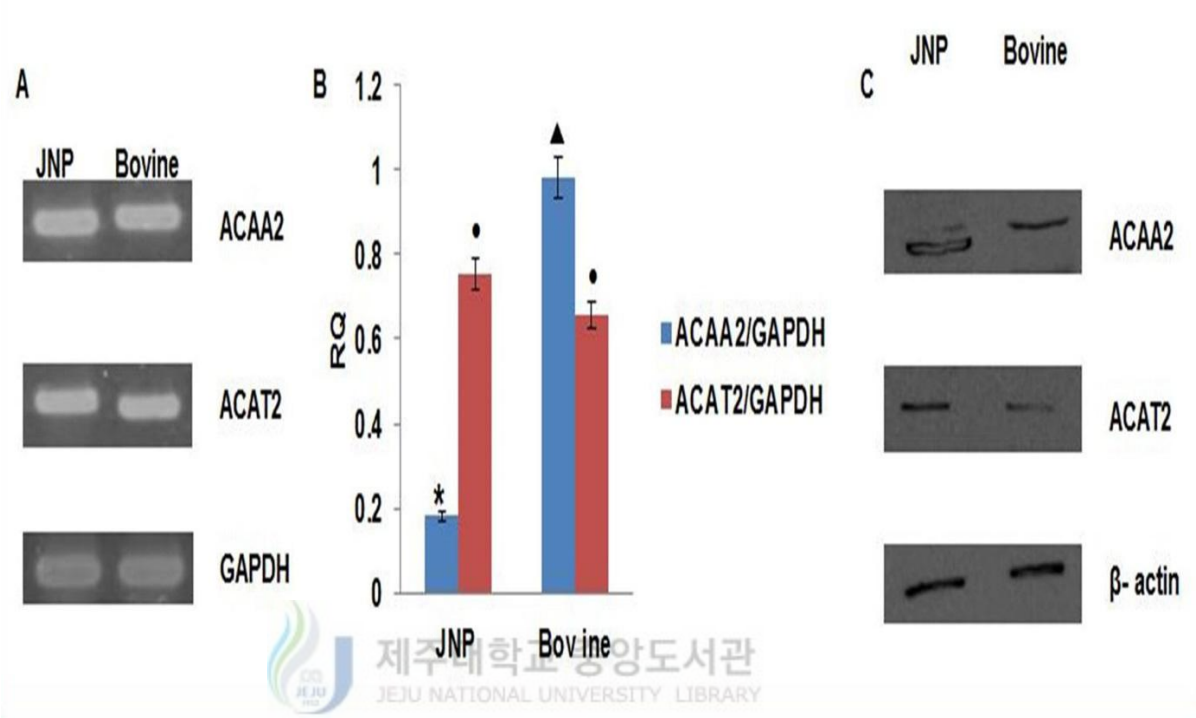



**Figure 5.1** Expression analysis of mRNA and protein levels of *ACAA2* and *ACAT2* in swine and bovine species. a) Expression of mRNA after RT-PCR using 1% agarose gel. b) Relative quantitative expression of *ACAA2* and *ACAT2* genes; bars with different superscripts show significantly different expression in the two species ($P < 0.05$). c) Representative blots expressions of ACAA2, ACAT2, and β-actin proteins obtained by western blotting.

**Table 5.4** List of the deleterious nsSNPs with their prediction tools and scores.

| rs id | Substitution native | Allele change | SIFT | PROVEAN | I-mutant (DDG*) |
|---|---|---|---|---|---|
| rs342236888 | I122V | T/C | 0.35 | –3.587 | –1.89 |
| rs324893294 | R139H | C/T | 1 | 5.246 | –1.14 |
| rs341690005 | A150T | C/T | 0.42 | 0.609 | –0.79 |
| rs321479200 | R281H | C/T | 0.15 | –4.012 | –1.31 |
| rs334695344 | E343D | T/G | 0.6 | 0.673 | –0.53 |

*****kcal/mol

activity, and transferring acyl groups other than amino-acyl groups (GO: 0016747). Therefore, changes of these residues can affect functions of the protein.

The wild-type and mutant amino acids of ACAT2 protein of *Sus scrofa* differ in size. The mutation in a nsSNP-rs342236888 of ACAT2 at position 122, is associated with a change of residue from isoleucine to valine, thus changes the charge (Figure 5.5). In addition, the mutant residue is smaller than the wild-type one, which leads to empty space in the core of the protein. These differences can disturb the core structure of this important domain and thereby affect the catalytic activity of a nsSNP-rs342236888 of ACAT2.

Similarly, for a nsSNP-rs321479200 of ACAT2 at position 281, mutation of Arg to His alters the charge between the wild-type and mutant amino acids (Figure 5.6). Specifically, this mutation causes a loss of charge, as well as interactions with molecules different from those that interact with the wild type. The difference in size between the mutant and the wild type would also possibly cause a loss of external interactions. The residue is located on the surface of ACAT2 and is not found to be in contact with other domains with known functions. However, contacts with other molecules or domains are possible and might be affected by this mutation.

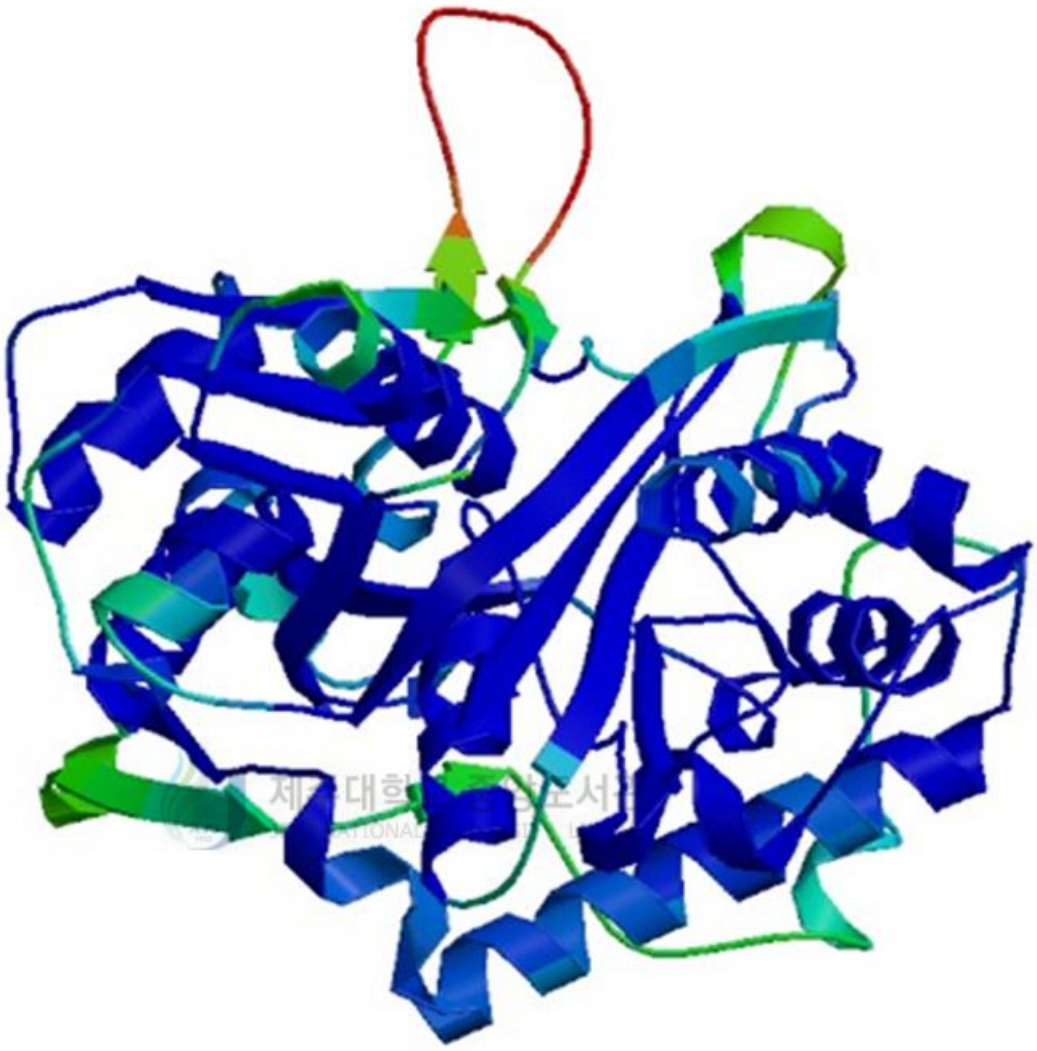

**Figure 5.2** The 3D homologous model of the native ACAT2 protein. A 3D structure of the native ACAT2 protein was prepared using SWISS-MODEL expasy and ESyPred3D. It depicts the native structure in 3D ribbon pattern of chain folding in ACAT2.

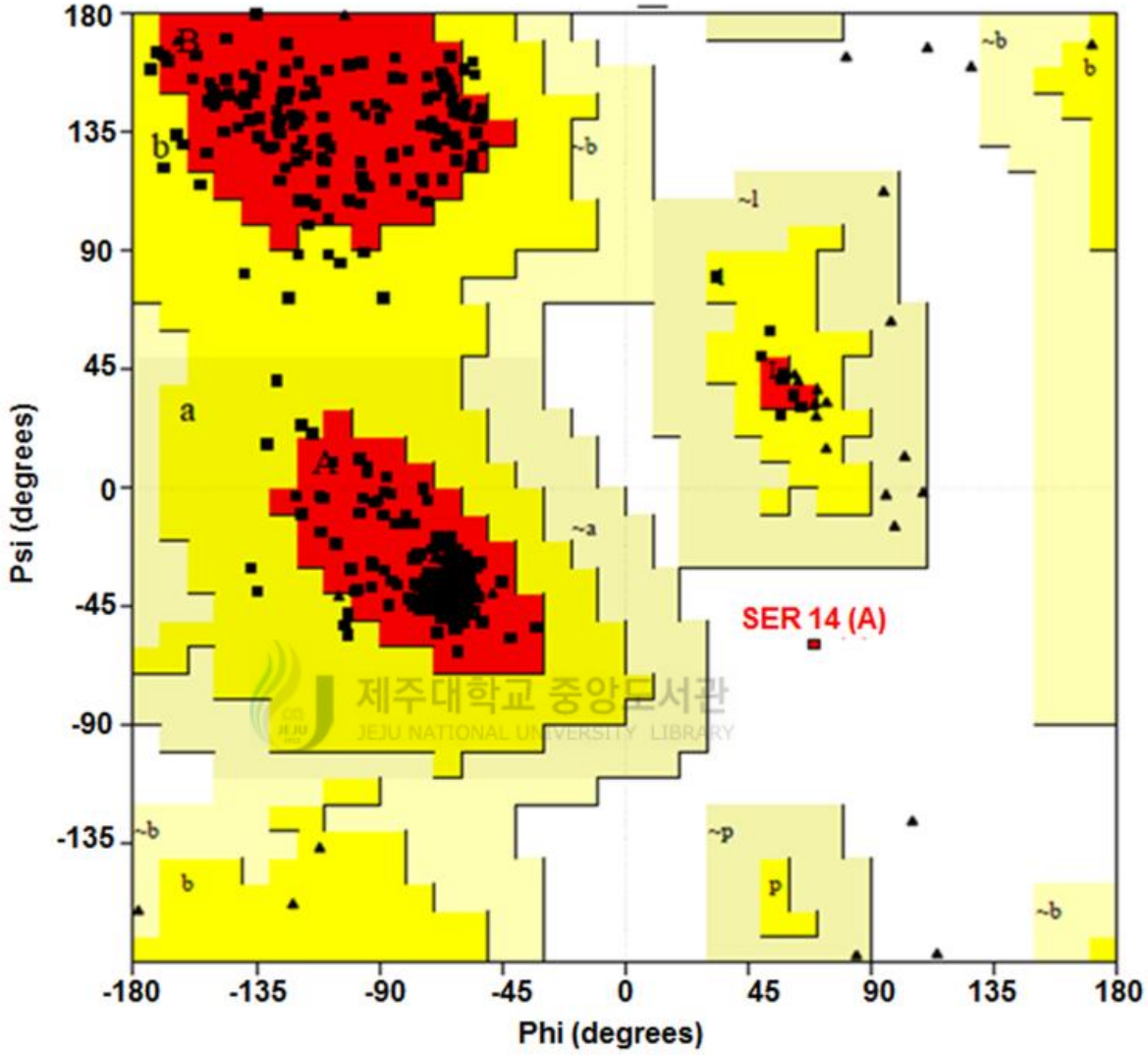


**Figure 5.3** The Ramachandran plot showing the empirical distribution of data points observed in the structure of ACAT2. This plot was used for quality checking and the structural validation of the modelled ACAT2 from *Sus scrofa*. The red, brown, and yellow regions represent the favoured, allowed and generously allowed regions as defined by Procheck.

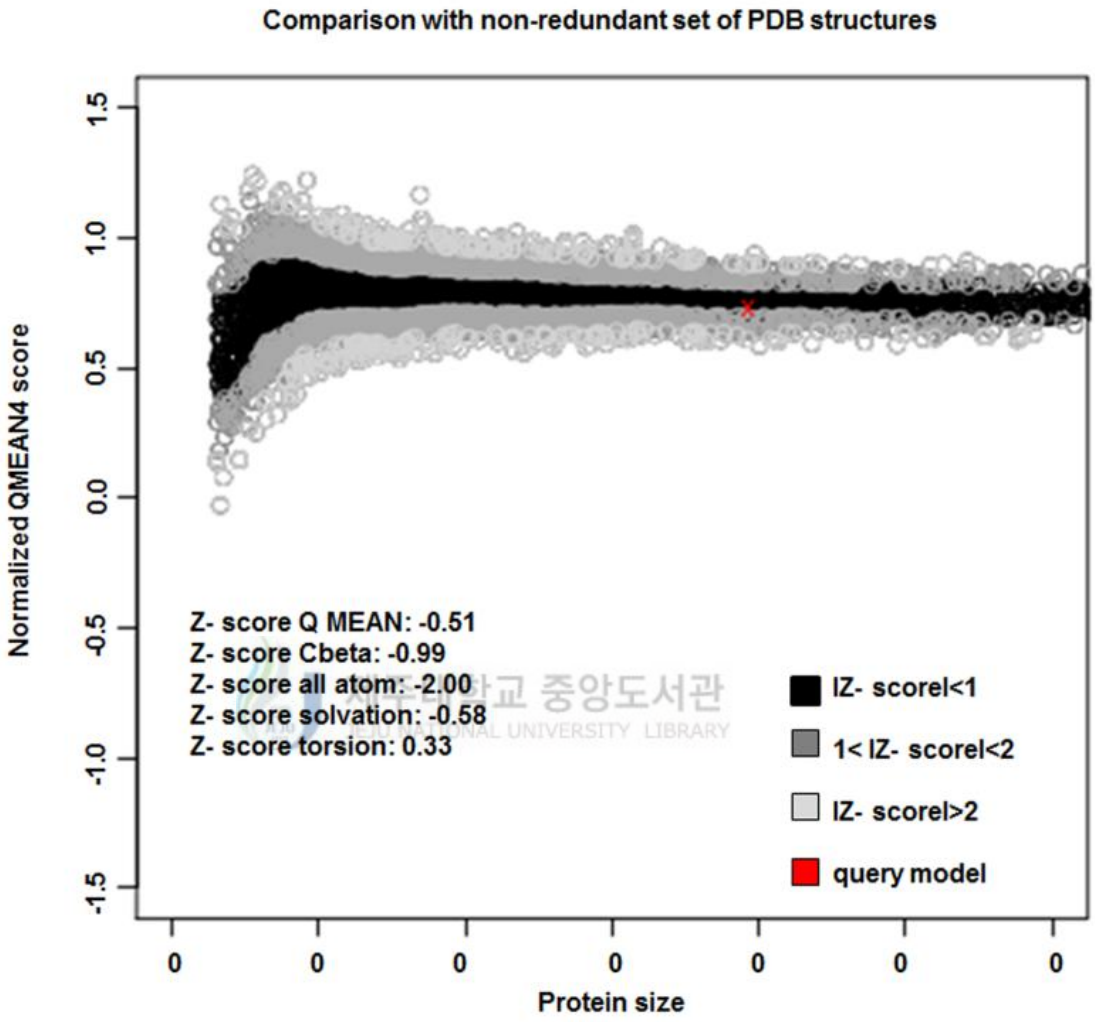


**Figure 5.4.** QMEAN-Z score of the model for *ACAT2* of *Sus scrofa*. The red legend indicates the QMEAN-Z score.

**Table 5.5** List of stabilising residues between the native and deleterious nsSNPs.

| Substitution | Existing residues | | | | | | | | | | | | | | Newly formed |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Native | VAL17 | VAL18 | ILE19 | VAL64 | ILE65 | VAL124 | ALA125 | GLY126 | ALA269 | MET273 | LYS274 | GLY387 | VAL388 | ---- | |
| I122V | VAL17 | VAL18 | ILE19 | VAL64 | ---- | VAL124 | ALA125 | ---- | ALA269 | MET273 | LYS274 | GLY387 | VAL388 | Phe57 | |
| R281H | ---- | ---- | ILE19 | VAL64 | ---- | VAL124 | ALA125 | ---- | ALA269 | ---- | LYS274 | ---- | ---- | Gly78, Ser232, Gly374 | |

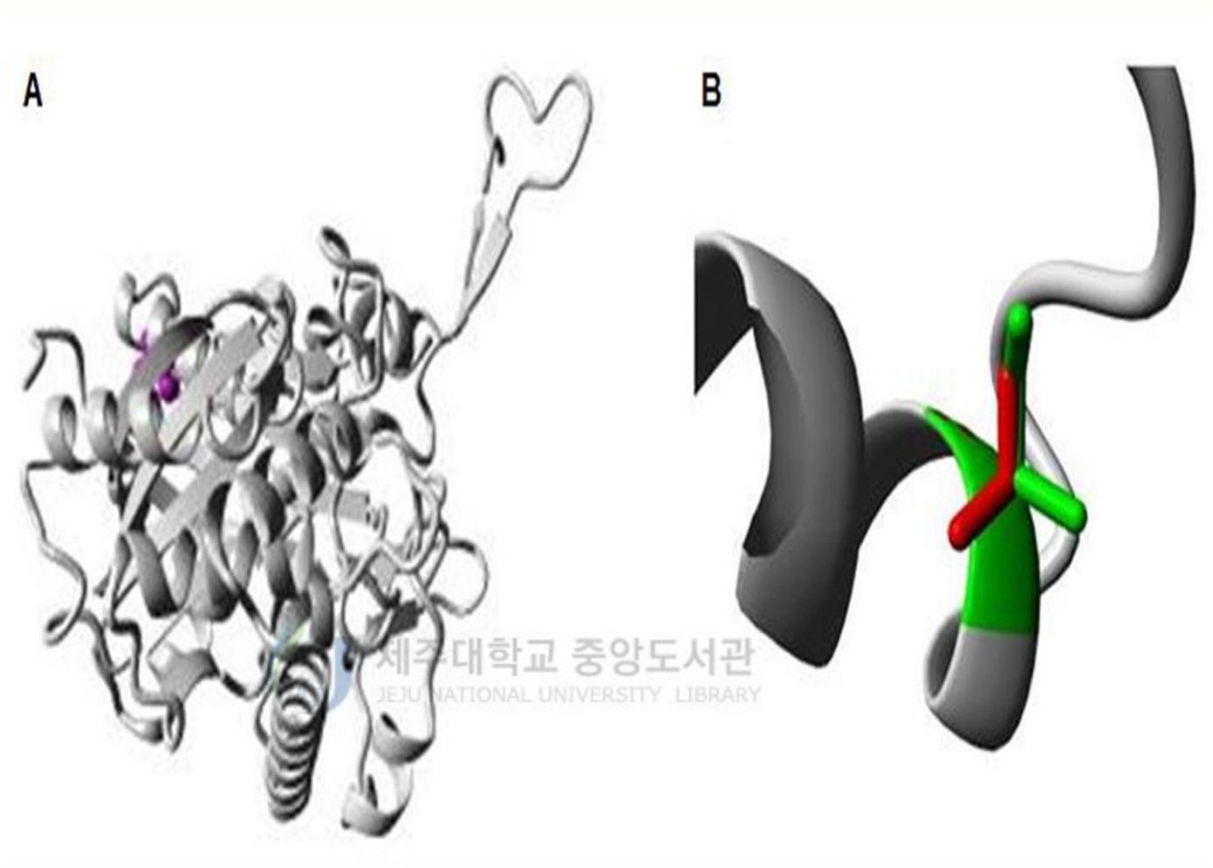



**Figure 5.5** Overview of the ACAT2 protein of *Sus scrofa* showing a mutation of isoleucine to valine (I>V) at position 122. A 3D ribbon pattern was chosen to explain the mutation. a) The protein is coloured grey; the side chain of the residue is shown as small magenta balls. b) Close-up view of the mutation of isoleucine to valine (I >V) at position 122 in ACAT2 of *Sus scrofa*. The side chains of both wild-type and mutant residues are shown in green and red, respectively.

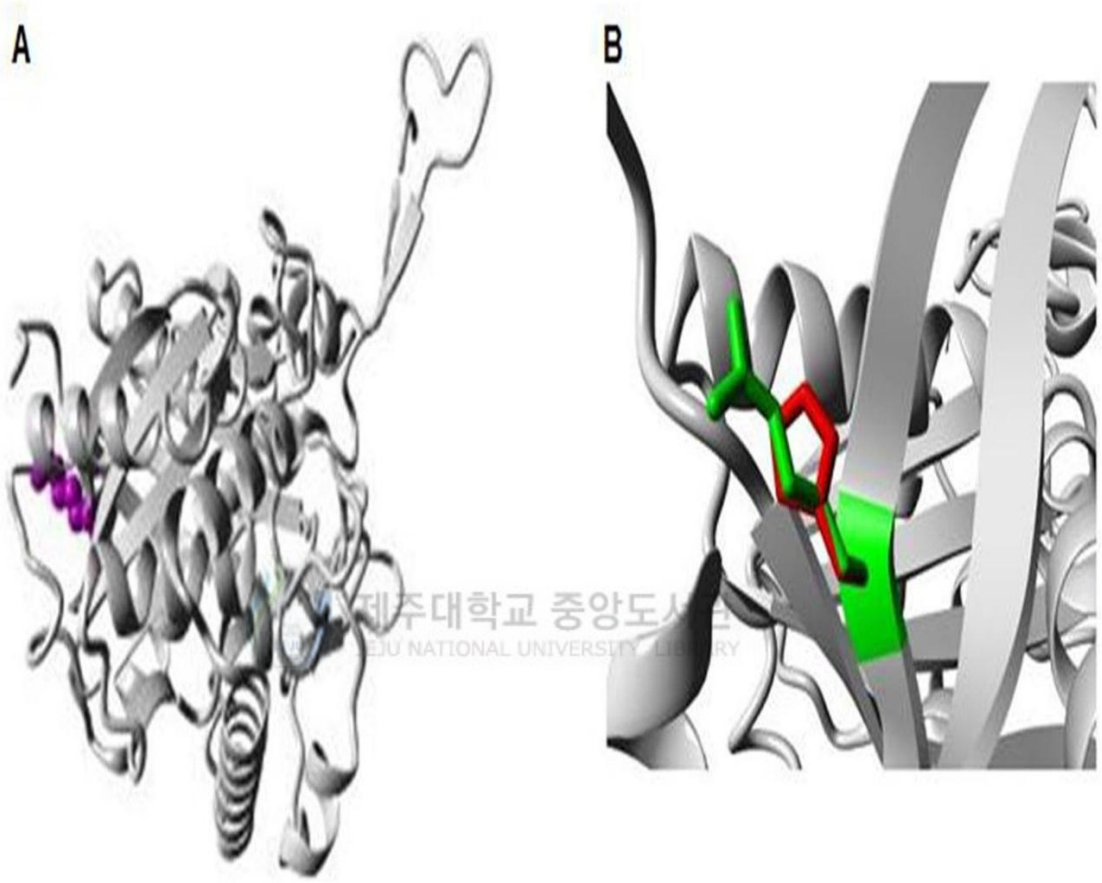


**Figure 5.6** Overview of the ACAT2 protein of *Sus scrofa* showing mutation of arginine to histidine (R>H) at position 281. A 3D ribbon pattern was chosen to explain the mutation. a) The whole protein is presented in grey, with the side chain of the residue as small magenta balls. b) Close-up view of the mutation of arginine to histidine (R>H) at position 281 in ACAT2 of *Sus scrofa*. The side chains of both wild-type and mutant residues are shown in green and red, respectively.

### 5.5. Discussion and conclusion

To our knowledge, this study is the first attempt to characterise nsSNPs that are related to the metabolism and expression of *ACAT2* in pig liver. Results of this study show that nsSNPs are considered to be potentially important for the functional and structural analysis of proteins. Such nsSNPs are of particular interest, given that they might provide markers of selection. These 5 nsSNPs were functionally annotated and were subjected to analysis using the SIFT and PROVEAN tools to predict their potential deleterious effects. The tolerance index of selected deleterious nsSNPs is inversely related to the functional impact of the corresponding amino acid residues (Ng and Henikoff, 2001; Xi *et al.*, 2004). rs342236888 (I122V) and rs321479200 (R281H) were selected as deleterious nsSNPs in this study (low tolerance indexes of -1.89 and -1.31, respectively). The *in silico* analysis indicated that misfolding and intermolecular interactions of the *ACAT2* gene, can have marked effects on the structure and functions of the protein. Thiolase domain has been enlisted as an important in the key enzymatic pathways such as fatty acid, steroid and polypeptide synthesis. It has also been reported to play a role in transferase activity and metabolic process. Therefore, it can be predicted that due to the mutations, I122V and R281H, nsSNPs might have a direct influence on the metabolical functions of ACAT2 in *S scarofa* (Panitz *et al.*, 2007).



Moreover, the food industry has been criticized in particular for the decline in the quality of pork from lean breeds (Wood, 2001). Further, it has been observed that the properties of lean meat are greatly influenced by its chemical composition. The amino acid contents of semimembranosus muscle (threonine, isoleucine, lysine, aspartic acid, serine, proline) increase with an increase in the proportion of lean meat (Okrouhla *et al.*, 2008). Moreover, in our study where isoleucine mutates to valine will result in the decrease in the number of residues of isoleucine in the native form of ACAT2 of *Sus scrofa*. Therefore, it can be one of the contributing factors for decrease in the amount of lean meat.

GO categories have been used previously to characterise protein function and to elucidate trends in protein datasets (Agarwal *et al.*, 2010). GO terms here helped to identify biological processes that could be affected by the mutations. Our finding that *ACAT2* is mainly responsible for metabolic processes involving lipid metabolism (Table 5.3) is supported by previous study (Nguyen *et al.*, 2012). It has been well established that both genes are involved in metabolic pathways (Cao *et al.*, 2008) and belong to the same protein family. It has also been observed that ACAT2 is mostly involved in lipid metabolism; in the present investigation, it has shown more metabolic activity than ACAA2. Validation in the laboratory reported empirical evidence of

actual biological similarities between ACAT2 and ACAA2. Later, this concept was supported by the similarities during GO annotation. Significantly higher expression of *ACAT2* in the JNP liver supports the metabolic role of *ACAT2* and its importance in physiological processes. In an earlier study, higher expression of *ACAT2* was also linked to cholesterol levels, which are in turn reported to be linked to increased lean mass (Jiang *et al.*, 2009).

ACAT2 plays an important role in efficient cholesterol absorption, lipoprotein metabolism, and atherosclerosis (Kushwaha *et al.*, 2005; Alger *et al.*, 2010; Nguyen *et al.*, 2012). Cholesterol plays a direct role in inflammation. It is also believed that inflammation plays a significant role in muscle building (Jiang *et al.*, 2009). Moreover, an earlier study has also reported significant role of cholesterol in the muscle building under the dose dependent manner (Riechman *et al.*, 2007). Therefore, it has been concluded that cholesterol levels are dose-dependently associated with increased muscle size and strength.

ACAT2 can esterify cholesterol and oxysterols in humans and other animals (Osada *et al.*, 1994; Cases *et al.*, 1998). Cholesterol is also essential for the formation of lipid rafts, which function in the assembly of the components of signalling pathways through sorting of proteins and construction of signalling complexes (Simons and Toomre, 2000). Earlier studies have reported that these signalling pathways contribute to the development of skeletal muscle (Smythe *et al.*, 2003; Freeman and Solomon, 2004; Lucero and Robbins, 2004; Bellot *et al*., 2005). Therefore, significantly high expression of *ACAT2* in *S scrofa* during the current study indicates that it can also be involved in the high muscle growth.



In conclusion, this study demonstrates the significant application of bioinformatic tools to clarify the changes in functions associated with the metabolic processes in which *ACAT2* participates. *In silico* analysis of the preferred nsSNPs of the *ACAT2* gene that are most strongly related to metabolic activity could explain the alteration in the metabolic activity caused by the mutations. Annotation of the effects of “I to V” substitution at residue 122 and “R to H” at residue 281 of ACAT2 might be useful as selection marker for molecular diagnosis. In terms of the importance of this loss of arginine (R), previous researchers suggested its role for improving meat quality (Ma *et al.*, 2010). Further evidence also indicates that arginine regulates the partitioning of dietary energy in favour of muscle protein accretion and fat reduction in animals (Wu *et al.*, 2008). Moreover, it was found that activation of endogenous arginine synthesis increased the average daily weight gain in piglets, compared with that in a control group (Frank *et al.*, 2007; Ma *et al.*, 2010).

Furthermore, our GO analysis and evaluation of mRNA and protein expression in the laboratory indicates that *ACAT2* plays more important role in lipid and cholesterol metabolism in

pig than bovine. Therefore, this study can be used as a stepping stone for planning future studies on the role of nsSNPs in metabolic processes, especially in pigs. The associations of SNPs in *ACAT2* with fertility and production traits can be used to plan studies for the improvement of growth performance and fertility traits in Jeju native pigs.

# Chapter 6

## Evaluation of body growth and immunity related differentially expressed genes through deep RNA Sequencing in the piglets of Jeju Native Pig and Berkshire

### 6.1. Abstract

The current study was carried out with the objective to investigate the differentially expressed genes (DEGs) between piglets of Jeju Native Pig (JNP) and Berkshire breeds. RNA-seq technique was used to investigate the transcriptomes in the fat, liver and *longissimus dorsi* muscle from these two breeds. Paired end reads of the sequences those passed the quality filters were aligned to *Sus scrofa* genome using Tophat2(v2.0.2). In the current study, 65% (muscle), 20% (fat) and 54% (liver) genes have shown higher expression in the piglets of JNP than Berkshire. Gene ontology and signaling pathway showed that immune response and lipid metabolisms were commonly enriched pathways in all the three tissues. It was found that the genes pertaining to body growth and immune system are significantly ($P$<0.01) more expressive in Berkshire piglets. DEGs explored between the piglets of two breeds might influence the identification of the genetic markers for further breed improvement programmes. Our findings give a new horizon for understanding and identification of the candidate genes that are involved in the altered biological functions. Moreover, transcriptome analysis makes it easier to understand the differences between genetic mechanisms of breeds.

***Key words*:** Differentially expressed genes, RNA-Seq, Transcriptome, Gene Ontology

## 6.2. Introduction

The indigenous pigs are important livestock animals and are major source of animal proteins (Sodhi *et al.,* 2014). Jeju Native Pig (JNP) is an indigenous breed found at Jeju-do. The consumer preference for JNP meat is attributed to its taste, tenderness and marbling quality than Western breeds (Cho *et al.,* 2011). However, the major snags for JNP are its lower feed efficiency and smaller litter size. Increased demand of pork has led the farmers to opt for commercial farming with more concentration on the muscular traits. Berkshire pigs have better growth rate and meat palatability traits than that of JNP (Sodhi *et al.,* 2014). Moreover a close relationship has been traced between Berkshire and the native pig breeds of Asia (Kim *et al.,* 2002). Therefore, Berkshire has been considered for the current comparative study.

Muscle tissue plays a major role in the economics of meat production. Still, the understanding of variation in expression of gene during muscle development in younger age is in its primitive stage. The liver, adipose tissue and skeletal muscle are the principal organs involved in the regulation of lipid metabolism (Corominas *et al.,* 2013). Fat assessed as marbling score is an important factor influencing the meat quality (Sheng *et al.,* 2014). Moreover, the weaker immune system in piglets will lead to the development of a herd with poor performance. Worldwide health challenges of the parent stock of the pigs are reported to adversely affect the performance, which causes huge economic loses to the pork industry. Therefore, identification of candidate genes of economic importance in piglets could be useful for the quick determination of the association of a specific genetic variant with their phenotype. Therefore, gaining knowledge of the porcine immune system will undoubtedly benefit many aspects of porcine disease research. Moreover, liver is a tissue of choice to study the immunity and detection of genes related to immune system which will confer an advantage in selection of breeding stock.



To obtain massively unbiased information on the differentially expressed genes (DEGs), next-generation sequencing (NGS) has come up as a revolutionary tool in genomic and epigenomics. RNA-sequencing (RNA-seq) has become the assay of choice for high-throughput studies of gene expression. RNA-seq uses NGS technology to sequence, map and quantify a population of transcripts (Mortazavi *et al.,* 2008). Moreover, NGS technology can analyze simultaneous expression of thousands of genes in a tissue. The different NGS platforms for RNA-seq have been used in different species especially in pigs (Metzger *et al.,* 2014; Molik *et al.,* 2014) but meager reports are available regarding the its application in JNP.

Therefore, the current study was planned with the objective to discover differentially expressed key regulatory genes which are significantly associated with pork quality, body growth

and immunity in purebred one day old piglets of JNP and Berkshire. To our knowledge, this study is first of its kind, which has been devoted to perform the transcriptome analysis for the divergence in gene expression in one day old piglets. Taken collectively, our integrated RNA-Seq approach aims to reveal differential expression profiles of the genes which may be useful for successful marker-assisted selection in pigs to prevent huge economic losses of pig industry due to animals with poor genetic makeup in the breeding stock.

## 6.3. Materials and Methods

### *6.3.1. Experimental design and Bioethics statement*

Five purebred one day old female piglets each from JNP and Berkshire breeds were selected. Piglets were weaned immediately after birth and moved to a common nursery maintained at an ambient temperature 25±1ºC. For one day, piglets were offered cow's milk to meet the nutrient allowances. The samples of fat, *longissimus dorsi* muscle and liver were collected immediately after slaughter. Samples were stored in dry ice and later were kept at -80℃ until RNA extraction. Piglets were slaughtered by electric stunning (250 volts for 3 seconds) following the guidelines of animal bioethics committee, Department of Animal Biotechnology, Jeju National University, Republic of Korea.

### *6.3.2. Extraction and quality analysis of RNA*

RNA was isolated from 100 mg of the fragmented frozen respective tissues of both breeds in five replicates. TRIzol™ (Invitrogen, USA) was used for the isolation of RNA. Tissue samples were homogenized in 2.0 ml each of TRIzol and chloroform, which were subsequently precipitated by using isopropanol (Junsei Chemical Ltd., Japan). Isolated RNA samples were stored at -80℃. To get rid of genomic DNA contamination, 25 μg of RNA from each sample was treated with the RNase-free DNase set (QIAGEN, Hilden, Germany) and purification of RNA was performed with the RNeasy mini kit according to the user guidelines (QIAGEN, Hilden, Germany). Quality of RNA was assessed by Bioanalyzer2100 with RNA 6000 Nano Labchips. To assess the quantity of RNA, automated capillary gel electrophoresis was performed following user guidelines (Agilent Technologies Ireland, Dublin).

