

# Investigation of Genetic Markers for Intramuscular Fat in the Hybrid Wagyu Cattle With Bulk Segregant Analysis

**Yun Zhu**

Ningxia University

**Liyun Han**

Ningxia Agricultural Reclamation Helanshan Dairy Co. Ltd,

**Peng Li**

Ningxia University

**Xiaolong Kang**

Ningxia University

**Xingang Dan**

Ningxia University

**Yun Ma**

Ningxia University

**Yuangang Shi** (✉ [shyga818@126.com](mailto:shyga818@126.com))

Ningxia University

---

## Research Article

**Keywords:** Bulk Segregant Analysis, Beef Cattle, QTLs, IMF

**Posted Date:** March 12th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-277670/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Bulked Segregant Analysis (BSA) is a rapid strategy for identifying genetic markers in specific regions of the phenotypical population and it has been widely used for QTLs mapping in smaller mixed F2 and F3 populations. We applied a modified BSA method to assess genome-wide homozygous and heterozygous linkage patterns in the Chinese Wagyu Beef Cattle F2/F3 mixed population. Two overlapped regions from F2 and F3 populations on autosomes were found with high-density heterozygote alleles between high and low intramuscular fat groups. Regions from 24.8M~29.6M of chromosome 23 were identified as most significantly correlated to the intramuscular fat in our samples. We also identified other 4 potential loci on chromosomes 5, 9, 15, and 21 correlated with Intramuscular fat. This study provided a novel low-cost method for QTLs mapping and identify molecular markers of phenotypical changes in a small mixed population.

## Introduction

In social production, cattle are one of the most important economic animals in the world. It serves the community by supplying milk, meat, leather, and nutrition to the human body. Throughout the years of development, this domestic animal has led to different cattle breeds with a variety of phenotypes in the different geographical regions. Natural selection has led to an increase in cattle diversity. At the same time, artificial manipulation has greatly been implemented to maximize the economic benefits of cattle. Many breeds have been successfully achieved to accommodate the demands of specific traits. To improve the economic benefit of Qinchuan cattle in China, we introduced the frozen sperm of Japanese Wagyu cattle and cross with Qinchuan cattle for the potential improvement of the quality of meat.

Intramuscular fat (IMF) of beef is distributed in the form of white marbling, commonly known as marbling, and is one of the important indicators for beef quality. The richer the marbling equal to the higher the Intramuscular fat content of beef and the better quality of meat. IMF deposition and distribution are regulated by heredity, age, and epigenetics. The word “marbling” is defined by the amount and distribution of visible intramuscular fat that accumulated within the muscle and muscle fiber bundles. It is an important economic trait in all cattle and stands out particularly in the Japanese cattle<sup>1</sup>. IMF composition has become one of the most important indexes for meat quality evaluation<sup>2</sup>. It affects the meat tenderness, hydraulic power, shearing force, flavor, and juiciness<sup>3</sup>. For the improvement of the IMF in Chinese beef cattle and to better understand its formation mechanism, the research on IMF has become a hot topic in recent years.

Both Qinchuan and Luxi cattle are the local breeds in China. They are both meat cattle and have decent intramuscular fat deposits on their longus dorsi muscles, which appeared to be the shape of snowflakes. However, in comparison with Japanese wagyu cattle, Qinchuan and Luxi cattle had higher intramuscular fat diffusion and smaller transection area of the longissimus dorsi muscle. Nevertheless, there is one disadvantage of Qinchuan cattle, the growth rate is relatively slow. Japanese wagyu cattle can begin to accumulate fat at 28 months old. But the fat deposits in Qinchuan cattle are around 40 months old.

Imported beef cattle breeds, such as wagyu, have made a significant contribution to beef cattle breeding in China. As a result, Japanese wagyu cattle are now being brought in to cross-breed with local Chinese cattle in the hope of improving the quality of meat. Simmental cattle were used as a control because of the lack of intramuscular fat deposits. The research of genes regulated IMF is ongoing. Therefore, in this study, we aim to elucidate the potential single nucleotide polymorphism (SNPs) by Bulk Segregant Analysis (BSA). It required crossbreeding of Wagyu and Qinchuan cattle for genetic heterogeneity. We observed a significant improvement in meat quality in the crossbreed offspring. We found that a semi-dominant phenomenon may occur in the genome of Chinese wagyu beef cattle. We hypothesize that the majority of IMF content in these offspring should be located in between Qinchuan and purebred Wagyu cattle, and some cattle could exhibit the extreme phenotypes on both ends. Due to the limitation of GWAS, it was difficult to locate the genes associated with IMF. Therefore, we analyzed gene mapping related to IMF content by mixed pool data after individual sequencing.

Bulk Segregant Analysis (BSA) is a rapid method for identifying markers in specific regions of the genome. BSA was mainly used to construct a family based on the parents with extreme phenotypic changes. The whole genome was then sequenced on the two sample pools with extreme traits and the DNA fragments between the two mixed pools were detected to be the candidate regions. The identification of genetic markers that drive the traits is highly thought-after. In recent years, BSA has been widely adopted for DNA and RNA analysis and has been successfully used in genetic mapping and single nucleotide polymorphism mapping<sup>4</sup>. BSA has a significant contribution in understanding animal genetics<sup>5</sup>, genomics, crop breeding, and improve productivity<sup>6,7,8</sup>. BSA analysis with high-throughput sequencing is an attractive tool to most scholars. Incomplete dominance, also known as semi-dominance, is the recessive genes that can express in the heterozygotes' progeny. The study of coat color demonstrated the semi-dominant inheritance in cattle<sup>9</sup>.

