Bulk segregant RNA-seq reveals expression and positional candidate genes and allele-specific expression for disease resistance against enteric septicemia of catfish

by

Ruijia Wang

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
December 14, 2013

Key word: Bulk segregant analysis (BSA); Bulk segregant RNA-seq (BSR-Seq); Bulk frequency ratio (BFR); Disease resistance; catfish; Allele-specific expression

Approved by

Zhanjiang Liu, Chair, Alumni Professor, School of Fisheries
David Rouse, Alumni Professor, School of Fisheries,
Nannan Liu, Professor, Department of Entomology and Plant Pathology
Jeffery Terhune, Associate Professor, School of Fisheries

Abstract

The application of RNA-seq has accelerated gene expression profiling and identification of gene-associated SNPs in many species. However, the integrated studies of gene expression along with SNP mapping have been lacking. Coupling of RNA-seq with bulked segregant analysis (BSA) should allow correlation of expression patterns and associated SNPs with the correlated phenotypes.

In this study, we demonstrated the use of bulked segregant RNA-seq (BSR-Seq) for the analysis of differentially expressed genes and associated SNPs with disease resistance against enteric septicemia of catfish (ESC). A total of 1,255 differentially expressed genes were found between resistant and susceptible fish. In addition, 56,419 SNPs were identified as significant SNPs between susceptible and resistant fish located on 4,304 unique genes. Mapping of the significant SNPs, along with analysis of differentially expressed genes, allowed identification of candidate genes underlining disease resistance against ESC disease.

This study demonstrated the use of BSR-Seq for the identification of genes involved in disease resistance against ESC through expression profiling and mapping of significantly associated SNPs. BSR-Seq is applicable to analysis of genes underlining various performance and production traits without significant investment in the development of large genotyping platforms such as SNP arrays.

Acknowledgments

The author is very grateful to Dr. Zhanjiang Liu for his constant guidance, support and patience in the past years. I have been very fortunate in having Dr. Zhanjiang Liu as my supervisors. His scientific attitude and philosophy have had a profound influence on me. I am also is grateful for time and expertise offered to me by my committee members: Dr. Nannan Liu, Dr. David Rouse and Dr. Jeffery Terhune, and the dissertation outside reader, Dr. Bernhard Kaltenboeck. I would also like to express my gratitude to Huseyin Kucuktas and Ludmilla Kaltenboeck for technical help, and the colleagues in the laboratory for their help, collaboration, and friendship. I am grateful to my parents and my beloved wife for their endless love and support.

Table of Contents

Abstract ii

Acknowledgments iii

List of Tables ii

List of Illustrations iii

List of Abbreviations iv

I. Introduction and literature review 1

- I.1 ESC disease in channel catfish 1
- I.2 SNPs studies in catfish and other species 2
- I.3 Deep sequencing of transcriptome response to bacterium infection 3
- I.4 BSA and BSR-seq 4

II Materials and methods 7

- II.1 ESC bacterial challenge 7
- II.2 Sampling and RNA isolation 8
- II.3 Illumina sequencing 8
- II.4 *De novo* assembly 9
- II.5 Gene identification and annotation 9
- II.6 Transcript-level gene expression analysis of different groups 10
- II.7 Sequencing mapping and significant SNP identification 10
- II.8 Analysis of bulk frequency ratios and allele specific expression 11

II.9 Genomic location of ESC resistance-related genes 12
III Results 12
III.1 Sequence assembly and analysis 12
III.2 Differentially expressed genes after infection 14
III.3 Comparison of gene expression in resistant fish and susceptible fish after infection 15
III.4 Identification of SNPs and significant SNPs 51
III.5 Bulk frequency ratios 52
III.6 Genes with large BFR caused by genetic segregation 63
III.7 Genes with large BFR caused by allele-specific expression 67
III.8 Location of genes with high bulk frequency ratio (BFR) 70
III.9 Parental origin of highly expressed alleles 76
IV Discussion 79
Involves in polyamine homeostasis 86
V Conclusions 89
Reference 90

List of Tables

Table 1 20

Table 2 21

Table 3 23

Table 4 24

Table 5 56

Table 6 59

Table 7 60

Table 8 73

Table 9 74

Table 10 92

List of Figures

Figure 1 58

Figure 2 71

Figure 3 72

Figure 4 79

Figure 5 81

Figure 6 82

Figure 7 83

Figure 8 85

Figure 9 87

Figure 10 90

List of Abbreviations

BSA	Bulk segregant analysis
BSR-Seq	Bulk segregant RNA-seq
BFR	Bulk frequency ratio
ESC	Enteric septicemia of catfish
ASE	Allele specific expression
RPKM	Reads Per Kilobase of exon model per Million mapped reads
FDR	False Discovery Rate

