

## 2. Animal Genomics & Bioinformatics Research

### (1) Construction of Standard Reference Genome Map in Hanwoo

The new sequencing technologies for whole genome of an organism have been developed. A proven Hanwoo bull based on the National Institute of Animal Science, Rural Development Administration, and the ongoing improvement in the beef was selected. High molecular weight genome library was constructed from the selected Hanwoo bull. In parallel, massive sequencing was done by next generation sequencing technology to produce specific genome information in Hanwoo. Genomic DNA was extracted from blood of a proven Hanwoo bull selected for improvement program. The color space sequencing method with high accuracy using SOLiD platform was applied to sequence Hanwoo genome. Large scale of sequencing data was mapped to a reference bovine (Hereford) genome produced by the international consortium of bovine genome sequencing with Bioscope software. After mapping, the specific genomic variations were discovered from mapping information. We used RefSeq data of Btau4.0 to annotate SNPs, and found 1,039,405 SNPs (17.2%; /total) in genic region. In addition, 31,724 SNPs were observed in exon. From the SNPs in exon, 17,717 SNPs were in CDS, and there were 7,167 synonymous SNPs and 10,550 non-synonymous SNPs(nsSNPs) respectively. Especially nsSNPs were classified into 2,968 homozygote and 4,199 heterozygote, which have potential to affect 1,775 and 2,161 genes respectively. All data were deposited to Hanwoo genome database. High molecular weight genome library (Bacterial Artificial Chromosome; BAC) was constructed from the same proven Hanwoo bull with six coverage or more of Hanwoo genome. And then, PCR(polymerase chain reaction)-based BAC clone screening system was prepared to validation of specific genomic regions in Hanwoo. Therefore, specificities of the Hanwoo genome was proven by sequencing BAC clones with shotgun strategy. All specificities were mapped on a draft of the Hanwoo genome map.

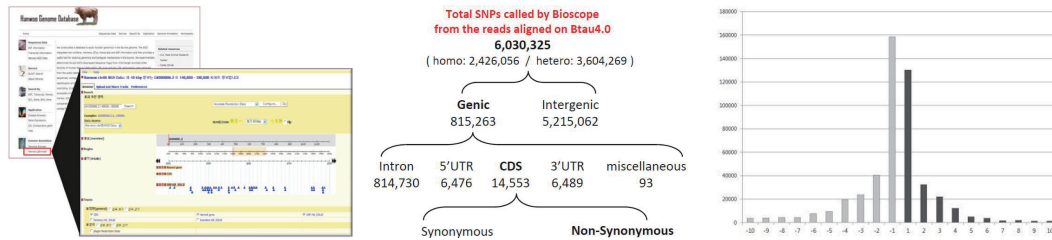


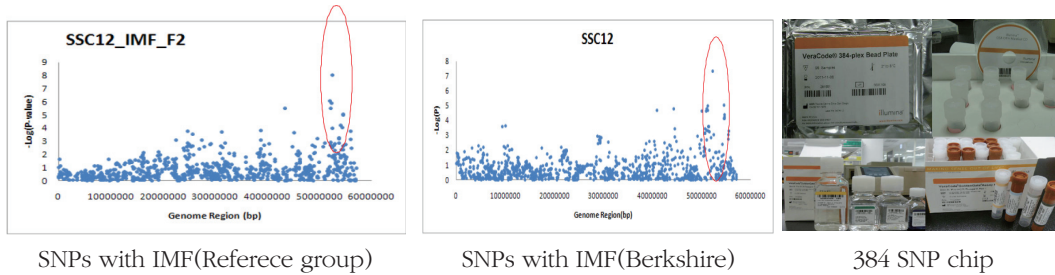
Fig. 2-1. Results of Hanwoo genome re-sequencing. (A) Example view of Hanwoo genome re-sequencing database, (B) distribution of the discovered single nucleotide polymorphic variations, (C) distribution of the detected indel variations

## (2) Porcine 384 SNP chip development for diagnosis of meat quality

Recently, livestock - as cattle, chicken, porcine etc. - in the genome project is completed, large amount of information of genome have been produced. When selection of animals, breeders hope to use genomic information, and studies of Genomic Selection method have been actively.

When the genetic factors will be found and controlled, customized pig breed should be developed by using marker assist selection (MAS). Commercial 60K SNP chip (iSelect Infinium BeadChip) was offered in consequence of Swine Genome Sequencing Consortium. It is contained porcine 62,163 SNPs.

A 594 reference family(Korean Native Pig X Landrace) and 704 purebred Berkshire were used SNP genotype analysis and association test with economic traits. Reference family group revealed 10 SNPs associated with belly, 30 SNPs with intramuscular fat contents, 40 SNPs with back fat thickness, 5 SNPs share force, 6 SNPs with meat color and 2 SNPs with daily weight gain, and these SNPs were patented. Berkshire group 12 SNPs associated with intramuscular fat contents, 8 SNPs with back fat thickness and 2 SNPs with share force. We made 384 SNP chip using these SNPs and these SNP markers may provide useful genetic information about meat quality of porcine.