In this study, a modified BSA-based method was applied for IMF QTLs mapping. Both F2 and F3 cross-population of Luxi × Qinchuan and Wagyu were used for bulk allele frequency difference analysis. Several loci were found correlated to extreme high and low IMF. 5Mb of chromosome 23 was identified as the most significant QTLs of IMF in two paired bulk (both F2 and F3) sequencing. There were six highly expressed fat genes and two highly expressed muscle genes in this region.

## Methods

### Population Samples

In total, 186 samples of Chinese wagyu beef cattle (Both F2 and F3 cross-population of Luxi × Qinchuan and Wagyu) were taken from three different farms in China (Shandong Kaiyuan Co. Ltd., Shandong Province, China; Ningxia Yijiayi Farming and Animal Husbandry Co. Ltd., Ningxia Hui Autonomous Region, China; Ningxia Xuanheyuan Agriculture and Animal Husbandry Co. Ltd., Ningxia Hui Autonomous Region, China). The background information of samples is shown in table 1.

Cattles are hybridized for four generations according to fig.1. The first generation is the cross between Qinchuan and Wagyu cattle, The offspring of the cow (F2) is used to continue the cross with Wagyu (All hybridization was performed by artificial insemination) for another two generations. The castration procedure was performed at 28 months of age before the cull. Blood and longissimus dorsi muscles were collected and intramuscular fat content was determined. The second and third generations were bred and raised in the same way as the first generationfig 1.

Subject to the limitation of cattle farms hybridization program, experimental samples from one single cattle farm are very difficult to achieve. Therefore, we selected three cattle farms that were certified for wagyu importation and sperm in-vitro fertilization. The source of wagyu beef sperm came from Beijing Dairy Cow CenterBDCC. The Simmental beef cattle used as control were obtained from Ningxia Muquan Halal Slaughterhouse in China. The animal use protocol listed below has been reviewed and approved by the Animal Ethical and Welfare CommitteeAEWC. NO. IACUC-NXU1014, and all cattles were electric shocked before slaughter, Electric shocks can make cattles painless when they are slaughtered, This test conform to the requirements of American Veterinary Medical Association (AVMA) Guidelines. Samples collected in this experiment have been given permission from respective farms in China from where of Chinese Wagyu Beef Cattle. I've provided three statements to confirm that permission was obtained from respective farms in China from where the of Chinese wagyu beef cattle. Three statements are uploaded in the Supplementary Materials and Related Files section.

**Phenotypes determination**

Phenotype determination was obtained after acid excretion and segmentation. In short, 100g of eye fillet muscle was cut out, and then intramuscular fat content was determined according to the national fat content determination standard. The classification of eye fillet meat is based on the Chinese beef grading standard, which includes the carcass and the yield quality rating. Both quality ratings were conducted after the cattle carcass was frozen and the acid was removed. The marbling of the 12th-13th rib longissimus dorsi muscles in cross-section and the top quality of the cattle and the meat and fat color was used as the reference standards. According to the amount of intramuscular fat in the cross-section, the grading of marbling is described as follows: extremely rich intermuscular fat is grade 5 and gradually reduce to grade 1 which has the least intramuscular fat. In addition to the marbling score, the samples of physiological maturity were determined by the ossification score. There is a tendency that cattle with more marbling often had less physiological maturity, which means the younger of the age and overall, the higher the beef grade score. It should be noted here that the final grade has been adjusted by the color of meat and fat and standardized across different farms.

The intramuscular fat content was determined following the guideline of Chinese national standard GB/T 9695.7-2008 in the booklet "Determination of Total Fat in Meat and Meat Products". The crude fat content was determined by Soxhlet method.

Table 1 The background information of samples

Dam	Male	Number	Name of the cattle farmer
Qinchuan	Wagyu	113	Ningxia Yijiayi Farming and Animal Husbandry Co. Ltd., Ningxia Hui Autonomous Region, China
Qinchuan×Wagyu	Wagyu	40	Ningxia Xuanheyuan Agriculture and Animal Husbandry Co. Ltd., Ningxia Hui Autonomous Region, China
Qinchuan×Wagyu×Wagyu	Wagyu	33	Shandong YuanLong Co. Ltd., Shandong Province, China

## DNA Sequencing

A total of 2 µg genomic DNA (35 ng/ml) was fragmented into 300 bp using the Bioruptor UCD-200 sonicator (Diagenode, Denville, NJ) for each sample. Library preparation was constructed using the Kapa Hyper DNA library prep kit for the Illumina platform. Fragmented DNA was end-repaired with an end-repair enzyme and a deoxyadenosine was added to the 3' ends of the fragments. Kapa barcoded DNA and Kapa indexed adapters were ligated to the sample libraries. The adapter-ligated libraries were selected for an average insert size of 300 bp using next-generation sequencing cleanup and size selection kit (NucleoMag, Macherey-Nagel, Duren, Germany) according to the manufacturer's protocols. The quality of libraries was assessed using the Bioanalyzer 4200 (Agilent Technologies, Santa Clara, CA). The libraries were then quantified by qRT-PCR and then sequenced were performed by Illumina Nova-seq platform to generate 150-bp paired-end reads.