I. Introduction and literature review

I.1 ESC disease in channel catfish

Catfish is the major aquaculture species in the United States, accounting for over 60% of all US aquaculture production. The two major cultivated catfish species are channel catfish *(Ictalurus punctatus)* which accounts for the majority of commercial aquaculture production and blue catfish *(Ictalurus furcatus)* which is closely related to the channel catfish and possesses several economically important traits. An inter-specific hybrid (channel catfish female × blue catfish male) has been available for commercial use, and displays strong heterosis and superior performance traits in disease resistance, growth rate, feed conversion efficiency and processing yields [1]. The immune system of catfish is perhaps the best characterized among teleost fish [2-8]. In recent years, efforts have been devoted to characterize the innate immune genes such as chemokines [9, 10], antimicrobial peptides [11-14], pathogen recognition receptors [15-20], lectins [21], and other innate immunity-related genes [22-26]. The decades of research on catfish immune genes make the identification and characterization of disease resistant genes become available. However, *E.ictaluri*, the causative agent of enteric septicemia in catfish (ESC), which is still the most significant disease of economic importance in the catfish industry, is responsible for US$60 million in annual losses to the cultured channel catfish industry in the US [27]. To identify ESC resistant genes in catfish, traditional positional cloning projects are often undertaken, although they are low-throughput and extend for many years. Many of the required resources have been developed in catfish, including a large number of polymorphic markers [28, 29], linkage maps [30, 31], bacterial artificial chromosome (BAC) libraries [32], physical maps [33], BAC end sequences (BES) [34] and the expressed sequence tag (EST) database [35].

However, even if the abundant resources are available, mapping of the genetic determinants of phenotypic variation is still a key and challenged step in the characterization of disease resistant gene and QTLs.

I.2 SNPs studies in catfish and other species

As the well knonwn molecule maker, single nucleotide polymorphisms (SNPs) are alternative bases at any given position of DNA. They are among the most abundant type of genetic variations and widely distributed within genomes. It is well known that SNPs can have four alleles in the population, but they most often exist as bi-allelic markers. Due to their potential for high genotyping efficiency, automation, data quality, genome-wide coverage and analytical simplicity [36], SNPs have rapidly become the marker of choice for many applications in genetics and genomics. As one of the biomarkers, SNPs are most suitable for whole genome association studies because linkage disequilibrium can be detected with high density SNP coverage of the genome when working with disease resistant traits and other production traits. By now, thousands of SNPs have enabled genome-wide association studies for even more complex traits in chicken [37], pig [38, 39] cattle [40-42] horse [43] and sheep [44, 45].

For model species which the whole genome has been sequenced, SNPs have been identified from genome sequencing efforts. In most cases, SNPs were identified by sequence variations between the two alleles of a single diploid individual whose genome was sequenced [46]. For nonmodel species, the identification of SNPs has been fuelled by mining large numbers of expressed sequence tags (ESTs) available in many species. Likewise, gene-associated SNPs derived from ESTs have been identified in several fish species, including Atlantic salmon [47], Atlantic cod [48] and catfish [49-51]. However, SNP identification from ESTs is limited by sequence coverage and depth. In channel catfish, 303,000 putative SNPs identified from EST database, but, only 48,594 were identified from contigs containing at least four ESTs and at least two sequences bearing the minor allele. [17]. In addition, putative SNPs identified from such contigs would have the minor alleles represented by only one sequence. Such SNPs could represent sequence errors and therefore, are not reliable [16].

Recently, more and more high-throughput expressed-sequence reads are used to increase coverage and depth and ensure sequence accuracy to identify larger numbers of gene-associated SNPs. Different next generation sequencing technologies such as Roche/454, Illumina/Solexa, and ABI/SOLiD sequencing platforms are particularly adapted to producing high coverage of expressed sequences within contigs [52]. Analysis in transcriptome level using next generation sequencing with multiple individuals has been demonstrated to be very effective for SNP identification [53]. For example, 454 sequencing was applied for the identification of gene-derived SNPs in a number of species such as eucalyptus grandis [19], pine tree [54], butterfly [55], lake sturgeon [56] and coral [57]. Althrough the 454 sequencing technology has been widely used for transcriptome analysis, Illumina sequencing technology is being gradually accepted for its dramatically improved sequencing throughput and quality [58, 59]. Especially when paired-end sequencing technology was used in Illumina sequencing make it possible to assemble contigs of transcripts from short reads [60].

I.3 Deep sequencing of transcriptome response to bacterium infection

The cost-effective and ultra-high-throughput sequencing technology, RNA-seq, is a revolutionary advance in genome-scale sequencing [61]. This transcriptome analysis method is fast and simple because it does not require bacterial cloning of the cDNAs. Direct sequencing of these cDNAs can generate short reads at an extraordinary depth. Following sequencing, the resulting reads can be assembled into a genome-scale transcription profile. It is a more comprehensive and efficient way to measure transcriptome composition, obtain RNA expression patterns, and discover new exons and genes [62, 63]. Recent transcriptomic studies on yeast [64], *Arabidopsis thaliana* [65], mouse, and human cells have demonstrated that this approach is well-suited for surveying the complexity of transcription in eukaryotes [66]. Since RNA-seq is not limited to detecting transcripts that correspond to existing genomic sequences, it is particularly attractive for non-model organisms with genomic sequences that are yet to be determined [67].

I.4 BSA and BSR-seq

Performance is related to the subtle variation in gene expression and this relationship differs among individuals [68]. In well-defined families, the first level of variation comes from genetic segregation and recombination of chromosomes. As a result of segregation and chromosomal recombination, each individual have different genetic makeup. Upon a given genetic background, genetic potential carried on DNA can only be realized when the genes are expressed. At the whole genome level, expression of each gene is affected by its genetic regulatory element as well as trans-acting factors including the impact of environment. A composite of genes, transcriptional regulation, post-transcriptional modification and regulation, translational regulation and post-translational modification and regulation, along with environmental impact and genotype environmental interactions eventually determines the phenotypes of individuals. When considered at the whole genome level, expression of tens of thousands of genes and combination of these genes both in the “quality” and quantity make the variation of performance traits extremely complex with huge variability. The task of modern agricultural genomics is to gain understanding of such variations and their relationship in determination of production and performance traits.