### *6.3.3. Preparation and sequencing of mRNA-seq Illumina library*

A library of template molecules suitable for subsequent cluster generation was constructed from the mRNA. Following the manufacturer's guidelines for the TruSeq RNA Sample Pre Kit, 0.1-4.0 µg of total RNA was used from the each sample for the construction and sequencing of the cDNA library. For a positive control sample, Agilent Technologies Human UHR total RNA has been used. A standard protocol provided by Illumina, Inc was followed to construct libraries by performing five technical replicates. Libraries with different indexes were pooled within a group and were sequenced in one lane. Libraries were sequenced in Illumina HiSeq2000 high-throughput sequencing instrument with 100 paired-end (PE) sequencing.

### *6.3.4. Sequence filtering and merging of reads*

The web based tool FASTQC was used to calculate the quality controlled statistical data generated from the Illumina second generation sequencing technology. The sequences were trimmed from both the ends and PE reads with low quality were removed prior to further analysis by Consensus Assessment of Sequence and Variation v 1.7 (CASAVA-v1.7) pipeline. PE reads those passed the quality filters were aligned to *Sus scrofa* genome (Sscrofa10.2) from the University of California Santa Cruz (UCSC) using Tophat2(v2.0.2). HTseq(v0.5.3p3) was used to count reads. The left over reads were later merged on the basis of the overlapping of PE reads. The correlation between two breeds was analyzed by Corrplot and DEGseq-R packages (Wang *et al.,* 2010). Further, statistical identification of DEGs was conducted by Fisher's exact test model (Bloom *et al.,* 2009). Samples from both the breeds were compared with the samples of their respective class by RNA-seq for detection of DEGs. Significant DEGs were selected at FDR<0.01 and *P<0.01*.



### *6.3.5. Analysis of functional annotations of cluster pathways*

DAVID (Database for Annotation,Visualization and Integrated Discovery) functional annotation cluster analysis was performed to enlist up- and down-regulated genes with four fold changes. The Expression Analysis Systematic Explorer (EASE) tool of DAVID was used for the investigation and representations of functional groups in three tissues of one day old piglets relative to the whole genome. The enrichment of gene ontology (GO) terms was measured by

EASE (Alterovitz and Ramoni, 2010). Functionally clustered genes were filtered to identify enriched GO terms according to EASE values <0.01.

Mapping of pathways was performed using the KEGG (Kyoto Encyclopedia of Genes and Genomes) tool to understand the functional meanings of higher-order and utilities of the cell or the organism from their genome information. ToppCluster tool was used to construct the network diagram of biological themes in data sets involving numerous sets of genes and system based dissection of biological states. Bonferroni correction method with cutoff $P<0.01$ was applied for the dissection of gene features.

### *6.3.6. quantitative Real-time PCR (qRT-PCR)*

Primers for qRT-PCR were designed using the online Primer-3 program (Rozen and Skaletsky, 2000; Table 6.1). qRT-PCR was conducted using Applied Biosystems, Step-One Real Time PCR system. The dye EvaGreen (Biotium,USA) was used to determine the quantity of transcripts of target genes present in each sample of Berkshire and JNP. The standard curve method was used to define the efficiency of qRT-PCR and each individual sample was quantified in triplicate. The efficiency of amplification of the target gene was compared with that of the endogenous *β-actin* control transcript (Wang *et al*., 2003). Quantification of transcript levels was performed by the comparative $C_T$ method. The results are represented as the relative expression normalised using the level of the transcript of the endogenous reference (Erkens *et al*., 2006).



### *6.3.7. Statistical analysis of the differential expression levels of genes*

The differential gene expression analysis was performed to understand the statistical values of the expression levels among the genes of two breeds. ArrayStarv4.1 (DNASTAR, Madison,WI) software was used to analyze data and to generate heat maps of selected genes. The level of gene expression for each gene was measured as reads per kilobase per million for the particular gene region (RPKMg) (Wagner *et al.*, 2012). The relative quantitative expressions of the genes by real-time qRT-PCR were analyzed by the analysis of variance (ANOVA). The significant differences between the mean expressions of different genes at $P<0.05$ were analyzed by Tukey's b-test. The values have been expressed as mean±SEM.

**Table 6.1** Primers used during the qRT-PCR to analyze the differential expression of genes in the fat, *longissimus dorsi* muscles and liver of Berkshire and JNP.

| Gene | Sequence | Product size | Annealing temperature ($T_a$) | Genebank ID |
|---|---|---|---|---|
| *COL2A1* | F- GATGTTGGTGAGAAAGGCCC<br>R- GTCCAGTCTCTCCACGTTCA | 181 | 61℃ | NC_010447 |
| *COL9A1* | F- TGTGCTCGGAAAACTTGTGG<br>R- TCACCGTCAATACCGTCGAT | 228 | 61℃ | NC_010443 |
| *SIK1* | F- CTTCGTCCTGAGGTCATGGA<br>R- AGCTTGATGTCCATGTTGCC | 233 | 61℃ | NC_010455 |
| *CXCL13* | F- CTGGGATCTCTGCTTCTCGT<br>R- CCCAGGAGGCCAGATTTGAA | 162 | 58.5℃ | NC_010450 |
| *SLA* | F- GCCAGAGTTTACCATGGCTG<br>R- TTCGTGTCTGGTAGCTGGAG | 191 | 61℃ | NW_003613535 |
| *IL7R* | F- GTCCGGACCACAAGCAAAAT<br>R- AATGATGTCCAGACGCTCCT | 188 | 61℃ | NC_010458 |
| *CRP* | F- TGGCTCAGCGGAAACAAATC<br>R-TTTGTCTTTGAGGGCTGCAC | 228 | 61℃ | NC_010446 |
| *PNPLA3* | F- CAGAACAAAGCACTGTCCCC<br>R- CGCAGATCTAGCGTGTAGGA | 247 | 61℃ | NW_003538349 |
| *ELOVL6* | F- AGCCAGAACCCATCACTTGT<br>R- ACACCAATCCTCTCCCCTTG | 213 | 61℃ | NC_010450 |
| *FASN* | F- ATGACTACGGGCCCTTCTTC<br>R- GATGCGGATGGAGGTGAAAC | 176 | 58.5℃ | NC_010454 |
| *β-actin* | F GACATCCGCAAGGACCTCTA<br>R ACACGGAGTACTTGCGCTCT | 157 | 60℃ | XM_003124280 |

## 6.4. Results

### *6.4.1. Quality analysis of RNA-seq data*

The data generated by Illumina HiSeq2000 consisted of highly reliable sequence reads. The sequencing produced clusters of sequence reads with maximum 100 base pair length. Some of the sequence reads (<5.0E-0.5%) could not pass the quality filtering analysis. On an average 90.7% (fat and *longissimus dorsi* muscle each) and 97.6% (liver) sequence reads from both the breeds passed quality control (Table 6.2). Later, the data was further analyzed by FASTQC tool to decide whether sequence trimming is required or not. The range of quality standards across all bases at every point in the FastQ file has been summarized (Figure 6.1) and good quality calls scattered across the fluorescent background of the plot have been represented. A warning was made when the lower quartile and the median for any base were less than 10 and 25, respectively.



**Table 6.2** RNA-seq reads and mapping rate of three tissues from JNP and Berkshire.

| **Breeds** | **No. of** | **Tissue** | | |
|---|---|---|---|---|
| | | **Fat** | **Liver** | ***Longissimus dorsi* muscle** |
| JNP | total reads | 41413088 | 45255914 | 35563538 |
| | reads after QC* | 41411459 | 45254078 | 35562180 |
| | accepted Hit | 37879239 (91.5%) | 44164596 (97.6%) | 31807200(89.4%) |
| Berkshire | total reads | 40310004 | 37618038 | 37310958 |
| | reads after QC* | 40308452 | 37616584 | 37309517 |
| | accepted Hit | 36254763(89.9%) | 36731445(97.6%) | 34331014 (92.0%) |

* Quality Check

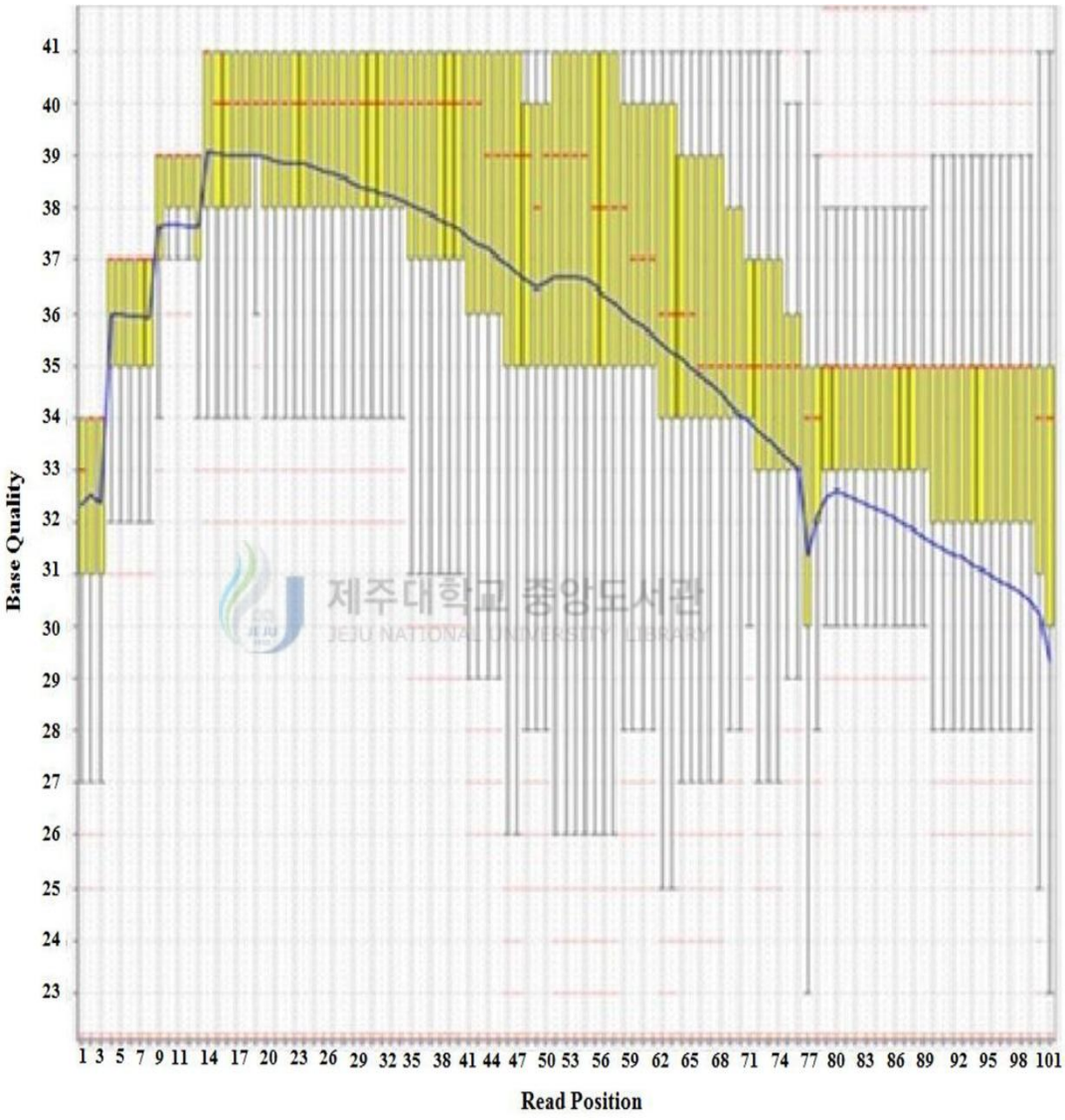



**Figure 6.1** Range of quality standards across the bases. The y-axis on the graph shows the quality scores with higher scores indicating better base calls. The background of the graph separates the y axis into calls of high-quality (florescent), reasonable quality (grey) and poor quality (white). In each of these findings, the red line is the median value, the florescent boxes correspond to the inter-quartile range (25–75%) and the blue line signifies the mean quality.

### *6.4.2. Analysis of DEGs*

Statistically significant ($P< 0.01$) differential gene expressions have been observed in the three tissues of the two breeds. Benhamini Hochberg correction was applied for the correction of the significant scores and each data point in the scatter plot represents an individual gene. Each scatter plot was constructed on the basis of expression levels of selected genes (Figure 6.2) and the data are visualized on a log scale. The dotted purple line in the scatter plot reflects the linear regression ($R^2$)/"best-fit" line which crosses near to as many data points as possible. The fold changes for muscle, fat and liver of the two breeds were calculated with log of two, four and eight folds (Table 6.3).

The DEGs from the three tissues of JNP and Berkshire piglets were selected at $P<0.01$ and FDR<0.01. A total number of 56 (*longissimus dorsi*), 143 (fat) and 110 (liver) DEGs respectively were identified in the three tissues. Out of these DEGs, 13 (muscle), 47 (fat) and 58 (liver) genes either matched un-named or hypothetical proteins or had no BLAST matches with database. Moreover, it has been observed during analysis that 65% (muscle), 20% (fat) and 54% (liver) genes have higher expression in JNP than Berkshire (Table 6.4, 6.5 and 6.6). These data in matrix form can be accessed through an open source online visualization at http://biopopdb.snu.ac.kr/PIG_DEG/. Further, the DEGs in three tissues of both breeds were analyzed through heat maps and the DEGs have been expressed by the grid of colored tiles (Figure 6.3). Each row in the grid corresponds to one gene and each column represents the respective breed under study. Clear significant ($P<0.01$) differential expressions of genes between the two breeds have been observed.



**Table 6.3** Summary of the Scatter plots for three tissues from JNP and Berkshire at the level of the three different fold changes.

| | **Fat** | ***Longissimus dorsi* muscle** | **Liver** |
|---|---|---|---|
| Total number of genes | 31,582 | 31,582 | 31,582 |
| Genes at 2 fold change | 12,132 | 5,478 | 5,677 |
| Genes at 4 fold change | 6,239 | 3,216 | 3,367 |
| Genes at 8 fold change | 4,422 | 2,751 | 2,755 |

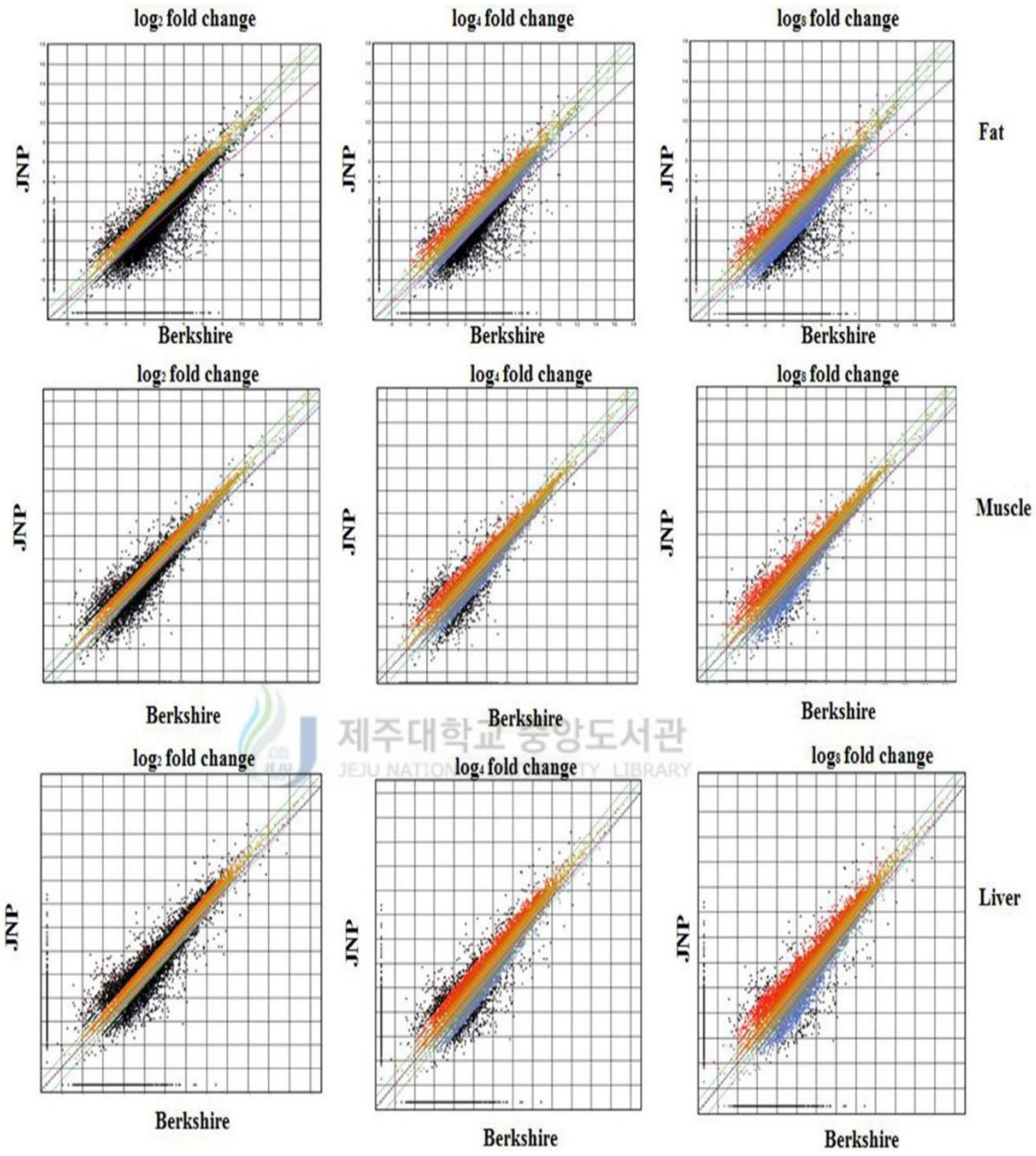




**Figure 6.2** Scatter plots for the visual comparison of gene expression levels. Scatter plots for muscle, fat and liver tissues of JNP and Berkshire were used for the correction of the significant scores. The plots were constructed with $\log_2$, $\log_4$ and $\log_8$ fold changes. Each data point on the plot represents an individual gene.

### *6.4.3. Analysis of the biological functions and pathways of DEGs*

The detailed GO analysis for molecular functions, biological processes and cellular components of the DEGs from the three different tissues have been summarized (Table 6.7, 6.8 and 6.9) to explain their role in the immunity, metabolic process, skeletal system morphogenesis and body growth. GO analysis of muscle revealed that DEGs are significantly associated with biological process like extracellular structure organization and skeletal muscle cell differentiation (Table 6.7). Similarly, the biological processes in fat tissue shared the functions such as inflammatory response, biological adhesion and immune response (Table 6.8), whereas, in liver tissue biological processes are associated with myeloid cell development and leukocyte aggregation (Table 6.9).

Gene's functions, revelation of common processes and pathways among the selected DEGs have been investigated through KEGG. The details of the pathways along with their confidence levels have been studied for respective tissues (Table 6.10). The DEGs are mainly involved in nine common pathways traced among the three tissues (Figure 6.4). The pathways are involved with the growth, metabolic processes and immune responses.



### *6.4.4. DEGs involved in body growth and skeletal system development*

Our study showed that four genes from the collagen family are differentially expressed. *Collagen type21-alpha-1* (*COL21A1*) and *Collagen type2-alpha-1* (*COL2A1*) are down regulated whereas *Collagen type 9-alpha-1* (*COL9A1*) and *Collagen type 9-alpha-2* (*COL9A2*) are found to be up regulated in JNP (Table 6.4; Figure 6.5). *Myosin binding protein-H* (*MYBPH*) related to myosin in muscles is up regulated in JNP. *Periostein* (*POSTN*) and *Serine/threonine-protein kinase-1* (*SIK1*) are reported to be down regulated in JNP (Table 6.4; Figure 6.5).

**Table 6.4** List of DEGs (FDR<0.01) and the log fold changes in *longissimus dorsi* muscle tissue of Berkshire and JNP piglets.

| ***Sus scrofa* Ensemble ID** | **Gene** | **Description** | **Expression level(RPKM)** | | | | |
|---|---|---|---|---|---|---|---|
| | | | ***P* value < 0.01** | **FDR < 0.01** | **JNP** | **Berkshire** | **Fold change in JNP** |
| ENSSSCG00000000918 | *EPYC* | Epiphycan | 2.14E-5 | 0.0075 | 2.47 | -8.98 | 2819.64 down |
| ENSSSCG00000001231 | *SLA-1* | Src-like-adaptor | 4.63E-6 | 0.0022 | 4.60 | 5.85 | 2.37 up |
| ENSSSCG00000001500 | *COL21A1* | Collagen, type XXI, alpha 1 | 2.33E-5 | 0.007 | 2.77 | 4. 71 | 3.84 down |
| ENSSSCG00000001832 | *ACAN* | Aggrecan | 5.25E-7 | 0.0003 | 2.03 | -0.95 | 7.94 up |
| ENSSSCG00000002264 | *ST8SIA2* | ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 2 | 2.62E-6 | 0.0013 | 4.02 | 1.01 | 8.09 up |
| ENSSSCG00000002294 | *ARG2* | Arginase 2 | 3.95E-7 | 0.0002 | 1.03 | 4.59 | 11.77 down |
| ENSSSCG00000002383 | *FOS* | FBJ murine osteosarcoma viral oncogene homolog | 7.44E-9 | 6.30E-6 | 3.15 | 5.42 | 4.83 down |
| ENSSSCG00000002831 | *IRX3* | Iroquois homeobox 3 | 3.95E-8 | 3.04E-5 | 0.75 | 3.62 | 7.29 down |
| ENSSSCG00000003006 | *CYP2B22* | *Sus scrofa* cytochrome P450 2B22 | 9.60E-10 | 1.01E-6 | 3.29 | 5.56 | 4.84 down |
| ENSSSCG00000003201 | *ATF5* | Activating transcription factor 5 | 2.77E-9 | 2.61E-6 | 3.42 | 6.28 | 4.61 down |
| ENSSSCG00000004094 | *PPP1R14C* | Protein phosphatase 1, regulatory subunit 14C | 1.26E-5 | 0.0049 | 1.97 | 3.38 | 2.65 down |
| ENSSSCG00000004274 | *COL9A1* | Collagen, type IX, alpha 1 | 1.38E-13 | 3.91E-10 | 2.80 | -3.50 | 79.24 up |
| ENSSSCG00000005385 | *NOR-1* | Neuron-derived orphan receptor-1 alfa | 2.35E-16 | 1.33E-12 | 0.13 | 3.48 | 10.26 down |

*Continued……*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ENSSSCG00000005636 | *SLC25A25* | *Sus scrofa* solute carrier family 25 | 1.60E-5 | 0.0059 | 3.23 | 4.73 | 2.83 down |
| ENSSSCG00000009585 | *GADD45G* | Growth arrest and DNA-damage-inducible, gamma | 1.68E-5 | 0.0060 | 3.47 | 5.42 | 3.84 down |
| ENSSSCG00000010461 | *ANKRD1* | Cardiac ankyrin repeat protein | 5.19E-6 | 0.0023 | 4.52 | 1.84 | 4.26 up |
| ENSSSCG00000010554 | *SCD* | Stearoyl-CoA desaturase | 3.76E-9 | 3.35E-6 | 4.54 | 7.50 | 7.77 down |
| ENSSSCG00000012077 | *MX1* | Myxovirus resistance 1 | 3.15E-5 | 0.0095 | 3.20 | 4.70 | 2.82 down |
| ENSSSCG00000013898 | *PDE4C* | Phosphodiesterase 4C, cAMP-specific | 1.110E-5 | 0.0044 | 3.25 | 4.02 | 1.71 down |
| ENSSSCG00000014137 | *HAPLN1* | Hyaluronan and proteoglycan link protein 1 | 9.01E-7 | 0.00054 | 6.48 | 1.42 | 11.32 up |
| ENSSSCG00000014156 | *ARRDC3* | Arrestin domain containing 3 | 1.35E-5 | 0.0051 | 4.13 | 5.98 | 3.59 down |
| ENSSSCG00000014725 | *HBB* | *Sus scrofa* hemoglobin beta | 2.22E-5 | 0.0075 | 8.80 | 6.81 | 3.99 up |
| ENSSSCG00000015595 | *ATF3* | Activating transcription factor 3 | 1.13E-10 | 1.48E-7 | 2.48 | 5.05 | 2.76 down |
| ENSSSCG00000017092 | *GPX3* | Glutathione Peroxidase 3 | 4.17E-6 | 0.0021 | 2.43 | 3.38 | 2.68 down |
| ENSSSCG00000017421 | *ACLY* | ATP citrate lyase | 2.99E-7 | 0.0002 | 3.06 | 5.17 | 4.30 down |
| ENSSSCG00000020953 | *ATP1A4* | ATPase, Na+/K+ transporting, alpha 4 polypeptide | 3.12E-14 | 1.05E-10 | -0.07 | 3.66 | 13.39 down |
| ENSSSCG00000021059 | *ADORA1* | Adenosine A1 receptor | 1.17E-9 | 1.16E-6 | 8.54 | 2.45 | 12.24 up |
| ENSSSCG00000022834 | *PRSS45* | Protease, serine, 45 | 2.81E-5 | 0.0088 | -8.98 | -2.85 | 70.27 up |
| ENSSSCG00000023215 | *MAOB* | Monoamine oxidase B | 2.40E-5 | 0.0078 | 0.75 | 2.30 | 2.92 down |
| ENSSSCG00000023322 | *COL9A2* | Collagen, type IX, alpha 2 | 1.04E-15 | 4.41E-12 | 2.48 | 1.05 | 1.24 up |
| ENSSSCG00000025128 | *CPN2* | Carboxypeptidase N, polypeptide 2 | 1.28E-5 | 0.0049 | -0.86 | 2.30 | 8.97 down |
| ENSSSCG00000025523 | *COL2A1* | Collagen, type II, alpha 1 | 3.27E-18 | 2.77E-14 | 1.87 | 5.75 | 14.72 down |

*Continued……*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ENSSSCG00000025541 | *ELOVL6* | ELOVL fatty acid elongase 6 | 1.018E-5 | 0.0043 | 4.41 | 2.97 | 8.87 up |
| ENSSSCG00000025698 | *SERPINE1* | Serpin peptidase inhibitor, clade E | 3.77E-8 | 3.04E-5 | 6.29 | 1.24 | 6.21 up |
| ENSSSCG00000027982 | *ISG15* | ISG15 ubiquitin-like modifier | 1.09E-5 | 0.00447 | 9.26 | 1.58 | 2.34 up |
| ENSSSCG00000028137 | *SIK1* | Salt-inducible kinase 1 | 1.73E-12 | 4.20E-9 | -1.16 | 2.64 | 13.99 down |
| ENSSSCG00000028326 | *NR4A1* | Nuclear receptor subfamily 4, group A, member 1 | 1.84E-7 | 0.00013 | 1.28 | 1.93 | 1.56 down |
| ENSSSCG00000029212 | *GDE1* | Glycerophosphodiester phosphodiesterase 1 | 5.008E-6 | 0.0023 | 4.39 | 5.93 | 2.89 down |
| ENSSSCG00000029515 | *PON3* | Paraoxonase 3 | 5.70E-6 | 0.0025 | -0.21 | 2.10 | 4.98 down |
| ENSSSCG00000029944 | *FASN* | *Sus scrofa* fatty acid synthase | 2.39E-11 | 3.37E-8 | 3.12 | 5.74 | 6.13 down |
| ENSSSCG00000030681 | *MYBPH* | Myosin binding protein H | 5.17E-10 | 6.26E-7 | 8.35 | 2.47 | 5.60 up |
| ENSSSCG00000009361 | *POSTN* | Periostin, osteoblast specific factor | 2.57E-8 | 2.08E-5 | 7.77 | 5.69 | 4.24 down |

**Table 6.5** List of DEGs (FDR<0.01) and the log fold changes in fat tissue of Berkshire and JNP piglets.

| ***Sus scrofa* Ensemble ID** | **Gene** | **Description** | **Expression level (RPKM)** | | | | |
|---|---|---|---|---|---|---|---|
| | | | ***P* value < 0.01** | **FDR < 0.01** | **JNP** | **Berkshire** | **Fold change in JNP** |
| ENSSSCG00000000080 | *GRAP2* | GRB2-related adaptor protein 2 | 7.80E-5 | 0.0096 | -3.84 | 2.63 | 89.06 down |
| ENSSSCG00000000257 | *ITGB7* | Integrin, beta 7 | 3.02E-5 | 0.0042 | -1.99 | 3.55 | 46.62 down |
| ENSSSCG00000000362 | *RDH5* | Retinol dehydrogenase 5 | 8.96E-6 | 0.0015 | 3.09 | -0.92 | 16.24 up |
| ENSSSCG00000000687 | *CD4* | *Sus scrofa* CD4 molecule | 1.95E-8 | 1.00E-5 | -9.30 | 4.09 | 108020.64 down |
| ENSSSCG00000000705 | *CD27* | *Sus scrofa* CD27 molecule | 1.87E-9 | 1.63E-6 | -3.41 | 6.25 | 815.41 down |
| ENSSSCG00000001551 | *CLPSL2* | Colipase-like 2 | 2.29E-5 | 0.003 | 5.50 | 1.62 | 14.62 up |
| ENSSSCG00000001844 | *PLIN1* | Perilipin 1 | 1.64E-5 | 0.0026 | 5.53 | 2.65 | 7.35 up |
| ENSSSCG00000002483 | *SERPINA3-3* | Serpin Peptidase Inhibitor, Clade A Member 3 | 7.23E-8 | 3.00E-5 | -0.55 | 6.33 | 118.70 down |
| ENSSSCG00000002487 | *SERPINA3-2* | Serpin Peptidase Inhibitor, Clade A Member 2 | 5.91E-7 | 0.00016 | -0.30 | 4.95 | 38.56 down |
| ENSSSCG00000002749 | *HP* | *Sus scrofa* haptoglobin | 5.75E-7 | 0.00016 | 2.06 | 8.71 | 100.70 down |
| ENSSSCG00000002821 | *CCL22* | Chemokine ligand 22 | 8.42E-7 | 0.00022 | -9.30 | 0.83 | 1130.42 down |
| ENSSSCG00000003041 | *CD79-ALPHA* | *Sus scrofa* CD79a molecule, immunoglobulin-associated alpha | 1.10E-5 | 0.0018 | -2.27 | 4.23 | 91.15 down |
| ENSSSCG00000003086 | *APOC2* | *Sus scrofa* apolipoprotein C-II | 1.22E-7 | 1.18E-6 | -1.18 | 7.36 | 375.58 down |
| ENSSSCG00000003608 | *LCK* | Lymphocyte-specific protein tyrosine kinase | 1.25E-10 | 2.13E-7 | -1.95 | 6.00 | 249.13 down |

*Continued……*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ENSSSCG00000004001 | *A1BG* | Alpha-1-B glycoprotein | 3.62E-6 | 0.00072 | -1.84 | 5.42 | 48.11 down |
| ENSSSCG00000004038 | *PLG* | Plasminogen | 3.43E-5 | 0.0047 | -1.37 | 4.80 | 72.43 down |
| ENSSSCG00000004454 | *ME1* | Malic enzyme 1 | 1.56E-9 | 1.43E-6 | 2.16 | 5.28 | 12.75 up |
| ENSSSCG00000004754 | *CHAC1* | ChaC, cation transport regulator homolog 1 | 6.27E-7 | 0.00017 | 4.45 | 0.95 | 11.32 up |
| ENSSSCG00000004779 | *PLCB2* | Phospholipase C, beta 2 | 4.24E-7 | 0.000129 | -2.77 | 2.71 | 44.88 down |
| ENSSSCG00000005312 | *SIT1* | Signaling threshold regulating transmembrane adaptor 1 | 6.21E-6 | 0.0011 | -9.73 | 6.05 | 19644.70 down |
| ENSSSCG00000006355 | *APOA2* | Apolipoprotein A-II | 2.22E-5 | 7.61E-5 | -1.17 | 5.98 | 142.72 down |
| ENSSSCG00000006372 | *ARHGAP30* | Rho GTPase activating protein 30 | 2.72E-8 | 1.28E-5 | -1.10 | 4.62 | 53.21 down |
| ENSSSCG00000006378 | *SLAMF7* | SLAM family member 7 | 9.91E-4 | 6.18E-6 | -9.30 | 0.76 | 1077.86 down |
| ENSSSCG00000006452 | *CD1D* | *Sus scrofa* CD1d molecule | 2.81E-9 | 2.34E-6 | -1.88 | 4.19 | 67.48 down |
| ENSSSCG00000006588 | *S100A9* | *Sus scrofa* S100 calcium binding protein A9 | 7.23E-5 | 0.0091 | 0.11 | 5.42 | 39.70 down |
| ENSSSCG00000006736 | *CD2* | *Sus scrofa* CD2 molecule | 5.70E-9 | 3.98E-6 | -0.60 | 5.51 | 69.64 down |
| ENSSSCG00000007082 | *OTOR* | Otoraplin | 1.51E-7 | 5.61E-5 | 4.57 | -0.43 | 32.14 up |
| ENSSSCG00000007656 | *PVRIG* | Poliovirus receptor related immunoglobulin | 3.12E-5 | 0.0043 | -9.23 | 5.49 | 104.62 down |
| ENSSSCG00000008128 | *DUSP2* | Dual specificity phosphatase 2 | 1.60E-6 | 0.00037 | -2.60 | 4.31 | 121.26 down |
| ENSSSCG00000008193 | *ZAP70* | Zeta-chain associated protein kinase 70k | 2.99E-5 | 0.0042 | -2.19 | 4.02 | 74.75 down |
| ENSSSCG00000008213 | *CD8B* | *Sus scrofa* CD8B molecule | 1.80E-10 | 2.42E-7 | -11.49 | 5.97 | 392.06 down |
| ENSSSCG00000008217 | *CD8A* | Sus scrofa CD8A molecule | 7.28E-11 | 1.59E-7 | -9.42 | 8.69 | 14087.56 down |