## Sequencing Genotypes Calling

Raw sequence reads were processed using Fastq(v0.12.4)<sup>10</sup> to remove low-quality reads and adapters. BWA (v0.7.16)<sup>11</sup> mem (-M -a) module was applied to align the high-quality reads to ARS-UCD1.2 cattle genome with default parameters. SAMtools (v1.9)<sup>12</sup> were used to sort BAM files, remove read of PCR duplications, and for sample alignment statistics. "BaseRecalibrator" and "ApplyBQSR" module of GATK (v4.0.10.1) were used for base quality score recalibration. "HaplotypeCaller" module with minimum-mapping-quality 30 for each sample GVCF. After that, raw cohort VCF was worked out with modules "CombineGVCFs" and "GenotypeGVCFs". Bcftools<sup>13</sup> (v1.9) was performed as variants quality filtering with "QUAL >=100". To annotate variants, ensemble gene annotations base on ARS-UCD1.2 were used for the database which was created by snpEff (v4.3T)<sup>14</sup>. The brief workflow of variants calling and annotation was shown as blue boxes in figure 2.

## BSA-based Allele Frequency Analysis

The allele frequency of each population was extracted from the VCF file.

One of the challenges of sequencing-based methods to identify causal mutation sites is the high background noise resulting from a large number of potential SNPs, and sequencing/mapping errors. We developed a new sequence data analysis model, named as DFScore, which calculates the variance of

allele frequency in a sliding window. When the allele frequency in a window obeys the frequency expectation of two bulk samples or the variance in each bulk sample close to or equal to zero, we treat it as the best window. The benefit of this calculation is that we can more accurately screen segments that meet genetic expectations. For instance, assumed the data is from an F2 population, the allele frequency expectation of dominant bulk should be 1/3 or 2/3 and the allele frequency expectation of recessiveness bulk should be 0 or 1. Similarly, if the data were generated with homozygous samples in the two extreme phenotype bulks, the allele frequency expectation of the two bulks will be either 0 or 1. The benefit of this calculation is that we can more accurately screen segments that meet genetic expectations.

The *Power AFD* can be calculated as follows:

$$\text{PowerAFD} = |\text{FREQ of High IMF Group} - \text{FREQ of Low IMF Group}|^4$$

## Gene Expression Analysis

Gene expression and annotation analysis were obtained on the cattle gene expression database from Bgee (<https://www.bgee.org/>). In total, 112 muscle and fat tissues of cattle RNA-seq datasets were downloaded from NCBI and processed to the FASTQ format using the NCBI sratoolkit (version 2.9.6). Then low-quality reads and adapters were removed using the Fastq program (version 0.12.4),<sup>10</sup>. Kallisto (version 0.45.0)<sup>15</sup> was used to quantized gene expression. Ensembl genome bos\_taurus.ARS-UCD1.2 and gene annotation version 100 were used as references gene. Gene expression heatmap was generated by R package “pheatmap” (version 1.0.12) with gene TPM.

# Results

## SNP

Extreme high and low IMF samples (16 samples from F1, 14 samples from F2, and 13 samples from F3) were sequenced with an average of 10X coverage with reference genome per sample. Total 1075G raw sequencing data were obtained. Cohort variants calling was performed with GATK4 best-practice pipelines. SNP and InDel were worked out with reference bos\_taurus.ARS-UCD1.2 for all samples. To apply BSA-based QTL mapping, samples from the same generation population were extracted into a sub VCF. Three paired-bulk datasets (F1 high bulk vs F1 low bulk; F2 high bulk vs F2 low bulk; F3 high bulk vs F3 low bulk) were used for allele frequency analysis in the next step.

## Allele Frequency Difference (AFD) Analysis

Calculate the SNP density with the allele frequency difference between the two paired-bulks is setting up with the filter premaster as  $AF \leq 0.45$  or  $AF \geq 0.55$ . Use 1Mb as slide window length on the genome to evaluate the SNP density with a difference in frequency of each allele. The highest density AFD SNP is 43/Mb of F2 bulks and 55/Mb of F3 bulks. There are 4 chromosomes in the F2 population and 5 chromosomes in F3 have AFD-SNP density greater than 30/Mb. Among them, chromosome 7 appeared in

both groups, however, the specific segments were inconsistent. Chromosome 23 also appeared in the two populations, and the segments are very close to each other. AFD-SNP appeared in the same segment on chromosome 19 in F2 and F3. Also, F3 had a high density of SNPs in the segment of 322M in chromosome 27. (Table 2)

Table 2 High density of SNP between two groups (>30 SNP/Mb).

Group	Chromosome	Start	End
F2	7	95600000	97000000
F2	15	50600000	52400000
F2	19	55600000	57400000
F2	23	25200000	29600000
F3	4	105200000	106200000
F3	7	5000000	6600000
F3	19	55200000	56600000
F3	23	24800000	29600000
F3	27	32200000	34000000

Figure 3 showed a further analysis of the density distribution of SNPs with  $AFD \geq 0.3$ . We observed that the high-density SNPs appeared in the middle of chromosome 23 in both populations. Among them, there is significant enrichment on chromosomes 18 and 19 in the F2 population. As depicted in the heat map of the F2 population, there are 3 segments with SNP density greater than 50/M. However, in the F3 population, there is only one significant SNP-enriched segment, but the values of SNPs were exceeded 90. (Figure 3. a)

To identify the molecular marker of IMF, we performed the QTL mapping with a statistical analysis by degree of freedom to the 10 to the power of 4 (Figure 3b). We also used with R package CMPlot (Figure 4) for the heatmap and found that there were several chromosomes with a significant score ( $DF \geq 0.3$ ) such as chromosomes 5, 15, 21, 23 in the F2 population. In the F3 population, chromosomes 9 and 23 are showing the highest statistical significance and DF score (Figure 3. b, table 3). Overall, there were 15 SNPs of statistical significance in the F3 ( $DF \geq 0.3$ ). Two alleles on chromosomes 21 and 23 are adjacent to each other.