Traditionally, genetic and molecular biological studies are conducted to dissect these variables at different levels. For instance, the effect of alleles can be dissected through genetic and QTL mapping analysis [69-72]. Gene expression can be analyzed using high throughput methodologies such as microarrays and RNA-seq analysis [73-75]. Various epigenetic regulations have also been studied to understand the differences in gene expression with similar genetic background. Such analyses have been very powerful in determination of genetic and epigenetic factors affecting performance and production traits [76-78].

However, performance and production traits are often highly complex and the outcome of agricultural operations is affected by variations at all levels. Takeing disease resistance as an example, genetic background is very important because disease resistance genes allow the organism to survive the serious infections [79, 80]. In most cases, disease resistance genes have been studied through genetic linkage and QTL analysis that allow the identification of genomic regions containing disease resistance genes to be identified. Even with the most powerful molecular approaches, analysis of complex traits such as disease resistance can be extremely challenging. In 1991, Michelmore et al. developed a method called bulked segregant analysis (BSA) to study the complex disease resistance in plants [80-87]. The basic idea of BSA was that phenotypic extremes should have drastic differences in genotypes. When samples are selected at the best and worst performing tails and their genotypes are analyzed in bulk, a correlation of genotypes with phenotypes can be attained. In other words, the variation among individuals may be quite subtle and difficult to detect; however, the pooled samples (bulk) of the phenotypic extremes should pose a strong contrast in their genotypes because of their strongest phenotypic differences. BSA has been evolving along with various types of molecular markers including Restriction Fragment Length Polymorphisms (RFLPs) [80], Random Amplified Polymorphic DNAs (RAPDs) [81], Simple Sequence Repeats (SSRs, or microsatellites) [82, 83], Amplified Fragment Length Polymorphisms (AFLPs) [83, 84] and Single Nucleotide Polymorphisms (SNPs) [85]. With the development of Next-generation sequencing (NGS) technologies, BSA was first enhanced by the application of sequence-based markers such as restriction-site associated DNA (RAD) markers [88] and whole genome sequencing [89].

In recent years, the application of RNA-seq [74, 75, 90-94] has allowed rapid and comprehensive understanding of transcriptome level of variations. Coupling of BSA with RNA-seq has led to the development of bulked segregant RNA-seq (BSR-Seq), and it has been successfully applied in plants [95, 96], but not yet demonstrated in animals. Apparently, BSR-Seq possesses the advantages of both BSA and RNA-seq, with the high throughput for deep coverage of the transcriptome as well as the strong ability to detect genetic differences underlining the traits. Such a technique is best suited to organisms with high fecundity such as many species of fish.

Catfish is the major aquaculture species in the United States, accounting for over 60% of all US aquaculture production. The two major cultivated catfish species are channel catfish *(Ictalurus punctatus)* and blue catfish *(Ictalurus furcatus)*. An inter-specific hybrid (channel catfish female × blue catfish male) has been popular for aquaculture because of strong heterosis [97]. Not only the interspecific hybrid is popular for aquaculture, it is also a superior system to study disease resistance because of their strong phenotypic difference. Blue catfish is extremely resistant against ESC disease while channel catfish is relatively susceptible. Genetic linkage analysis of F2 generation of the interspecific hybrids [30, 98, 99] would allow identification of disease resistance/susceptibility genes. In this study, we take advantage of BSR-Seq for the analysis of disease resistant genes using the F2 generation backcross progenies (F1 hybrid backcrossed with the susceptible channel catfish) of the interspecific hybrids. Here we demonstrate that BSR-Seq is capable of 1) revealing differentially expressed genes; 2) revealing positional candidates containing genes related to disease resistance; and 3) revealing allele-specific expression after bacterial infection.

II Materials and methods

II.1 ESC bacterial challenge

Four families of backcross progenies which generated by F1 hybrid male backcross with channel catfish female (Marion strain) were reared at the Auburn University Fish Genetics Research Unit prior to challenge. One year fish were challenged in 500 L (400 L water) tank with control group containing 400 fish (100 per family) and treatment group containing 1200 fish (300 per family). The MS-S97-773 isolate of *E. ictaluri* bacteria was obtained from a natural outbreak and utilized in the experimental challenge. Bacteria were re-isolated from a single symptomatic fish and biochemically confirmed by appearance (small, punctate white colonies) and through biochemical assay (oxidase negative, fermentative in O/F glucose or glucose motility deeps (GMD), triple sugar iron (TSI) slant reaction K/A with no H_2S, and negative for indole production in tryptone broth). The confirmed bacteria were then cultured in Brain Heart Infusion broth (BHI) and incubated in a shaker incubator at 28℃ overnight. The concentration of the bacteria was determined using colony forming unit (CFU) per mL by plating 10 ml of 10-fold serial dilutions onto BHI agar plates. During challenge, 1000 mL bacterial culture with a concentration of 4×10^8 CFU/ml was added into the tank. Water was turned off in the tank for 2 h of immersion exposure, and then continuous water flow-through resumed for the duration of the challenge experiment. During 3-5 days after challenge, all dying fish with classical ESC clinic signs were collected as susceptible fish from two families. After two weeks of the challenge, all surviving fish were collected as resistant fish from the same two families.