*Continued……*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ENSSSCG00000008997 | *FGB* | *Sus scrofa* fibrinogen beta chain | 8.60E-8 | 3.49E-5 | -9.30 | 1.19 | 1448.72 down |
| ENSSSCG00000009148 | *LEF1* | *Sus scrofa* lymphoid enhancer-binding factor 1 | 1.47E-10 | 2.13E-7 | -3.15 | 5.54 | 418.19 down |
| ENSSSCG00000009237 | *HPSE* | Heparanase | 6.25E-5 | 0.0080 | 2.72 | 6.27 | 22.97 down |
| ENSSSCG00000009261 | *CSN2* | *Sus scrofa* casein beta | 7.55E-5 | 0.0094 | -11.56 | 6.87 | 10880.08 down |
| ENSSSCG00000009412 | *LCP1* | Lymphocyte cytosolic protein 1 | 1.03E-5 | 0.0017 | 1.69 | 6.45 | 27.18 down |
| ENSSSCG00000010273 | *TBATA* | Thymus, brain and testes associated | 1.65E-11 | 7.49E-8 | -2.40 | 6.34 | 419.24 down |
| ENSSSCG00000010554 | *SCD* | *Sus scrofa* stearoyl-CoA desaturase | 3.47E-5 | 0.0047 | 5.61 | 1.13 | 22.25 up |
| ENSSSCG00000010780 | *CYP2E1* | Cytochrome P450, family 2, subfamily E, polypeptide 1 | 1.95E-7 | 6.81E-5 | -2.31 | 4.81 | 139.99 down |
| ENSSSCG00000010908 | *PTPRC* | Protein tyrosine phosphatase, receptor type, C | 4.36E-6 | 0.00084 | -0.77 | 4.68 | 43.87 down |
| ENSSSCG00000011579 | *PPARG* | Peroxisome proliferator-activated receptor gamma | 5.61E-5 | 0.00737 | 5.05 | 2.23 | 7.04 up |
| ENSSSCG00000011640 | *TF* | Serotransferrin; Transferrin | 8.77E-7 | 0.00023 | 1.84 | 8.13 | 78.15 down |
| ENSSSCG00000011800 | *FETUB* | Fetuin B | 1.06E-9 | 1.15E-6 | 1.43 | 4.28 | 452.84 down |
| ENSSSCG00000011801 | *HRG* | Histidine-rich glycoprotein | 2.80E-6 | 0.00058 | 1.69 | 65.28 | 32.21 down |
| ENSSSCG00000012347 | *ALAS2* | Aminolevulinate, delta-, synthase 2 | 1.02E-5 | 0.0017 | 1.12 | 5.72 | 24.36 down |
| ENSSSCG00000012397 | *IL2RG* | *Sus scrofa* interleukin 2 receptor, gamma | 2.80E-6 | 0.00058 | -1.47 | 4.68 | 28.81 down |
| ENSSSCG00000012480 | *TNMD* | Tenomodulin | 5.25E-11 | 1.50E-7 | 5.75 | -0.55 | 79.72 up |
| ENSSSCG00000012652 | *SASH3* | SAM and SH3 domain containing 3 | 3.56E-9 | 2.82E-6 | -2.50 | 4.14 | 100.51 down |
| ENSSSCG00000012909 | *PTPRCAP* | Protein tyrosine phosphatase, receptor type, C-associated protein | 2.81E-5 | 0.0040 | 1.68 | 6.44 | 27.11 down |

*Continued…..*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ENSSSCG00000012912 | *TBC1D10C* | TBC1 domain family, member 10C | 3.10E-6 | 0.00063 | -3.30 | 3.633 | 122.33 down |
| ENSSSCG00000012970 | *CTSW* | Cathepsin W | 1.07E-8 | 6.27E-6 | -0.43 | 6.14 | 96.16 down |
| ENSSSCG00000013115 | *CD5* | *Sus scrofa* CD5 molecule | 6.02E-11 | 1.50E-7 | -3.42 | 4.72 | 283.29 down |
| ENSSSCG00000013252 | *F2* | Coagulation Factor II (Thrombin) | 3.12E-6 | 0.00063 | -1.99 | 4.26 | 76.80 down |
| ENSSSCG00000013578 | *CLEC4G* | *Sus scrofa* C-type lectin domain family 4 | 4.73E-6 | 0.0009 | -2.70 | 3.98 | 103.01 down |
| ENSSSCG00000013593 | *CCL25* | *Sus scrofa* chemokine (C-C motif) ligand 25 | 2.29E-8 | 1.11E-5 | -1.88 | 5.34 | 149.54 down |
| ENSSSCG00000013612 | *ACP5* | Acid phosphatase 5 | 1.78E-5 | 0.0028 | 0.82 | 6.10 | 39.09 down |
| ENSSSCG00000013916 | *COMP* | Cartilage oligomeric matrix protein | 1.02E-8 | 6.18E-6 | 3.30 | -0.97 | 19.51 up |
| ENSSSCG00000013973 | *LYPD8* | LY6/PLAUR domain containing 8 | 1.11E-6 | 0.00028 | -11.56 | 6.42 | 174.91 down |
| ENSSSCG00000014626 | *HPX* | Hemopexin | 1.70E-5 | 0.0027 | -1.36 | 4.46 | 56.60 down |
| ENSSSCG00000015092 | *CD3E* | CD3e molecule, epsilon | 5.56E-12 | 4.86E-8 | -5.19 | 4.18 | 574.64 down |
| ENSSSCG00000015093 | *CD3D* | *Sus scrofa* CD3d molecule, delta | 1.88E-8 | 9.99E-6 | -9.30 | 4.63 | 15752.36 down |
| ENSSSCG00000015094 | *CD3G* | *Sus scrofa* CD3g molecule, gamma | 4.25E-9 | 3.09E-6 | -8.76 | 4.64 | 319.24 down |
| ENSSSCG00000015287 | *TMCC2* | Transmembrane and coiled-coil domain family 2 | 6.80E-6 | 0.0012 | -2.49 | 3.38 | 58.80 down |
| ENSSSCG00000015481 | *MYOC* | Myocilin, trabecular meshwork inducible glucocorticoid response | 5.19E-6 | 0.00097 | 3.78 | 0.50 | 9.67 up |
| ENSSSCG00000015656 | *FAIM3* | Fas apoptotic inhibitory molecule 3 | 1.62E-6 | 0.00037 | -9.30 | 3.78 | 8689.73 down |
| ENSSSCG00000015688 | *CXCR4* | *Sus scrofa* chemokine receptor 4 | 1.94E-5 | 0.0029 | 0.82 | 5.22 | 21.15 down |

*Continued……*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ENSSSCG00000015716 | *MARCO* | Macrophage receptor with collagenous | 3.65E-8 | 1.63E-5 | 6.80 | 7.05 | 70.01 down |
| ENSSSCG00000016657 | *AOAH* | Acyloxyacyl hydrolase | 1.51E-7 | 5.61E-5 | -1.47 | 5.15 | 98.75 down |
| ENSSSCG00000016823 | *C1QTNF3* | C1q and tumor necrosis factor related protein 3 | 1.32E-5 | 0.00219 | 0.33 | -2.53 | 7.34 up |
| ENSSSCG00000016832 | *IL7R* | Interleukin 7 receptor | 2.43E-6 | 0.0005 | -9.30 | 0.91 | 1188.00 down |
| ENSSSCG00000016855 | *FYB* | FYN binding protein | 1.50E-7 | 5.61E-5 | -2.20 | 4.25 | 2.47 down |
| ENSSSCG00000017037 | *FABP6* | Fatty acid binding protein 6 | 1.14E-9 | 1.17E-6 | -9.30 | 6.55 | 59287.92 down |
| ENSSSCG00000017358 | *SLC4A1* | Solute carrier family 4 member 1 | 1.53E-8 | 8.37E-6 | -2.25 | 3.22 | 44.53 down |
| ENSSSCG00000017368 | *PPY* | Pancreatic polypeptide | 1.54E-7 | 5.61E-5 | 8.57 | 4.40 | 17.98 up |
| ENSSSCG00000017433 | *KRT14* | Keratin 14 | 1.06E-6 | 0.00027 | -3.42 | 3.933 | 163.67 down |
| ENSSSCG00000017466 | *CCR7* | Chemokine receptor 7 | 8.16E-6 | 0.00145 | -4.08 | 3.38 | 178.05 down |
| ENSSSCG00000017498 | *PPP1R1B* | Protein phosphatase 1, regulatory subunit 1B | 4.29E-5 | 0.0057 | 0.72 | 1.75 | 6.44 up |
| ENSSSCG00000017943 | *ACAP1* | ArfGAP with coiled-coil, ankyrin repeat and PH domains 1 | 1.35E-7 | 5.37E-5 | -1.45 | 4.45 | 60.18 down |
| ENSSSCG00000021602 | *TTC36* | Tetratricopeptide repeat domain 36 | 1.80E-5 | 0.002 | 5.36 | 2.15 | 9.26 up |
| ENSSSCG00000021944 | *RAC2* | Ras-related C3 botulinum toxin substrate 2 | 2.06E-6 | 0.00045 | 5.08 | 6.96 | 3.68 down |
| ENSSSCG00000022500 | *IL20RB* | Interleukin 20 receptor beta | 2.30E-7 | 7.64E-5 | -9.30 | 3.03 | 5176.13 down |
| ENSSSCG00000023497 | *CAMK4* | Calcium/calmodulin-dependent protein kinase IV | 1.18E-8 | 6.64E-6 | -4.62 | 3.62 | 304.76 down |
| ENSSSCG00000023772 | *SPN* | Sialophorin | 2.10E-8 | 1.049 | 2.86 | 6.42 | 2.14 down |
| ENSSSCG00000024447 | *SPOT14* | Thyroid hormone responsive | 7.64E-5 | 0.0095 | 5.84 | 1.24 | 28.46 up |
| ENSSSCG00000025042 | *ICOS* | Inducible T-cell costimulator | 4.19E-5 | 0.0056 | -9.30 | -1.54 | 216.85 down |

*Continued……*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ENSSSCG00000025410 | *PRSS16* | Protease, serine, 16 | 1.80E-14 | 3.15E-10 | -9.30 | 7.62 | 125124.46 down |
| ENSSSCG00000025523 | *COL2A1* | Collagen, type II, alpha 1 | 2.77E-7 | 8.97E-5 | -3.68 | -1.12 | 5.90 down |
| ENSSSCG00000026689 | *CCDC114* | Coiled-coil domain containing 114 | 2.26E-5 | 0.0033 | -2.48 | 4.28 | 99.24 down |
| ENSSSCG00000027609 | *GC* | Group-specific component | 1.48E-6 | 0.00035 | -0.01 | 5.85 | 58.50 down |
| ENSSSCG00000028525 | *SAA4* | Serum amyloid A4 | 1.06E-5 | 0.00180 | -9.30 | -2.42 | 117.67 down |
| ENSSSCG00000028731 | *CXCL13* | Chemokine ligand 13 | 2.94E-5 | 0.0042 | -0.73 | 4.30 | 32.90 down |
| ENSSSCG00000029515 | *PON3* | Paraoxonase 3 | 1.85E-5 | 0.0028 | -1.84 | 6.24 | 9.00 up |
| ENSSSCG00000029691 | *CORO1B* | Coronin, actin binding protein, 1B | 2.03E-6 | 0.00045 | 4.42 | 5.20 | 1.71 down |
| ENSSSCG00000029782 | *CELA1* | Chymotrypsin-like elastase family, member 1 | 1.93E-5 | 0.0029 | 5.12 | 2.01 | 8.61 up |
| ENSSSCG00000029944 | *FASN* | Fatty acid synthase | 8.12E-5 | 0.0099 | 5.27 | 1.26 | 6.12 up |
| ENSSSCG00000030680 | *TCF7* | Transcription factor 7 | 1.211E-6 | 0.00029 | 2.63 | 2.22 | 1.32 up |

**Table 6.6** List of DEGs (FDR<0.01) and the log fold changes in liver tissue of Berkshire and JNP piglets.

| ***Sus scrofa* Ensemble ID** | **Gene** | **Description** | **Expression level (RPKM)** | | | | |
|---|---|---|---|---|---|---|---|
| | | | ***P* value < 0.01** | **FDR < 0.01** | **JNP** | **Berkshire** | **Fold change in JNP** |
| ENSSSCG00000000046 | *CYP2D6* | Vitamin D3 25-Hydroxylase | 1.15E-5 | 0.0024 | 8.41 | 5.63 | 6.87 up |
| ENSSSCG00000000289 | *NFE2* | Nuclear factor, erythroid 2 | 5.18E-6 | 0.0012 | 0.19 | 3.99 | 13.89 down |
| ENSSSCG00000003006 | *CYP2B22* | Cytochrome P450 2B22 | 1.02E-9 | 5.70E-7 | 5.71 | 1.75 | 15.56 up |
| ENSSSCG00000003099 | *PGLYRP* | Peptidoglycan recognition protein 1 | 8.22E-6 | 0.0018 | -1.18 | 3.89 | 33.81 down |
| ENSSSCG00000003547 | *NCMAP* | Noncompact myelin associated protein | 2.68E-5 | 0.0048 | -9.38 | 1.96 | 2620.58 down |
| ENSSSCG00000003590 | *PTPRU* | Protein tyrosine phosphatase, receptor type, U | 2.77E-5 | 0.0049 | 2.28 | -0.04 | 5.02 up |
| ENSSSCG00000003669 | *MFSD2A* | Major facilitator superfamily domain containing 2A | 5.60E-14 | 6.91E-11 | 3.69 | -1.39 | 33.89 up |
| ENSSSCG00000004167 | *MYB* | V-myb avian myeloblastosis viral oncogene homolog | 6.78E-7 | 0.0002 | -2.53 | 1.94 | 22.19 down |
| ENSSSCG00000004492 | *SLC14A1* | Solute carrier family 14 (urea transporter), member 1 | 7.08E-6 | 0.0016 | -9.38 | 0.83 | 1194.42 down |
| ENSSSCG00000004509 | *LIPG* | Lipase, endothelial | 2.31E-5 | 0.004 | 5.31 | 2.96 | 5.08 up |
| ENSSSCG00000004714 | *EPB42* | Erythrocyte membrane protein band 4.2 | 1.06E-13 | 1.22E-10 | -1.32 | 3.84 | 35.99 down |
| ENSSSCG00000006142 | *CA1* | Carbonic anhydrase I | 5.35E-8 | 2.20E-5 | -9.38 | 3.31 | 6644.97 down |

*Continued……*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ENSSSCG00000006403 | *CRP* | C-reactive protein, pentraxin-related | 7.94E-11 | 6.52E-8 | 9.52 | 4.83 | 25.86 down |
| ENSSSCG00000006431 | *SPTA1* | Spectrin, alpha, erythrocytic 1 | 1.64E-10 | 1.09E-7 | -2.00 | 3.23 | 37.76 down |
| ENSSSCG00000006588 | *S100A9* | S100 calcium binding protein A9 | 9.36E-8 | 3.59E-5 | 2.30 | 5.74 | 10.84 down |
| ENSSSCG00000006589 | *S100A12* | S100 calcium binding protein A12 | 6.03E-9 | 2.97E-6 | 3.42 | 7.10 | 12.79 down |
| ENSSSCG00000006590 | *S100A8* | S100 calcium binding protein A8 | 1.03E-6 | 0.0002 | 1.80 | 5.77 | 15.62 down |
| ENSSSCG00000006719 | *HSD3B* | 3 beta-hydroxysteroid dehydrogenase | 1.23E-5 | 0.002 | 3.45 | -0.40 | 14.50 up |
| ENSSSCG00000007022 | *ANK1* | Ankyrin 1, erythrocytic | 1.17E-10 | 8.13E-8 | -1.84 | 3.31 | 18.07 down |
| ENSSSCG00000007528 | *PHACTR3* | Phosphatase and actin regulator 3 | 2.36E-6 | 0.0006 | -3.14 | 1.74 | 29.58 down |
| ENSSSCG00000009144 | *ETNPPL* | Ethanolamine-phosphate phospho-lyase | 8.26E-12 | 8.39E-9 | 7.59 | 3.29 | 19.62 up |
| ENSSSCG00000009585 | *GADD45G* | Growth arrest and DNA-damage-inducible, gamma | 1.53E-7 | 5.77E-5 | 5.27 | 2.79 | 5.59 up |
| ENSSSCG00000009708 | *AADAT* | Aminoadipate aminotransferase | 1.34E-21 | 1.16E-17 | 5.01 | -9.38 | 21569.52 up |
| ENSSSCG00000010233 | *DNAJC12* | DnaJ (Hsp40) homolog, subfamily C, member 12 | 1.99E-5 | 0.003 | 4.29 | 6.60 | 4.98 down |
| ENSSSCG00000011437 | *ALAS1* | Aminolevulinate, delta-, synthase 1 | 1.28E-5 | 0.002 | 2.17 | 5.73 | 6.06 up |
| ENSSSCG00000011800 | *FETUB* | Fetuin B | 8.41E-6 | 0.0018 | 6.77 | 10.08 | 9.96 down |
| ENSSSCG00000012147 | *S100G* | S100 calcium binding protein G | 1.43E-11 | 1.37E-8 | 3.72 | 7.83 | 17.28 down |
| ENSSSCG00000012265 | *CHST7* | Carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 7 | 2.30E-7 | 8.10E-5 | 1.33 | -3.59 | 30.52 up |
| ENSSSCG00000012347 | *ALAS2* | Aminolevulinate, delta-, synthase 2 | 9.60E-10 | 5.52E-7 | 7.94 | 5.34 | 11.76 down |

*Continued……*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ENSSSCG00000027266 | *PNPLA3* | Patatin-like phospholipase domain containing 3 | 2.45E-3 | 0.0018 | 2.48 | 8.24 | 11.26 down |
| ENSSSCG00000012600 | *AGTR2* | Angiotensin II receptor, type 2 | 1.37E-5 | 0.0027 | -3.66 | 1.55 | 37.33 down |
| ENSSSCG00000012852 | *CDHR5* | Cadherin-related family member 5 | 3.37E-5 | 0.005 | 5.58 | 3.38 | 4.61 up |
| ENSSSCG00000013087 | *DAK* | Dihydroxyacetone kinase 2 | 4.35E-5 | 0.007 | 6.08 | 3.75 | 5.03 up |
| ENSSSCG00000013180 | *YPEL4* | Yippee-like 4 | 3.17E-5 | 0.005 | -9.38 | 1.97 | 2635.93 down |
| ENSSSCG00000013378 | *ABCC8* | ATP-binding cassette, sub-family C member 8 | 1.73E-6 | 0.0004 | -9.38 | 0.69 | 1083.20 down |
| ENSSSCG00000013415 | *AZU1* | Azurocidin 1 | 7.73E-6 | 0.0017 | -1.52 | 3.31 | 28.64 down |
| ENSSSCG00000013730 | *KLF1* | Kruppel-like factor 1 | 2.16E-10 | 1.33E-7 | -9.38 | 3.26 | 5427.73 down |
| ENSSSCG00000013958 | *TRIM58* | Tripartite motif containing 58 | 4.04E-6 | 0.0010 | -9.38 | 2.58 | 4028.03 down |
| ENSSSCG00000014725 | *HBB* | Hemoglobin, beta | 1.55E-5 | 0.003 | 8.53 | 11.72 | 9.11 down |
| ENSSSCG00000015287 | *TMCC2* | Transmembrane and coiled-coil domain family 2 | 8.84E-11 | 6.94E-8 | -1.00 | 3.70 | 26.06 down |
| ENSSSCG00000016247 | *TM4SF20* | Transmembrane 4 L six family member 20 | 6.76E-7 | 0.0002 | 3.56 | -0.17 | 13.34 up |
| ENSSSCG00000016381 | *SNED1* | Sushi, nidogen and EGF-like domains 1 | 5.90E-5 | 0.009 | 1.95 | -0.67 | 6.16 up |
| ENSSSCG00000017262 | *SLC16A6* | Solute carrier family 16, member 6 | 1.02E-6 | 0.0002 | 4.86 | 2.36 | 11.29 up |
| ENSSSCG00000017637 | *MPO* | Myeloperoxidase | 1.16E-10 | 8.13E-8 | -0.74 | 3.70 | 21.88 up |
| ENSSSCG00000020736 | *ADD2* | Adducin 2 | 1.90E-6 | 0.0005 | -1.15 | 2.43 | 12.04 down |
| ENSSSCG00000021847 | *SAA4* | Serum amyloid A4, constitutive | 6.85E-22 | 1.16E-17 | 6.77 | -9.38 | 73373.01 up |

*Continued……*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ENSSSCG00000022284 | *DUSP6* | Dual specificity phosphatase 6 | 5.60E-5 | 0.0091 | 5.18 | 3.64 | 2.90 up |
| ENSSSCG00000022427 | *RGS1* | Regulator of G-protein signaling 1 | 6.14E-5 | 0.0096 | 3.60 | 0.78 | 7.01 up |
| ENSSSCG00000023772 | *SPN* | Sialophorin | 3.80E-5 | 0.0065 | 0.06 | 2.34 | 4.84 down |
| ENSSSCG00000023773 | *ARG1* | Arginase-1 | 1.50E-5 | 0.0029 | 8.56 | 6.41 | 4.43 down |

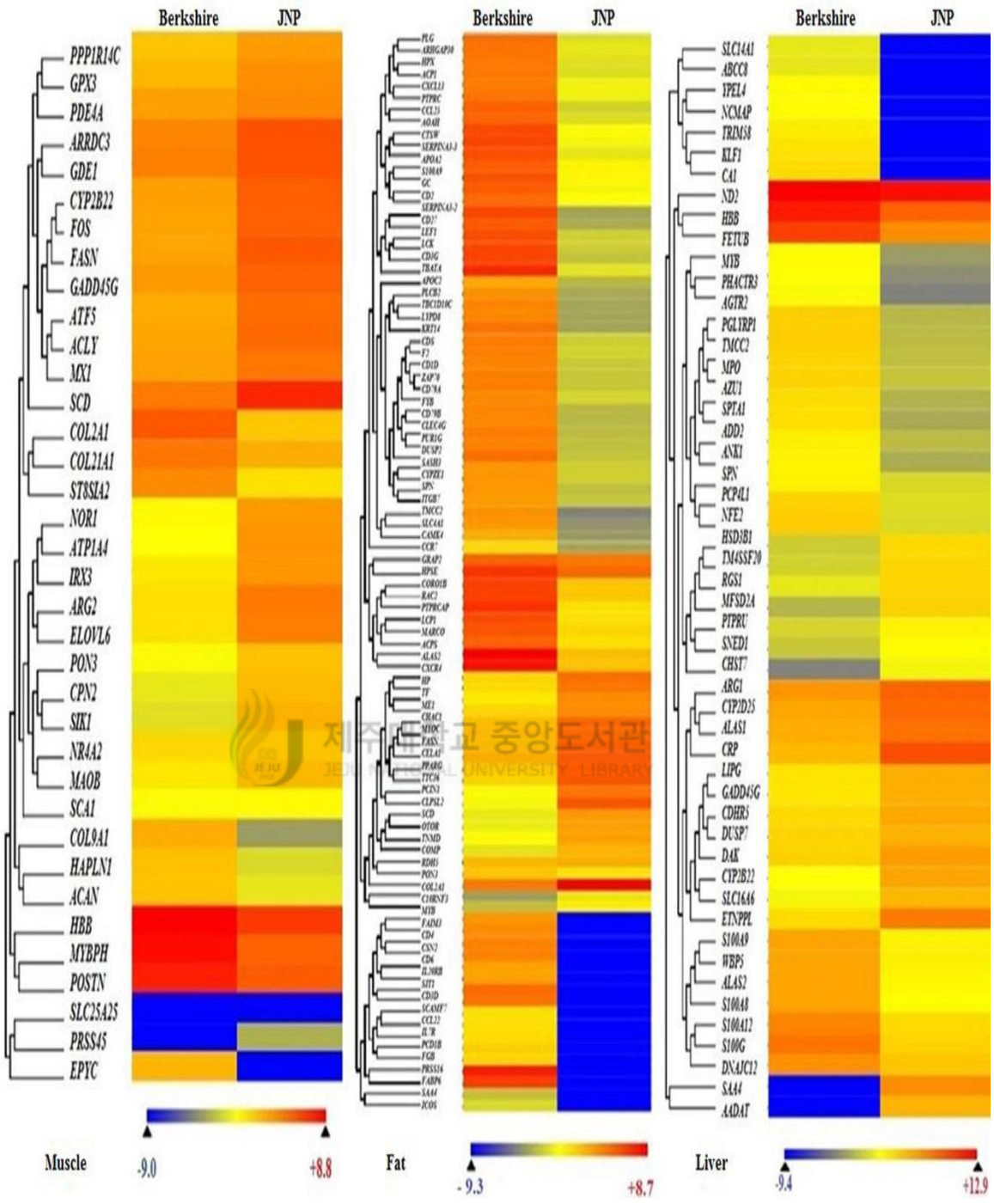


**Figure 6.3** Heatmap of DEGs in *longissimus dorsi* muscle, fat and liver tissues. Each row of colored tiles in the grid corresponds to one gene and each column represents a specific breed. The Gene tree drawn to the left of the heat map displays the clustering of genes.

**Table 6.7** Gene Ontology (GO) categories enriched for co-expressed genes modules from *longissimus dorsi m*uscle tissue ($P$<0.01).

| Term description | GO ID | Involved Genes number |
|---|---|---|
| ***Biological process*** | | |
| Extracellular matrix organization | GO:0030198 | *ACAN, COL21A1, COL2A1, COL9A1, COL9A2, HAPLN1, POSTN, SERPINE1* |
| Extracellular structure organization | GO:0043062 | *ACAN, COL21A1, COL2A1, COL9A1, COL9A2, HAPLN1, POSTN, SERPINE1* |
| Triglyceride biosynthetic process | GO:0019432 | *ACLY, ELOVL6, FASN, SIK1* |
| Long-chain fatty-acyl-coa biosynthetic process | GO:0035338 | *ACLY, ELOVL6, FASN* |
| Neutral lipid biosynthetic process | GO:0046460 | *ACLY, ELOVL6, FASN, SIK1* |
| Acylglycerol biosynthetic process | GO:0046463 | *ACLY, ELOVL6, FASN, SIK1* |
| Fatty-acyl-coa biosynthetic process | GO:0046949 | *ACLY, ELOVL6, FASN* |
| Skeletal muscle cell differentiation | GO:0035914 | *ANKRD1, ATF3, FOS, NR4A1* |
| Long-chain fatty-acyl-coa metabolic process | GO:0035336 | *ACLY, ELOVL6, FASN* |
| Fatty-acyl-CoA metabolic process | GO:0035337 | *ACLY, ELOVL6, FASN* |
| ***Molecular function*** | | |
| Extracellular matrix structural constituent conferring tensile strength | GO:0030020 | *COL2A1, COL9A1, COL9A2* |
| Extracellular matrix structural constituent | GO:0005201 | *ACAN, COL2A1, COL9A1, COL9A2* |
| Fatty acid synthase activity | GO:0004312 | *ELOVL6, FASN* |
| ***Cellular component*** | | |
| Proteinaceous extracellular matrix | GO:0005578 | *ACAN, COL21A1, COL2A1, COL9A1, COL9A2, EPYC, HAPLN1, POSTN* |

*Continued……*

| | | |
|---|---|---|
| Extracellular matrix | GO:0031012 | *ACAN, COL21A1, COL2A1, COL9A1, COL9A2, EPYC, HAPLN1, POSTN, SERPINE1* |
| Extracellular region part | GO:0044421 | *ACAN, COL21A1, COL2A1, COL9A1, COL9A2, EPYC, GPX3, HAPLN1, ISG15, PON3, POSTN, SERPINE1* |
| Collagen type IX | GO:0005594 | *COL9A1, COL9A2* |
| Collagen | GO:0005581 | *COL21A1, COL2A1, COL9A1, COL9A2* |
| Extracellular matrix part | GO:0044420 | *ACAN, COL21A1, COL2A1, COL9A1, COL9A2* |
| FACIT collagen | GO:0005593 | *COL9A1, COL9A2* |
| Anchoring collagen | GO:0030934 | *COL9A1, COL9A2* |

**Table 6.8** Gene Ontology (GO) categories enriched for co-expressed genes modules from Fat tissue ($P$<0.01).

| Term description | GO ID | Involved Genes number |
|---|---|---|
| ***Biological process*** | | |
| Regulation of interleukin-4 production | GO:0032673 | *CD1D, CD3E, IL20RB, LEF1, SASH3* |
| Inflammatory response | GO:0006954 | *ACP5, AOAH, APOA2, C1QTNF3, CAMK4, CCL22, CCL25, CCR7, CELA1, CXCL13, CXCR4, F2, HP, IL20RB, PPARG, S100A9, SAA4, SPN* |
| Interleukin-4 production | GO:0032633 | *CD1D, CD3E, IL20RB, LEF1, SASH3* |
| Antigen receptor-mediated signaling pathway | GO:0050851 | *CCR7, CD3D, CD3E, CD3G, CD4, FYB, GRAP2, LCK, PTPRC, ZAP70* |
| Positive thymic T cell selection | GO:0045059 | *CD1D, CD3D, PTPRC, ZAP70* |
| Regulation of lymphocyte differentiation | GO:0045619 | *CAMK4, CCR7, CD1D, CD2, CD27, IL2RG, IL7R, LCK, PTPRC, SASH3, ZAP70* |
| Alpha-beta T cell activation | GO:0046631 | *CCR7, CD1D, CD3E, IL2RG, LEF1, PTPRC, SASH3, TCF7, ZAP70* |
| T cell mediated immunity | GO:0002456 | *CD1D, CD8A, IL20RB, IL7R, PTPRC, SASH3, SPN* |
| Defense response | GO:0006952 | *ACP5, AOAH, APOA2, C1QTNF3, CAMK4, CCL22, CCL25, CCR7, CD1D, CD3G, CD4, CD8A, CD8B, CELA1, CXCL13, CXCR4, F2, FAIM3, GRAP2, HP, HPX, HRG, IL20RB, LCK, MARCO, PPARG, PTPRC, PTPRCAP, S100A9, SAA4, SCD, SLAMF7, SPN* |

*Continued……*

| | | |
|---|---|---|
| Immune response | GO:0006955 | *APOA2, CAMK4, CCL22, CCL25, CCR7, CD1D, CD27, CD3D, CD3E, CD3G, CD4, CD8A, CD8B, CTSW, CXCL13, FYB, GRAP2, HPX, HRG, ICOS, IL20RB, IL2RG, IL7R, ITGB7, LCK, LCP1, LEF1, MARCO, PPARG, PTPRC, S100A9, SASH3, SLAMF7, SPN, TCF7, ZAP70* |
| Immune system development | GO:0002520 | *ALAS2, CAMK4, CCR7, CD1D, CD2, CD27, CD3D, CD3E, CD4, CD8A, CXCL13, ICOS, IL2RG, IL7R, LCK, LEF1, PPARG, PTPRC, SASH3, SPN, TCF7, TF, ZAP70* |
| Acute Inflammatory response | GO:0002526 | *APOA2, CCR7, F2, HP, IL20RB, PPARG, SAA4, SPN* |
| ***Molecular Function*** | | |
| Glycoprotein binding | GO:0001948 | *CD2, CD4, CD5, COMP, HRG, LCK, PTPRC* |
| Receptor binding | GO:0005102 | *APOA2, CCL22, CCL25, CD1D, CD2, CD3E, CD3G, CD4, CD8A, CD8B, CHAC1, CXCL13, F2, FGB, FYB, LCK, LEF1, PPARG, PPP1R1B, PPY, S100A9* |
| T cell receptor binding | GO:0042608 | *CD1D, CD3E, CD3G* |
| Coreceptor activity | GO:0015026 | *CD4, CD8A, CD8B, CXCR4* |
| Interleukin-7 receptor activity | GO:0004917 | *IL2RG, IL7R* |
| ***Cellular Component*** | | |
| External side of plasma membrane | GO:0009897 | *CCR7, CD1D, CD2, CD27, CD3E, CD4, CD5, CD8A, CD8B, CXCR4, FGB, ICOS, IL2RG, IL7R, PLG, PTPRC, SLAMF7, SPN* |
| Extracellular region part | GO:0044421 | *APOA2, APOC2, C1QTNF3, CCL22, CCL25, COL2A1, COMP, CXCL13, F2, FETUB, FGB, GC, HP, HPSE, HPX, MARCO, MYOC, PLG, PON3, S100A9, SAA4, SPN, TF* |
| T cell receptor complex | GO:0042101 | *CD3D, CD3E, CD3G, CD4, CD8A, CD8B, ZAP70* |
| Alpha-beta T cell receptor complex | GO:0042105 | *CD3D, CD3E, CD3G* |

**Table 6.9** Gene Ontology (GO) categories enriched for co-expressed genes modules from liver tissue ($P$ <0.01).

| Term description | GO id | Involved Genes numbers |
|---|---|---|
| ***Biological process*** | | |
| Defense response to bacterium | GO:0042742 | *AZU1, CRP, MPO, S100A12, S100A8, S100A9, SPN* |
| Defense response to fungus | GO:0050832 | *MPO, S100A12, S100A8, S100A9* |
| Erythrocyte development | GO:0048821 | *ANK1, EPB42, HBB, KLF1* |
| Porphyrin-containing compound biosynthetic process | GO:0006779 | *ALAS1, ALAS2, ANK1, SPTA1* |
| Myeloid cell development | GO:0061515 | *ANK1, EPB42, HBB, KLF1* |
| Leukocyte aggregation | GO:0070486 | *S100A8, S100A9, SPN* |
| ***Molecular Function*** | | |
| RAGE receptor binding | GO:0050786 | *S100A12, S100A8, S100A9* |
| N-succinyltransferase activity | GO:0016749 | *ALAS1, ALAS2* |
| 5-aminolevulinate synthase activity | GO:0003870 | *ALAS1, ALAS2* |
| Pyridoxal phosphate binding | GO:0030170 | *AADAT, ALAS1, ALAS2, ETNPPL* |
| Succinyltransferase activity | GO:0016748 | *ALAS1, ALAS2* |
| Antioxidant activity | GO:0016209 | *HBB, MPO, S100A8, S100A9* |
| Toll-like receptor 4 binding | GO:0035662 | *S100A8, S100A9* |
| ***Cellular component*** | | |
| Spectrin-associated cytoskeleton | GO:0014731 | *ANK1 SPTA1* |

**Table 6.10** Pathways analysis of the DEGs in *longissimus dorsi* muscle, fat and liver tissues.

| Pathway | *p*-Value | Gene List |
|---|---|---|
| ***Longissimus dorsi* muscle** | | |
| ECM proteoglycans | $2.54E^{-7}$ | *ACAN,COL9A1,COL9A2,HAPLN1,SERPINE1* |
| Protein digestion and absorption | $3.04E^{-6}$ | *ATP1A4,COL21A1,COL2A1,COL9A1,COL9A2* |
| Extracellular matrix organization | $1.44E^{-5}$ | *ACAN,COL21A1,COL2A1,COL9A1,COL9A2,HAPLN1, SERPINE1* |
| Fatty Acyl-CoA Biosynthesis | 2.88E-4 | *ACLY,ELOVL6,FASN* |
| Fatty Acid Biosynthesis | $3.15E^{-6}$ | *ACLY,FASN,SCD* |
| Collagen biosynthesis and modifying enzymes | $2.02E^{-4}$ | *COL21A1,COL2A1,COL9A1,COL9A2* |
| **Fat** | | |
| Primary immunodeficiency | 1.94E-08 | *CD3D,CD3E,CD4,CD8A,CD8B,ICOS,IL2RG,IL7R,LCK, PTPRC,ZAP70* |
| T-Cell Receptor (TCR) Signaling Pathway | 3.95E-09 | *CD3D,CD3E,CD3G,CD4,CD8A,FYB,GRAP2, ICOS,PTPRC,LCK, ZAP70* |
| CXCR4-mediated signaling events | 1.67E-2 | *CD3D,CD3E,CD3G,CD4,CXCR4,LCK,PLCB2,PTPRC* |
| PPAR signaling pathway | 2.67E-7 | *APOA2,FABP6,ME1,PLIN1,PPARG,SCD* |
| **Liver** | | |
| Hemoglobin's Chaperone | 1.49E-2 | *ALAS1,ALAS2,HBB* |
| Vitamin B12 Metabolism | $4.87\times 10^{-6}$ | *HBB,MPO,SAA4* |
| Folate Metabolism | 6.06E-9 | *HBB,MPO,SAA4* |
| Tetrapyrrole biosynthesis | 2.18E-5 | *ALAS1,ALAS2* |
| Selenium Pathway | $4.19\times 10^{-2}$ | *CRP,HBB,MPO,SAA4* |
| Heme biosynthesis | $3.89\times 10^{-5}$ | *ALAS1,ALAS2* |

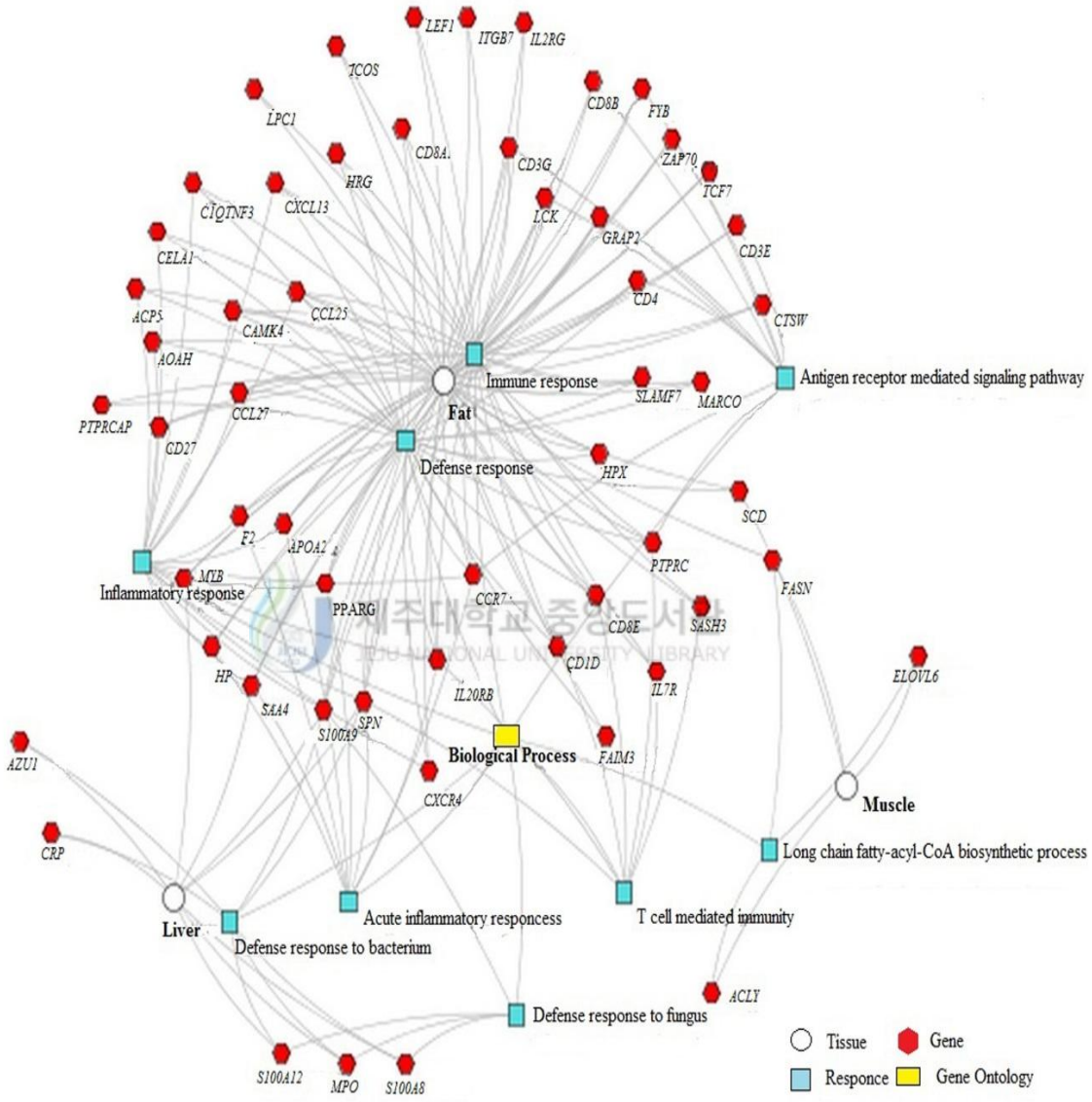


**Figure 6.4** An abstracted gene network of DEGs of *longissimus dorsi* muscle, fat and liver tissues. The gene network is based on the involvement of genes in significant biological pathways.