Table 3 Loci and statistical power ( $DF \geq 0.3$ )

Population	Chrome	Pos	Power4(DF)
F2	5	58695696	0.4358063
	7	40433553	0.3164062
	15	81148712	0.4822531
	21	34817429	0.4822531
F3	2	19661262	0.3811172
	2	65195647	0.3426863
	2	121333197	0.3659503
	3	31979076	0.3099656
	5	29006626	0.4096
	9	34246206	0.4978714
	11	46714219	0.342135
	13	42035405	0.3048815
	18	63302729	0.3226693
	19	32997278	0.3016062
	19	41594141	0.3458889
	21	34817458	0.4822531
	21	34817459	0.4822531
	23	24479417	0.3465916
	23	25583794	0.3659503
	27	33107854	0.3064806

## Target Region Gene Expression

As chromosome 23 consistently appeared in two generations, we investigated this target region with further analysis. There were 212 genes on this target region from 24.8M~29.6M. Among them, there were 177 genes associated with muscle or fat tissue (At least one of 112 data based showed  $\log_2(\text{TPM}) > 1$ ). The heatmap in figure 4 depicted those 46 genes that were identified with AFD SNP between two phenotypes. Six genes (ENSBTAG00000001476, ENSBTAG00000048364, ENSBTAG00000051047, ENSBTAG00000053664, ENSBTAG00000048304, ENSBTAG00000053433) were found highly expressed in fat and two genes (ENSBTAG00000013919 and ENSBTAG00000037605) were found highly expressed in muscle ( $\max \log_2(\text{TMP}) > 8$ ) (figure 4). There were 35 out of 112 genes in the region that were not associated with muscle and fat tissues ( $\max \log_2(\text{TPM}) < 1$ ).



We further investigated the conservation of muscle and fat genes in the candidate region. We annotated 46 genes with a database from Bgee (<https://www.bgee.org>) and found 33 genes with variants matched with records. 5 genes were no data in this database. Most importantly, 19 genes were found conserved in skeletal muscle tissue, 14 genes in gluteal muscle, 13 genes in scapular muscle, 11 genes in intercostal muscle, and 10 genes in longissimus thoracis muscle (Table 4).

Table 4 33 gene expression annotated in the database (<https://www.bgee.org>).

Anatomical entities	Conservation score	Max expression score	Genes with the presence of expression	Genes with the absence of expression	Genes with no data
colon	0.71	96.97	24	4	5
endometrium	0.64	93.12	23	5	5
spleen	0.43	99.13	20	8	5
lung	0.36	98.51	19	9	5
adult mammalian kidney	0.36	91.53	19	9	5
skeletal muscle tissue	0.36	80.81	19	9	5
liver	0.29	97.5	18	10	5
conceptus	0.21	52.43	17	11	5
heart	0.14	89.2	16	12	5
testis	0.14	80.58	16	12	5
placenta	0.07	89.66	15	13	5
brain	0.07	74.83	15	13	5
gluteal muscle	0	94.91	14	14	5
scapular muscle	-0.07	94.31	13	15	5
prefrontal cortex	-0.14	82.5	12	16	5
intercostal muscle	-0.21	93.69	11	17	5
longissimus thoracis muscle	-0.29	91.75	10	18	5
spermatid	-0.29	86.54	10	18	5
spermatocyte	-0.29	80.71	10	18	5

## Discussion

Traditional SNP GWAS provides limited genetic information. To increase the sensitivity of region/segments with SNP linkage in our F2/F3 populations, we used the haplotypes which contained 10 adjacent SNPs as a window, and calculated individual SNP frequency based on the sliding window. In general, chromosome recombination and exchange occur randomly and independently and can be very different in each individual. The sliding windows that were significant associated with the population may be related to our target traits. During the haplotypes analysis, we found more segments were associated with target traits in the mixed pools of the two groups. However, only a few segments were overlapped in two groups. To rationale our finding, it may be related to small sample cohorts and some background noise. Nevertheless, our finding showed the chromosomes 7, 10, 13, 19, 21, and 23 are relatively important to the IMF (Figure 5). It was confirmatory to our finding in Figure 3a—chromosome 19, 21, 23, etc.

It has been shown that most of the F1 should inherit 50% of genes from each parent. The genotype of indel from the hybrid F1 is biased as homozygous, and such genotype cannot be used to calculate AFD in the two extremes traits.

In the classical BSA analysis, researchers usually first mix individuals with the same traits in equal amounts of DNA for next-generation sequencing and allele frequency analysis<sup>[7] [16]</sup>. By mixing the samples, it can minimize the costs of next-generation sequencing. On the contrary, we sequenced all individuals with extreme traits. The overall data of the same phenotype is later mixed to ensure the consistency and coverage of sequencing. Our data suggested that sequencing individual samples may be a better option to ensure data consistency and sequencing depth.