### (3) Genome research of immunity and fat through nutrigenomics and gene network in pigs

The objective of the present study(2010) was to investigate the effects of dietary lysine and vitamin A level on the growth performance, carcass and meat quality parameters in finishing pigs. Regarding the growth performance, final weight, average daily gain and feed conversion ratio were no effect of lysine level and vitamin A level. Dressing was higher in lysine 0.70% level than in lysine 0.48% level(P=0.01). Backfat thickness, meat color, marbling score, meat grade and carcass grade were no effect of lysine and vitamin A level. For meat quality parameters, it didn't show any significant difference in chemical composition (% moisture, crude fat, crude ash) by lysine and vitamin A level.

Objective of this gene expression study (2010) was to obtain the effects of dietary fat sources on gene expression. Gene expression of INSR, FOSO1, PDE3 were higher(P<0.01) in lysine 0.75% level than in lysine 0.45% level. For improvement of porcine immunity system, CASP8 and IL2 genes were found in higher gene expressoion in spleen of pig which was supplemented with quercetin diet.

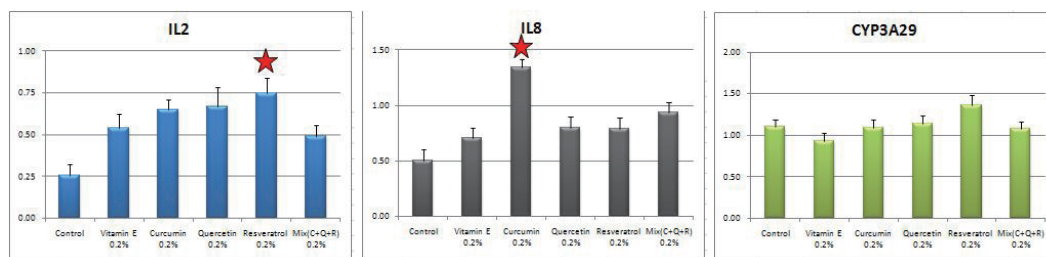


Fig. 2-2. Gene expression related with immunity system)



#### (4) Genetic markers development for genetic disease in dogs

Canine Hip Dysplasia (CHD) is a common inherited trait in dogs and the secondary debilitating hip osteoarthritis which results in pain, lameness, and physical disability. Because CHD is a complex trait, the genotype of a dog cannot be judged from its phenotype. To determine the genetic variation of traits associated with canine hip dysplasia, we used 48 dogs. DNA of dog was extracted with Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA) and genotyped Canine SNP20 Beadchip (Illumina, San Diego, CA, USA). In this canine pedigree, traits associated with canine hip dysplasia are heritable. The associated SNP was 25 loci in CFA 4, CFA 38 and X chromosome (genome-wide  $P < 0.001$ ). These data suggest that the QTL which control CHD may be expressed differently in different breed and their crossbreed offspring.



SNPs related with CHD in CFA 4

Canine hip dysplasia

Public information for CHD

#### (5) Discovery of cellulase genes from ruminal microbial metagenome in Korean Black Goat

The rumen is an unknown treasury of microorganisms including fungi, protozoa, bacteria and archaea that are mostly obligate anaerobic. The bacterial and fungal components in rumen digest the plant fiber. Unfortunately, rumen microorganisms are not able to be cultured in vitro in many cases. Therefore, there are many kinds of unknown microorganisms in rumen. In this study, we investigated the diversity of ruminal bacteria in three Korean Black Goats (KBGs) that had been fed only with rice straw for one month. First, rumen digests were obtained via the rumen fistula after 90 minutes after morning feeding. Samples were partitioned into liquid phase (Lq) and adherent phase (Ad) before DNA extraction. Metagenomic DNAs were sequenced using 454 GS FLX Titanium.

Approximately, NGS reads of 2 Gb were assembled with Newbler 2.5.3 software. ORF finding, COG and Pfam analysis were performed in RAMMCAP pipeline. A total of 1,298 GH (glycosyl hydrolase) family genes and 255 CBM (carbohydrate-binding module) family genes could be found. Additionally, fosmid library was constructed and 150 fosmid clones with CMCase activity were screened. Of those clones, five fosmid clones were sequenced with shotgun strategy and 16 cellulase genes were found. Therefore, this approach may produce a great chance to find more effective cellulase genes.