### *6.4.5. Exposition of genes involved in immune response and inflammation*

*Chemokine ligand-13* (*CXCL*13), a potent chemo-attractant and neutrophil activator is reported to be down regulated and *Src-like-adaptor* (*SLA*) an associate of the variation in the immune response is found to have slight higher expression (2.37 fold) in JNP (Table 6.4; Figure 6.5). 11 different types of *CD* genes have significantly expressed more in Berkshire (Table 6.5). *Interleukin-7 receptor* (*IL7R*) and *C-reactive protein* (*CRP*) are significantly ($P$<0.05) down regulated in JNP (Table 6.6; Figure 6.5).

### *6.4.6. Exposition of genes bounded with metabolic process*

The 12 DEGs of fatty-acyl-CoA metabolic process, shared by JNP and Berkshire have been included in the current study. *Patatin-like phospholipase domain containing-3* (*PNPLA3*) is reported to be down-regulated (11.26 folds, Table 6.6; Figure 6.5) whereas fatty acid *Elongase-6* (*ELOVL6*) is up-regulated (8.87 fold) in JNP (Table 6.4). An insight into the DEGs indicated that the *Fatty acid synthase* (*FASN)* is also up-regulated by 6.13 folds in JNP (Figure 6.5).



## 6.5. Discussion and conclusion

Piglets with higher growth performance and better immune system can be the stronger base for the breeding stock. For genetically improved breeding stock, there is dire need to select the piglets with marker genes for immune system and growth performance. Selection of such markers will help in the foundation of an improved breeding stock. To our knowledge current study is the first comprehensive insight into the transcriptome analysis of three tissues of JNP and Berkshire, one day old piglets using RNA-seq. In the current study, more than 90% seq reads generated by Illumina Hiseq2000 passed quality control. It suggests that the differences in annotated reads and mapping percentage might be due to the GC content, dinucleotide fragmentation sites, primer bases and the protocols followed at laboratories (Sendler *et al*., 2011). The extent of variation for the expression of genes in JNP and Berkshire using scatter plot analysis has also facilitated the visual comparison of the genes. The visual comparison indicated that genes on identity line in both species may have the same total number of SNPs and same class distribution.

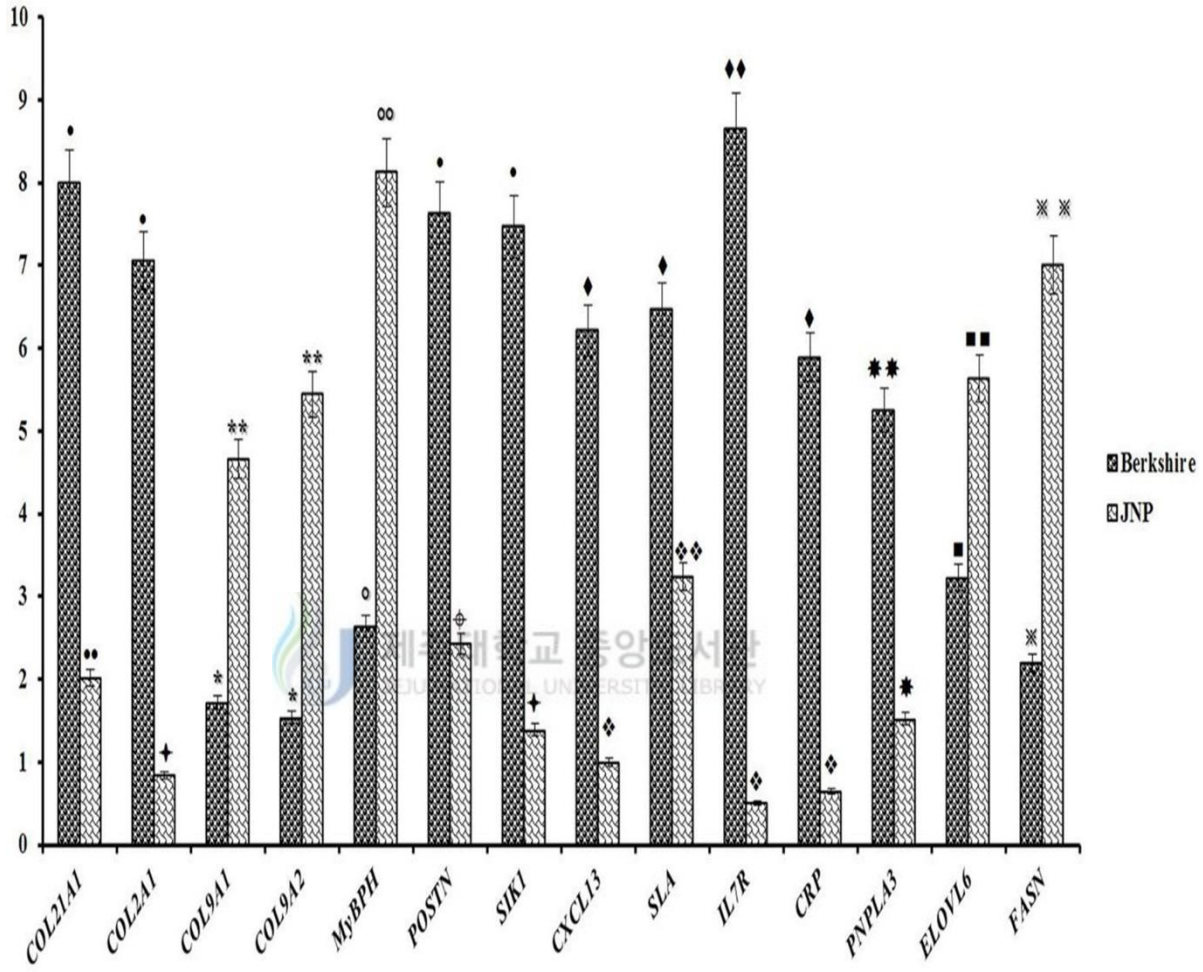


**Figure 6.5** Expression analyses of mRNA transcript levels of DEGs in JNP and Berkshire. Relative transcript levels of DEGs in the three tissues of JNP and Berkshire. Bars with different superscripts show significant differential transcript levels between the two breeds ($P$<0.05). Values are expressed as mean ± SEM.

An open source online visualization of matrix of the data at http://biopopdb.snu.ac.kr/PIG_DEG/ showed the differential expression of genes at *P<0.01* and FDR<0.01. At the same time the findings of the heat map analysis have also significantly (*P*<0.05) complimented the earlier observations. The results of GO and the KEGG analysis have suggested that the most of DEGs are involved in immune responses and metabolic processes. PPARs play a vital role in the clearance of cellular lipids via the regulation of the gene expression involved in lipid oxidation and cell proliferation (Mehla *et al*., 2014). The genes from collagen family are related to the skeletal system morphogenesis and growth of cartilage (Sylvie, 2011). In the current study, down regulation of the genes of collagen family indicates the lesser growth rate in JNP. *SIK1* a down regulated gene in JNP, after its activation is reported to phosphorylate muscle cells by inhibiting class-II histone deacetylases and it promotes expression of *MEF2* genes in myocytes. Moreover, *MEF2* is reported to regulate differentiation of vertebrate skeletal muscles in conjunction with bHLH transcription factors (Wang *et al*., 2001). Earlier studies proved that palatability of pork depends upon the type of muscle fibers. Myofibrils contain actin and myosin as the predominant protein in thin and thick filaments respectively. Therefore, higher transcript levels of *MyBPH* in JNP (Table 6.4; Figure 6.5) could be an indicative for luscious taste, tenderness and marbling characters. Therefore, it is suggested that these up- and down-regulated genes in the one day old piglets of JNP may act as markers for the growth improvement breeding programmes.



Periostein (*POSTN*) which regulates the fat deposition (Chen *et al*., 2011) is down regulated in JNP piglets (Table 6.4; Figure 6.5). It was also reported that *POSTN* gene enhance the incorporation of bone morphogenetic proteins which induces bone development and signaling pathways known for skeletal development (Rolfe *et al*., 2014). Current study has revealed highly enriched DEGs in the piglets that are involved in muscle growth, cytoskeletal rearrangement and development. With the identification of such kind of DEGs, piglets can act as a source for an improved breeding stock.

The variants of the genes related to immune response are considered as potential source of innate immune response (Kubistova *et al*., 2009). In current investigation, GO categories related to immune response and inflammation mainly corresponded with the fat and liver tissues. It is quite an interesting fact that most of the genes related with immune response expressed more in Berkshire piglets which indicated better immune system of the breed. *CXCL13,* a down regulated gene in JNP piglets is reported to induce mobilization of neutrophils from bone marrow (Akha *et al*., 2013). An earlier study reported that the expression of *CXCL13* could enhance or suppress immune functions depending on the infection and environmental conditions (Mehla *et*

*al*., 2014). *CRP* is secreted by the liver in response to a variety of inflammatory cytokines (DU, 2000) and its measurement is widely used to monitor various inflammatory states. The differences in the baseline immunity among different breeds have been observed, which would be affected by age of the pig. Moreover, Berkshire pigs of all age groups are reported to have the highest ($P$<0.01) neutrophil and lymphocyte counts as compared to other breeds (Sutherland *et al*., 2005). This recognition provides early defense through adaptive immune system. Significant representation of immune-related DEGs in the current study reflects an underlying expansion of immune surveillance in JNP and Berkshire piglets.

Lipids along with their metabolites function as transcriptional regulators for gene expression and fundamentally affect physiological processes. *PNPLA3* mainly expresses in porcine adipose and liver tissue (Chen *et al*., 2011). Fatty acid composition is a critical aspect for meat quality and its variation effects flavor, color, firmness and softness of the fat in meat (Wood *et al*., 2004). *ELOVL6* can be a promising positional and functional candidate gene for the meat quality. *ELOVL6* is suggested to affect the percentages of palmitic and palmitoleic fatty acids in pigs (Corominas *et al*., 2013) which in-turn affects the quality parameters of pork. *FASN* is associated with the meat quality and fatty acid composition in pig. A study in Korean native pig reported that *FASN* would be another candidate gene for producing high quality pork (Kim *et al*., 2011).



The present investigation is among the pioneer reports on comparative study of the genes related to skeletal system morphogenesis, body growth and immune system of piglets of JNP and Berkshire. Earlier studies suggested larger fraction of the transcriptome variability with respect to tissue differences rather than the breeds. In the current study, the expression profiles of DEGs between JNP and Berkshire along with their most profuse biological pathways have been investigated. The identification of specialized biological functions and regulatory gene/s through transcriptomic analysis can be used as selection markers for further breed improvement programmes. Moreover, transcriptome analysis makes it easier to understand the differences between genetic mechanisms of breeds. Our findings can be source for further extensive *in-vitro* and *in-vivo* studies for the regulated expression of genes enlisted in the current study. Therefore, it can be concluded here that the current findings would add up in the process of improvement in the pig genome annotations.

# Chapter 7

## Systematic *in-silico* annotation of single nucleotide polymorphisms of *MyBPH* for Jeju Native Pig and Berkshire and *in-vitro* acquisition of functional analysis of the correlation of *MyBPH* with the muscle regulatory factors associated with the skeleton muscle growth

### 7.1. Abstract

The demand of pig meat is increasing day by day and during the last decade the quality of pork has become one of the main selection criteria. *MyBPH* has significantly higher transcript level in the *longissimus dorsi* muscle of JNP as compare to Berkshire. The current study targets the role of the *MyBPH* in the muscle growth of JNP. The systematic *in-silico* annotations of SNPs and *in-vitro* functional analysis of *MyBPH* and other muscle regulatory factors have been performed. Different levels of polymorphism among the SNP sites were confirmed in *MyBPH* by sequence- and structure-based algorithms. More than 80% of total nsSNPs have been predicted to be highly deleterious in JNP. *In silico* analysis by I-Mutant3 and HOPE tool has indicated that misfiling and intermolecular interaction of the *MyBPH* can significantly ($P$<0.05) affect the structures and functions of the protein. Therefore, such kind of changes due to nsSNPs target the functioning of *MyBPH* and inspite of higher transcript levels, *MyBPH* is still not able to potentiate the myogenesis in JNP. The sharing of specific myogenic features associated with *MyBPH* and MRFs through ToppCluster indicates that *MyBPH* can be a potent candidate gene to potentiate biological processes. The mutations in *MyBPH* can alter the muscle-specific biological and pathophysiological processes involved in muscle growth and metabolism. The current study is one of the pioneer transcriptomic studies being conducted in JNP. Our findings reveal *MyBPH* as a candidate gene for the muscle specific biological and pathophysiological processes.



***Keywords:*** *MyBPH*, nsSNPs, Jeju Native Pig (JNP), Berkshire

## 7.2. Introduction

The *Sus scrofa* is considered an economical source of animal protein. Fresh pork with high marbling score is the most preferred meat. During the last decade much attention has been given on the quality of pork and it has become one of the main selection criteria. Therefore, such kinds of quality parameters give thrust to the increased quality production of pork (Gorni *et al*., 2008; Luo *et al*., 2012). JNP is an indigenous breed of Jeju Island and consumers have preference for black coat color than white coat color in pigs. That is why it is one of the reasons the pork from JNP has been regarded as the most expensive and high-quality pork in Korea (Hur *et al*., 2013). A drastic decrease in the population of JNP since early 20th century is attributed to their slower growth rate, lower productivity (Sodhi *et al*., 2014) and poor genetic traits as compared to the commercial breeds (Kim *et al*., 2002). Therefore JNP has been crossbred with commercial breeds to upgrade its economic traits (Park *et al*., 2007; Kim *et al*., 2013). Among all of the commercial breeds, Berkshire is closely related with the Asian breeds of pig (Kim *et al*., 2002). Suzki *et al*. (2003) reported that in Japan, the retail price of purebred Berkshire meat is 50% higher than that of regular commercial pork meat. High marbling score, drip loss and water holding capacity of the pork from the crossbreds of JNP and Berkshire make it the most preferred meat in Korea (Lee *et al*., 2011). Currently, it has been reported that in Korea, pork produced from crossbreds of Berkshire and Jeju Black pig, is sold at a higher price than regular commercial pork (Oh *et al*., 2014).



The molecular basis of the phenotypic expression of skeletal muscles along with their interaction with different environmental factors is providing valuable insight to improve our understanding of myogenesis and muscle growth. One of the recent approaches in vogue relies on the examination of association between the SNPs and the qualitative traits. Moreover, to satisfy the consumer's curiosity the use of DNA- based methods for the identification of meat has been extensively explored (Dalvit *et al*., 2007). During the last decade various types of markers like mitochondrial DNAs (Alves *et al*., 2009), microsatellites (Kim *et al*., 2007) and SNPs (Shim *et al*., 2010) have been tried for their use in traceability systems in livestock. Since 2009, such kind of traceability systems are quiet popular for Korean beef. Recently, the Korean government has formed a committee to consider suitable methods for implementing a genetic identification system of pork to enhance the production and quality of pork from JNP. Our earlier studies on comparative transcriptomic analysis to identify differentially expressed genes in fat, liver and muscle tissue of JNP and Berkshire breeds have highlighted the genes which are expressing differentially between these two breeds. Based on our previous findings, the current study was

planned to identify a genetic marker for the pork production. The current study aimed to identify markers with sufficient genetic information to discriminate between the two main black pig populations of Korea.

Identification and selection of key genes as molecular markers for muscle growth may aid in the planning of swine breeding research trials. The earlier studies reported that muscle regulatory factors (MRFs), control formation of muscles fibers (myogenesis). Myogenic determination factor (*MyoD*), myogenic factor-6 (*Myf6*) and myogenic factor-5 (*Myf5*) *are* the versatile members of MRF and play vital role in differentiation of skeletal muscles (Wyszynska and Kuryl, 2004; Patruno *et al*., 2007). Similarly, the myosin filaments of striated muscles contain a family of enigmatic myosin- binding proteins (MyBP); MyBP-C and MyBP-H (Gruen and Gautel, 1999). Myosin binding protein H (*MyBPH*), mainly expresses in the skeletal muscles during the different stages of the life. It is also reported to be significantly involved in the metabolisms and development of skeletal muscles in *Sus scrofa*. Moreover, *MyBPH* is reported to have significantly higher transcript levels in the *longissimus dorsi* muscle of JNP as compare to Berkshire (Sodhi *et al*., 2014).

The earlier studies which focused on the identification of candidate markers to enhance the growth rate and meat quality, had limited success in JNP. With the aim to identify a genetic marker for the pork production, this study was planned with the sufficient genetic information to discriminate between JNP and Berkshire. This study was carried out to investigate the role of the *MyBPH* in the muscle growth of JNP. To our knowledge, the present study is the first attempt to characterize *MyBPH* gene in JNP. This investigation aids in the generation of large amount of scientific knowledge that helps to improve the understanding of myogenesis and muscle growth in JNP.



## 7.3. Materials and methods

### *7.3.1. Selection of animals and collection of tissue samples*

In the present study, five adult (5 months old) and one day old female animals each from JNP and Berkshire breeds were selected. Animals were kept under the uniform environmental and nutritional conditions. *ad-libitum* supply of the commercial feed (Seoul Feed, Jeju- Si, South Korea) and water was offered to the adult pigs whereas piglets were offered cow's milk. The experiment was approved by the animal ethics committee of the Faculty of Biotechnology, Jeju

National University, Jeju and animals were slaughtered according to the standard protocols of Jeju National University. The *longissimus dorsi* muscle between the 12$^{th}$ and 13$^{th}$ rib space from each animal were collected immediately after the slaughter. Samples were stored immediately in dry ice and later were stored at -80℃ in the lab until used for RNA extraction.

### *7.3.2. RNA and protein extraction from longissimus dorsi muscle tissue*

The fragmented frozen *longissimus dorsi* muscle (120 mg) of both the breeds in five replicates was used for the RNA isolation. TRIzol (Invitrogen, USA) reagent was used for the extraction of RNA from the tissues. Isolated RNA samples were immediately stored at -80℃. From each isolated RNA samples, 25 µg of RNA was treated with the RNase-free DNase set (QIAGEN, Hilden, Germany) to get rid of genomic DNA contamination and it was purified with the RNeasy mini kit according to the user guidelines (QIAGEN, Hilden, Germany). The quality and quantity of RNA was assessed by the Bioanalyzer 2100 with RNA 6000 Nano Labchips and automated capillary gel electrophoresis following user guidelines (Agilent Technologies Ireland, Dublin, Ireland). 28S/18S ratios for the RNA samples ranged from 1.8 to 2.0 and integrity of RNA had ranged from 8.0 to 10.0. Further, the first- strand cDNA was synthesized from 1µg RNA by using SuperScript® III First- Strand kit (Invitrogen, USA) following the user guidelines.



Protein from the homogenized *longissimus dorsi* muscle was extracted by "Radio Immuno Precipitation Assay (RIPA) buffer". The concentration of protein was quantified using "Pierce™ BCA" protein assay kit (Thermo Scientific, USA) in Bio-Rad Micro-plate Reader (Model-680) following the manufacturer's guide lines. All the RNA and protein samples were stored at -80 ℃ until their further use.

### *7.3.3. Cloning of MyBPH*

Primers for the cloning of *MyBPH* coding sequence were designed on the basis of the homologous regions of porcine (GenBank accession No. NC_010451) and Humans (GenBank accession No.NM_004997). Primers were designed using primer 3.0 tool for 5'-RACE (rapid amplification of cDNA ends) and 3'RACE to obtain the C-terminal coding region of the *MyBPH* gene. The polymerase chain reaction (PCR) was performed using cDNA from both the breeds and Prime Taq Premix (2X), (GenetBio, S. Korea) in a total volume of 20 µl mixture. The amplified DNA fragments were subsequently cloned into pUC57 (Figure 7.1).

### *7.3.4. Identification of the gene and retrieval of SNPs from MyBPH of JNP and Berkshire*

Purified PCR products of *MyBPH* from both the breeds were sequenced and compared by "Cosmo genetech", South Korea. The DNA sequence was confirmed by using Basic Local Alignment Search Tool for nucleotide (BLASTn) of National center for biotechnology information (NCBI) to confirm its high similarity with the sequence of porcine *MyBPH*. The amino acid substitutions in *MyBPH* from both the breeds were compared with the MyBPH protein sequence available in NCBI database. The data related to SNPs were obtained from the NCBI dbSNP (http://www.ncbi.nlm.nih.gov/snp/) and SWISSProt databases (http://expasy.org/) for computational analysis.

#### *7.3.4.1. Annotation of variations due to amino acid substitution in MyBPH using SIFT tool*

SIFT (Sorting Intolerant from Tolerant) (http://sift.jcvi.org/www/SIFT_seq_submit2.html) uses sequence homology to predict whether an amino acid substitution affects protein function or not. Output of the SIFT scores was classified as damaging when in the range from 0.00 to 0.05; the potentially damaging between 0.051-0.10 and 1.00 being neutral. The threshold has been fixed at 0.05 in our study. If the score will be less than the cutoff value, then the respective amino acid substitution would be predicted as deleterious.

#### *7.3.4.2. Functional annotation of nsSNPs using protein variation effect analyzer (PROVEAN)*

The protein sequences of MyBPH from JNP and Berkshire associated with the identification of nsSNPs through SIFT were submitted as queries in PROVEAN. The protein variant would be deleterious if the final score is less than the threshold score of -2.5 and the protein variant with score above the threshold was considered as the neutral one.

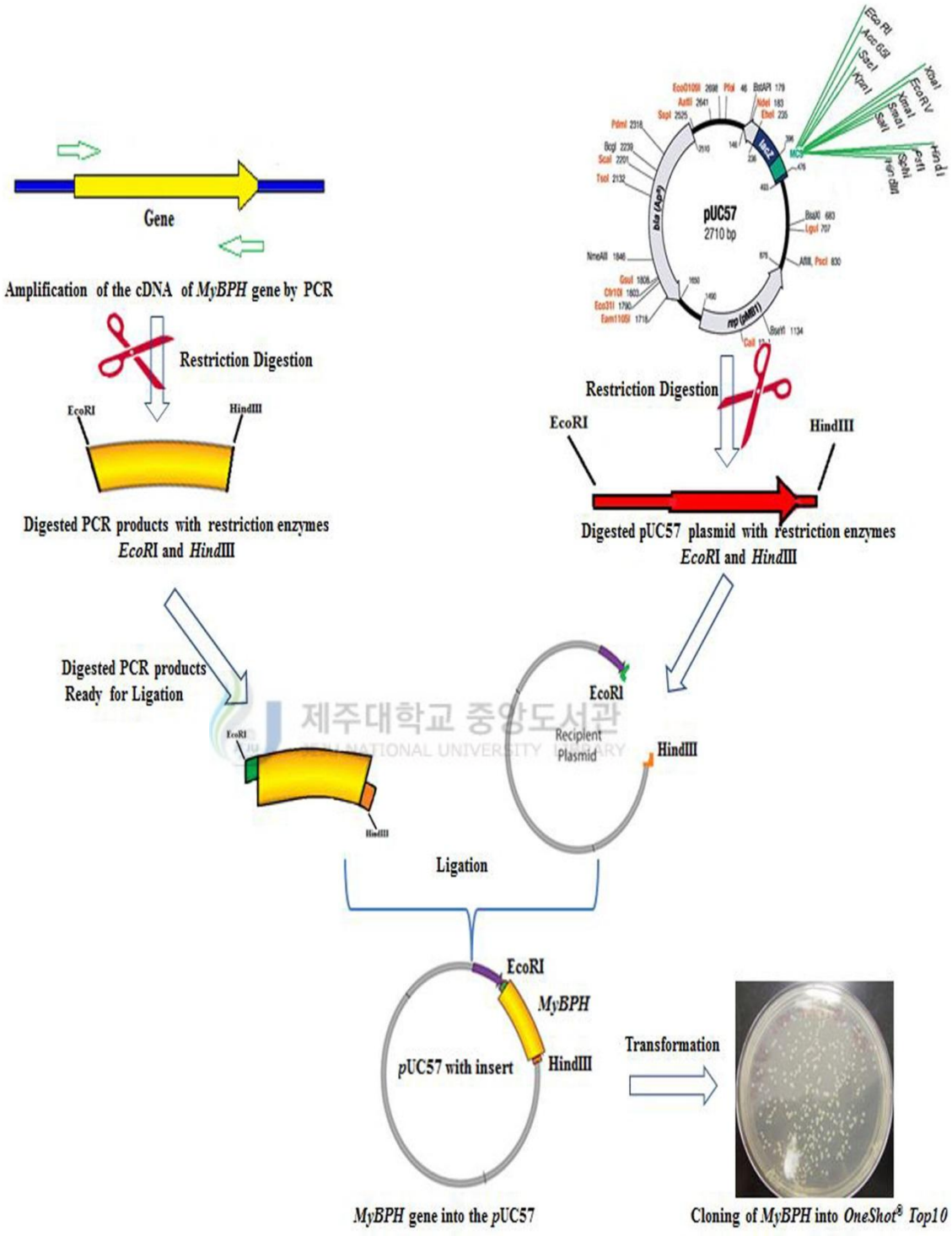


**Figure 7.1** The schematic representation of ligation of target gene in the pUC57 and cloning of the *MyBPH* in the OneShot® Top 10 competent cells.

#### *7.3.4.3. Potential consequences of the nsSNPs on MyBPH using PANTHER*

The likelihood of a particular non-synonymous coding SNP to cause a functional impact on the protein was estimated by PANTHER. The input protein sequence is scored against the Hidden Markov Model (HMMs) in the PANTHER library. The subPSEC (substitution position-specific evolutionary conservation) score estimates the likelihood of a functional effect from a single amino acid substitution. PANTHER subPSEC scores are of continuous range from 0 (neutral) to -10 (most likely to be deleterious). A cutoff of -3 corresponds that there is 50% probability that the amino acid substitution will be deleterious. It means that a given variant with a subPSEC score of -3 will cause 50% deleterious effect ($P_{deleterious}$) on protein function.

#### *7.3.4.4. Identification of the effect of mutation on the protein using Project HOPE(Have yOur Protein Explained)*

The Project HOPE, a mutant analysis server was used to combine the information to give analytical vision regarding the effect of a mutation on MyBPH protein. The protein sequence and the mutated amino-acid were used as the input. 3D analysis of a mutated protein structure of MyBPH was conducted by the project HOPE (http://www.cmbi.ru.nl/hope) to provide an insight on the structural effects of a mutation.



### *7.3.5. Gene ontology (GO) analysis of MyBPH*

GO analysis has become a commonly used approach for the functional study of gene of interest. In the current study, the biological processes associated with the functioning of the *MyBPH* gene in *S. scrofa* have been analyzed with the GO analysis. The shared and specific features associated with any number of genes involved in biological processes have been recorded in the current study. ToppCluster with a cutoff *P*-value of 0.05 was selected and no false discovery correction method was applied.

### *7.3.6. Cloning of MyBPH into expression vector PiggyBac*

For the cloning of *MyBPH*, the plasmid vector PiggyBac was procured (Clontech, USA). For the propagation of plasmid and as a maintenance host, *E. coli* Oneshot® Top10 (Invitrogen,

USA) competent cells were used. HindIII restriction enzyme was used to linearize the vector and later, the respective target genes of JNP (165 bp) and Berkshire (161 bp) were picked by PCR. Sub-cloning was performed with cloning primers (Table 7.1) having restriction enzyme sites for EcoRI and HindIII incorporated in forward and reverse primers respectively. PCR products were separated on 1.2% agarose gel and were extracted by Expin Gel (GeneAll Biotech, South Korea) extraction kit by following the user guidelines.

The amplified PCR expression vector PiggyBac and insert (1µg each), were digested with 20U/µl of restriction enzymes (EcoRI and HindIII) alongwith 10X EZ-One buffer in 20µl reaction volume. Later, the digested samples which were ready for ligation were stored at -20℃. The strategic approach was followed to construct the expression vector of the recombinant *MyBPH* gene of both breeds (Figure 7.2). The respective digested PCR products from both the species were cloned into the PiggyBac (6613bp) with 1:5 ratios of vector and insert for ligation at EcoRI and HindIII sites. The ligation was performed with T4 DNA ligase (400U/µl) kit (Invitrogen, USA). Immediately after the completion of ligation processes, ligation mixtures were transformed in *E. coli* Oneshot® Top10 competent cells with 100µg/ml Ampicillin as selection marker. Later, the positive transformants were selected by PCR screening.

The recombinant plasmid (PiggyBac-MyBPH) was subjected to ligation confirmation by single and double digestion. The end products of recombinant plasmids were subjected 10% Triton X-114 to make it endotoxin free and later were used for transfection in the target C2C12 myoblast cells.



### *7.3.7. Cell culture, transfection and selection of stable cell line*

The C2C12, mouse myoblast cells were maintained in the Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 20% fetal bovine serum (FBS), 100 unit/ml penicillin, 100 µg/ml streptomycin and 100 µg/ml gentamycin. The cells were seeded at the rate of $1x10^5$ cells/$25cm^2$. At 70% confluency, cells were sub-cultured and media was once again changed 24 hrs before transfection. Later, the trypsnized cells were washed twice with phosphate buffer saline (PBS, pH 7) and 2 x $10^6$ cells were transfected following the manufacturer's guidelines with recombinant PiggyBac-MyBPH plasmid (10µg) having gene from both the breeds respectively. The nucleofection was performed by using Amaxa® Cell Line Nucleofector® Kit V through "Program B-032". pmaxGFP® (green fluorescent protein, 2µg) vector of Amaxa Nucleofector-II system was used as a positive control. The transfected cells were subjected to selection with 250µg/ml puromicin from 48 hrs onwards after transfection upto for 3 weeks.

### *7.3.8. Extraction of RNA and protein from the cells*

The RNA from the non-transfected and transfected cells C2C12carrying genes from both the breeds were extracted by easy-Blue$^{TM}$ total RNA extraction kit (Intron Biotechnology, South Korea). The extracted RNA was further processed to synthesize cDNA by using the same protocol as mentioned earlier in the section 7.3.2. of current manuscript. The cells were washed twice with cold 1X PBS before the extraction of proteins. The protein extraction was performed by using 1X phenyl methyl sulfonyl fluoride (PMSF) and Radio Immuno Precipitation Assay (RIPA) buffer (T & I, South Korea) following user guidelines. Cells were pelleted at 1.5 x 1000 x g. The concentrations of proteins were measured in a Bio-Rad Micro-plate Reader (Model-680) by following the manufacturer's guidelines of PierceTM BCA protein assay kit (Thermo Scientific, USA). The protein samples were stored at -80 ℃ until their further use.