II.2 Sampling and RNA isolation

Equal amounts of liver tissue was used from each of the 72 fish (12 fish/family, 24 fish/group each for resistant, susceptible, and control) used for RNA isolation. The tissue samples were ground separately to a fine powder in the presence of liquid nitrogen. Total RNA was extracted using the RNeasy Universal Tissue Kit (Qiagen, USA). The samples belong to the same groups were then diluted to the same concentration (100 ng/μl)and pooled together prior to library construction.

II.3 Illumina sequencing

Sequencing libraries were prepared with 2.14-3.25 μg of starting total RNA and processed using the Illumina TruSeq RNA Sample Preparation Kit, as dictated by the TruSeq protocol. The libraries were amplified with 15 cycles of PCR and contained TruSeq indexes within the Illumina adaptors, specifically indexes barcode 1-3 to label the three groups. The final, amplified library yields were 30 μl of double-stranded product (19.8-21.4 ng/μl) with an average length of 275 base pair (bp), indicating a concentration of 110-140 nM. After quantitation performed using KAPA Library Quant Kits (Kapa Biosystems, USA) and dilution, the libraries were clustered 3 per lane and sequenced on a HiSeq 2500 instrument with 100 bp paired end (PE) reads at the HudsonAlpha Genomic Services Lab (Huntsville, AL, USA). The image analysis, base calling and quality score calibration were processed using Illumina Pipeline Software v1.5, and FASTQ reads files containing the sequencing read, quality scores and paired reads information were exported for the following trimming and assembly process. Raw reads were processed for initial trimming by CLC Genomics Workbench (version 5.5.2; CLC bio, Aarhus, Denmark). Adaptor sequences, ambiguous nucleotides ('N' in the end of reads), and low quality sequences (quality scores < 30 or read length < 30 bp) were removed.

II.4 *De novo* assembly

As the primary algorithm used in RNA-seq assembly, the *de bruijn graph* method was utilized in this study. Trinity version 2013-02-25 was chosen in this study due to its good performance [100]. Briefly, the raw reads were assembled into the unique sequences of transcripts in Inchworm via greedy K-mer extension (k-mer 25). After mapping of reads to Inchworm contigs, Chrysalis incorporated reads into *de bruijn* graphs and the Butterfly module processed the individual graphs to generate full-length transcripts. And then CD-hit [99] and CAP3 [101] were used to remove assembly redundancy by setting global sequence identity in CD-hit to 1, the minimal overlap length and percent identity in CAP3 to 100 bp and 99%.

II.5 Gene identification and annotation

The final Trinity assembly contigs were used as queries against the NCBI non-redundant (nr) protein database and the zebrafish protein database using BlastX by setting the cut-off value (E-value, the likelihood that the matching sequence is obtained by chance) of $1e^{-5}$ and returning only the top 10 hit results of each query. The top gene identifications and names were initially assigned to each contig. "Hypothetical" or "uncharacterized" top BLAST results were replaced by more informative hits from the top ten lists when available.

II.6 Transcript-level gene expression analysis of different groups

The modified assembly of Trinity was used as the pseudo-reference against which trimmed reads were mapped for gene expression analysis. Reads per kilobase of exon model per million mapped reads (RPKM) [102] were calculated as the original expression value. The original expression values were normalized in order to ensure that samples were comparable [103]. The expression fold change was calculated based on the modified expression value between the infection and control groups. The Kal's test [104] was used to test the significance of the expression fold change. Analysis was performed using the RNA-seq module and the expression analysis module in CLC. The threshold of gene expression selection was set to: p-value <0.05, mapped reads >5, weighted proportions of fold change ≥|2|, or unless otherwise clarified for more stringent analysis.

II.7 Sequencing mapping and significant SNP identification

Sequencing mapping for SNP identification analysis was performed using CLC Genomics Workbench (version 5.5.2; CLC bio, Aarhus, Denmark). Trimmed sequence reads were aligned against the Trinity assembly contigs. Mapping of reads from each group to the reference sequence was performed with mismatch cost of 2, deletion cost of 3 and insertion cost of 3. The highest scoring matches that shared ≥ 95% similarity with the reference sequence across ≥ 90% of their length should be included in the alignment. Finally, all mapping results were converted to BAM format.

The initial SNP identification was conducted by the mpileup function of SAMtools version 0.1.18 [102]. PoPoolation2 version 1.201 [103] was used to call the genotype at each variant with the lowest criteria setting in order to get all possible real SNPs, SNP loci that has more than 2 allele variants were discarded. Two factors that are important for increasing quality of putative

SNPs were set based on the sequence error rate and total coverage: 1) minimum reads in each group≥6 , and 2) Total minor allele reads count >3. SNPs passed all the optimal factors were considered as initial SNPs.

Significant SNPs were identified between resistant catfish group and susceptible catfish group. SNPs which displayed heterozygous genotype (allele variants=2 and minor allele reads ≥2 in each group) were retrieved to test the difference level of allele frequencies between resistant catfish group and susceptible catfish group using two-tailed Fisher's Exact test [104]. The threshold was set as FDR p-value ≤0.05.