## Conclusions

This study provided a BSA-based QTL mapping method to analyze the genetic marker of IMF in the different generation (F2/F3) cross populations between Luxi—Qinchuan and Wagyu. Following Mendelian's law of inheritance, the allele frequencies of the hybrid populations were also being investigated, and those AFD may be associated with the phenotypic trait of IMF which had a potentially high economical value of cattle. Multiple potential QTL loci were found. For better association analysis and marker validation, it is recommended that more sequence of extreme traits is needed to obtain better resolution.

## Declarations

**Author Contributions:** Conceptualization, Yun Zhu and Liyun Han.; methodology, Yun Zhu—Liyun Han and Peng Li.; software, Yun Zhu—Liyun Han and Peng Li; validation, Yun Zhu and Liyun Han.; formal analysis, Yun Zhu—Liyun Han and Peng Li; investigation, Yun Zhu and Peng Li. data curation, Yun Zhu.; writing—original draft preparation, Yun Zhu. writing—review and editing, Yun Zhu. supervision, Xiaolong Kang ,Xingang Dan,Yun Ma and Yuangang Shi; project administration, Xiaolong Kang ,Xingang Dan,Yun Ma

and Yuangang Shi.; funding acquisition, Ningxia Hui Autonomous Region key research and development program2017BY078. All authors have read and agreed to the published version of the manuscript

**Acknowledgments** :We are grateful to Ningxia Hui Autonomous Region for its key research and development project, to Ningxia University for its continuous instruction on our study, and to our tutor for his constant guidance in finishing this paper. Research supported by Ningxia Hui Autonomous Region key research and development program(2017BY078)

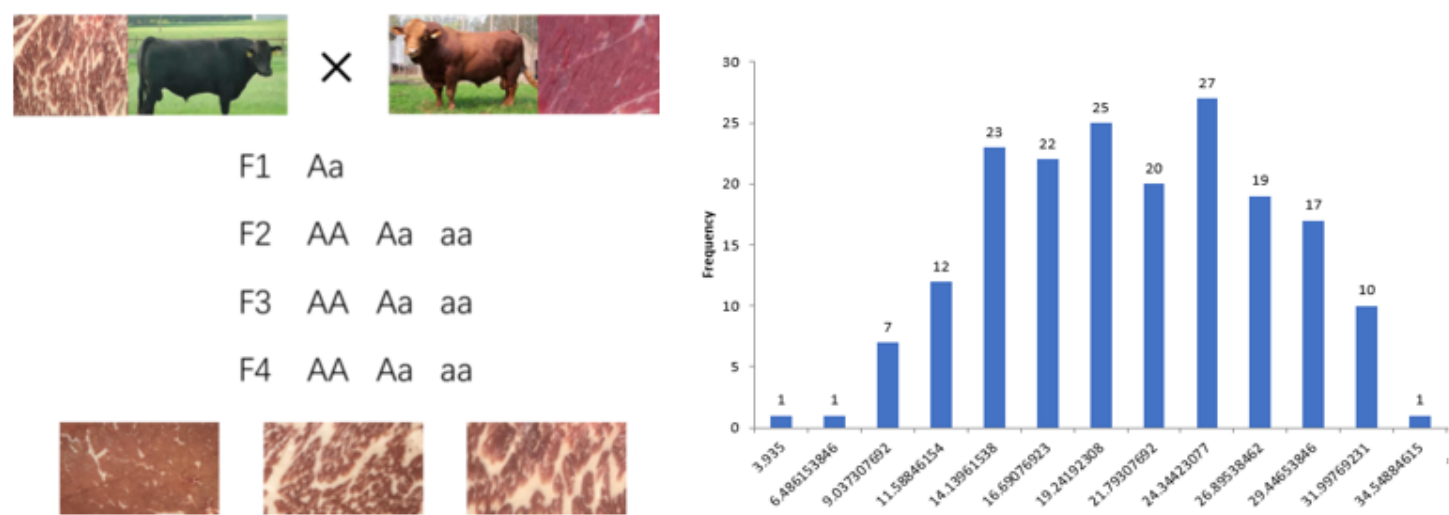
**Conflicts of Interest:** The authors declared that they have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

## References

1. Yamada, T. *et al.* Association of a single nucleotide polymorphism in titin gene with marbling in Japanese Black beef cattle. *BMC Res. Notes.* **2**, 1–6 (2009).
2. Fernandez, X., Monin, G., Talmant, A., Mourot, J. & Lebret, B. Influence of intramuscular fat content on the quality of pig meat – 2. Consumer acceptability of m. longissimus lumborum. *Meat Sci.* **53**, 67–72 (1999).
3. Gao, S. Z. & Zhao, S. M. Physiology, affecting factors and strategies for control of pig meat intramuscular fat. *Recent Pat. Food. Nutr. Agric.* **1**, 59–74 (2009).
4. Wu, S., Qiu, J., Gao, Q. & QTL-BSA: A Bulked Segregant Analysis and Visualization Pipeline for QTL-sEq. *Interdiscip. Sci. Comput. Life Sci.* **11**, 730–737 (2019).
5. Camus, M. F., Piper, M. D. W. & Reuter, M. Sex-specific transcriptomic responses to changes in the nutritional environment. *Elife.* **8**, e47262 (2019).
6. Chen, X., Hedley, P. E., Morris, J., Waugh, N. R. & E, H. L. R. & Combining genetical genomics and bulked segregant analysis-based differential expression: An approach to gene localization. *Theor. Appl. Genet.* **122**, 1375–1383 (2011).
7. Gorski, B. *et al.* MMAPP: Mutation Mapping Analysis Pipeline for Pooled RNA-seq. *Genome Res.* **23**, 687–697 (2013).
8. Sun, J. *et al.* Identification of a cold-tolerant locus in rice (*Oryza sativa* L.) using bulked segregant analysis with a next-generation sequencing strategy. *Rice.* **11**, 1–12 (2018).
9. Schmutz, S. M. & Dreger, D. L. Interaction of MC1R and PMEL alleles on solid coat colors in Highland cattle. *Anim. Genet.* **44**, 9–13 (2013).
10. Chen, S., Zhou, Y., Chen, Y., Gu, J. & Fastp An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* **34**, i884–i890 (2018).
11. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics.* **26**, 589–595 (2010).