### *7.3.9. Relative expression analysis of MyBPH*



#### *7.3.9.1. Expression analysis in the stable cell line by immuno- cytochemistry*

The transfected and non-transfected C2C12 cells at 70% confluency were seeded in 6 well plate for immunostaining. The cells were fixed in 4% paraformaldehyde at 4℃, followed by the treatment with 0.1% Triton X-100 (pH 7) at room temperature. The non- specific binding was blocked by using washing buffer (5% v/v FCS in PBS with 0.05% v/v Triton X-100) for 30 mins. The fixed cells were incubated overnight at 4℃ in dark with primary antibodies: MyBPH (sc-165049), MyoD (sc-304), Pax-7 (sc-7748), MyoG (sc- 576), Myf6 (sc-301), Myf5 (sc- 302) and BMP2 (sc- 9003) specific for the proteins. All the antibodies were procured from Santa Cruz, Heidelberg, Germany. The cells were further incubated with appropriate Alexa Flour secondary antibodies (Invitrogen, Cergy Pontoise, France) successively. The cells were counterstained with 4, 6-diamindino-2-phenylindole (DAPI, @0.3µg/ml) (Sigma, USA) for 5 min for the visualization of nuclei of cells. Subsequent image analysis of the labelled cells was done through Olympus, IX 70 fluorescent microscope.

**Table 7.1** Details of detection and cloning primers used for the confirmation of the gene sequence in the vectors.

| S No. | Primer | Sequence | Tm | GC% |
|---|---|---|---|---|
| 1. | Bk*, Cloning MyBPH-2 (Forward) | ATGAATTCAGTCTTACTGGACCCGGAGTG | 60.14℃ | 60.00% |
| 2. | Bk, Cloning MyBPH -2 (Reverse) | CGGGATCCCTAGTGGGTGATAGCTACGGTGA | 60.90 ℃ | 55.00% |
| 3. | JNP**, Cloning MyBPH -2 (Forward) | ATGAATTCAGGAGTGGCTAGGGAAGCTG | 62.78 ℃ | 63.16% |
| 4. | JNP, Cloning MyBPH -2 (Reverse) | CGGGATCCCTAGTGGGTGATAGCTACGGTGA | 60.90 ℃ | 55.00% |
| 5. | Bk, Detection MyBPH-2 (F) | GTCTTACTGGACCCGGAGTG | 60.14 ℃ | 60.00% |
| 6. | Bk, Detection MyBPH-2(R) | GTGGGTGATAGCTACGGTGA | 60.90 ℃ | 55.00% |
| 7. | JNP, Detection MyBPH-2(F) | GGAGTGGCTAGGGAAGCTG | 62.78 ℃ | 63.16% |
| 8. | JNP, Detection MyBPH-2(R) | GTGGGTGATAGCTACGGTGA | 60.90 ℃ | 55.00% |



Bk* Berkshire

JNP** Jeju Native Pig

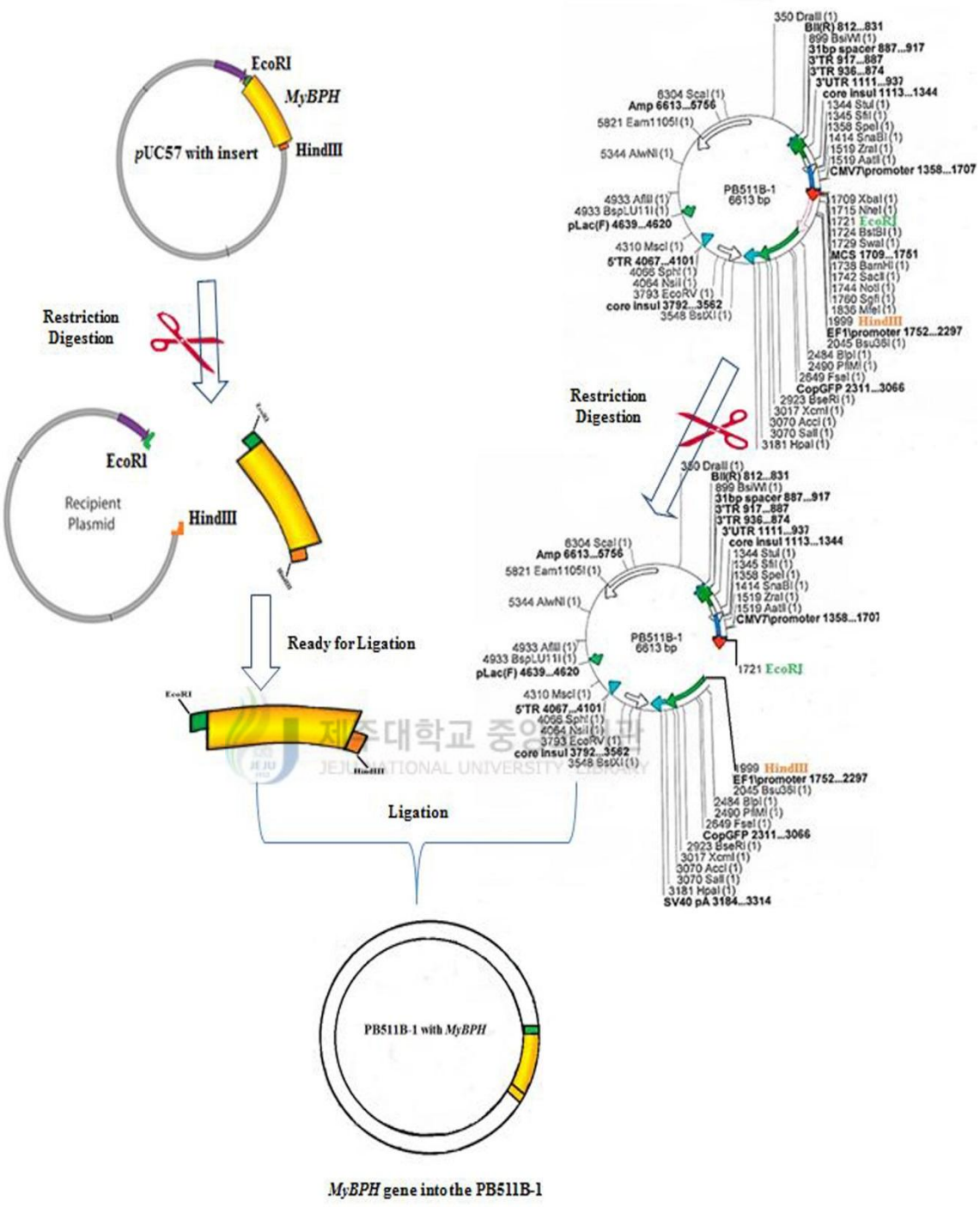


**Figure 7.2** The schematic representation of construction of recombinant PiggyBac-*MyBPH* vector (6613bp) by inserting target gene between the sites of EcoRI and HindIII digestive enzymes.

#### *7.3.9.2. Quantitative Real-Time PCR (qRT- PCR) assay*

Primers used in quantitative real-time PCR were designed using the online Primer3 program (Rozen and Skaletsky, 2000). The primers have been enlisted in table 7.2. These primers were used to evaluate the quantitative transcript levels in *longissimus dorsi* muscle, non-transfected and transfected cells. The dye, EvaGreen (Biotium, USA) was used to determine the quantity of transcript of target genes and each individual sample was quantified in triplicate. Standard curve method was used to define the efficiency of real-time PCR. The efficiency of amplification of the target gene was compared with that of the endogenous *β-actin* control transcript (Wang *et al*., 2003). Quantification of mRNA transcripts was performed using the comparative $C_T$ method. The results are reported as the relative expression normalised using the level of the transcript of the endogenous reference (Van Poucke *et al*., 2001; Erkens *et al*., 2006).

#### *7.3.9.3. Western Blotting*



Sixty micrograms of proteins extracted from the *longissimus dorsi* muscle and transfected cells were diluted as 1:1 with 2X SDS loading buffer. Protein extracts were denatured by boiling for 10 min before loading onto a 12% SDS-PAGE gel. Later, proteins were transferred to 0.45µm nitrocellulose membrane (Invitrogen, USA) after electrophoresis. The membranes were blocked with 5% milk in 1X PBS containing 0.05% Tween-20 (TBS-T), at room temperature and incubated with primary antibodies: MyBPH (sc-165049), MyoD (sc-304), Pax-7/3 (sc-7748), MyoG (sc- 576), Myf6 (sc-301), Myf5 (sc- 302), BMP2/4 (sc- 9003) and β-actin (sc- 47778) specific for the proteins overnight at 4℃. Next day, membranes were washed with 1X PBS and incubated with appropriate secondary antibodies at room temperature for 1 hour followed by washing with 1X PBS. Immunoreactive bands were detected using EZ-Western reagent (DoGen, Life Science, S. Korea). The immunoreactive bands with chemiluminescent signals were detected using a Luminescent Image Analyzer (LAS-4000 mini) instrument. The results are depicted as the relative band intensities normalised relative to the band intensities of *β-actin* bands for each gene using Image J software (National Institute of Health, Bethesda, Maryland, USA). The means were compared between JNPs and Berkshire.

### *7.3.10. Statistical analysis to access the expression levels of genes*

Analysis was designed to understand the statistical value of the expression levels among the genes between the Berkshire and JNP. The data was analyzed by the analysis of variance (ANOVA) for the relative quantitative expressions of the genes by real-time qRT-PCR and western blotting. The significant differences between the mean expressions of different genes at $P<0.05$ were analyzed by Tukey's b- test. The values have been expressed as mean±SEM.

## 7.4. Results

### *7.4.1. Differential expression of MyBPH and MRFs in one day old and adult animals of JNP and Berkshire*

Relative quantitative mRNA transcript levels of *MyBPH*, *MyoD*, *Pax7*, *Myf6* and *Myf5* normalized with the transcript levels of the endogenous reference, β-actin were investigated through Real Time qRT- PCR. In our study, the relative quantitative transcript level of *MyoD* was significantly ($P<0.05$) higher in Berkshire during both the stages than in JNP (Figure 7.3A). Significantly ($P<0.05$) higher transcript levels of *Pax7* also has been recorded in the adult Berkshire breed as compare to the JNP whereas no significant difference was observed in the day old piglets of both breeds (Figure 7.3B). The blot expressions of proteins also complimented the transcript levels of the genes under study. Blot expressions of MyoD and Pax7 proteins are significantly ($P<0.05$) higher in Berkshire as compared to JNP during the post natal development (Figure 7.4(a) and Figure 7.4(b)). Relative transcript levels and blot expressions of *MyBPH* have shown contrary results to relative expression analysis of *MyoD* and *Pax7*. Within the breed analysis has shown that *MyoD* and *Pax7* have significantly higher transcript levels in adult animals than the one day old piglets of Berkshire whereas no significant difference was observed in JNP. *MyBPH* has given significantly high transcript levels and blot expressions during both age groups in JNP than Berkshire (Figure 7.3C, Figure 7.4(a) and Figure 7.4(b)). Moreover, within the breed for *MyBPH*, no significant differences have been observed in the day-old piglets and adult animals. Both *Myf6* and *Myf5* did not indicate any statistically significant difference between the breeds during postnatal development (Figure 7.3C, Figure 7.4(a) and Figure 7.4(b)).

**Table 7.2** List of primers used for quantitative real time PCR to analyze the relative quantity of transcripts in the *longissimus dorsi* muscle and stable transfected cell line carrying gene sequence of Berkshire and JNP.

| Gene name | Primer Sequences | $T_m$ | Product size | Gene bank ID |
|---|---|---|---|---|
| *MyBPH* | F: AGTGCAGAAGGCAGACAAA<br>R :AGACCCGGAAGGAGTAAGA | 62℃ | 117 | XM_001033014 |
| *MyOG* | F: GGTGAGGGAGTGCAGATTGT<br>R: CAGTGAATGCAGTTCCCACA | 56℃ | 130 | NM_001012406.1 |
| *MyOD* | F: TGCAAACGCAAGACCACTAA<br>R: GCTGATTCGGGTTGCTAGAC | 55℃ | 127 | NM_001002824.1 |
| *MYF6* | F: ATCTTGAGGGTGCGGATTTC<br>R: CAATGTTTGTCCCTCCTTCCT | 62℃ | 108 | XM_003481764 |
| *MYF5* | F: CCGACACAGCTTGTGGAATA<br>R: GCCAATCAACTGATGGCTTT | 55℃ | 128 | XM_001924362.2 |
| *PAX3* | F: ATCGGCTAATCCTGACATGC<br>R: ACGGTGGGAAACTTTTGATG | 54℃ | 130 | AY579430.1 |
| *PAX7* | F: GGCAGAGGATCTTGGAGACA<br>R: TGGGTGGGGTTTTCATCAAT | 55℃ | 144 | AY653213.1 |
| *BMP2* | F: CTTGACGCTTTTCCCTTTTG<br>R: AAGAGGCATGTGCGGATTAG | 55℃ | 130 | NM_001195399.1 |
| *BMP4* | F: ACGGTGGGAAACTTTTGATG<br>R: ATCGGCTAATCCTGACATGC | 55℃ | 140 | NM_001101031.1 |
| *GAPDH* | F: ACCCAGAAGACTGTGGATGG<br>R: ACGCCTGCTTCACCACCTTC | 60℃ | 247 | AF017079.1 |

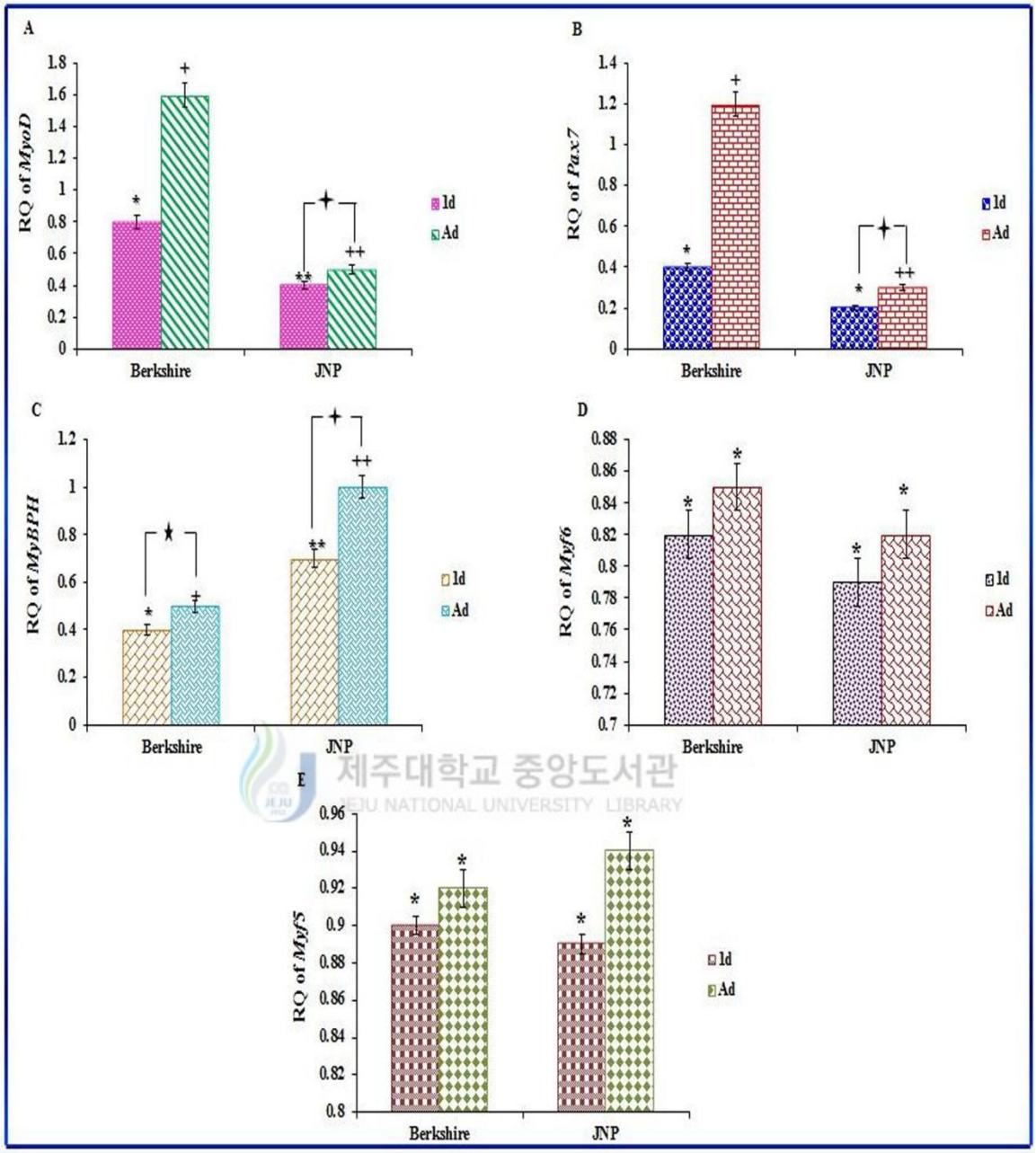




**Figure 7.3** Relative Quantity (RQ) of transcripts in Berkshire and JNP. RQ of (A) *MyoD*, (B) *Pax7*, (C) *MyBPH*, (D) *Myf6* and (E) *Myf5* has been presented in the one day old piglets and adult animals from both the breeds. Bars with different superscripts are significantly different ($P<0.05$).

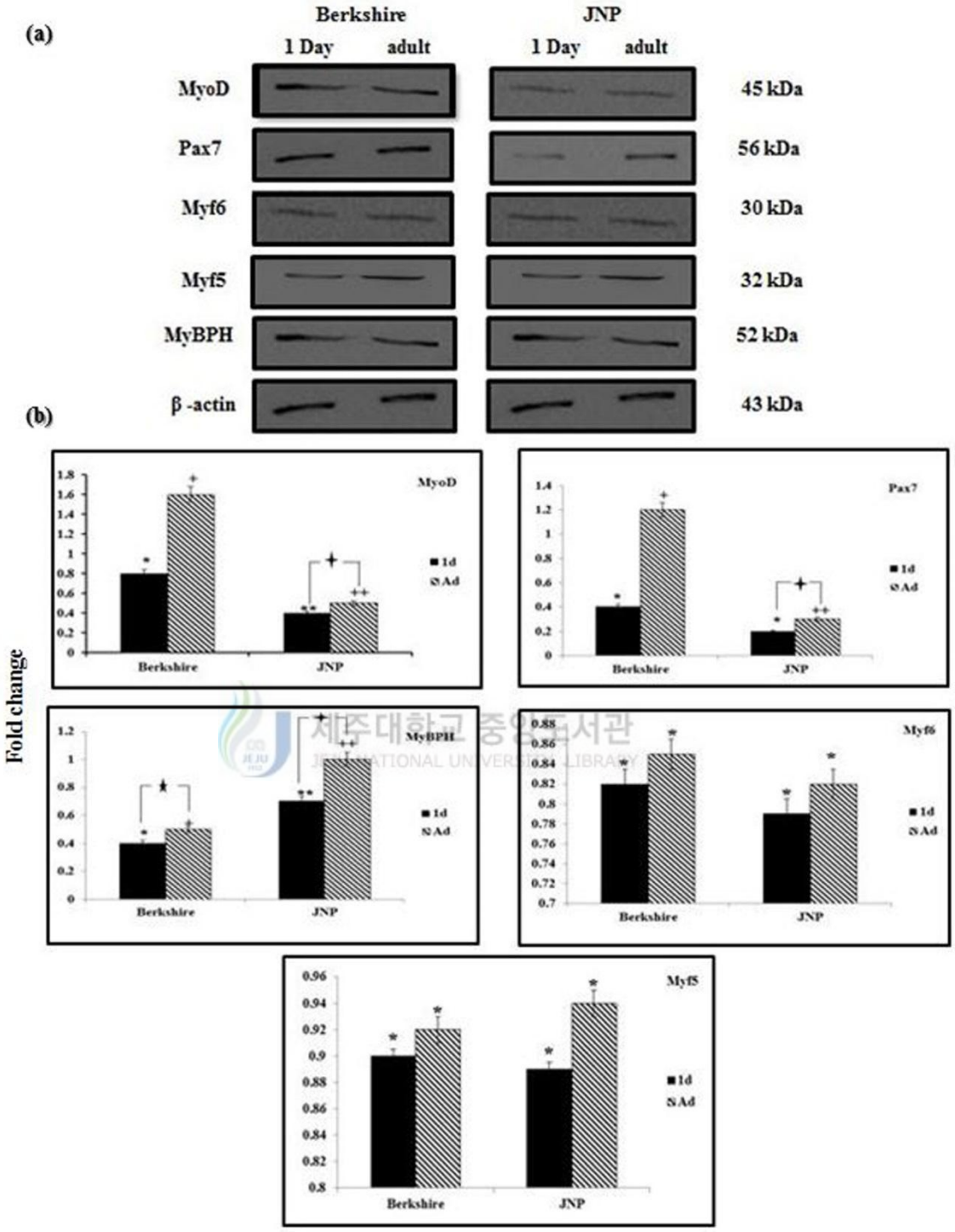


**Figure 7.4** Expression levels of MyoD, Pax7, Myf6, Myf5, MyBPH and β- actin proteins isolated from the *longissimus dorsi* muscles of one day old piglets and adult animals of Berkshire and JNP. Protein levels were normalized for the levels of β- actin.

### *7.4.2. Retrieval of SNPs and analysis of functional consequences of nsSNPs*

The maximum numbers of variations were observed from 629 to 973 bps located at the second coding region of *MyBPH* gene in both JNP and Berkshire breeds. Our preliminary SNPs analysis showed highly deleterious effects at second coding region in comparison with other coding regions of *MyBPH* of *Sus scrofa*. Polymorphism data retrieved from the *MyBPH* gene in *longissimus dorsi* of JNP and Berkshire revealed that total number of 61 and 11 amino acids variants respectively from JNP and Berkshire were found in the second coding region (Table 7.3 and 7.4). To predict the deleterious effects of mutations, five *in silico* SNP prediction algorithms: SIFT, PROVEAN, I-MutantDDG, PANTHER and Project HOPE were used in the current study. The functional changes due to mutations of amino acids were categorized as deleterious in both the breeds. The four nsSNP each from JNP and Berkshire respectively at the constitutive amino acids positions of W93T, D218R, D230W, W315V (Table 7.3) and G100L, P158K, L248R, T312W were selected on the basis of the highest deleterious score (Table 7.4).

Further 3D structure of the protein with mutated amino acids (Figure 7.5 and 7.6) provide an insight on the effect of mutation and the better explanation for the substantial changes in the mechanism of the protein. The wild type and mutant amino acids differ in their charges and sizes. The charge of the respective buried wild type residue is lost during mutation. It has been observed that the wild type residue is bigger than the mutant residues for W93T and W315V whereas smaller than the mutant residues for D218R and D230W in JNP. Further the nsSNPs analysis in Berkshire confirmed the bigger mutated residues than the wild type residues. The wild type residues are varyingly conserved whereas the mutant residues were not like the other residue types which were observed at the respective positions in other homologous proteins. The residues are located on the surface of the proteins and mutations of those residues can disturb interaction between the molecules or with other parts of the protein. The hydrophobicity of the wild type and mutant residues also differs. Therefore mutation might cause loss of the hydrophobic interactions with the other molecules on the surface of the protein.

**Table 7.3** The effect of single amino acid variation in *MYBPH* for JNP.

| Position and Substitution | SIFT | PROVEAN (cutoff= -2.5) | I-Mutant3 (DDG Kcal/mole) | PANTHER | |
|---|---|---|---|---|---|
| | | | | subPSEC | $P_{deleterious}$ |
| S87C | 0.05 | -2.789 | -0.61 | -3.24623 | 0.56125 |
| S88C | 0.00 | -3.516 | -0.52 | -4.12907 | 0.75567 |
| V89W | 0.00 | -5.014 | -0.69 | -4.64081 | 0.83764 |
| T90P | 0.00 | -4.238 | -0.30 | -3.256 | 0.56365 |
| V91W | 0.00 | -2.481 | -0.86 | -4.27764 | 0.78205 |
| S92R | 0.01 | 0.267 | 0.00 | -1.52967 | 0.18689 |
| **W93T** | **0.00** | **-11.374** | **-1.31** | **-8.3632** | **0.99534** |
| P121I | 0.06 | -1.968 | -0.60 | -1.70706 | 0.21536 |
| V122A | 0.29 | 1.818 | -1.37 | -1.24993 | 0.14804 |
| N123R | 0.00 | -4.091 | 0.42 | -2.84648 | 0.4617 |
| T124L | 0.12 | -2.195 | -0.13 | -1.51733 | 0.18502 |
| R125A | 0.05 | -0.676 | -0.50 | -1.87316 | 0.24474 |
| P126G | 0.01 | -4.209 | -1.64 | -2.88962 | 0.47243 |
| M127S | 0.12 | -0.759 | -1.06 | -0.58387 | 0.08195 |
| M128P | 0.05 | -0.631 | -0.59 | -2.77808 | 0.44475 |
| V129R | 0.09 | 2.418 | -1.06 | -2.61999 | 0.40612 |
| V145S | 0.00 | -3.245 | -2.20 | -3.0683 | 0.51707 |
| R146G | 0.01 | -6.064 | -1.32 | -7.15513 | 0.98456 |
| V147G | 0.00 | -6.056 | -2.69 | -4.76604 | 0.85396 |
| T148S | 0.55 | -0.438 | -2.69 | -0.95549 | 0.11461 |
| A149W | 0.00 | -5.804 | -0.70 | -5.50923 | 0.92479 |
| S152R | 0.02 | -1.205 | -0.29 | -1.5572 | 0.19111 |
| H212C | 0.00 | -3.469 | -0.23 | -4.2253 | 0.77299 |
| N213W | 0.00 | -5.783 | 0.41 | -5.61575 | 0.93187 |
| G214N | 0.04 | -4.831 | -1.16 | -2.20784 | 0.3117 |
| H215W | 0.00 | -2.308 | 0.21 | -3.58245 | 0.64163 |
| L217E | 0.00 | -5.228 | -1.83 | -4.76256 | 0.85353 |
| **D218R** | **0.00** | **-7.358** | **-0.28** | **-4.42503** | **0.80612** |
| S219E | 0.03 | -1.278 | -0.17 | -1.65736 | 0.20708 |
| Q220P | 0.02 | -1.679 | -0.63 | -2.63667 | 0.41015 |
| R221W | 0.00 | -5.035 | -0.63 | -4.82149 | 0.86074 |
| V222S | 0.00 | -4.852 | -0.35 | -3.06015 | 0.51503 |

*Continued……*

| | | | | | |
|---|---|---|---|---|---|
| N223G | 0.07 | -2.270 | -1.80 | -2.57022 | 0.39418 |
| V224C | 0.00 | -4.054 | -1.21 | -4.07902 | 0.74631 |
| R225L | 0.00 | -6.560 | -1.68 | -5.04245 | 0.88518 |
| G227P | 0.02 | -2.000 | -0.72 | -3.14942 | 0.53729 |
| D228G | 0.01 | -4.451 | -1.03 | -2.99477 | 0.49869 |
| Q229P | 0.02 | -1.836 | -1.70 | -2.64696 | 0.41265 |
| **D230W** | **0.00** | **-9.268** | **-0.83** | **-6.91751** | **0.9805** |
| S231P | 0.00 | -3.646 | -0.32 | -3.50858 | 0.62447 |
| R240S | 0.00 | -5.442 | -1.30 | -3.96018 | 0.72316 |
| S241R | 0.04 | -2.125 | -0.49 | -1.69087 | 0.21263 |
| D242P | 0.00 | -4.381 | -1.30 | -4.4409 | 0.80859 |
| S243C | 0.00 | -4.848 | -0.54 | -5.19129 | 0.89946 |
| V305W | 0.00 | -6.404 | -0.70 | -5.60913 | 0.93145 |
| Q306L | 0.00 | -6.764 | 0.23 | -4.85907 | 0.86519 |
| K307W | 0.00 | -7.754 | 0.55 | -6.39295 | 0.96748 |
| A308V | 0.00 | -3.757 | 0.11 | -3.31362 | 0.57777 |
| D309T | 0.00 | -6.530 | 0.03 | -5.01178 | 0.88203 |
| K310S | 0.00 | -4.068 | -0.29 | -2.92575 | 0.48144 |
| K311S | 0.00 | -4.717 | -0.29 | -3.22473 | 0.55595 |
| G313S | 0.00 | -1.154 | -1.36 | -3.9381 | 0.71872 |
| Q314C | 0.00 | -5.270 | -0.40 | -3.61771 | 0.6497 |
| **W315V** | **0.00** | **-13.510** | **-0.97** | **-5.65694** | **0.93444** |
| F316L | 0.00 | -5.307 | -0.99 | -3.00398 | 0.50099 |
| A317R | 0.00 | -0.533 | -0.11 | - | - |
| V318A | 0.06 | -3.642 | -1.54 | - | - |
| E320Q | 0.00 | -2.733 | -0.37 | -3.74714 | 0.67856 |
| R321G | 0.01 | -3.077 | -1.21 | -3.21139 | 0.55265 |
| Y322L | 0.22 | -5.287 | -0.32 | -2.4465 | 0.36505 |
| H323A | 0.05 | -5.087 | -0.47 | -2.19454 | 0.30886 |

**Table 7.4** The effect of single amino acid variation in *MYBPH* for Berkshire.

| Position and Substitution | SIFT | PROVEAN (cutoff= -2.5) | I-Mutant3 (DDG Kcal/mole) | PANTHER | |
|---|---|---|---|---|---|
| | | | | subPSEC | $P_{deleterious}$ |
| W98P | 0.06 | 1.014 | -1.05 | -2.64682 | 0.40612 |
| L99V | 0.42 | 0.472 | -1.23 | -2.81408 | 0.32475 |
| **G100L** | **0.00** | **-8.372** | **-0.28** | **-3.92575** | **0.36505** |
| K101L | 0.09 | -1.356 | 0.28 | -1.99324 | 0.80144 |
| **P158K** | **0.00** | **-6.104** | **-1.28** | **-3.42149** | **0.37529** |
| E247S | 0.00 | -3.292 | -1.4 2 | -1.87246 | 0.47243 |
| **L248R** | **0.00** | **-4.812** | **-1.55** | **-3. 6308** | **0.17507** |
| V250C | 0.00 | -4.388 | -1.01 | -2.6472 | 0.40126 |
| **T312W** | **0.00** | **-6.054** | **-0.10** | **-3. 33162** | **0.55277** |
| G313L | 0.00 | -1.908 | -0.23 | -3.1427 | 0.59348 |
| Q314W | 0.00 | -4.681 | 0.24 | 1.24318 | 0.17412 |

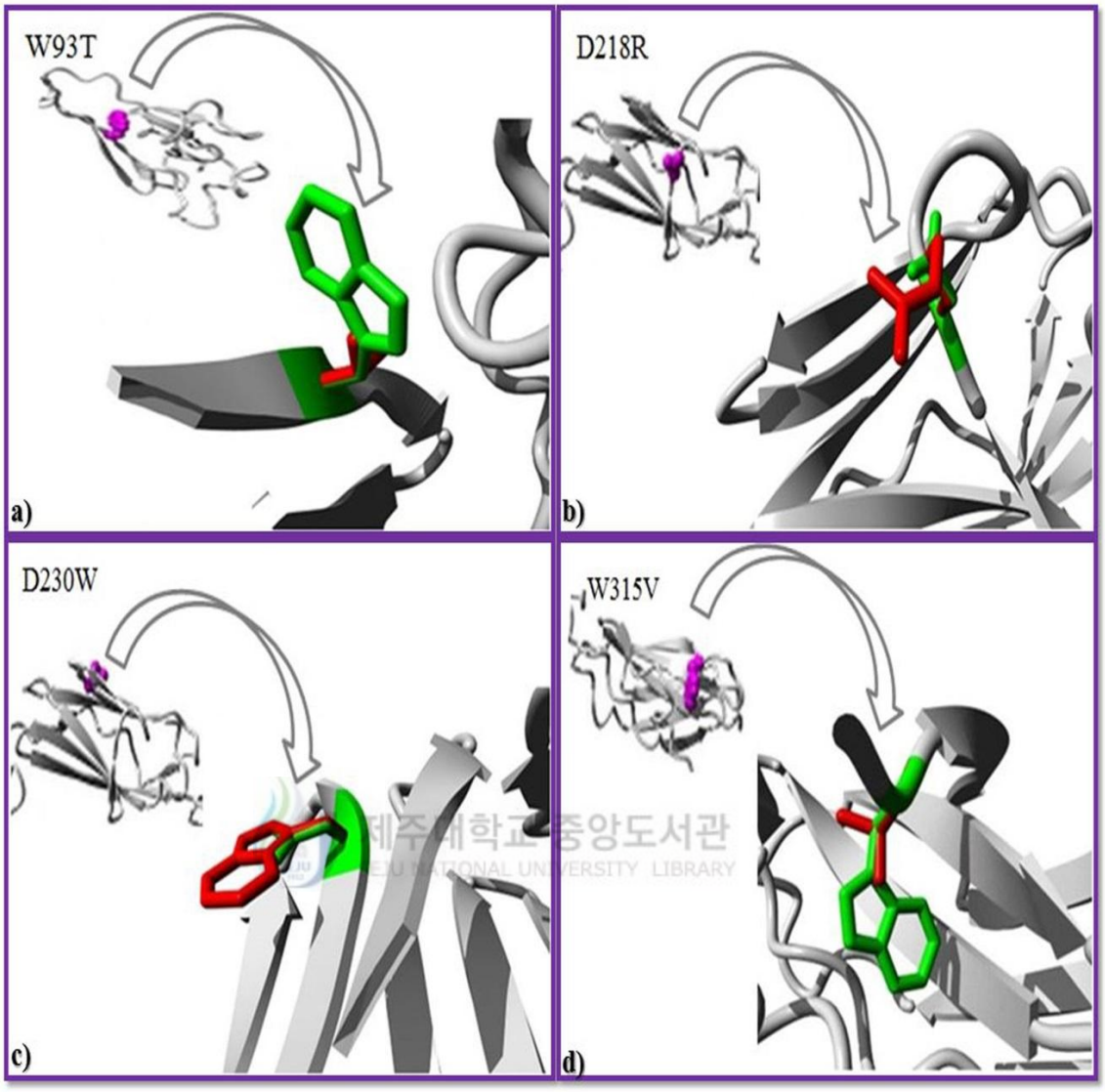


**Figure 7.5** Overview of the MyBPH protein of JNP showing a mutation of amino acids. A 3D ribbon pattern was chosen to explain the mutation. The protein is coloured grey and the side chain of the residue is shown as small magenta balls. In the close-up views the mutations have been shown; a) Tryptophan to Threonine (W>T) at position 93 b) Aspartic acid to Argenine (D>R) at position 218 c) Aspartic acid to Tryptophan (D>W) at position 230 d) Tryptophan to Valine (W>V) at position 315 in MyBPH of JNP. The side chains of both wild-type and mutant residues are shown in green and red, respectively.