II.8 Analysis of bulk frequency ratios and allele specific expression

In order to compare the SNP allele frequencies more directly, bulk frequency ratios (BFR) were generated from the RNA-seq data between the two bulks, the resistant fish and the susceptible fish following process: the frequency of informative base (the major allele of resistant fish) was first calculated for resistant and susceptible bulks and then BFR was equal to the ratio of these frequencies between the bulks. The ratio of the two alleles was also analyzed with combined bulks of both resistance and susceptible fish to help classify the genes containing significant SNPs into several categories: 1) The genes containing significant SNPs with BFR≥4 and allele ratio≤9; 2) The genes containing significant SNPs with BFR≥4 and allele ratio≥14; 3) The genes containing significant SNPs with BFR≤4 and allele ratio≥14; 4) The genes containing significant SNPs with BFR≤4 and allele ratio≤9; 5)The genes containing significant SNPs with allele ratio from 9 to 14. Analysis of parental origin of the alleles for the genes with high allele ratio was then performed. The inter-species SNPs database of blue and channel catfish [105] and the parents' genotype of two target families on SNP chip (unpublished data) were used. The inter-species SNPs were mapped to the genes containing significant SNPs with high allele ratio (allele ratio>14) to check whether they shared the same position with the significant SNPs. The major allele origin of mapped SNPs was labeled based on the genotype information of the inter-specific SNPs.

II.9 Genomic location of ESC resistance-related genes

Genes harboring significant SNPs with BFR ≥5 and combined allele ratio ≤4 were used as query to map to the whole genome scaffold and linkage groups (unpublished data) by BLASTN with e-value of $1e^{-20}$.The mapped scaffolds were then located to the linkage groups by 2nd generation catfish linkage map [99].The linkage groups contain more than 10 genes with significant SNPs and at least one gene harboring significant SNPs with BFR ≥ 10 were identified as potential genomic regions harboring candidate genes for ESC resistance.

III Results

III.1 Sequence assembly and analysis

RNA-seq was conducted using Illumina sequencing with three pooled samples of resistant fish, susceptible fish, and control fish. Each pooled sample contained equal amount of RNA collected from 24 individuals. Each pooled sample was barcoded such that reads from each of the three samples can be traced and analyzed separately. A total of over 400 million reads were generated with 151 million reads from the resistant fish, 116 million from the susceptible fish, and 132 million from the control fish. After quality trimming, a total of 374 million reads were carried forward for analysis, with an average read length of 95.9 bp (Table 1). Raw read data are archived at the NCBI Sequence Read Archive (SRA) under Accession SRX327159.

Table 1. Summary of Illumina sequencing of the catfish liver transcriptome with extreme phenotypes after ESC infection

	Control	**Susceptible**	**Resistant**	**Total**
Number of reads	132,406,228	116,652,262	151,256,664	400,315,154
Read length (bp)	100 bp	100 bp	100 bp	100 bp
Number of reads after trimming	123,437,613	110,344,258	140,669,530	374,451,401
Percentage kept after trimming	93.2%	94.6%	93.0%	93.5%
Average read length after trimming	95.8 bp	96.0 bp	95.8 bp	95.9 bp

The reads after quality trimming were assembled *de novo* into 232,338 non-redundant contigs with a N50 contig length of 1,900 bp and an average contig length of 825 bp. Of the assembled contigs, over 51,000 had a length of over 1,000 bp. The assembled contigs were analyzed by BLAST searches to determine the gene identities of the contigs. Of the 232,338 contigs, 55,130 had hits to 22,126 unigenes, of which 15,599 were known genes and 6,527 were hypothetical genes (Table 2).

Table 2. Summary of *de novo* assembly of the catfish liver transcriptome with infection of ESC generated by Illumina sequencing and assembled with Trinity. Contigs were used as queries for BLAST searches against Zebrafish and NR protein databases. The cut-off value for gene identity was set at e-value ≤ 1e^{-5}.

Number of non-redundant contigs	232,338
Large contigs (≥1,000bp)	51,601
Length of the largest contig	18,759 bp
N50 size	1900 bp

Average length of non-redundant contigs	825 bp
% reads mapped to the final reference	91.68%
Number of contigs with hits	55,130
Unigene matches	22,126
Known gene matches	15,599
Unknown hypothetical gene matches	6,527

III.2 Differentially expressed genes after infection

Differentially expressed genes after infection in resistant fish and susceptible fish were determined by comparing their expression levels in RPKM (reads per kilobase of exon model per million mapped reads) with that of the control group. As summarized in Table 3, a total of 224 genes were differentially expressed in the resistant group as compared with the control group, of which 130 were up-regulated, and 94 genes were down-regulated. A significant fraction of the differentially expressed genes in resistant fish were highly regulated, with 42 genes (18.8%) being up- or down-regulated 10-fold or more, and 21 additional genes up- or down-regulated 5-10 fold (Table 3).

Although a relatively small number of genes were differentially expressed in resistant fish, a large number of genes were differentially expressed in susceptible fish, with a total of 1,240 genes being differentially expressed (Table 3). Of the 1,240 differentially expressed genes in the susceptible fish, 771 were up-regulated, and 469 were down regulated. Not only the number of differentially expressed genes was drastically more in susceptible fish, the number of highly regulated genes was also much greater in susceptible fish, with 233 genes being up- or down-regulated 10 folds or more, and additional 287 genes were up- or down-regulated 5-10 folds (table 3).