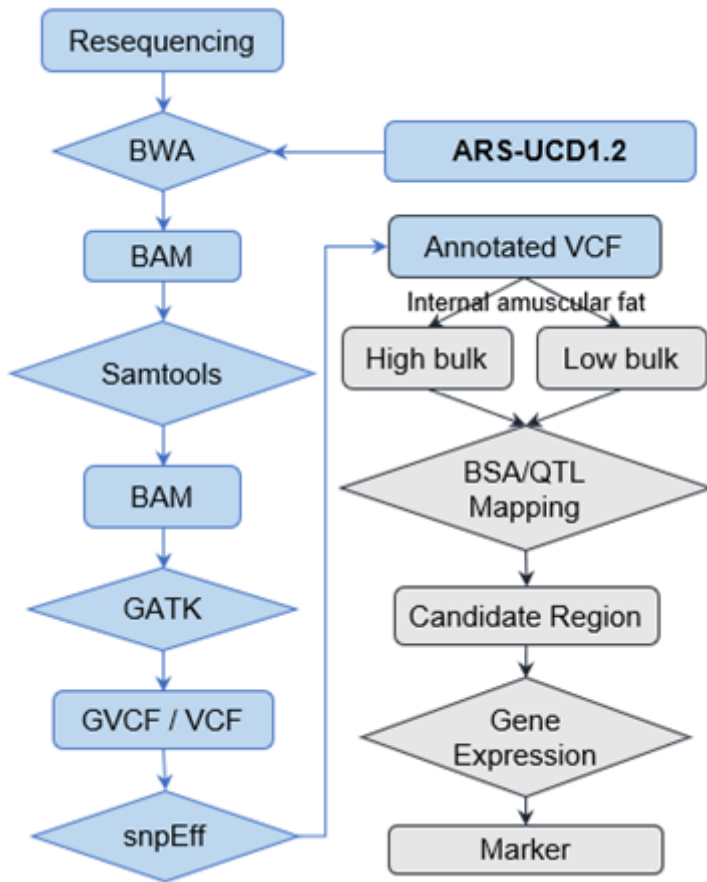
12. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. **25**, 2078–2079 (2009).
13. Narasimhan, V. *et al.* BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics*. **32**, 1749–1751 (2016).
14. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. **6**, 80–92 (2012).
15. Weijers, S. R. *et al.* KALLISTO: cost effective and integrated optimization of the urban wastewater system Eindhoven. *Water Pract. Technol.***7**, (2012).
16. Nettleton, D., Tang, H. M., Schnable, P. S., Liu, S. & Yeh, C. T. Gene Mapping via Bulk Segregant RNA-Seq (BSR-Seq). *PLoS One*. **7**, e36406 (2012).