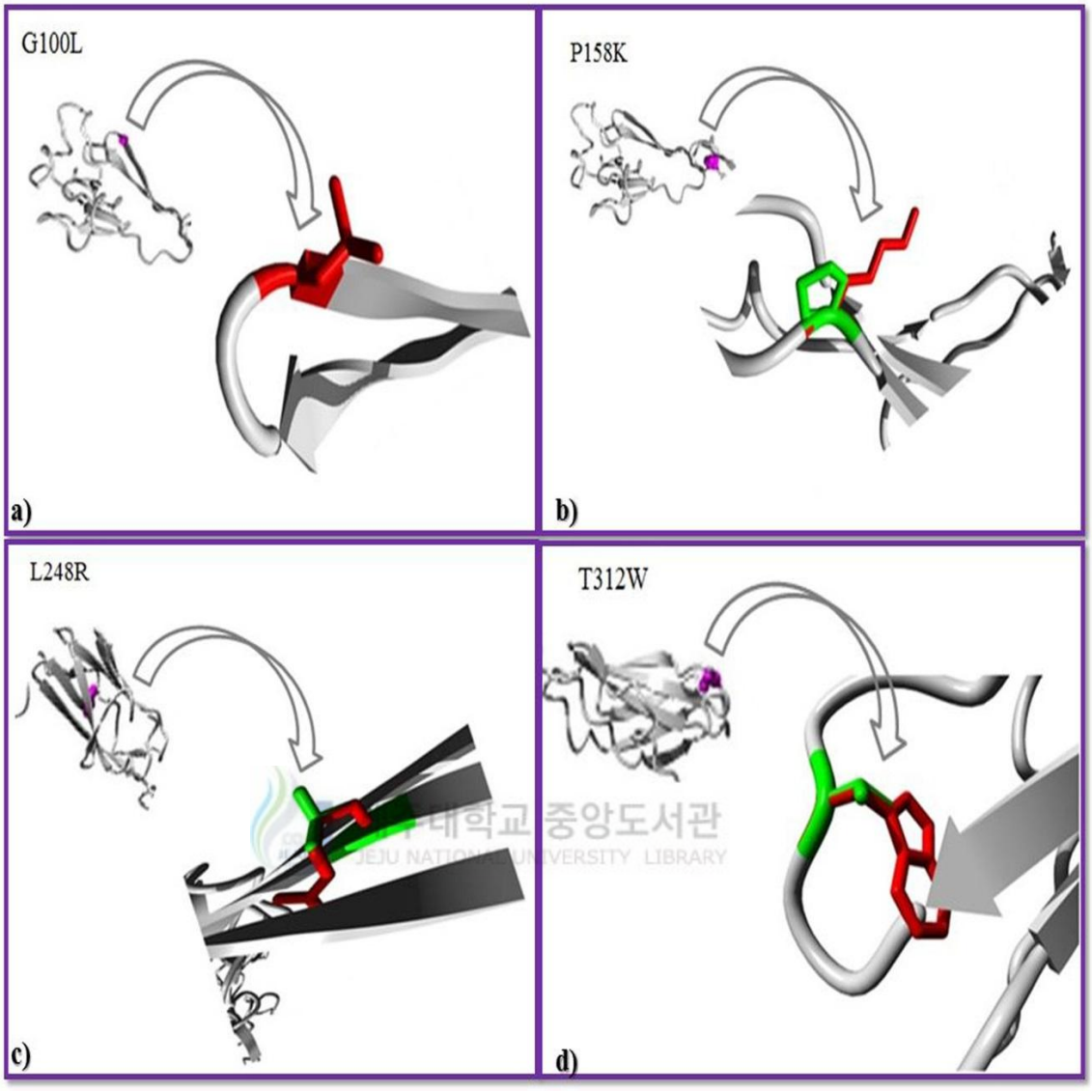


**Figure 7.6** Overview of the MyBPH protein of Berkshire showing a mutation of amino acids. A 3D ribbon pattern was chosen to explain the mutation. The protein is coloured grey and the side chain of the residue is shown as small magenta balls. In the close-up views the mutations have been shown; a) Glycine to Leucine (G>l) at position 100 b) Proline to Lysine (P>K) at position 158 c) Leucine to Arginine (L>R) at position 248 d) Threonine to Tryptophan (T>W) at position 312 in MyBPH of Berkshire. The side chains of both wild-type and mutant residues are shown in green and red, respectively.

### *7.4.3. GO analysis of MyBPH*

The details of the biological ontology of *MyBPH* and the pathways in term of the potentially associated genes were studied. The biological process of *MyBPH* significantly associated with skeletal muscle organ development, skeletal muscle tissue development and striated muscle cell differentiation. Furthermore, ToppCluster was used to visualize the shared and specific features associated with *MyBPH* and other MRFs involved in biological processes. To provide a dissected gene-level view, we chose the gene ontology of the biological process which is shared by *MyBPH* and MRFs expressing in the muscle tissue (Figure 7.7). Although, we were able to observe the clear functional separation through the abstracted network analysis, but a dense network depicted the close relationship among the *MyBPH* and the MRFs. The abstracted network analysis will be very useful, especially for the researchers working on further exploration of specific enriched terms in the output.

### *7.4.4. Construction of recombinant cell expression vectors containing MyBPH gene from JNP and Berkshire:*



The pUC57 plasmid containing *MyBPH* from both the breeds were identified using *MyBPH* detection primers for JNP and Berkshire, respectively (Table 7.1). The synthetic fragments in both breeds were observed on 1.5% agarose gel electrophoresis (Figure 7.8). The respective gene products popped out by the restriction digestion confirmed the ligation and presence of gene in the respective vectors (Figure 7.9). Further, the ligation of both the genes within the vector was confirmed by PCR amplification of the gene fragments from the recombinant plasmids (Figure 7.10) and confirmed the successful insertion of the genes from both the species in the PiggyBac vector. Sequencing chromatogram later confirmed the quality of the recombinant product. The results have been shown by Finch TV software and the sequences were evaluated again by BLAST. The results showed that the sequence data has 100% similarity with original gene sequence and confirmed the good quality score of the cloned product.

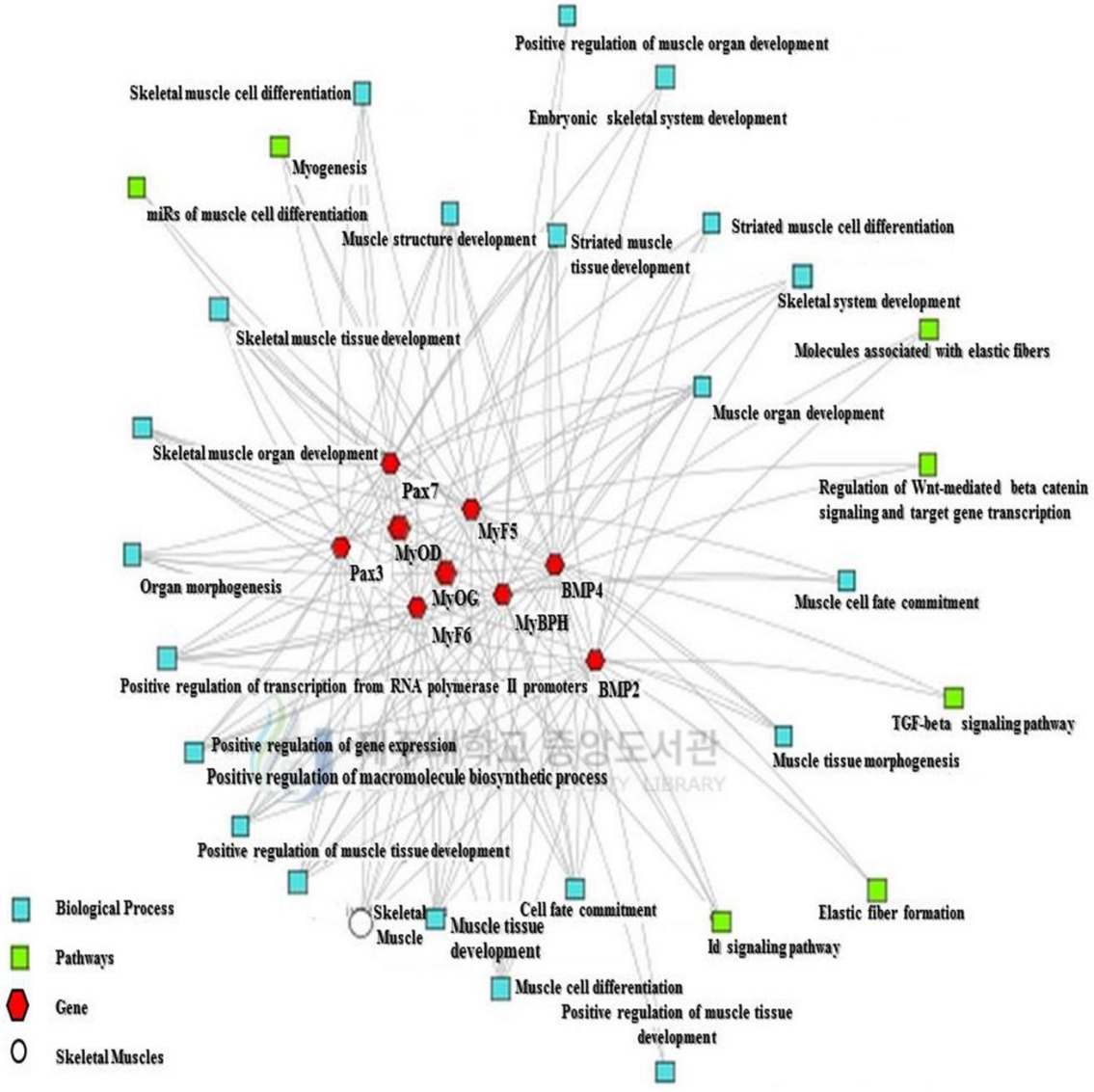


**Figure 7.7** Dissected gene level view sharing GO of biological process shared by *MyBPH* and MRFs in *longissimus dorsi* muscle. An abstracted network is showing biological pathways enriched with the cluster of genes specific for skeletal muscle.

### *7.4.5. Establishment of distinct transfected stable cell lines of MyBPH- JNP and Berkshire respectively*

The transfected cells showed the expression of green fluorescent protein (GFP) of recombinant *MyBPH* from JNP and Berkshire respectively. After 24 hrs of transfection, more than 50% transfection efficiency was observed in the cells with recombinant genes with respect to 70% efficiency in the positive control Pmax GFP (Figure 7.11). After the selection with puromycin, within 15 days purified colonies of the transfected cells were obtained (Figure 7.11).

### *7.4.6. Relative expression of MyBPH in the stable transfected cell lines by immunocytochemistry*

Non-transfected C2C12 cells and transfected cells, carrying *MyBPH* of JNP and Berkshire respectively were subjected to immunochemistry analysis to see the relative expression of genes in both the breeds. The signals after staining with respective primary antibodies showed relative expression of the genes in the two breeds (Figure 7.12). Non transfected cells did not show any significant staining with primary antibodies. A significantly ($P$<0.05) higher expression of *MyBPH* has been observed in the JNP transfected cells. Apart from *MyBPH*, other MRFs i.e. *MyOG*, *MyoD* and *Myf6* have shown significantly ($P$<0.05) relative higher expression in transfected cells carrying the gene sequence from Berkshire as compare to JNP. No significant difference in the expression of *Myf5* and *BMP2* has been observed in both the breeds. A significantly higher expression of *Pax7* has been observed in case of Berkshire.



### *7.4.7. Relative quantitative expression of MyBPH in the C2C12 transfected cells*

The relative quantitative mRNA transcript levels of *MyBPH* with respect to MRFs under study, normalized with the transcript levels of the endogenous reference (*β-actin*) and keeping the non-transfected cells as reference sample were investigated through Real Time qRT- PCR. *MyBPH* has shown significantly ($P$<0.05) almost 2.5 times more transcript levels in JNP as compare to Berkshire transfected mouse myoblasts (Figure 7.13). On the contrary, other MRFs under study like *MyOG*, *MyoD*, and *Myf6* have shown significantly lesser relative quantity in JNP transfected cells. Pax3 and Pax7 are known to induce self-renewal of satellite cells. A significantly higher transcript level of *Pax7* in Berkshire transfected cells has been recorded whereas no significant difference in the transcript levels of *Pax3* has been recorded. Similarly, no

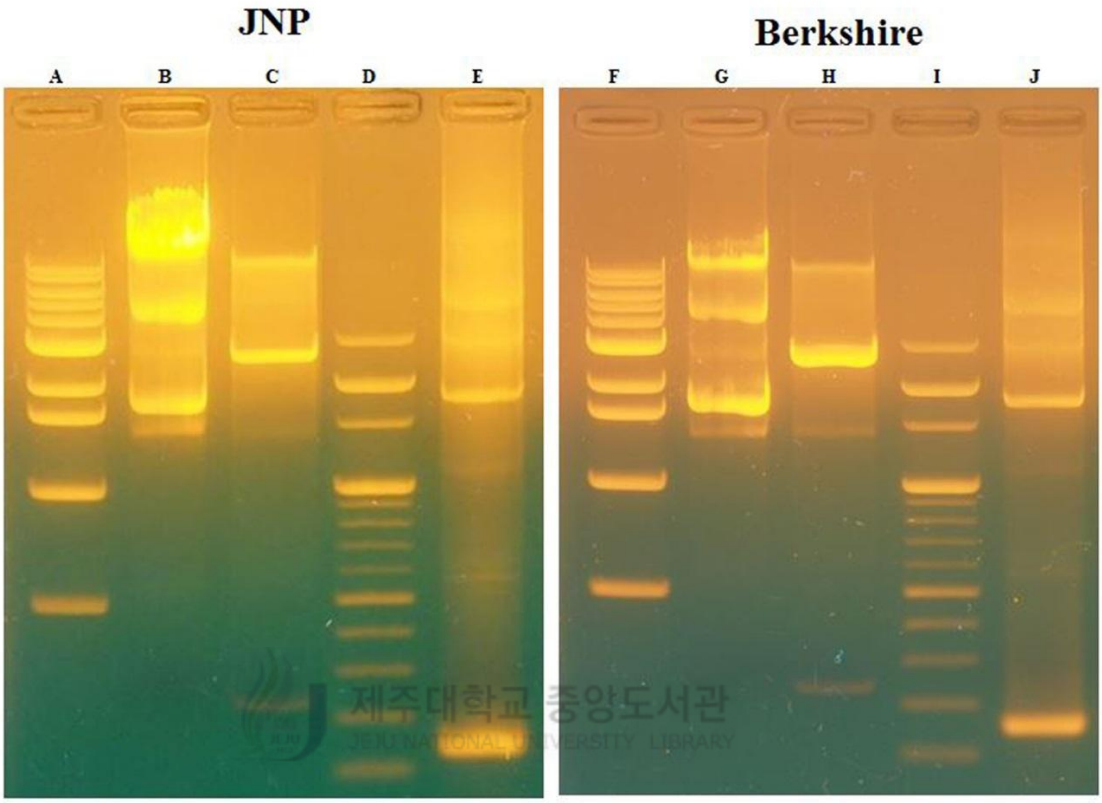



**Figure 7.8** PCR detection and amplification of genes from JNP and Berkshire respectively in pUC57 plasmid. Column "A" and "F"- 1Kb DNA ladder; Column "B" and "G"- pUC57 vector with *MyBPH* gene from JNP and Berkshire respectively; Column "C" and "H"- PCR detected *MyBPH* genes from JNP and Berkshire respectively; Column "D" and "I"- 100 bp DNA ladder. Column "E" and "J"- PCR amplification of *MyBPH* gene of JNP and Berkshire respectively from pUC57 vector using cloning primers.

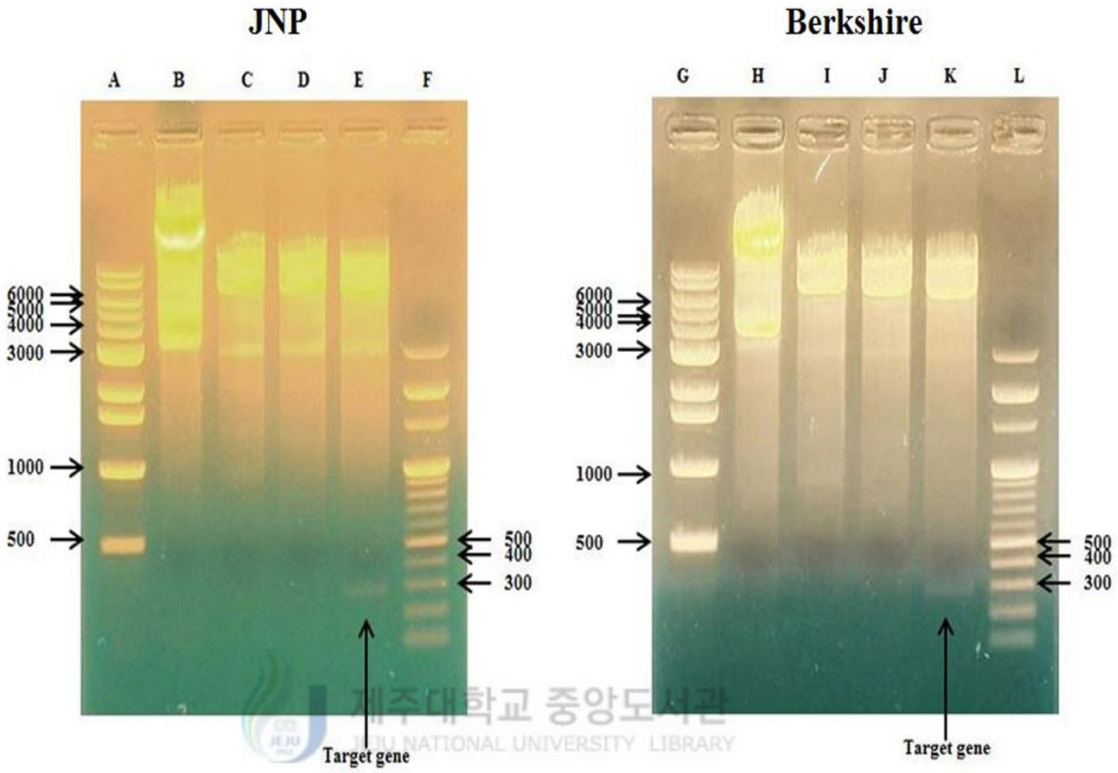




**Figure 7.9** Confirmation of ligation of *MyBPH* genes from JNP and Berkshire respectively in the recombinant PiggyBac-*MyBPH* plasmid by restriction enzymes. Column "A" and "G"- 1Kb DNA ladder; Column "B" and "H"- Undigested recombinant plasmids (PiggyBac-*MyBPH*); Column "C" and "I"- Single digestion of recombinant plasmids with EcoRI; Column "D" and "J"- Single digestion of recombinant plasmids with HindIII; Column "E" and "K"- Double digestion of recombinant plasmids with EcoRI and HindIII digestive enzymes; Column "F" and "L"- 100bp ladder, respectively for JNP and Berkshire.

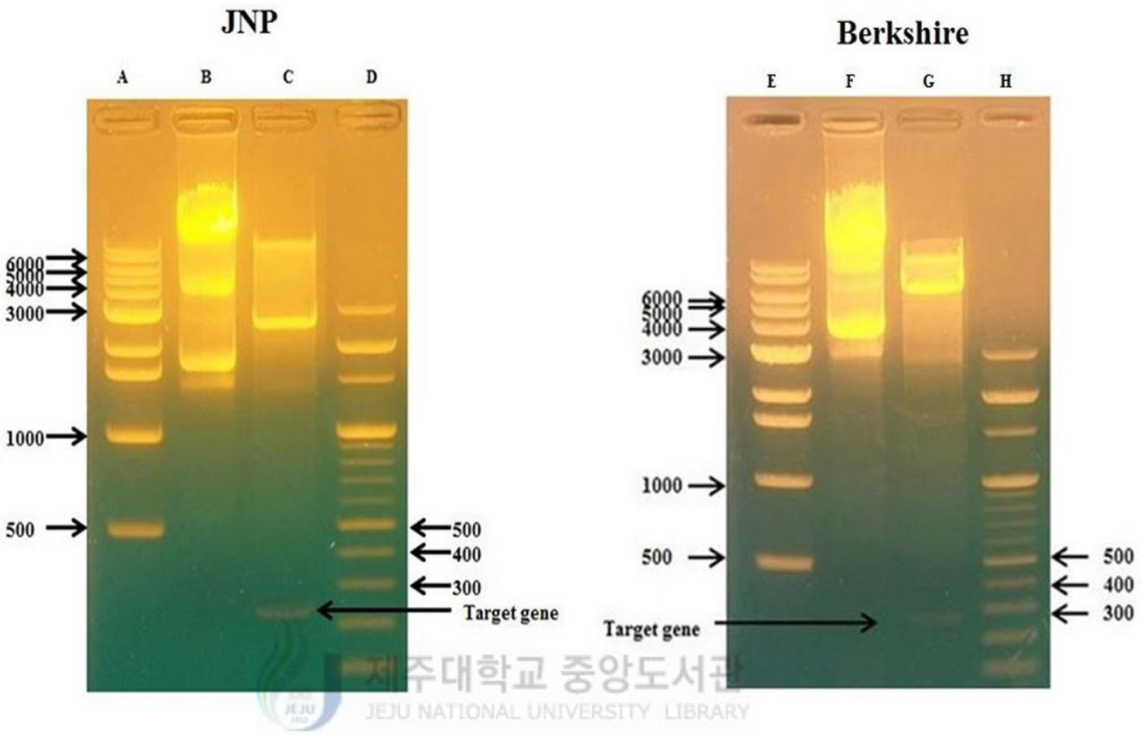



**Figure 7.10** Confirmation of ligation by PCR with detection primers. Column "A" and "E"- 1Kb DNA ladder; Column "B" and "F"- Undigested recombinant plasmid (PiggyBac-*MyBPH*); Column "C" and "G"- PCR detected *MyBPH* genes; Column "D" and "H"- 100bp DNA ladder respectively in the samples from JNP and Berkshire.

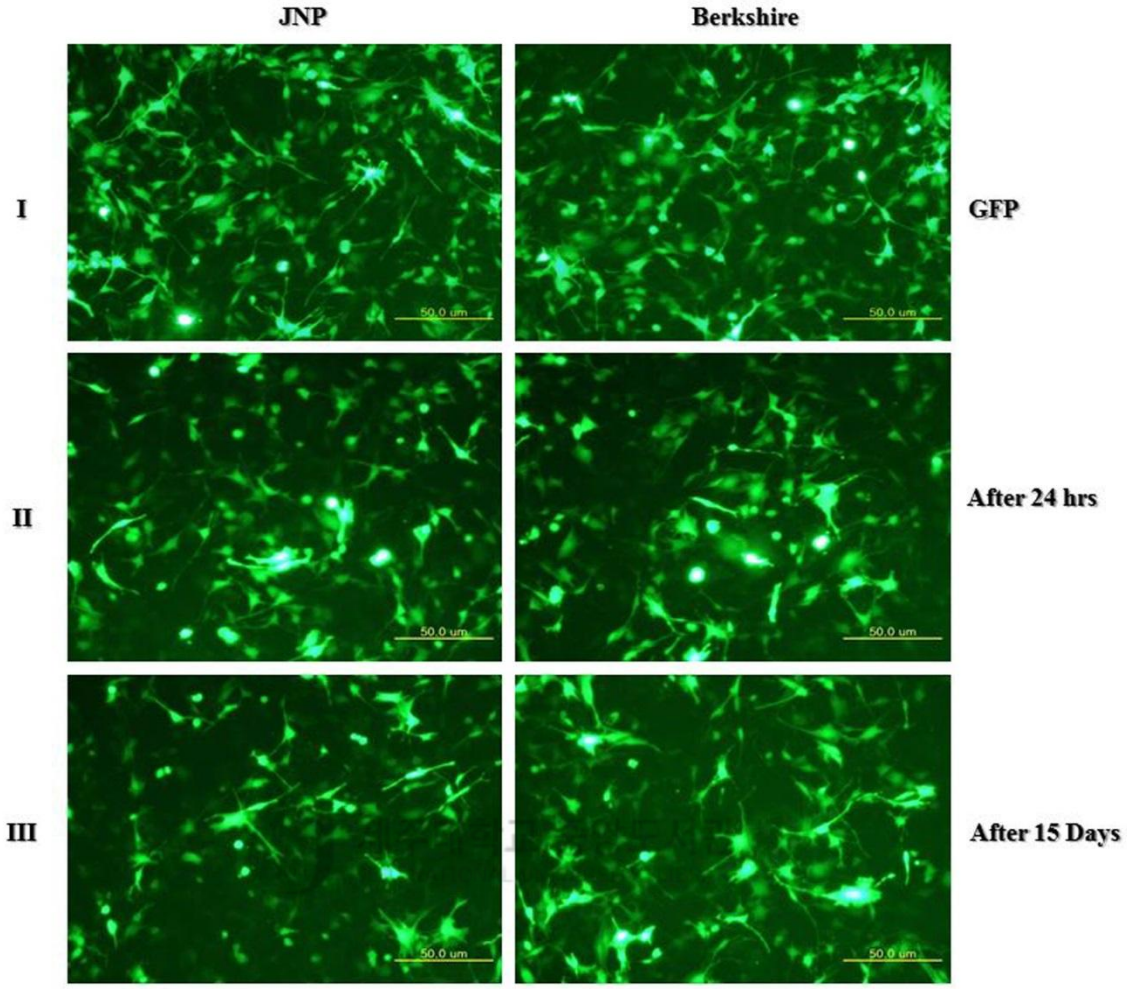


**Figure 7.11** Transfection and establishment of C2C12 stable cell lines with recombinant Piggybac-MyBPH in JNP and Berkshire respectively. Panel-I, Expression of $P_{max}$ GFP (positive control) after 24hrs of transfection; Panel-II, Expression of recombinant MyBPH GFP after 24hrs of transfection; and panel-III, Purified transfected C2C12 cell line after 15days of transfection in puromycin as selection marker.

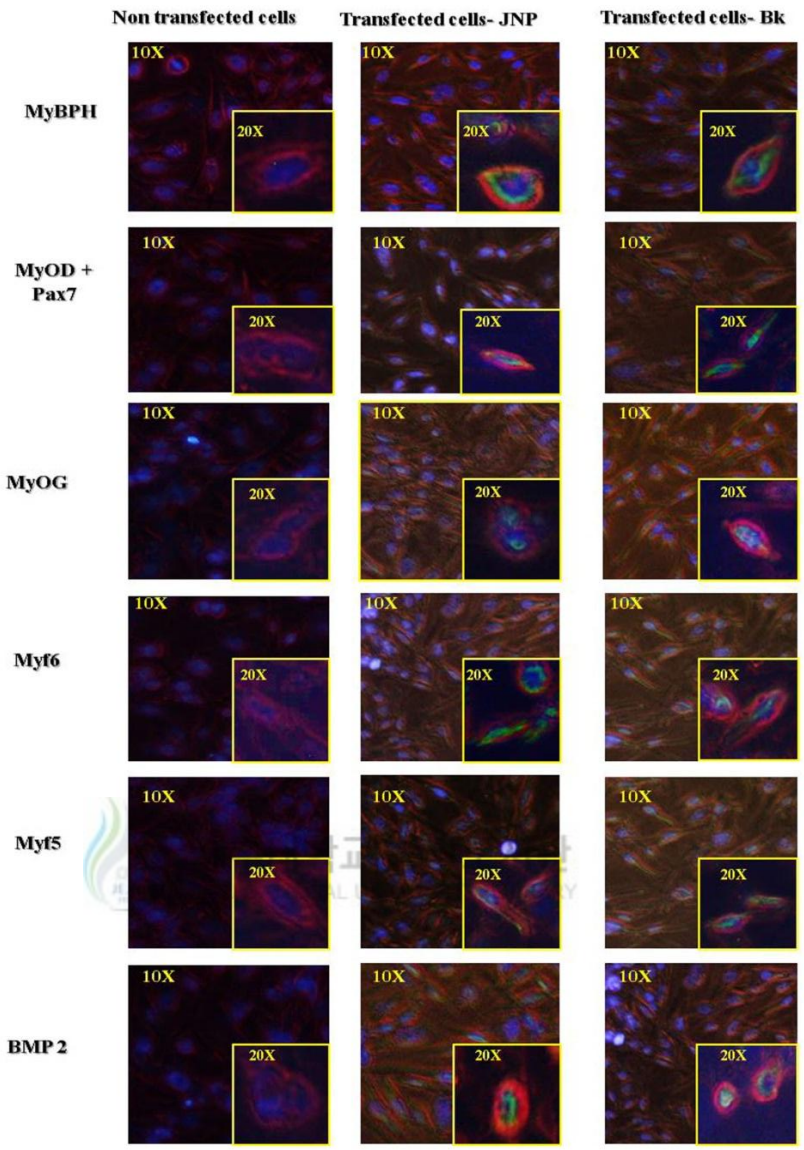


**Figure 7.12** Relative expression of *MyBPH* and MRFs in the non- transfected and stable transfected cell lines by immunocytochemistry.

significant differences have been observed in the transcript levels of bone morphogenic proteins 2 (*BMP2*) among both the breeds (Figure 7.13).

### *7.4.8. Relative expression of proteins in the transfected stable C2C12 cell lines*

The relative blot expressions of proteins have been presented as relative band intensities. The relative band intensities of the candidate proteins were normalized with the area densities of endogenous β-actin bands through image-J software. The findings of relative protein expression analysis through blotting have complimented the earlier findings of immunocytochemistry and relative quantitative mRNA transcript levels. A significantly ($P<0.05$) higher expression of MyBPH protein followed by significantly lesser expressions of MyOG, MyoD and Myf6 proteins have been recorded in the JNP transfected myoblast cell line. On the contrary, no significant differences have been recorded in the blot expressions of Myf5, Pax3, BMP2 and BMP4 proteins (Figure 7.14).

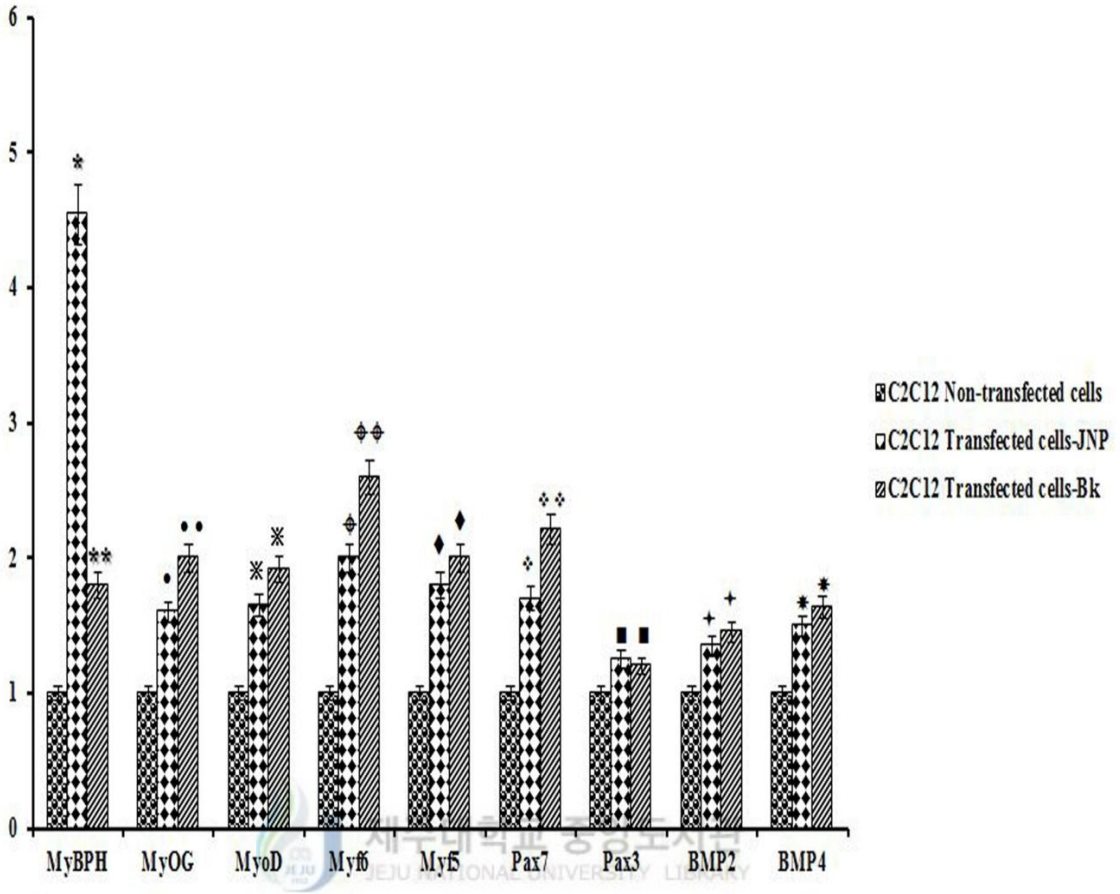


**Figure 7.13** Relative transcript levels (RQ) of *MyBPH* and MRFs in the stable transfected cell lines carrying JNP and Berkshire gene sequences. RQ has been normalized with the transcript levels of the endogenous reference (β-actin) and keeping the non-transfected cells as reference sample. Bars with different superscripts show significant differences in the transcript levels between the two breeds ($P < 0.05$). Values are expressed as mean ± SEM.

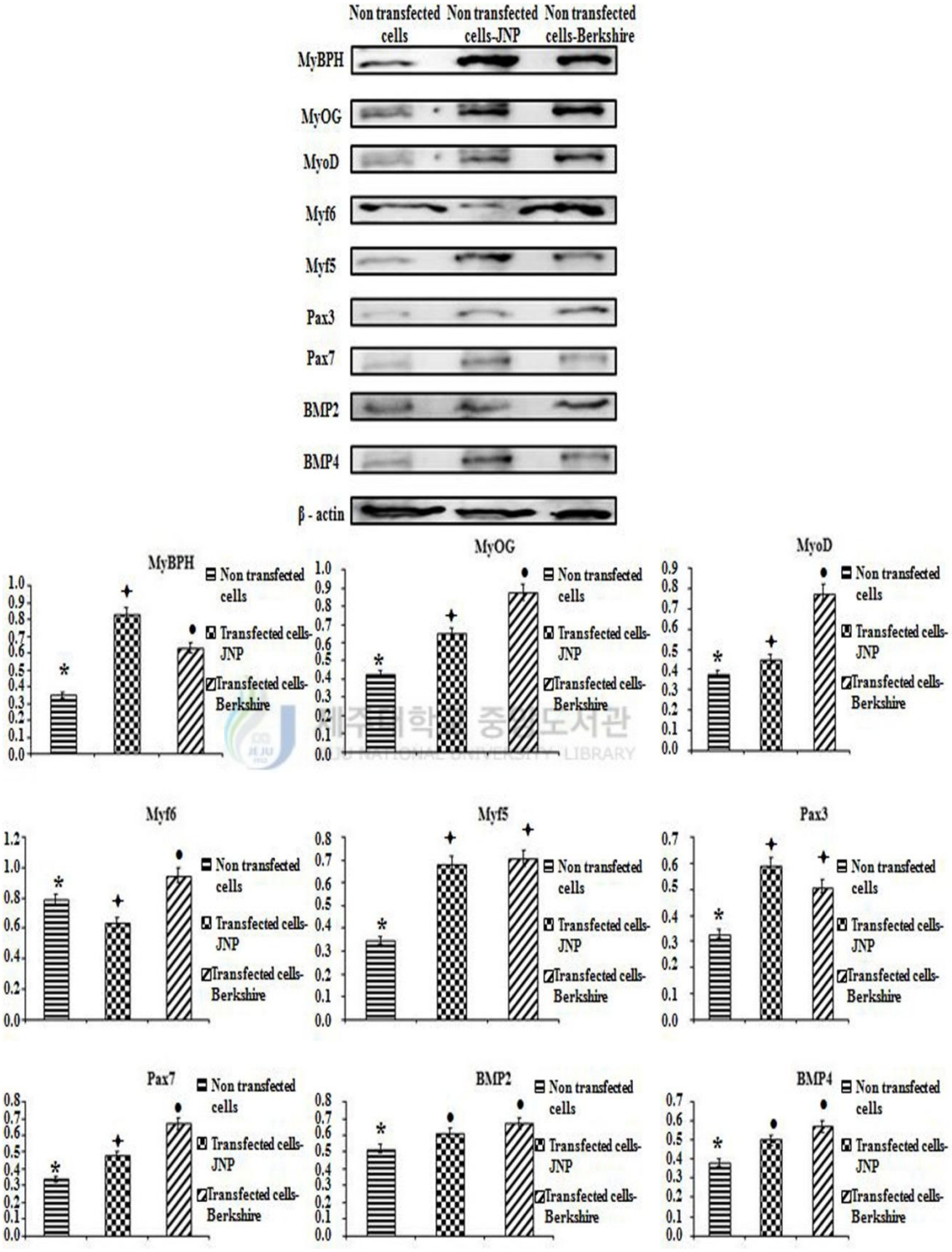


**Figure 7.14** Relative differential blot expression analysis of proteins expressed in the non-transfected and stable transfected cell lines carrying JNP and Berkshire gene sequences.