Table 3. Analysis of differentially expressed genes (fold change ≥2, p≤0.05) after infection with *Edwardsiella ictaluri*

	Total	**Up-regulated**	**Down-regulated**	**Fold change>5**	**Fold change>10**
Number of genes differentially expressed in resistant fish	224	130	94	63	42
Number of genes differentially expressed in susceptible fish	1,240	771	469	520	233
Regulated genes only in resistant fish	89	57	32	43	35
Regulated genes only in susceptible fish	1,093	682	411	464	200

III.3 Comparison of gene expression in resistant fish and susceptible fish after infection

Although a small number of genes were differentially expressed in resistant fish when being compared with the control fish, a large number of genes exhibited differential expression in resistant fish when being compared with the susceptible fish. A total of 1,255 genes were differentially expressed in resistant and susceptible fish, with 528 genes expressed significantly higher in resistant fish than in susceptible fish and 727 genes expressed significantly lower in resistant fish than in susceptible fish (Table 4). Of the genes expressed significantly higher in resistant fish, 4 genes were expressed over 100 times more than in susceptible fish; 19 were expressed 50-100 times more than in susceptible fish; 86 were expressed 10-50 times more than in susceptible fish; 94 were expressed 5-10 times more than in susceptible fish; and 325 were expressed 2-5 times more than in susceptible fish (Table 5, Figure 1). Of the genes expressed significantly lower in resistant fish, 2 genes were expressed over 100 times less in resistant fish than in susceptible fish; 10 were expressed 50-100 times less in resistant fish than in susceptible fish; 86 were expressed 10-50 times less in resistant fish than in susceptible fish; 159 were expressed 5-10 times less in resistant fish than in susceptible fish; and 470 were expressed 2-5 times less in resistant fish than in susceptible fish (Table 5, Figure 1).

Table 4. Differentially expressed genes between the resistant and susceptible fish.

ID	Gene name	Fold change	Normalized P-value
NP_001070921.1	cytochrome P450, family 2, subfamily X, polypeptide 9	151.80	1.21E-14
NP_001038880.1	prostaglandin-H2 D-isomerase precursor	134.30	6.13E-05
NP_571007.1	muscle creatine kinase a	129.03	0.00E+00
XP_002667408.2	PREDICTED: sodium- and chloride-dependent GABA transporter 2-like, partial	117.69	2.07E-02
NP_956290.1	calmodulin 2a	104.29	3.53E-08
NP_001073465.1	cytochrome P450, family 4, subfamily V, polypeptide 7	96.74	2.20E-03
NP_571638.1	troponin C type 2 (fast)	95.97	1.07E-09
NP_001004661.1	aquaporin 8a, tandem duplicate 1	91.25	2.66E-07
NP_001007030.1	ATPase, Ca++ transporting, fast twitch 1	83.36	4.58E-03
NP_999944.1	nuclear receptor subfamily 5, group A, member 5 isoform 1	81.30	7.56E-05
NP_899191.2	cytosolic sulfotransferase 3	73.58	1.71E-04
XP_002663536.1	PREDICTED: inositol hexakisphosphate and diphosphoinositol-pentakisphosphate kinase 2	72.04	8.59E-03
NP_001108586.1	glyceraldehyde-3-phosphate dehydrogenase	68.95	1.02E-02
NP_857636.1	troponin T3b, skeletal, fast isoform 2	68.61	9.10E-06
NP_571404.1	myosin regulatory light chain 2, atrial isoform	67.51	8.88E-16
NP_001002682.1	calsequestrin-2 precursor	62.78	1.45E-02
NP_001096096.2	slow myosin heavy chain 2	57.63	9.44E-04
NP_001001409.2	actin, alpha, cardiac muscle 1a	56.01	1.10E-09
NP_991136.1	parvalbumin 6	54.20	8.82E-05
NP_001020354.1	slow skeletal ventricular myosin alkali light chain 3	52.49	2.60E-02

NP_001002371.1	elongation factor 1-alpha 2	51.46	2.76E-02
NP_001007366.1	troponin I, skeletal, fast 2a.1	50.77	1.52E-04
NP_001073641.1	2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase	50.68	1.24E-05
XP_001923800.1	PREDICTED: titin	49.40	3.11E-02
XP_002666938.1	PREDICTED: UDP-glucuronosyltransferase 2A3-like	49.40	2.30E-03
XP_003453579.1	PREDICTED: fetuin-B-like	49.19	1.52E-06
NP_919358.2	fructose-bisphosphate aldolase A	47.34	2.88E-03
XP_689158.5	DENN/MADD domain containing 2Db	46.93	3.58E-02
BAE46429.1	ORF2-encoded protein	45.80	3.42E-03
NP_571720.2	fatty acid desaturase 2	45.54	3.66E-05
XP_001919815.3	PREDICTED: acetyl-CoA carboxylase isoform 1	45.28	3.96E-02
NP_001108032.1	synaptic vesicle membrane protein VAT-1 homolog-like	43.10	4.49E-02
NP_001093499.1	si:ch73-252g14.4	41.96	0.00E+00
NP_571180.1	tropomyosin alpha-1 chain	38.08	1.16E-03
NP_956506.1	parvalbumin isoform 1b	38.08	1.86E-08
NP_001038953.1	type IV antifreeze protein precursor	37.87	1.11E-16
NP_571666.1	actin, alpha cardiac muscle 1	37.20	4.29E-12
NP_001117196.1	FXYD domain containing ion transport regulator 5b precursor	30.67	1.94E-04
XP_003198186.1	PREDICTED: ig heavy chain V-III region HIL-like	29.84	8.13E-10
NP_001002434.1	hepatitis A virus cellular receptor 1 homolog precursor	29.33	2.15E-02
XP_003200261.1	PREDICTED: serine/threonine-protein kinase pim-2-like	28.82	2.28E-02
NP_001038401.1	carboxylesterase 3 precursor	27.93	0.00E+00
BAE46430.1	reverse transcriptase	27.53	1.69E-03
NP_001002367.1	solute carrier family 25 member 48	26.76	2.89E-02
NP_001001946.3	alcohol dehydrogenase 8a	26.17	0.00E+00
NP_001018363.1	complement C1q subcomponent subunit A precursor	25.23	0.00E+00
NP_001074161.1	UDP glucuronosyltransferase 5 family, polypeptide B4 precursor	25.21	3.45E-02
NP_853632.3	actin, cytoplasmic 2	25.10	8.44E-04
NP_001008651.1	prostaglandin reductase 1	25.08	8.26E-11
NP_001183988.1	hepatic leukemia factor b	24.70	3.67E-02
XP_002662598.2	Tubulin beta-1 chain	23.95	3.98E-02
NP_571322.1	L-lactate dehydrogenase B-A chain	23.16	0.00E+00
AAP49009.1	transposase	22.64	5.70E-04
NP_919349.1	hepatocyte nuclear factor 4-alpha	22.13	4.96E-02
NP_001029357.2	interferon alpha/beta receptor 1 precursor	22.03	4.98E-02
NP_001035384.1	rhamnose binding lectin-like precursor	21.95	1.67E-02
XP_004072944.1	PREDICTED: FK506-binding protein 15-like	21.27	1.87E-02
NP_571742.1	GTP cyclohydrolase 1	21.10	6.89E-03
NP_059333.1	COX1_15414 cytochrome c oxidase subunit I	20.79	2.74E-03