## Figures



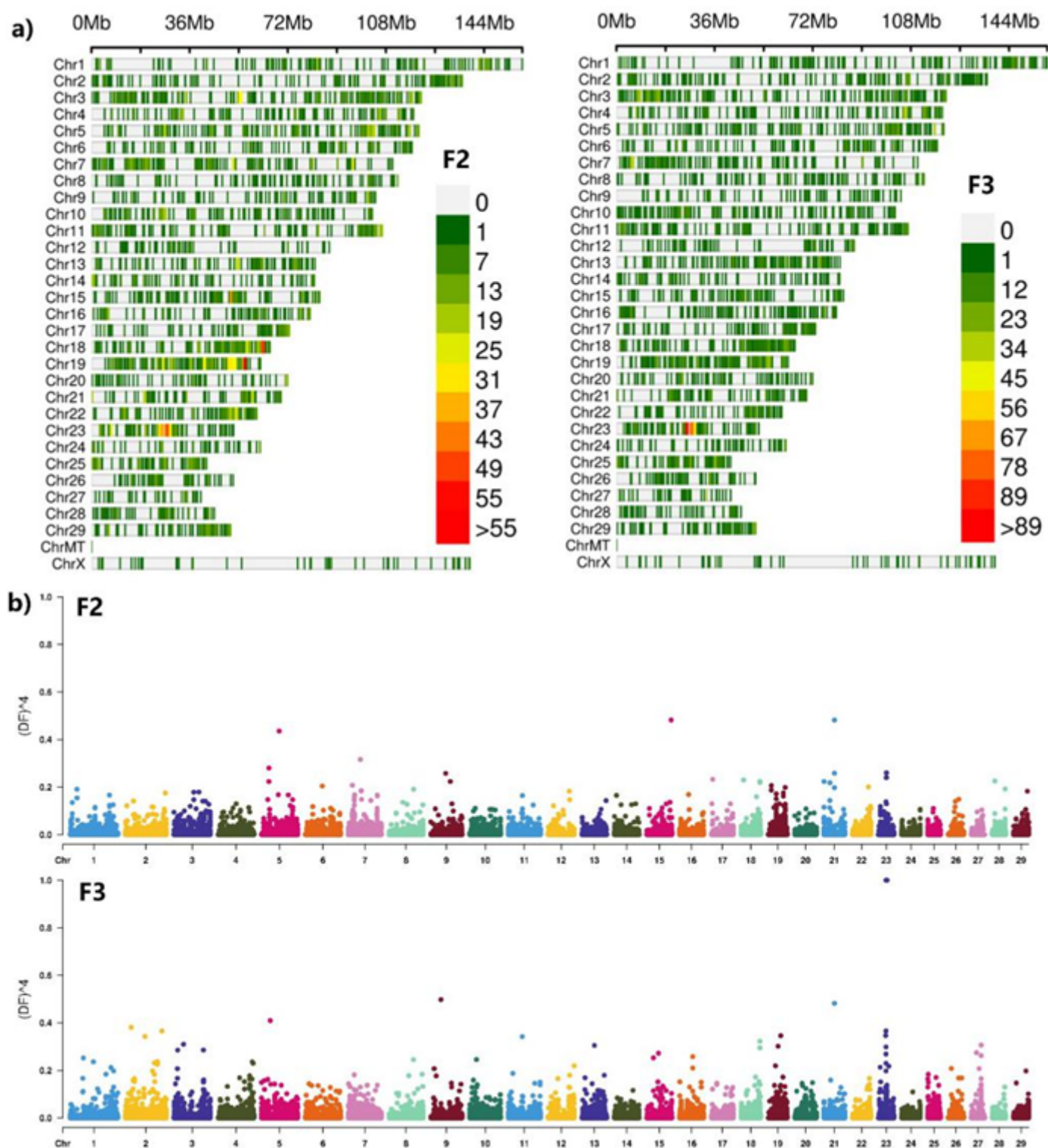
**Figure 1**

Population distribution of samples (left) and each sample frequency distribution of intramuscular fat (right).