### 7.5. Discussion and conclusion

Being an economical source of animal proteins, the demand of pig meat is increasing day by day. In the tropical Asian and Southern European continents, pigs also play a fundamental role in the social and economic status of meat industry (Sodhi *et al*., 2014). Muscle growth rate is the most important economic trait for meat production. Genes for skeletal muscle development are potentially functional candidates in livestock production and meat quality (Xu *et al*., 2013). Therefore research on the relationships between the characteristics of skeletal muscle and meat quality is important to improve our understanding of the molecular basis of the phenotypic expressions in skeletal muscle (Karlsson *et al*., 1999).

Recent studies in neonatal and adult skeletal muscles of pig have revealed the developmental patterns of gene expression including gene associated with the myogenesis (Sollero *et al*., 2011). Reports from our earlier study on comparative transcriptomic analysis to identify differentially expressed genes in JNP and Berkshire breeds have reported relative higher transcript level of *MyBPH* in the *longissimus dorsi* muscle of JNP (Sodhi *et al*., 2014). Therefore, based on our previous findings, the current study targeted the identification of *MyBPH* as a genetic marker for improved myogenesis in JNP. *MyBPH* is a vital segment of myosin protein present in the skeletal muscles and is found at the cross bridges of C zones of striated muscle sarcomere. Myofibrils contain actin and myosin as the predominant proteins in thin and thick filaments (Klont *et al*., 1998; Zhang 2009). Major protein in the thick filaments which affects the muscle fiber type is myosin (MHC: myosin heavy chain). The *MyBPH* gene in pig encodes the functional protein with 479 amino acids and plays vital role in skeletal muscle development. *MyBPH* is principally associated with fibers of fast twitch muscles (Vaughan *et al*., 1993). It has been reported that *MyBPH* is quite homologous to *MyBP-C* and is a significant constituent of vertebrate myofiber skeleton. It has been reported that mutations in *MyBP-C* with deletions at the C terminus inhibit myofibrinogenesis (Gilbert *et al*., 1996). Similarly, the mutation in *MyBPH* may be inhibiting its co-expression with MRFs. Therefore the current study aimed to identify the deleterious SNPs of *MyBPH* gene in adult JNP and Berkshire breeds by *in silico* annotation as well as gene expression profiling through the stable transfected cell line. Our extend work also focused on the co-expression profiles of the other MRFs, paired box transcription factors and bone morphogenetic proteins with *MyBPH* gene in both the breeds.

Screening of functional genetic variants of *MyBPH* in the coding region through sequence- and structure-based algorithms confirmed the different levels of polymorphism among the SNP sites. However, good coherence among all the tools has predicted more than 80% of total

nsSNPs to be highly deleterious in JNP. The structure-based predication using I-Mutant3 also confirmed that all the nsSNPs of the variants are highly destabilizing. The SNPs differentially affect the growth in pig (Tian *et al*., 2014; Wang *et al*., 2013). For the *MyBPH* gene the highly deleterious polymorphic sites have been observed in JNP and in Berkshire (Table 7.3 and 7.4). Therefore, the significant ($P$<0.05) association of nsSNPs within the gene affect the function of *MyBPH* gene. *In silico* analysis by HOPE tool has indicated that misfiling and intermolecular interaction of the *MyBPH* gene can significantly ($P$<0.05) affect the structures and functions of the proteins. The four nsSNP sites in JNP significantly ($P$<0.05) affected the growth then Berkshire. Due to mutations the changes in charge, size and hydrophobicity of protein all together lead to destabilization of the protein through nsSNP. Being deleterious, these nsSNPs affect the normal functionality of the protein. Therefore, such kind of changes due to nsSNPs can be one of the reasons of altered functioning of *MyBPH* in JNP which inspite of its higher expression is still not able to potentiate the myogenesis in the JNP. The mutations in *MyBPH* gene can change the gene expression which in-turn leads to disrupt muscle-specific biological and pathophysiological processes those control muscle growth and metabolism. Moreover, these four nsSNPs may be used as breeding selection markers for JNP. In the current study, a significant ($P$<0.05) higher expression of *MyBPH* through immuno cytochemistry in JNP as compare to Berkshire has been observed. The differences in transcript level also support the concept of different genetic origin of two breeds. The existence of polymorphisms specific for breeds and their coding regions significantly target the correlation between the body weight and daily weight gain.



Further, the transcript levels of MRFs, paired box transcription factors, bone morphogenetic proteins and their co-expression profiling with *MyBPH* in both the breeds would act as a base for the deep understanding of molecular mechanisms regulating muscle growth, especially for meat production and quality (Ropka-Molik *et al*., 2011). Moreover, it has been reported that co-expression of *MyoD* and *Pax-7* promotes myogenesis and close vicinity of these two genes in our results through abstracted gene analysis in ToppCluster (Figure 7.7) compliments the finding of Braun and Gautel (2011). The sharing of specific myogenic features associated with *MyBPH* and MRFs through ToppCluster indicates that *MyBPH* can be a potent candidate gene to potentiate biological processes.

*MyoD* and *Pax7*, co- express in the activated satellite cells (Mesires and Doumit, 2001). *Pax7* is reported to induce self-renewal of satellite cells and can be one of the candidate genes to affect the spirited stages of post-natal muscle growth in the pig (Patruno *et al*., 2007). Earlier studies have enlightened that the satellite cells play a significantly important role in the muscle hypertrophy. Satellite cells are one of the oldest adult stem cells. These play vital role in the

growth and regeneration of muscles (Siegel *et al*., 2011). It has been reported that *Pax7* co-expresses with *MyoD* in both inactive and activated satellite cells. After proliferation, the activated satellite cells differentiate into myotubes or form new myotubes (Patruno *et al*., 2007). Further, it is also observed that without *MyoD* gene transcripts activated cells return to their inactive state (Zammit *et al*., 2006). In pig, it was observed that *Pax7* plays a pivitol role in satellite cell physiology and may represent one of the candidate genes influencing the dynamic stages of early post- natal growth (Patruno *et al*., 2007). Such mechanism would allow muscle progenitor expansion and/or maintains precautious differentiation (Figure 7.15). Various studies indicate that regulation of *MyoD* activity by *Pax7* represents a nodal point for the regulation of myogenic progression (Olguin and Pisconti, 2012).

In conclusion, this investigation in JNP and Berkshire pig breeds reveal *MyBPH* as a candidate gene associated with the muscle specific biological and pathophysiological processes. The investigation of polymorphisms showed that amino acids variations in *MyBPH* cause the functional disruption. The association of the polymorphism in *MyBPH* with growth and meat quality traits aids in further understanding of MRFs in different pig breeds. The type of the muscle fibers is reported to be responsible for the color, stability and tenderness of meet. The higher transcript levels of *MyBPH* in JNP during the current study indicated that it may contribute to the meat quality of JNP. Moreover the expression of *MyBPH* revealed its association with MRFs. Although *MyBPH* is highly expressed in JNP but the low expression of other MRFs might be one of the causes for retarded growth performance in JNP. Our findings on co-expression patterns of *MyBPH* with MRFs provide a base to understand the genetic basis of postnatal myogenesis process and its correlation with growth pattern in JNP and Berkshire.

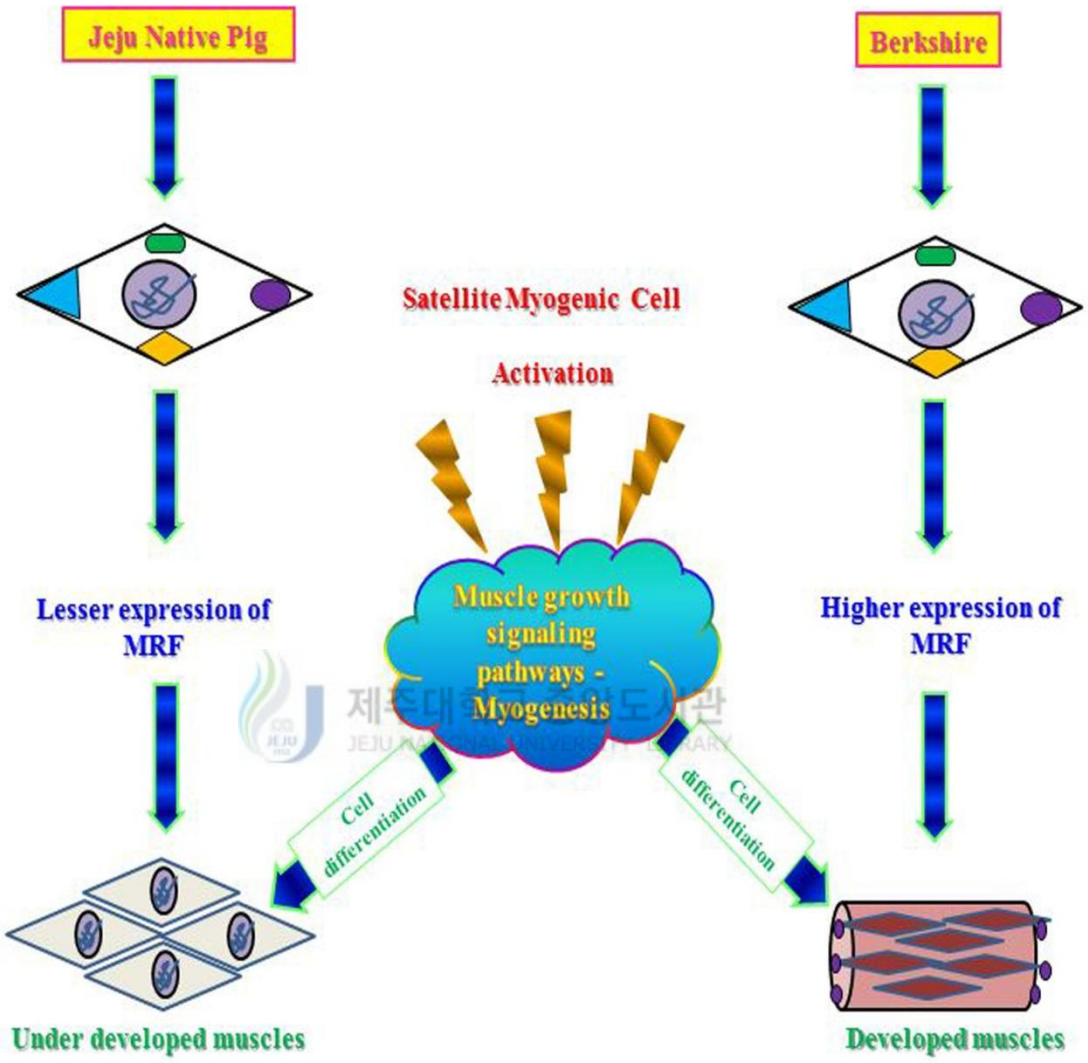


**Figure 7.15** Depiction of a bridging avenue for the understanding of the pattern of myogenesis in JNP and Berkshire**.**

# REFERENCES


Abdelmegeed M.A., Banerjee A., Yoo S.H., Jang S., Gonzalez F.J. and Song B.J. (2012) Critical role of cytochrome P450 2E1 (CYP2E1) in the development of high fat-induced non-alcoholic steatohepatitis. J. Hepatol. 57: 860–866.

Adessi C., Matton G., Ayala G., Turcatti G., Mermod J.J., Mayer P. and Kawashima E. (2000) Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. Nucleic Acids Res. 28: e87.

Adler M., Murani E., Ponsuksili S. and Wimmers K. (2013) PBMC Transcription Profiles of Pigs with Divergent Humoral Immune Responses and Lean Growth Performance. Int. J. Biol. Sci. 9: 907–916.

Agarwal S.M., Raghav D., Singh H. and Raghava G.P.S. (2010) CCDB: a curated database of genes involved in cervix cancer. Nucleic Acids Res. 39: D975–D979.

Akbar H., Schmitt E., Ballou M.A., Correa M.N., DePeters E.J. and Loor J.J. (2013) Dietary lipid during late-pregnancy and early-lactation to manipulate metabolic and inflammatory gene network expression in dairy cattle liver with a focus on PPARs. Gene. Regul. Syst. Bio. 7: 103–123.



Akha A.A.S., Theriot C.M., Erb-Downward J.R., McDermott A.J., Falkowski N.R., Tyra H.M., Rutkowski D.T., Young V.B. and Huffnagle G.B. (2013) Acute infection of mice with Clostridium difficile leads to eIF2a phosphorylation and pro-survival signaling as part of the mucosal inflammatory response. Immunology 140: 111–122.

Alger H.M., Brown J.M., Sawyer J.K., Kelley K.L., Shah R, Wilson M.D., Willingham M.C. and Rudel L.L. (2010) Inhibition of acyl-Coenzyme A: Cholesterol acyltransferase 2 (ACAT2) prevents dietary cholesterol-associated steatosis by enhancing hepatic triglyceride mobilization. J. Biol. Chem. 285: 14267–14274.

Alkan C., Kidd J.M., Marques-Bonet T., Aksay G., Antonacci F., Hormozdiari F., Kitzman J.O., Baker C., Malig M., Mutlu, O., Sahinalp S.C., Gibbs R.A. and Eichler E.E. (2009) Personalized copy number and segmental duplication maps using next-generation sequencing. Nat. Genet. 41: 1061-1067.

Alterovitz G. and Ramoni M.F. (2010) Knowledge based bioinformatics. 1st ed. Wiley Online Library. TJ International Ltd, Padstow, Cornwall.

Alves E., Fernandez A.I., Fernandez- Rodriguwz A., Perez- Montarelo D., Benitez R., Ovilo C., Rodriguez C., Perez- Montarelp D., Benitez R., Ovilo C., Rodriguez C. and Silio L.

(2009) Identification of mitochondrial markers for genetic traceability of European wild boars and Iberian and Duroc pigs. Animal 3:1216-1223.

Anderson RA, Joyce C, Davis M, Reagan JW, Clark M, Shelness G.S. and Rudel L.L. (1998) Identification of a form of acyl- CoA:cholesterol acyltransferase specific to liver and intestine in nonhuman primates. J. Biol. Chem. 273: 26747–26754.

Baik M., Etchebarne B.E., Bong J. and Vande-Haar M.J. (2009) Gene expression profiling of liver and mammary tissues of lactating dairy cows. Asian-Aust. J. Anim. Sci. 22: 871–884.

Bekhit A.E.A., Hopkins D.L., Fahri F.T. and Ponnampalam E.N. (2013) Oxidative Processes in Muscle Systems and Fresh Meat: Sources, Markers, and Remedies. Compr. Rev. Food Sci. F. 12: 565-597.

Bellot A.C., Patel K.C. and Burkholder T.J. (2005) Reduction of caveolin-3 expression does not inhibit stretch induced phosphorylation of ERK2 in skeletal muscle myotubes. J. Appl. Physiol. 98: 1554–1561.

Bilek K., Knoll A., Stratil A., Svobodova K., Horak P., Bechynova R., Van Poucke M. and Peelman L. (2008) Analysis of mRNA expression of CNN3, DCN, FBN2, POSTN, SPARC and YWHAQ genes in porcine foetal and adult skeletal muscles. Czech. J. Anim. Sci. 53: 181.



Bloom J.S., Khan Z., Kruglyak L., Singh M. and Caudy A.A. (2009) Measuring differential gene expression by short read sequencing: Quantitative comparison to 2-channel gene expression microarrays. BMC Genomics 10: 221.

Boulatnikov I. and Popov K.M. (2003) Formation of functional heterodimers by isozymes 1 and 2 of pyruvate dehydrogenase kinase. Biochim. Biophys. Acta. 1645: 183–192.

Braun T. and Gautel M. (2011) Transcriptional mechanisms regulating skeletal muscle differentiation, growth and homeostasis. Nat. Rev. Mol. Cell Biol. 12: 349-361.

Brenner S., Johnson M., Bridgham J., Golda G., Lloyd D.H., Johnson D., Luo S., McCurdy S., Foy M., Ewan M., Roth R., George D., Eletr S., Albrecht G., Vermaas E., Williams S.R., Moon K., Burcham T., Pallas M., DuBridge R.B., Kirchner J., Fearon K., Mao J. and Corcoran K. (2000) Gene expression by massively parallel signature sequencing (MPSS) on microbead arrays. Nat. Biotechnol. 18:630-634.

Buckingham M., Bajard L., Chang T., Daubas P., Hadchouel J., Meilhac S., Montarras D., Rocancourt D. and Relaix F. (2003) The formation of skeletal muscle: from somite to limb. J Anat 202: 59-68.

Cameron N.D. and Enser M.B. (1991) Fatty acid composition of lipid in longissimus dorsi muscle of Duroc and British Landrace pigs and its relationship with eating quality. Meat Sci. 29: 295-307.

Cao W., Liu N., Tang S., Bao L., Shen L., Yuan H., Zhao X. and Lu H. (2008) Acetyl-Coenzyme A acyltransferase 2 attenuates the apoptotic effects of BNIP3 in two human cell lines. Biochim. Biophys. ACTA 1780: 873–880.

Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N., Oyama R., Ravasi T., Lenhard B., Wells C., et al. (2005) The transcriptional landscape of the mammalian genome. Science 309: 1559–1563.

Cases S., Novak S., Zheng Y.W., Myers H.M., Lear S.R., Sande E., Welch C.B., Lusis A.J., Spencer T.A. and Krause B.R., Erickson S.K. and Farese Jr R.V. (1998) ACAT-2, a second mammalian acyl-CoA:cholesterol acyltransferase. J. Biol. Chem. 273: 26755-26764.

Chamoun Z., Vacca F., Parton R.G. and Gruenberg J. (2013) PNPLA3/adiponutrin functions in lipid droplet formation. Biol. Cell. 105: 219-233.

Chen C., Ai H., Ren J., Li W., Li P., Qiao R., Ouyang J., Yang M., Ma J. and Huang L. (2011) A global view of porcine transcriptome in three tissues from a full-sib pair with extreme phenotypes in growth and fat deposition by paired-end rna sequencing. BMC Genomics 12: 448.



Chen H., Centola M., Altschul S.F. and Metzger H. (1998) Characterization of gene expression in resting and activated mast cells. J. Exp. Med. 188: 1657-1668.

Chen Z., Gao X., Lei T., Chen X., Zhou L., Yu A., Lei P., Zhang R., Long H. and Yang Z. (2011) Molecular characterization, expression and chromosomal localization of porcine PNPLA3 and PNPLA4. Biotechnol. Lett. 33: 1327-1337.

Chen Z., Gerhold-Ay A., Gebhard S., Boehm D., Solbach C., Lebrecht A., Battista M., Sicking I., Cotarelo C., Cadenas C., Marchan R., Stewart J.D., Gehrmann M., Koelbl H., Hengstler J.G. and Schmidt M. (2012) Immunoglobulin kappa C predicts overall survival in node-negative breast cancer. PLoS One 7: e44741.

Cho I., Park H., Yoo C., Lee G., Lim H., Lee J., Jung E., Ko M., Lee J. and Jeon J. (2011) QTL analysis of white blood cell, platelet and red blood cell‐related traits in an F2 intercross between Landrace and Korean native pigs. Anim. Genet. 42: 621-626.

Clapperton M., Bishop S. and Glass E. (2006) Selection for lean growth and food intake leads to correlated changes in innate immune traits in Large White pigs. Anim. Sci. 82: 867–876.

Cloonan N., Forrest A.R., Kolle G., Gardiner B.B., Faulkner G.J., Brown M.K., Taylor D.F., Steptoe A.L., Wani S., Bethel G., Robertson A.J., Perkins A.C., Bruce S.J., Lee C.C., Ranade S.S., Peckham H.E., Manning J.M., McKernan K.J. and Grimmond S.M. (2008) Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nat. Methods 5: 613-619.

Coates E.E., Riggin C.N. and Fisher J.P. (2012) Matrix Molecule Influence on Chondrocyte Phenotype and Proteoglycan 4 Expression by Alginate-Embedded Zonal Chondrocytes and Mesenchymal Stem Cells. J. Orthop. Res. doi: 10.1002/jor.22166.

Cochran S.D., Cole J.B., Null D.J. and Peter J.H. (2013) Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. BMC Genet. 14: 49.

Corino C., Mourot J., Magni S., Pastorelli G. and Rosi F., (2002) Influence of dietary conjugated linoleic acid on growth, meat quality, lipogenesis, plasma leptin and physiological variables of lipid metabolism in rabbits. J. Anim. Sci. 80: 1020–1028.

Corney D.C. (2013) RNA-seq Using Next Generation Sequencing. Mater. Methods 3: 203.

Corominas J., Caldas Y.R., Oliveras A.P., Estelle J., Castello A, Alves E, Pena R.N., Ballester M. and Folch J.M. (2013) Analysis of porcine adipose tissue transcriptome reveals difference in de novo fatty acid synthesis in pigs with divergent muscle fatty acid composition. BMC Genomics 14: 843.

Costa D., Lazzarini E., Canciani B., Giuliani A., Spano R., Marozzi K., Manescu A., Cancedda R. and Tavella S. (2013) Altered bone development and turnover in transgenic mice over-expressing lipocalin-2 in bone. J. Cell. Physiol. 228: 2210-2221.

Dalvit C., Marchi M.D. and Cassandro M. (2007) Genetic traceability of livestock products: A review. Meat Sci. 77:437-449.

Damon M., Wyszynska-Koko J., Vincent A., Herault F. and Lebret B. (2012) Comparison of Muscle Transcriptome between Pigs with Divergent Meat Quality Phenotypes Identifies Genes Related to Muscle Metabolism and Structure. PLoS One 7: e33763.

Daniels T.F. (2009) Progress toward genomically optimized beef: cholesterol transport pathways and lipid homeostasis. Ph.D. Thesis, Washington State University.

Darwin C. (1868) The variation of animals and plants under domestication. John Murray, Albemarle Street, London.

Datson N.A., Jong J.V.D.P., Berg M.P.van-den, Kloet E.R.de and Vreugdenhil E. (1999) MicroSAGE: a modified procedure for serial analysis of gene expression in limited amounts of tissue. Nucleic Acids Res. 27: 1300-1307.

Dennis Jr G., Sherman B.T., Hosack D.A., Yang J., Gao W., Lane H.C. and Lempicki R.A. (2003) DAVID: Database for annotation, visualization, and integrated discovery. Genome Biol. 4: P3.

Ding S., Wu X., Li G., Han M., Zhuang Y. and Xu T. (2005) Efficient transposition of the PiggyBac (PB) transposon in mammalian cells and mice. Cell 122:473-483.

Du Clos T.W. (2000) Function of C-reactive protein. Ann. Med. 32: 274-278.

Epstein J. and Bichard M. (1984) In Evolution of domesticated animals, (ed. IL Mason), pp. 145-162. Longman, London; New York.

Erkens T., Van Poucke M., Vandesompele J., Goossens K., Van Zeveren A. and Peelman L.J. (2006) Development of a new set of reference genes for normalization of real-time RT-PCR data of porcine backfat and longissimus dorsi muscle, and evaluation with PPARGC1A. BMC Biotechnol. 6: 41–48.

Eveland A.L., McCarty D.R. and Koch K.E. (2008) Transcript profiling by 3-untranslated region sequencing resolves expression of gene families. Plant Physiol 146: 32–44.

Ferraz A.L.J., Ojeda A., Bejar M.L., Fernandes L.T., Castello A., Folch J.M. and Enciso M.P. (2008) Transcriptome architecture across tissues in the pig. BMC Genomics. 9: 173.

Frank J.W., Escobar J., Nguyen H.V., Jobgen S.C., Jobgen WS, Jobgen W.S., Davis T.A. and Wu G. (2007) Oral N carbamylglutamate supplementation increase protein synthesis in skeletal muscle of piglets. J Nutr. 137: 315–319.

Freeman M.R. and Solomon K.R. (2004) Cholesterol and prostate cancer. J. Cell Biochem. 91: 54–69.

Friendly M. (2002). Corrgrams: Exploratory displays for correlation matrices. Am. Stat. 56: 316–324.

Gilbert R., Kelly M.G., Mikawa T. and Fischman D.A. (1996) The carboxyl terminus of myosin binding protein C (MyBP-C, C-protein) specifies incorporation into the A-band of striated muscle. J. Cell Sci. 109: 101-111.

Giuffra E., Kijas J.M., Amarger V., Carlborg O., Jeon J.T. and Andersson L. (2000) The origin of the domestic pig: independent domestication and subsequent introgression. Genetics 154: 1785-1791.

Glenn K.L., Ramos A.M. and Rothschild M.F. (2007) Analysis of FMO genes and off flavour in pork. J. Anim. Breed. Genet. 124: 35-38.

Goh K.I., Cusick M.E., Valle D., Childs B., Vidal M. and Barabasi A.L. (2007) The human disease network. Proc. Natl. Acad. Sci. USA 104:8685-8690.

Gorni C., Iacuaniello S., Castiglioni B., Pagnacco G. and Mariani P. (2008) Pig KALRN, MYH1, MLC2V, SNX13, AK1, and PPIA loci RH mapping and chromosome position refining. Genet. Mol. Res. 7(4): 982-985.

Graugnard D., Moyes K., Trevisi E., Khan M., Keisler D., Drackley J., Bertoni G. and Loor J. (2012) Liver lipid content and inflammometabolic indices in peripartal dairy cows are altered in response to prepartal energy intake and postpartal intramammary inflammatory challenge. J. Dairy Sci. 96: 918-935.

Gruen M. and Gautel M. (1999) Mutations in beta-myosin S2 that cause familial hypertrophic cardiomyopathy (FHC) abolish the interaction with the regulatory domain of myosin-binding protein-C. J. Mol. Biol. 286: 9339- 9349.

Guryev V., Eugene B., Malik R., Ronald H., Plasterk A and Cuppen E. (2004) Single nucleotide polymorphisms associated with rat expressed sequences. Genome Res. 14: 1438-1443.

Hardcastle T.J. and Kelly K.A. (2010) Bayseq: empirical bayesian methods for identifying differential expression in sequence count data. BMC Bioinformatics 11: 422.

Harismendy O., Ng P.C., Strausberg R.L., Wang X., Stockwell T.B., Beeson K.Y., Schork N.J., Murray S.S., Topol E.J., Levy S. and Frazer K.A. (2009) Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome Biol. 10:R32.

Hashimoto S., Suzuki T., Dong H.Y., Nagai S., Yamazaki N. and Matsushima K. (1999) Serial analysis of gene expression in human monocyte-derived dendritic cells. Blood 94: 845-852.

Heid C.A., Stevens J., Livak K.J. and Williams P.M. (1996) Real time quantitative PCR. Genome Res. 6:986-994.

Higuchi R., Fockler C., Dollinger G. and Watson R. (1993) Kinetic PCR: Real time monitoring of DNA amplification reactions. Biotechnol. 11:1026-1030.

Hitzemann R., Bottomly D., Darakjian P., Walter N., Iancu O., Searles R., Wilmot B. and McWeeney S. (2013) Genes, behavior and next-generation RNA sequencing. Genes Brain Behav. 12:1-12.

Ho C.S., Rochelle E.S., Martens G.W., Schook L.B. and Smith D.M. (2006) Characterization of swine leukocyte antigen polymorphism by sequence-based and PCR-SSP methods in Meishan pigs. Immunogenetics 58: 873-882.

Hong D., Rhie A., Park S.S., Lee J., Ju Y.S., Kim S., Yu S.B., Bleazard T., Park H.S. and Rhee H. (2012) FX: An RNA-seq analysis tool on the cloud. Bioinformatics 28: 721-723.

Hosack D.A., Dennis Jr G., Sherman B.T., Lane H.C. and Lempicki R.A. (2003) Identifying biological themes within lists of genes with EASE. Genome Biol. 4: R70.

Hu Zhi-Liang B.J. and Reecy J.M. (2008) CateGOrizer: A Web-Based Program to Batch Analyze Gene Ontology Classification Categories. Online J. Bioinform. 9: 108–112.

Hubank M. and Schatz D.G. (1994) Identifying differences in mRNA expression by representational difference analysis of cDNA. Nucleic Acids Res. 22: 5640-5648.

Hur S.J., Jeong T.C., Kim G.D., Jeong J.Y. , Cho I.C., Lim H.T., Kim B.W. and Joo S.T. (2013) Comparison of Live Performance and Meat Quality Parameter of Cross Bred (Korean Native Black Pig and Landrace) Pigs with Different Coat Colors. Asian Australas. J. Anim. Sci. 26: 1047-1053.

Hussain M.R.M., Shaik N.A., Yousuf Al-Aama J., Asfour H.Z., Khan F.S., et al. Masoodi T.A., Khan M.A. and Shaik N.S. (2012) In silico analysis of single nucleotide polymorphisms (SNPs) in human BRAF gene. Gene 508: 188–196.

Hyman E.D. (1988) A new method of sequencing DNA. Anal. Biochem. 174: 423–436.

Irie M., Ohmoto K. and Nishimura K. (1985) Standard measurements for determination of porcine soft fat. Final reports for research grants for meat and meat products. The Ito Foundation, Tokyo, Japan. 3: 156-164.

Izsvak Z., Hackett P.B., Copper L.J. and Ivics Z. (2010) Translating sleeping beauty transposition into cellular therapies: Victories and challenges. Bioessays 32:756- 767.

Jassar A.C., Bernado P., Chidgey M. and Overduin M. (2013) Hinged plakin domains provide specialized degrees of articulation in envoplakin, periplakin and desmoplakin. Plos One 8: e69767 .

Jett J.H., Keller R.A., Martin J.C., Marrone B.L., Moyzis R.K., Ratliff R.L., Seitzinger N.K., Shera E.B. and Stewart C.C. (1989) High-speed DNA sequencing: an approach based upon fluorescence detection of single molecules. J. Biomol. Struct. Dyn. 7: 301-309.

Jiang Z., Michal J.J., Chen J., Daniels T.F., Kunej T, Garcia M.D., Gaskins C.T., Busboom J.R., Alexander L.J., Wright Jr. R.W. and MacNeil M.D. (2009) Discovery of novel genetic networks associated with 19 economically important traits in beef cattle. Int. J. Biol. Sci. 5: 528-542.

Jin S., Kim C., Song Y., Jang W., Kim Y., Yeo J., Kim J. and Kang K. (2001) Physicochemical characteristics of longissimus muscle between the Korean native pig and Landrace, Korean J. Food Sci. Ani. Resour. 21: 142-148.

Jones D., Ruzzo W., Peng X. and Katze M. (2012) A new approach to bias correction in RNA-Seq. Bioinformatics 28: 921-928.

Jones, G. F. (1998) Genetic aspects of domestication, common breeds and their origin, pp. 17–50 in The Genetics of the Pig, edited by A. Ruvinsky and M. F. Rothschild. CAB International, Oxon, UK.

Joyce C.W., Shelness G.S., Davis M.A., Lee R.G., Skinner K, Anderson R.A. and Rudel L.L. (2000) ACAT1 and ACAT2 membrane topology segregates a serine residue essential for activity to opposite sides of the endoplasmic reticulum membrane. Mol. Biol. Cell. 11: 3675–3687.

Jung W.Y., Kwon S.G., Son M., Cho E.S., Lee Y., Kim J.H., Kim B.W., Hwang J.H., Kim T.W. and Park H.C. (2012) RNA-Seq approach for genetic improvement of meat quality in pig and evolutionary insight into the substrate specificity of animal carbonyl reductases. Plos One 7: E42198.

Kaimal V., Bardes E.E., Tabar S.C., Jegga A.G. and Aronow B.J. (2010) ToppCluster: a multiple gene list feature analyzer for comparative enrichment clustering and network-based dissection of biological systems. Nucl. Acids Res. 38: W96-W102.

Karlssona A.H., Klontb R.E. and Fernandezc X. (1999) Skeletal muscle fibers as factors for pork quality. Livest. Prod. Sci. 60: 255-269.

Kavathas P., Sukhatme V.P., Herzenberg L.A. and Parnes J.R. (1984) Isolation of the gene encoding the human T-lymphocyte differentiation antigen Leu-2 (T8) by gene transfer and cDNA subtraction. Proc. Natl. Acad. Sci. USA 81: 7688-7692.



Kawaida H. (1993). Studies on the performance of meat production and meat quality in pigs. Report of Kagoshima Prefectural Animal Husbandry Experiment Stations. 26: 1–195.

Kijas J.M. and Andersson L. (2001) A phylogenetic study of the origin of the domestic pig estimated from the near-complete mtDNA genome. J. Mol. Evol. 52: 302-308.

Kim B.W., Cho I.C., Park M.S., Zhong T., Lim H.T., Lee S.S., Park H.B., Ko M.S., Lee J.H. and Jeon J.T. (2011) Characterization of the European type of maternal lineage evident in extant Jeju native pigs. Genes Genom. 33: 111-117.

Kim D., Seong P., Cho S., Kim J., Lee J., Jo C. and Lim D. (2009) Fatty acid composition and meat quality traits of organically reared Korean Native Black Pigs. Livest. Sci. 120: 96-102.

Kim G.D., Kim B.W., Jeong J.Y., Hur S.J., Cho I.C., Lim H.T. and Joo S.T. (2013) Relationship of Carcass Weight to Muscle Fiber Characteristics and Pork Quality of Crossbred (Korean Native Black Pig × Landrace) F2 Pigs. Food Bioprocess Technol. 6:522-529.

Kim K.I., Lee J.H., Li K., Zhang Y.P., Lee S.S., Gongora J. and Moran C. (2002) Phylogenetic relationships of Asian and European pig breeds determined by mitochondrial DNA D-loop sequence polymorphism. Anim. Genet. 33: 19-25.

Kim K.S., Yeo J.S. and Kim J.W. (2002) Assessment of genetic diversity of Korean native pig (Sus scrofa) using AFLP markers. Genes Genet. Syst. 77:361-368.

Kim M.J., Li G.H., Oh J.D., Cho K.H., Jeon G.J., Choi B.H., Lee J.H., Hong Y.S., Kong H.S. and Lee H.K. (2007) Characterization of a Korean traditional porcine breed using microsatellite markers and the establishment of an individual identification system. Korean J. Food Sci. Anim. Resour. 27:150-156.

Kim N.K., Lim J.H., Song M.J., Kim O.H., Park B.Y., Kim M.J., Hwang I.H. and Lee C.S. (2008) Comparisons of Longissimus muscle metabolic enzymes and muscle fiber types in korean and western pig breeds. Meat Sci. 78: 455-460.