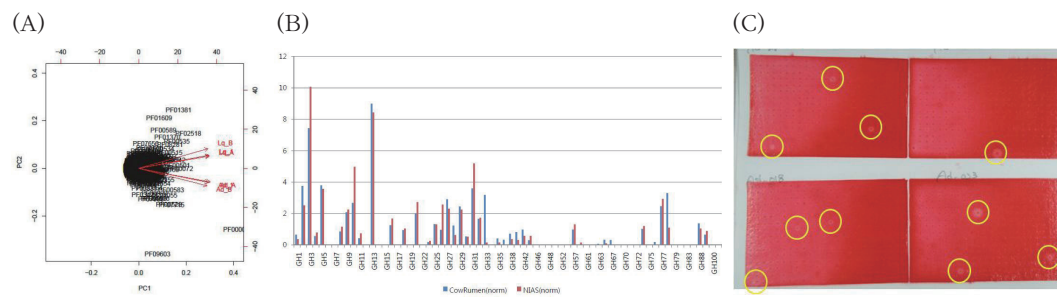


Fig. 2-3. Metagenome analysis and screening of fosmid clones with cellulase activity. (A) Difference between Lq and Ad phases by Pfam analysis, (B) Comparison with cattle ruminal metagenome in frequency of GH family genes. (C) Examples of screening of fosmid clones with cellulase activity.

#### (6) Identification of genes related to Hanwoo marbling using protein- protein interaction network

Complex traits are determined by the combined effects of many loci and are affected by gene networks or biological pathways. Systematic biology approaches have an important role in the identification of candidate genes related to complex diseases or traits at the system level. The present study systemically analyzes genes that are associated with bovine marbling score and identifies the relationships of those genes. The candidate nodes were obtained using MedScan text-mining tools and linked by protein-protein interaction (PPI) from the Human Protein Reference Database (HPRD). To determine key node of marbling, the degree and betweenness centrality (BC) were used. The





hub nodes and biological pathways of our network are consistent with the previous reports about marbling traits, and also suggest unknown candidate genes associated with intramuscular fat. Five nodes were identified as hub genes, which were consistent with the network analysis using quantitative reverse-transcription PCR (qRT-PCR). Key nodes of the PPI network have positive role (RUNX1T1) and negative roles (RXRA, CAMK2A) in the development of intramuscular fat by several adipogenesis-related pathways. This study provides genetic information for identifying candidate genes for the marbling trait in bovine.

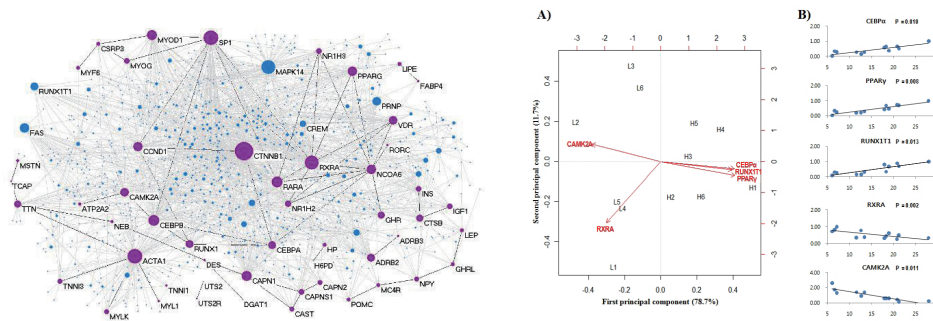


Fig. 2-4. Bovine marbling PPI(Protein-Protein Interaction) network and Analysis of gene expression data by PCA and regression model)

### (7) Prediction of GEV with SNP tagging using RJMCMC approach in Hanwoo

Recent advances in molecular biotechnology facilitates not only incorporation of DNA information into conventional animal breeding but also detection of genes that contribute to genetic variation of quantitative traits. The incorporated DNA information into estimating of genomic breeding value may achieve an improvement of estimated breeding value (EBV) and selection accuracy in cattle populations. Genomic breeding values (GEV) predicted from dense marker panels are now being used in the dairy cattle for bull selection. However, the cost of genotyping for the dense SNP panel is still too high to apply genomic selection for many animals. If potential selection candidate bulls were genotyped for a standard panel of less expensive loose SNP panel that

traces the chromosome segments having QTL effects in selection candidate, it would be practiced cost effective genomic selection. In this study, the accuracy of GEBV was evaluated based on SNP markers, that were selected by the QTL effect derived from single marker regression analysis( $P < 0.0001$ ). GEBV was predicted using RJMCMC. The RJMCMC predicted a prior QTL distribution based on each of SNP allele frequencies and phenotypes and then least squares were estimated to assume known distribution of QTL effects.

The result was quite successful for the correlation between estimated GEBV and EBV for carcass traits(MAR, CWT, EMA, BF) using tagged SNP having QTL effect and SNP plus polygenic.

	MAR	CWT	EMA	BF
Var(G)	0.60	370.4	24.1	5.82
Var(E)	0.61	754.3	42.0	6.13
Var(P)	1.22	1124.8	66.1	11.95
Heritability	0.49	0.33	0.37	0.49
Accuracy	0.90	0.93	0.90	0.81

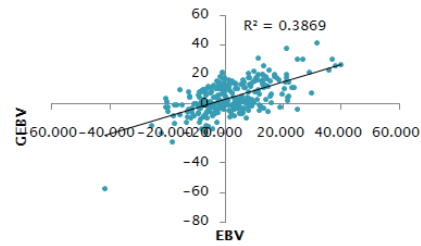


Fig. 2-5. Accuracy of GEBV with EBV