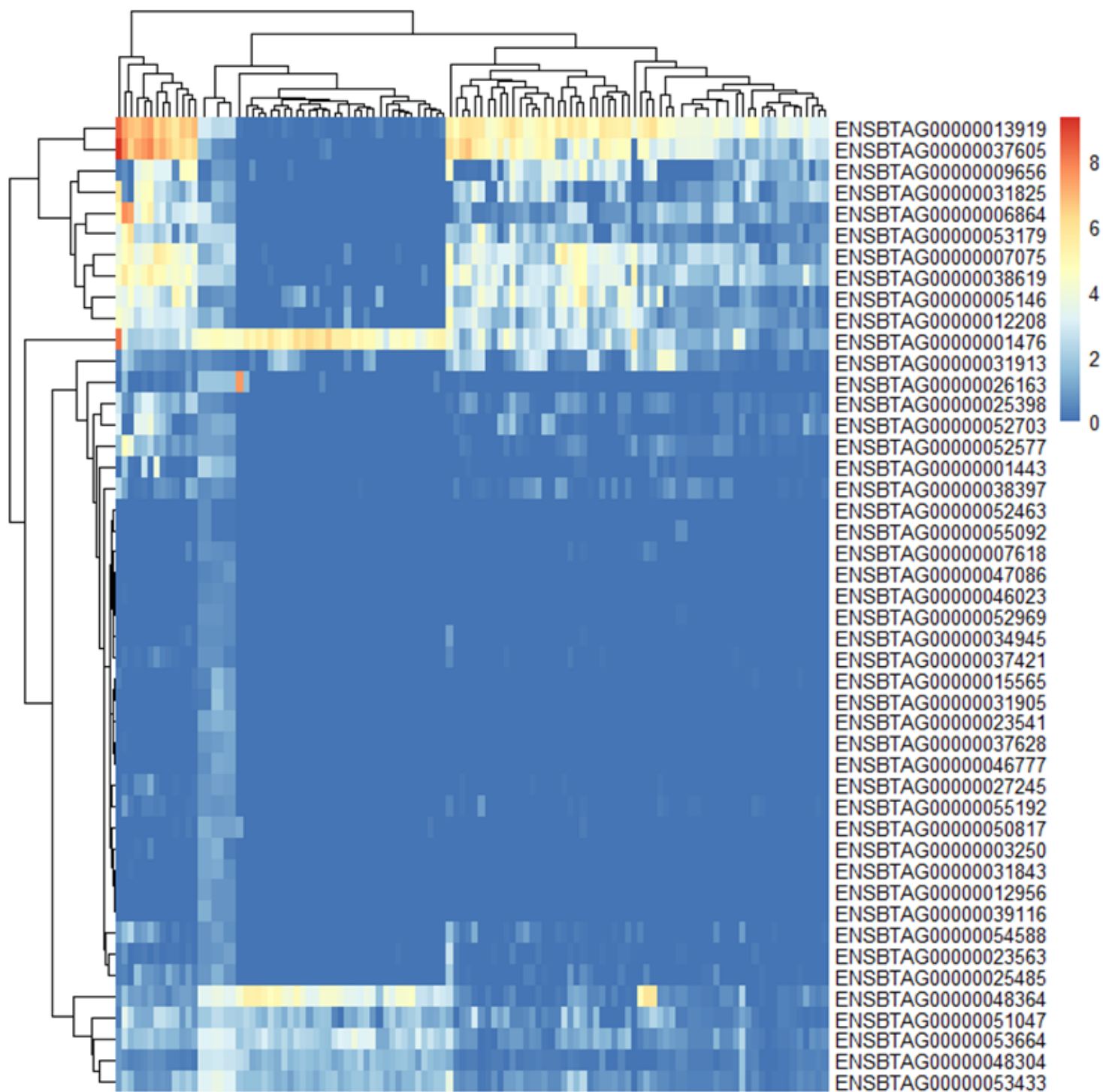
**Figure 2**

Workflow of variants calling (blue boxes) and QTL mapping (grey boxes).



**Figure 3**

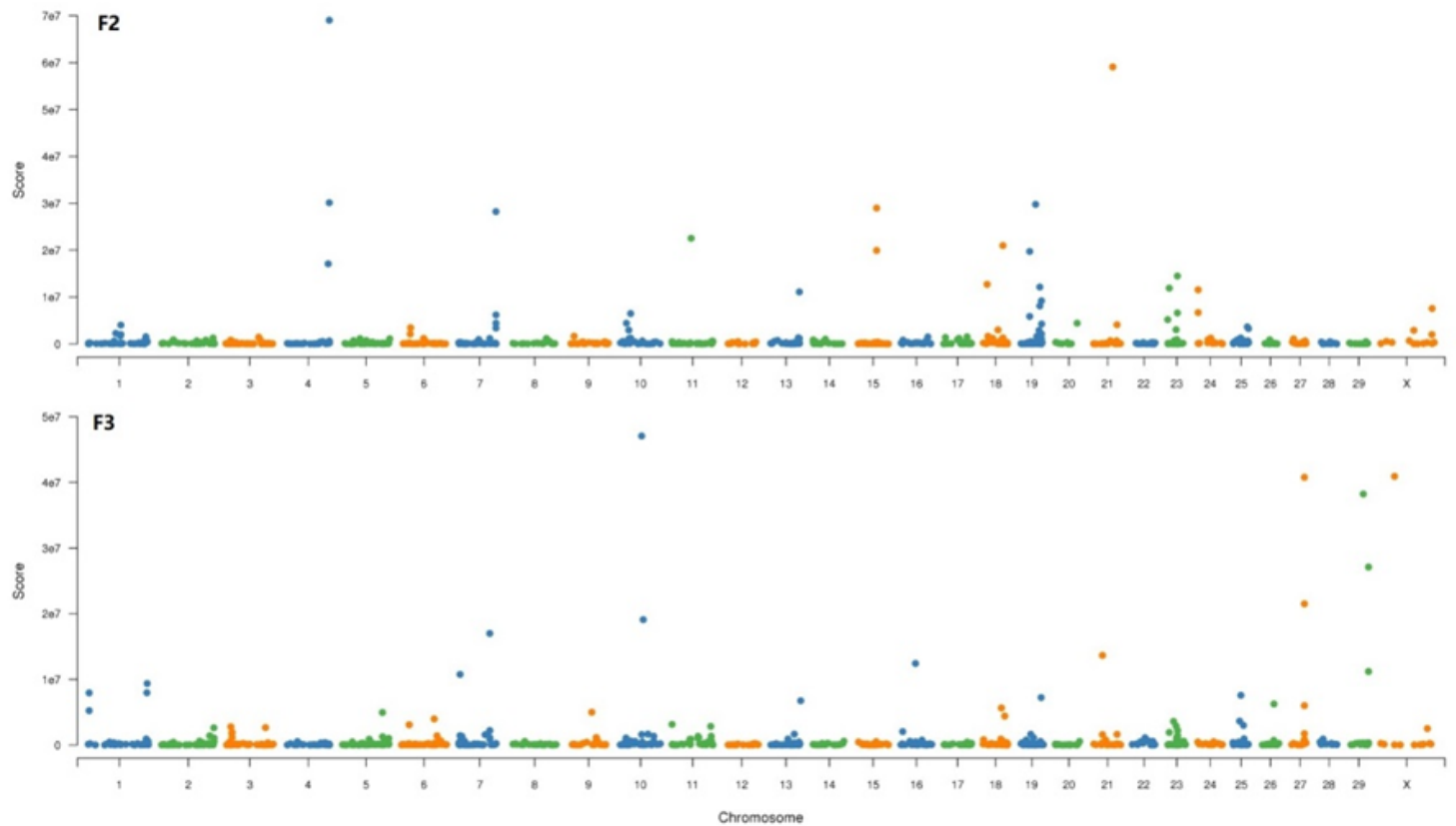
a) SNP density plot (Filter SNP with allelic frequency difference  $\geq 0.3$  in two groups). Figures from left to right were the population F2 and F3; b) QTL-mapping of allele frequency difference of F2, and F3 population with the degree of freedom.



**Figure 4**

Expression pattern of 46 gene in chromosome 23: 24.8M~29.6M.





**Figure 5**

Variance analysis of allele frequency in 10 SNP slide windows.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial.xlsx](#)