Kim S.W., Choi Y., Choi J.S., Kim J.J., Choi, B.H., Kim, T.H. and Kim, K.S. (2011) Porcine fatty acid synthase gene polymorphisms are associated with meat quality and fatty acid composition. Korean J. Food Sci. Anim. Res. 31: 356-365.

Kim S.W., Jung J.H., Do K.T., Kim K.S., Do C.H., Kyu P.J., Joo Y.K., Kim T.S., Choi B.H. and Kim T.H. (2007) Investigation of single nucleotide polymorphisms in porcine candidate gene for growth and meat quality traits in the Berkshire breed. J Life Sci. 17: 1622-1626.

Kim T.H., Kim K.S., Choi B.H., Yoon D.H., Jang G.W., Lee K.T., Chung H.Y., Lee H.Y., Park H.S. and Lee J.W. (2005) Genetic structure of pig breeds from Korea and China using microsatellite loci analysis. J. Anim. Sci. 83:2255-2263.

Klont R., Brocks L. and Eikelenboom G. (1998) Muscle fibre type and meat quality. Meat Sci. 49: S219- S229.

Kouba M., Enser M., Whittington F., Nute G. and Wood J. (2003) Effect of a high-linolenic acid diet on lipogenic enzyme activities, fatty acid composition, and meat quality in the growing pig. J. Anim. Sci. 81: 1967-1979.

Kubistova Z., Mrazek F. and Petrek M. (2009) Polymorphisms of the immune response genes: selected biological, methodical and medical aspects. Biomedical papers of the Medical Faculty of the University Palacky, Olomouc, Czechoslovakia Republic 153: 93-102.

Kushwaha R.S., Rosillo A., Rodriguez R. and McGill Jr. H.C. (2005) Expression levels of ACAT1 and ACAT2 genes in the liver and intestine of baboons with high and low lipemic responses to dietary lipids. J. Nutr. Biochem. 16: 714–721.

Lage K., Hansen N.T., Karlberg E.O., Eklund A.C., Roque F.S., Donahoe P.K., Szallasi Z., Jensen T.S. and Brunak S. (2008) A large-scale analysis of tissue-specific pathology and

gene expression of human disease genes and complexes. Proc. Natl. Acad. Sci. USA 05:20870-20875.

Lan J., Lei M.G., Zhang Y.B., Wang J.H., Feng X.T., Xu D.Q., Gui J.F. and Xiong Y.Z., 2009. Characterization of the porcine differentially expressed PDK4 gene and association with meat quality. Mol. Biol. Rep. 36: 2003–2010.

Larson G., Dobney K., Albarella U., Fang M., Matisoo-Smith E., Robins J., Lowden S., Finlayson H., Brand T., Willerslev E., Conwy P.R., Andersson L. and Cooper A. (2005) Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. Science 307: 1618-1621.

Lee S.H., Choe J.H., Choi Y.M., Jung K.C., Rhee M.S., Hong K.C., Lee S.K., Ryu Y.C. and Kim B.C. (2012) The influence of pork quality traits and muscle fiber characteristics on the eating quality of pork from various breeds. Meat Sci. 90: 284–291.

Lee Y.H., Kwon S.G., Park D.H., Kwon E.J., Cho E.S., Bang W.Y., Park H.C., Park B.Y., Choi J.S. and Kim C.W. (2011) Development of high meat quality using microsatellite markers in Berkshire pigs. J. Anim. Sci. Technol. 53:89-97.

Lefebvre V. and Smits P. (2005) Transcriptional control of chondrocyte fate and differentiation. Birth Defects Res., Part C 75: 200-212.



Li A.M., Pettitt S.J., Eckert S., Ning Z., Rice S., Cadinanos J., Yusa K., Conte N. and Bradley A. (2013) The PiggyBac transposon displays local and distant reintergration prefernces and can cause mutations at noncanonical integration sites. Mol. Cell Biol. 33: 1317-1330.

Li B., Ruotti V., Stewart R.M., Thomson J.A. and Dewey C.N. (2010) RNA-Seq gene expression estimation with read mapping uncertainty. Bioinformatics 26: 493-500.

Li B., Ruotti V., Stewart R.M., Thomson J.A. and Dewey C.N. (2010) RNA-Seq gene expression estimation with read mapping uncertainty. Bioinformatics 26: 493–500.

Li H.D. (2008) I: Segregation, Cloning, SNP Detection and Its Association with Economic Traits of 5 Bovine Genes II: Quantitative Trait Loci Analysis of Meat Quality Traits in Pigs. Ph.D. thesis. Dissertation ID: 1647914. http://www.dissertationtopic.net/doc/1647914.

Li M., Tian S., Yeung C.K.L., Meng X., Tang Q., Niu L., Wang X., Jin L., Ma J., Long K., Zhou C., Cao Y., Zhu L., Bai L., Tang G., Gu Y., Jiang A., Li X. and Li R. (2014) Whole-genome sequencing of Berkshire (European native pig) provides insights into its origin and domestication. Scientific Reports 4:4678.

Li M.A., Turner D.J., Ning Z., Yusa K., Liang Q., Eckert S., Rad L., Fitzgerald T.W., Craig N.L. and Bradley A. (2011) Mobilization of giant PiggyBac transposons in the mouse genome. Nucleic Acids Res. 39: 148.

Liang P. and Pardee A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257: 967-971.

Liu M., Ling X. and Xiong Y. (2012) Molecular characterization of differentially expressed TXNIP gene and its association with porcine carcass traits. Mol. Biol. Rep. 39: 10439-10446.

Lockhart D.J., Dong H., Byrne M.C., Follettie M.T., Gallo M.V., Chee M.S., Mittmann M., Wang C., Kobayashi M., Horton H. and Brown E.L. (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat. Biotechnol. 14: 1675-1680.

Looft C., (2013) Trancriptome analysis of testis and liver for androstenone by using RNA sequencing. Proc. of Plant & Anim. Genome XXI Conference, San Diego, CA.

Lucero H.A. and Robbins P.W. (2004) Lipid rafts-protein association and the regulation of protein activity. Arch. Biochem. Biophys. 426: 208–224.

Lunney J.K. (2007) Advances in swine biomedical model genomics. Int. J. Biol. Sci. 10: 179-184.

Luo W., Cheng D., Chen S., Wang L., Li Yong., Ma X., Song X., Liu X., Li W., Liang J., Yan H., Zhao K., Wang C., Wang L. and Zhang L. (2012) Genome-wide association analysis of meat wuility trails in a Procine Large White X Minzhu Intercross population. Int. J. Biol. Sci. 8:580-595.



Luo W., Diaz F.J. and Wiltbank M.C. (2011) Induction of mRNA for chemokines and chemokine receptors by prostaglandin F2α- is dependent upon stage of the porcine corpus luteum and intraluteal progesterone. Endocrinology 152: 2797-2805.

Ma X., Lin Y., Jiang Z., Zheng C., Zhou G., Yu D., Cao T., Wang J. and Chen F. (2010) Dietary arginine supplementation enhances antioxidative capacity and improves meat quality of finishing pigs. Amino Acids 38: 95-102.

Magi A., Benelli M., Gozzini A., Girolami F., Torricelli F. and Brandi M.L. (2010) Bioinformatics for Next generation Sequencing Data. Gene 1: 294-307.

Majer M., Popov K.M., Harris R.A., Bogardus C. and Prochazka M. (1998) Insulin down regulates pyruvate dehydrogenase kinase (PDK) mRNA: potential mechanism contributing to increased lipid oxidation in insulin-resistant subjects. Mol. Genet. Metab. 65: 181-186.

Malek M., Dekkers J.C., Lee H.K., Baas T.J., Prusa K., Lonergan E.H. and Rothschild M.F. (2001). A molecular genome scan analysis to identify chromosomal regions influencing economic traits in the pig. II. meat and muscle composition. Mamm. Genome. 12: 637-645.

Mallard B.A., Wilkie B.N. and Kennedy B.W. (1998) Immune responsiveness in swine: eight generations of selection for high and low immune response in Yorkshire pigs. Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, Armidale, Australia. 1998. pp. 257-264.

Mardis E.R. (2008) Next-generation DNA sequencing methods. Annu. Rev. Genomics. Hum. Genet. 9: 387-402.

Marioni J.C., Mason C.E., Mane S.M., Stephens M. and Gilad Y. (2008) RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. Genome Res. 18:1509-1517.

Mehla K., Magotra A., Choudhary J., Singh A.K., Mohanty A.K., Upadhyay R.C., Srinivasan S., Gupta P., Choudhary N., Antony B. and Khan F. (2014) Genome-wide analysis of the heat stress response in Zebu (Sahiwal) cattle. Gene 10, 500-507.

Mesires N.T. and Doumit M.E. (2001) Satellite cell proliferation and differentiation during postnatal growth of porcine skeletal muscle. Am. J. Physiol. Cell Physiol. 282: 899-902.

Metzger J., Tonda R., Beltran S., Agueda L., Gut M. and Distl O. (2014) Next generation sequencing gives an insight into the characteristics of highly selected breeds versus non-breed horses in the course of domestication. BMC Genomics 15: 562.

Molik K.R., Eckert R. and Piorkowska K. (2011) The expression pattern of myogenic regulatory factors MyoD, Myf6 and Pax7 in postnatal porcine skeletal muscles. Gene Exp. Pat. 11: 79- 83.

Molik K.R., Zukowski K., Eckert R., Gurgul A., Piorkowska K. and Oczkowicz M. (2014) Comprehensive analysis of the whole transcriptomes from two different pig breeds using RNA-Seq method. Animal Genetics 45: 674-684.

Moon J.K., Kim K.S., Kim J.J., Choi B.H., Cho B.W., Kim T.H. and Lee C.K. (2009) Differentially expressed transcripts in adipose tissue between korean native pig and yorkshire breeds. Anim. Genet. 40: 115-118.

Morlein D., Lungershausen M., Steinke K., Sharifi A.R. and Knorr C. (2012) A single nucleotide polymorphism in the cyp2e1 gene promoter affects skatole content in backfat of boars of two commercial duroc-sired crossbred populations. Meat Sci. 92: 739- 744.

Mortazavi A., Williams B.A., Mccue K., Schaefferm L. and Wold B., (2008) Mapping and quantifying mammalian transcriptomes by RNA-seq. Nat. Methods. 5: 621-628.

Mourot J., Kouba M. and Peiniau P. (1995) Comparative study of in vitro lipogenesis in various adipose tissues in the growing domestic pig (Sus Domesticus). Comp. Biochem. Physiol., B: Biochem. Mol. Biol. 111: 379-384.

National Pork Producers Council (1995) Genetic Evaluation, Terminal Line Program Results. Des Moines: NPPC.

Ng P.C. and Henikoff S. (2001) Predicting deleterious amino acid substitutions. Genome Res. 11: 863-874.

Nguyen T.M., Sawyer J.K., Kelley K.L., Davis M.A., Kent C.R. and Rudel L.L. (2012) ACAT2 and ABCG5/G8 are both required for efficient cholesterol absorption in mice: evidence from thoracic lymph duct cannulation. J. Lipid Res. 53: 1598-1609.

Novak R.F. and Woodcroft K.J. (2000) The alcohol inducible form of cytochrome P450 (Ctyp 2E1): role in toxicology and regulation of expression. Arch. Pharma. Res. 23: 267–282.

Nowrousian M. (2010) Next-generation sequencing techniques for eukaryotic microorganisms: sequencing-based solutions to biological problems. Eukaryot. Cell 9:1300-1310.

Nyren P. and Lundin A. (1985) Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis. Anal. Biochem. 151: 504-509.

Oh J.D., Song K.D., Seo J.H., Kim D.K., Kim S.H., Seo K.S., Lim H.T., Lee J.B., Park H.C., Ryu Y.C., Kang M.S., Cho S., Kim E.S., Choe H.S., Kong H.S. and Lee H.K. (2014) Genetic Traceability of Black Pig Meats Using Microsatellite Markers. Asian Australas. J. Anim. Sci. 27: 926-931.

Oh S.H. and Whitley N. C. (2011). Pork Production in China, Japan and South Korea. Asian-Aust. J. Anim. Sci. 24:1629-1636.

Okrouhla M., Stupka R., Citek J., Sprysl M., Trnka M and Kluzakova E. (2008) Effect of lean meat proportion on the chemical composition of pork. Czech. J. Food. Sci. 26: 464–469.

Okubo K., Hon N., Matoba R., Niiyama T., Fukushima A., Kojima Y. and Matsubara K. (1992) Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. Nat. Genet. 2: 173-179.

Okumura N., Ishiguro N., Nakano M., Hirai K., Matsui A. and Sahara M. (1996) Geographic population structure and sequence divergence in the mitochondrial DNA control region of the Japanese wild boar (Sus scrofa leucomystax), with reference to those of domestic pigs. Biochem. Genet. 34: 179-189.

Olguin H.C. and Pisconti A. (2012) Marking the tempo for myogenesis: Pax 7 and the regulation of muscle stem cell fate decisions. J. Cell. Mol. Med. 16: 1013- 1025.

Orford M., Hadjipavlou G., Tzamaloukas O., Chatziplis D., Koumas A., Mavrogenis A., Papachristoforou C. and Miltiadou D. (2012) A single nucleotide polymorphism in the acetyl-coenzyme A acyltransferase 2 (ACAA2) gene is associated with milk yield in Chios sheep. J Dairy Sci. 95: 3419-3427.

Osada K., Sasaki E. and Sugano M. (1994) Lymphatic absorption of oxidized cholesterol in rats. Lipids 29: 555-559.

Panitz F., Stengaard H., Hornshoj H., Gorodkin J., Hedegaard J, Cirera S., Thomsen B, Madsen L.B., Hoj A., Vingborg R.K., Zahn B., Wang X., Wang X., Wernersson R., Jorgensen C.B., Knudsen K.S., Arvin T., Lumholdt S., Sawera M., Green T., Nielsen B.J., Havgaard J.H., Brunak S., Fredholm M. and Bendixen C. (2007) SNP mining porcine ESTs with MAVIANT, a novel tool for SNP evaluation and annotation. Bioinformatics 23: 387-391.

Parini P., Davis M., Lada A.T., Erickson S.K., Wright T.L., Gustafsson U., Sahlin S., Einarsson C., Eriksson M., Angelin B., Tomoda H., Omura S., Willingham M.C. and Rudel L.L. (2004) ACAT2 is localized to hepatocytes and is the major cholesterol-esterifying enzyme in human liver. Circulation 110: 2017-2023.

Park B.Y., Kim N.K., Lee C.S. and Hwang I.H. (2007) Effect of fiber type on postmortem proteolysis in longissimus muscle of Landrace and Korean native black pigs. Meat Sci. 77: 482-491.

Pashaj A., Yi X., Xia M., Canny S., Riethoven J.J. and Moreau R. (2013) Characterization of genome-wide transcriptional changes in liver and adipose tissues of ZDF (fa/fa) rats fed R-α-lipoic acid by next-generation sequencing. Physiol. Genomics. 45: 1136-1143.

Paszek A.A., Flickinger G.H., Fontanesi L., Beattie C.W., G.A. Rohrer, Alexander L. and Schook L.B. (1998) Evaluating evolutionary divergence with microsatellites. J. Mol. Evol. 46: 121-126.

Patruno M., Caliaro F., Martinello T. and Mascarello F. (2007) Expression of the paired box domain Pax7 in myogenic cells isolated from the porcine semitendinosus muscle after birth. Tissue Cell 40: 1- 6.

Petkov S.G., Marks H., Klein T., Garcia R.S., Gao Y., Stunnenberg H. and Hyttel P. (2011) In vitro culture and characterization of putative porcine embryonic germ cells derived from domestic breeds and Yucatan mini pig embryos at days 20–24 of gestation. Stem Cell Res, 6: 226-237.

Prather R. S. (2013) Pig Genomics for biomedicine. Nature Biotech. 31: 122-124.

Pushkarev D. and Neff N.F. (2009) Quake, S.R. Single-molecule sequencing of an individual human genome. Nat. Biotechnol. 27: 847-852.

Rehfeldt C. and Kuhn G. (2006) Consequences of birth weight for postnatal growth performance and carcass quality in pigs as related to myogenesis. J. Anim. Sci. 84: E113-E123.

Riechman S.E., Andrews R.D., Maclean D.A. and Sheather S. (2007) Statins and dietary and serum cholesterol are associated with increased lean mass following resistance training. J. Gerontol. A. Biol. Sci. Med. Sci. 62: 1164-1171.

Rischkowsky B. and Pilling D. (2007) Commission on Genetic Resources for Food and Agriculture. The state of the world's animal genetic resources for food and agriculture. Commission on Genetic Resources for Food and Agriculture, Food and Agriculture Organization of the United Nations, Rome.

Robinson M.D. and Oshlack A. (2010) A Scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 11: R25.

Robinson M.D., Mccarthy D.J. and Smyth G.K. (2010) Edger: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139-140.

Rolfe R.A., Nowlan N.C., Kenny E.M., Cormican P., Morris D.W., Prendergast P.J., Kelly D. and Murphy P. (2014) Daniel Kelly2 and Paula Murphy1, Identification of mechanosensitive genes during skeletal development: alteration of genes associated with cytoskeletal rearrangement and cell signaling pathways. BMC Genomics 15:48.

Ropka-Molik K., Eckert R. and Piorkowska K. (2011) The expression pattern of myogenic regulatory factors MyoD, Myf6 and Pax7 in postnatal procine skeletal muscles. Gene Expr. Patterns 11: 79-83.

Rosenvold K. and Andersen H.J. (2003) Factors of significance for pork quality - a review. Meat Sci. 64: 219-237.

Rozen S. and Skaletsky H. (2000) Primer3 on the www for general users and for biologist programmers. Methods Mol. Biol. 132: 365-386.

Rustemeyer S., Lamberson W., Ledoux D., Wells K., Austin K. and Cammack K. (2011) Effects of dietary aflatoxin on the hepatic expression of apoptosis genes in growing barrows. J. Anim. Sci. 89: 916-925.

Ruvinsky A. and Rothschild M.F. (1998) Systematics and evolution of the pig, pp. 1–16 in The Genetics of the Pig, edited by Ruvinsky A. and Rothschild M.F.. CAB International, Oxon, UK.

Ryu Y.C. and Kim B.C. (2006) Comparison of histochemical characteristics in various pork groups categorized by postmortem metabolic rate and pork quality. J. Ani. Sci. 84: 894-901.

Ryu Y.C., Choi Y.M., Lee S.H., Shin H.G., Choe J.H., Kim J.M., Hong K.C. and Kim, B.C., (2008) Comparing the histochemical characteristics and meat quality traits of different pig breeds. Meat Sci. 80: 363-369.

Samborski A., Graf A., Krebs S., Kessler B. and Bauersachs S. (2013) Deep sequencing of the porcine endometrial transcriptome on day 14 of pregnancy. Biol. Reprod. 88: 84.

Schena M., Shalon D., Davis R.W. and Brown P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467-470.

Schibli D.J., Hunter H.N., Aseyev V., Starner T.D., Wiencek J.M., McCray Jr. P.B., Tack B.F. and Vogel H.J. (2002) The solution structures of the human beta-defensins led to a better understanding of the potent bactericidal activity of HBD3 against Staphylococcus aureus. J. Biol. Chem. 277: 8279-8289.

Schwab C.R., Baas T.J., Stalder K.J. and Mabry J.W. (2006) Effect of long-term selection for increased leanness on meat and eating quality traits in Duroc swine. J. Anim. Sci. 84: 1577-1583.

Sendler E., Johnson G.D. and Krawetz S.A. (2011) Local and global factors affecting RNA sequencing analysis. Anal. Biochem. 419: 317-322.

Seo H., Choi Y., Shim J., Choi Y. and Ka H. (2012) Regulatory mechanism for expression of IL1B receptors in the uterine endometrium and effects of IL1B on prostaglandin synthetic enzymes during the implantation period in pigs. Biol. Reprod. 87: 1-11.



Serao N.V., Veroneze R., Ribeiro A.M., Verardo L.L., Braccini N.J., Gasparino E., Campos C.F., Lopes P.S. and Guimaraes S.E. (2011) Candidate gene expression and intramuscular fat content in pigs. J. Anim. Breed. Genet. 128: 28-34.

Shendure J. and Ji H. (2008) Next-generation DNA sequencing. Nat. Biotechnol. 26:1135-1145.

Shendure J., Porreca G.J., Reppas N.B., Lin X., McCutcheon J.P., Rosenbaum A.M., Wang M.D., Zhang K., Mitra R.D. and Church G.M. (2005) Accurate multiplex colony sequencing of an evolved bacterial genome. Science 309: 1728-1732.

Sheng X., Ni H., Liu Y., Li J., Zhang L. and Guo Y. (2014) RNA-seq analysis of bovine intramuscular, subcutaneous and perirenal adipose tissues. Molecular Biology Reports 41: 1631–1637.

Shim J.M., Seo D.W., Seo S., Kim J.J., Min D.M., Kim J.K., Jeon J.T. and Lee J.H. (2010) Discrimination of Korean cattle (Hanwoo) with imported beef from USA based on the SNP markers. Korean J. Food Sci. Anim. Resour. 30:918-922.

Shin D.J., Joshi P., Hong S.H., Mosure K., Shin D.G. and Osborne T.F. (2012) Genome-wide analysis of FoxO1 binding in hepatic chromatin: Potential involvement of FoxO1 in

linking retinoid signaling to hepatic gluconeogenesis. Nucleic Acids Res. 40:11499-11509.

Shiraki T., Kondo S., Katayama S., Waki K., Kasukawa T., Kawaji H., Kodzius R., Watahiki A., Nakamura M., Arakawa T., Fukuda S., Sasaki D., Podhajska A., Harbers M., Kawai J., Carninci P. and Hayashizaki Y. (2003) Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proc. Natl. Acad. Sci. USA 100: 15776–15781.

Shyamsundar R., Kim Y.H., Higgins J.P., Montgomery K., Jorden M., Sethuraman A., Rijn M.V.D., Botstein D., Brown P.O. and Pollack J.R. (2005) A DNA microarray survey of gene expression in normal human tissues. Genome Biol. 6: R22.

Siegel A.L., Kuhlmann P.K. and Cornelison D.D. (2011) Muscle satellite cell proliferation and association: New insights from myofiber time-lapse imaging. Skelet. Muscle 1: 7.

Simons K. and Toomre D. (2000) Lipid rafts and signal transduction. Nat. Rev. Mol. Cell Biol. 1: 31-39.

Smythe G.M., Eby J.C., Disatnik M.H. and Rando T.A. (2003) A caveolin-3 mutant that causes limb girdle muscular dystrophy type 1C disrupts Src localization and activity and induces apoptosis in skeletal myotubes. J. Cell Sci. 116: 4739-4749.

Sodhi S.S., Park W.C., Ghosh M., Kim J.N., Sharma N., Shin K.Y., Cho I.C., Ryu Y.C., Oh S.J., Kim S.H., Song K.D., Hong S.P., Cho S., Kim H.B. and Jeong D.K. (2014) Comparative transcriptomic analysis to identify differentially expressed genes in fat tissue of adult Berkshire and Jeju Native Pig using RNA-seq. Mol. Biol. Rep. 41:6305-6315.

Sodhi S.S., Song K.D., Ghosh M., Sharma N., Lee S.J., Kim J.H., Kim N., Mongre R.K., Adhikari P., Kim J.Y., Hong S.P., Oh S.J. and Jeong D.K.. (2014) Comparative transcriptomic analysis by RNA-seq to discern differential expression of genes in liver and muscle tissues of adult Berkshire and Jeju Native Pig. Gene 546: 233-242.

Sollero B.P., Guimaraes S.E.F., Rilington V.D., Tempelman R.J., Raney N.E., Steibel J.P., Guimaraes J.D., Lopes P.S., Lopes M.S. and Ernst C.W. (2011) Transcriptional profiling during foetal skeletal muscle development of Piau and Yorkshire-Landrace cross bred pigs. Anim. Genet. 42:600-612.

Soto M.I.N., Jozani R.J., Bridle B. and Mallard B.A. (2008) Analysis of gene expression patterns by microarray hybridization in blood mononuclear cells of SLA-DRB1 defined Canadian Yorkshire pigs. BMC Res. Notes 31: 1-7.

Sousounis K., Looso M., Maki N., Ivester C.J., Braun T. and Tsonis P.A. (2013) Transcriptome analysis of newt lens regeneration reveals distinct gradients in gene expression patterns. Plos One 8: e61445.

Stewart T.S., Lofgren D.L., Harris D.L., Einstein M.E. and Schinckel A.P. (1991) Genetic improvement programs in livestock: swine testing and genetic evaluation system (stages). J. Anim. Sci. 69: 3882-3890.

Sutherland M.A. Rodriguez-Zas S.L., Ellis M. and Salak J.J.L. (2005) Breed and age affect baseline immune traits, cortisol, and performance in growing pigs. J. Anim. Sci. 83: 2087-2095.

Suzuki K., Abe H., Ogawa Y., Ishida M., Shimizu T. and Suzuki A. (1997) Effect of terminal sire on the meat quality of three-way crossbred pigs. Anim. Sci. Technol. 68: 310-317.

Suzuki K., Shibata T., Kadowaki H., Abe H. and Toyoshima T. (2003). Meat quality comparison of Berkshire, Duroc and crossbred pigs sired by Berkshire and Duroc. Meat Sci. 64: 35-42.

Suzuki K., Shimizu Y., Abe H., Tonai K. and Suzuki A. (2001). Comparison of meat quality between breeds, sex and site of longissimus thoracis muscle in pigs. Anim. Sci. J. 72: J215-J223.



Swindle M.M., Makin A., Herron A.J., Clubb F.J. and Frazier K.S. (2012) Swine as models in biomedical research and toxicology testing. Vet. Pathol. 49: 344-356.

Switonski M., Stachowiak M., Cieslak J., Bartz M. and Grzes M. (2010) Genetics of fat tissue accumulation in pigs: a comparative approach. J. Appl. Genet. 51:153-168.

Sylvie R.B. (2011) The Collagen Family. Cold Spring Harb. Perspect. Biol. 3: a004978.

Takahashi H., Lassmann T., Murata M. and Carninci P. (2012) 5' end-centered expression profiling using cap-analysis gene expression and next-generation sequencing. Nat. Protoc. 7: 542-561.

Tarazona S., Alcalde F.G., Dopazo J., Ferrer A. and Conesa A. (2011) Differential expression in RNA-seq: A matter of depth. Genome Res. 21: 2213-2223.

te Pas M.F.W., Hulsegge I., Coster A., Pool M.H., Heuven H.H. and Janss L.L.G. (2007) Biochemical pathways analysis of microarray results: regulation of myogenesis in pigs. BMC Dev. Biol. 7: 66.

Tian Y.G., Yue M., Gu Y., Gu W.W. and Wang Y.J. (2014) Single-nucleotide polymorphism analysis of GH, GHR, and IGF-1 genes in minipigs. Braz. J. Med. Biol. Res. 47: 753-758.

Trapnell C., Roberts A., Goff L., Pertea G., Kim D., Kelley D.R., Pimentel H., Salzberg S.L., Rinn J.L. and Pachter L. (2012) Differential gene and transcript expression analysis of RNA-seq Experiments with tophat and cufflinks. Nature Prot. 7: 562-578.

Tribout T., Caritez J.C., Gruand J., Bouffaud M., Guillouet P, et al. Billon Y., Pery C., Laville E. and Bidanel J.P. (2010) Estimation of genetic trends in French Large White pigs from 1977 to 1998 for growth and carcass traits using frozen semen. J. Anim. Sci. 88: 2856-2867.

Uddin M.J., Duy D.N., Cinar M.U., Tesfaye D., Tholen E., Juengst H., Looft C. and Schellander K. (2011) Detection of quantitative trait loci affecting serum cholesterol, LDL, HDL, and triglyceride in pigs. BMC Genet. 12: 62.

Ushio H. and Nagasaka R. (2013) Utilization of biological responses of fish and shellfish for improving seafood qualities. Aqua. Bio. Sci Mono. 6: 91-98.

Van Poucke M., Yerle M., Tuggle C., Piumi F., Genet C., Van Zeveren A. and Peelman L.J. (2001) Integration of porcine chromosome 13 maps. Cytogenet. Cell Genet. 93: 297-303.

Vaughan K.T., Weber F.E., Reinach F.C., Ried T., Ward D. and Fischman D.A. (1993) Human myosin-binding protein H (MyBP-H): Complete primary sequence, repeat structure, genomic organization and chromosomal localization. Genomics 16: 34-40.

Velculescu V.E., Zhang L., Vogelstein B. and Kinzler K.W. (1995) Serial analysis of gene expression. Science 270:484-487.

Venkatesh B., Hettwer U., Koopmann B. and Karlovsky P. (2005) Conversion of cDNA differential display results (DDRT-PCR) into quantitative transcription profiles. BMC Genomics 6:51.

Verbeke W., Van Oeckel M.J., Warnants N., Viaene J. and Boucque C.V. (1999) Consumer perceptions, facts and possibilities to improve acceptability of health and sensory characteristics of pork. Meat Sci. 53: 77-99.

Wagner G.P., Kin K. and Lynch V.J. (2012) Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. 131: 281-285.

Wang D.Z., Valdez M.R., McAnally J., Richardson J. and Olson E.N. (2001) The Mef2c gene is a direct transcriptional target of myogenic bHLH and MEF2 proteins during skeletal muscle development. Development 128: 4623-4633.

Wang H.L., Zhu Z.M., Wang H., Yang S.L., Zhao S.H. and Li K. (2003) Molecular characterization and association analysis of porcine CA3. Cytogenet. Genome Res. 115: 129-133.

Wang L., Feng Z., Wang X., Wang X. and Zhang X. (2010) Degseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 26: 136-138.

Wang L., Liu X., Zhang L., Yan H., Luo W., Liang J, et al. Cheng D., Chen S., Ma X., Song X., Zhao K. and Wang L. (2013) Genome-wide copy number variations inferred from SNP genotyping arrays using a Large White and Minzhu intercross population. PLoS One 8: e74879. doi: 10.1371/journal.pone.0074879.

Wang M.C., Bohmann D. and Jasper H. (2003) JNK signalling confers tolerance to oxidative stress and extends lifespan in Drosophila. Dev. Cell 5: 811-816.

Wang X., Carre W., Rejto L. and Cogburn L.A. (2005) Transcriptional profiling in liver of hormonally-manipulated chickens. In: Dawson A, Sharp PJ, editors. Functional avian endocrinology. Narosa Publishing House, New Delhi, India.

Wang Z., Gerstein M. and Snyder M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. 10: 57–63.

Watanabe T., Hayashi Y., Kimura J., Yasuda Y., Saitou N., Tomita T. and Ogasawara N. (1986) Pig mitochondrial DNA: polymorphism, restriction map orientation, and sequence data. Biochem. Genet. 24: 385-396.

Welsh J., Chada K., Dalal S.S., Cheng R., Ralph D. and McClelland M. (1992) Arbitrarily primed PCR finger printing of RNA. Nucleic Acid Res. 20: 4965-4970.



Whittemore C. (2006) Development and Improvement of Pigs by Genetic Selection In Whittemore's Science and Practice of Pig Production, (ed. I Kyriazakis, C Whittemore), pp. 184-205. Blackwell Publishing Ltd.

Wilhelm B.T. and Landry J.R. (2009) RNA-seq quantitative measurement of expression through massively parallel RNA-sequencing. Methods 48: 249-257.

Wittwer C.T., Herrmann M.G., Moss A.A. and Rasmussen R.P. (1997) Continuous fluorescence monitoring of rapid cycle DNA amplification. Bio. Techniques. 22:130-138.

Wong G.K., Liu B., Wang J., Zhang Y., Yang X, et al. (2004) International chicken polymorphism map consortium. A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. Nature 432: 717-722.

Wood J., Richardson R., Nute G., Fisher A., Campo M., Kasapidou E., Sheard P. and Enser M. (2004) Effects of fatty acids on meat quality: A review. Meat Sci. 66: 21-32.

Wood J.D. (2001) Meat quality and the designer pig. In: Lyons TP, Coles DJA, editors. Concepts of meat science. Nottingham Nutrition International, Nottingham. 17-31.

Wu G., Bazer F.W., Davis T.A., Kim S.W., Li P., Rhoads J.M., Satterfield M.C., Smith S.B., Spencer T.E. and Yin Y. (2008) Arginine metabolism and nutrition in growth, health and disease. Amino Acids 37: 153-168.

Wu T., Zhang Z., Yuan Z., Lo L.J., Chen J., Wang Y. and Peng J. (2013) Distinctive genes determine different intramuscular fat and muscle fiber ratios of the longissimus dorsi muscles in Jinhua and Landrace Pigs. PLoS One 8: e53181.

Wyszynska-Koko J. and Kuryl J. (2004) Porcine MYF6 gene: sequence, homology analysis, and variation in the promoter region. Anim. Biotechnol. 15: 159- 173.

Xu H., Xu G., Wang D., Ma J. and Wan L. (2013) Molecular cloning, sequence identification and expression analysis of novel caprine MYLPF gene. Mol. Biol. Rep. 40: 2565-2572.

Yaffe D. and Saxel O. (1977) Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. Nature 270: 725-727.

Yamamoto M., Wakatsuki T., Hada A. and Ryo A. (2001) Use of serial analysis of gene expression (SAGE) technology. J. Immunol. Methods 250: 45-66.

Yu G., Xiang H., Wang J. and Zhao X. (2013) The phylogenetic status of typical Chinese native pigs: analyzed by Asian and European pig mitochondrial genome sequences. J. Anim. Sci. Biotechnol. 4: 9.



Yuan H.X., Xiong Y. and Guan K.L. (2013) Nutrient sensing, metabolism and cell growth control. Mol. Cell 49: 379-387.

Zammit P.S., Partridge T.A. and Yablonka-Reuveni Z. (2006) The skeletal muscle satellite cell: the stem cell that came in from the cold. J. Histochem. Cytochem. 54: 1177-1191.

Zhang W. (2009) Involvement of protein degradation, calpain autolysis and protein nitrosylation in fresh meat quality during early postmortem refrigerated storage. Ph. D. Thesis, Iowa State University, Ames, Iowa.

Zhang W., Morris Q.D., Chang R., Shai O., Bakowski M.A., Mitsakakis N., Mohammad N., Robinson M.D., Zirngibl R., Somogyi E., Laurin N., Eftekharpour E., Sat E., Grigull J., Pan Q., Peng W.T., Krogan N., Greenblatt J., Fehlings M., Kooy D.V., Aubin J., Bruneau B.G., Rossant J., Blencowe B.J., Frey B.J. and Hughes T.R. (2004) The functional landscape of mouse gene expression. J. Biol. 3: 21.

Zhou J. and Shephard E.A. (2006) Mutation, polymorphism and perspectives for the future of human flavin containing mono- oxygenase 3. Mutat. Res. 612:165-171.

Zhou Q.Y., Fang M.D., Huang T.H., Li C.C., Yu M. and Zhao S.H. (2009) Detection of differentially expressed genes between Erhualian and Large White placentas on day 75 and 90 of gestation. BMC Genomics 10:337. doi: 10.1186/1471-2164-10-337.

Zimdahl H., Nyakatura G., Brandt P., Schulz H., Ummel O, Fartmann B., Brett D., Droege M., Monti J., Lee Y.A., Sun Y., Zhao S., Winter E.E., Ponting C.P., Chen Y., Kasprzyk A., Birney E., Ganten D., N. Hubner. (2004) A SNP map of the rat genome generated from cDNA sequences. Science 303: 807.