XP_002664804.1	PREDICTED: sarcolipin-like	20.18	4.91E-02
NP_001018424.1	inter-alpha (globulin) inhibitor H3	19.73	3.59E-03
NP_571173.1	apolipoprotein Eb precursor	19.72	8.16E-11
NP_694549.1	retinol-binding protein 2	19.59	0.00E+00
XP_695347.1	PREDICTED: pulmonary surfactant-associated protein D-like	19.35	4.05E-03
XP_001335383.3	PREDICTED: olfactomedin-4-like	18.97	2.00E-06
XP_001634362.1	predicted protein NADPH 1	18.27	1.29E-02
XP_001919624.3	PREDICTED: paired box protein Pax-3	17.84	2.68E-03
NP_001003482.1	complement component 1, q subcomponent, B chain precursor	17.67	5.88E-15
NP_571202.1	lipoprotein lipase precursor	17.50	1.54E-02
NP_001005397.1	emopamil-binding protein-like	16.91	1.68E-03
NP_571933.1	secreted frizzled-related protein 5 precursor	16.47	1.94E-02
XP_697618.4	PREDICTED: beta-2-glycoprotein 1	16.37	1.11E-04
XP_003454537.1	PREDICTED: centromere protein F	16.36	1.76E-10
NP_694492.1	fatty acid-binding protein 10-A, liver basic	15.85	0.00E+00
NP_956897.1	ectonucleotide pyrophosphatase/phosphodiesterase family member 2 precursor	15.44	5.85E-03
NP_059334.1	COX2_15414 cytochrome c oxidase subunit II	15.44	1.19E-02
NP_059343.1	CYTB_15414 cytochrome b	15.44	1.19E-02
NP_001002332.1	uricase	15.44	2.53E-07
NP_001093575.1	ribonuclease like 2 precursor	15.41	0.00E+00
ACO14339.1	HERV-H LTR-associating protein 2 precursor	14.84	1.67E-06
NP_001230313.1	WD repeat and SOCS box-containing protein 2-like	14.61	1.49E-02
XP_002666846.1	PREDICTED: olfactomedin-like	14.26	3.65E-12
XP_001339480.1	PREDICTED: monoacylglycerol lipase abhd6-A	14.24	2.06E-04
NP_001071272.2	ubuquitin c	14.23	4.51E-03
NP_001038846.1	phosphotriesterase-related protein	13.72	1.03E-02
NP_001025339.1	ectonucleotide pyrophosphatase/phosphodiesterase family member 1	13.64	3.69E-02
NP_001166113.1	arylacetamide deacetylase-like 4	13.14	2.72E-12
AAA16654.1	immunoglobulin light chain	12.84	0.00E+00
NP_059331.1	ND1_15414 NADH dehydrogenase subunit 1	12.61	4.68E-02
XP_003454878.1	PREDICTED: caspase-1-like	12.61	4.68E-02
NP_001038299.2	flavin monooxygenase	12.58	2.90E-03
NP_059337.1	COX3_15414 cytochrome c oxidase subunit III	12.35	4.97E-02
XP_003197888.1	PREDICTED: transposon TX1 uncharacterized 149 kDa protein-like	11.96	6.52E-03
NP_001006081.1	DNA replication complex GINS protein PSF1	11.91	0.00E+00
NP_997902.1	macrophage expressed 1 precursor	11.83	6.89E-03
NP_001186657.1	apolipoprotein M precursor	11.42	0.00E+00
ADF97628.1	14 kDa apolipoprotein	11.32	0.00E+00
NP_001136246.1	alanine aminotransferase 2-like	11.32	8.62E-03

NP_001076281.1	chemokine-like receptor 1	11.00	1.01E-03
NP_001005976.1	complement C1q subcomponent subunit C	10.98	3.61E-09
NP_997974.2	glycogen phosphorylase, brain form	10.88	1.66E-02
XP_692090.3	PREDICTED: l-serine dehydratase/L-threonine deaminase-like	10.88	1.66E-02
NP_001013279.1	fibronectin precursor	10.83	0.00E+00
NP_001103843.1	hyaluronan-binding protein 2	10.82	0.00E+00
XP_003198238.1	PREDICTED: ig lambda chain V-III region LOI-like	10.70	4.52E-02
NP_001082926.1	histidine ammonia-lyase	10.63	5.50E-04
NP_001073481.1	hedgehog-interacting protein precursor	10.50	4.79E-02
NP_001184218.1	dimethylanaline monooxygenase-like	10.40	6.92E-11
NP_001116749.1	sterol 26-hydroxylase, mitochondrial	10.38	2.35E-03
NP_955893.1	hydroxysteroid dehydrogenase-like protein 2	10.00	2.34E-02
XP_689652.2	PREDICTED: tyrosine-protein kinase STYK1	9.95	3.64E-02
XP_003198731.1	PREDICTED: protein FAM189A2-like	9.88	7.18E-03
NP_001013547.1	isoaspartyl peptidase/L-asparaginase	9.71	7.72E-04
XP_003198184.1	PREDICTED: ig heavy chain V-I region HG3-like	9.61	4.10E-02
NP_835234.2	selenoprotein H	9.61	3.84E-03
NP_996951.1	peroxisomal bifunctional enzyme	9.58	7.24E-06
NP_001002372.1	transmembrane 4 L6 family member 5	9.46	7.16E-07
NP_001007152.1	sex hormone binding globulin precursor	9.12	0.00E+00
AER39749.1	inositol-3-phosphate synthase	9.09	4.88E-02
XP_002660413.1	unnamed protein product	9.00	1.12E-02
NP_571809.1	glutathione S-transferase pi	8.98	3.37E-10
XP_003201806.1	PREDICTED: protein FAM20A-like, partial	8.92	1.72E-02
NP_001082849.1	glycerol-3-phosphate acyltransferase 2, mitochondrial	8.82	3.71E-02
NP_001103819.2	vascular endothelial growth factor A-A isoform 2 precursor	8.81	1.24E-02
NP_999897.1	D-amino acid oxidase	8.77	6.20E-04
XP_003201204.1	PREDICTED: opioid growth factor receptor	8.70	3.72E-02
NP_001132951.1	alpha-2-macroglobulin-like precursor	8.61	2.28E-08
XP_003200436.1	PREDICTED: urokinase plasminogen activator surface receptor-like	8.59	0.00E+00
ACN11475.1	Transposable element Tcb1 transposase	8.48	1.57E-03
NP_956000.1	glycerol-3-phosphate dehydrogenase 1b	8.41	1.30E-14
XP_002666918.1	PREDICTED: papilin-like	8.36	3.17E-02
NP_954684.1	collagen alpha-1(I) chain precursor	8.23	1.94E-03
ACO10049.1	Coagulation factor XI precursor	8.17	2.55E-15
NP_001003512.1	glucose-6-phosphatase a, catalytic, tandem duplicate 1	8.17	2.35E-05
NP_998135.1	microsomal triglyceride transfer protein large subunit precursor	8.14	1.34E-02
NP_001107052.1	NEDD8 ultimate buster 1	8.07	4.89E-02
XP_003201396.1	PREDICTED: ig kappa chain V-I region Scw-like, partial	8.07	0.00E+00

BAI49518.1	skin mucus lectin	8.03	5.56E-12
NP_958886.1	collagen, type I, alpha 1b precursor	8.01	5.79E-03
NP_998584.2	dual specificity mitogen-activated protein kinase kinase 1	7.98	2.67E-02
NP_001002579.1	mitochondrial glycine cleavage system H protein	7.85	1.70E-05
NP_001092242.1	folate receptor alpha precursor	7.76	2.21E-02
NP_892013.2	collagen alpha-2(I) chain	7.66	2.43E-03
XP_003198203.1	PREDICTED: si:ch211-66h3.1	7.64	2.22E-16
NP_958879.1	aldehyde dehydrogenase family 9 member A1-A	7.63	1.90E-05
XP_002156036.1	PREDICTED: similar to zinc finger, MYM domain containing 1, partial	7.61	6.06E-08
NP_997814.1	fatty aldehyde dehydrogenase	7.46	4.73E-02
NP_001007773.1	actin-related protein 2	7.34	7.27E-03
XP_003201131.1	glycerophosphodiester phosphodiesterase domain-containing protein 1	7.30	4.98E-02
NP_991309.1	regucalcin	7.20	6.24E-03
NP_001167083.1	Peptide-N4-N-acetyl-beta-D-glucosaminylasparagine amidase F precursor	7.18	3.02E-02
NP_998151.1	hypoxanthine-guanine phosphoribosyltransferase	7.04	1.42E-04
NP_001187282.1	glutathione S-transferase kappa 1	7.00	3.43E-02
NP_001005981.1	kininogen 1 precursor	6.95	0.00E+00
XP_001921577.2	PREDICTED: alpha-2-macroglobulin-like	6.90	0.00E+00
NP_001019822.1	fatty acid binding protein 1-B.1	6.89	0.00E+00
NP_571780.1	major histocompatibility complex class I UEA precursor	6.84	1.93E-03
XP_001919767.2	PREDICTED: b-cell receptor CD22-like	6.79	3.83E-02
NP_001070930.1	fatty-acid amide hydrolase 2-B	6.74	3.78E-11
AEM44666.1	T cell receptor alpha	6.70	4.56E-04
XP_003446632.1	PREDICTED: kynureninase-like	6.69	1.54E-02
NP_997776.1	very long-chain specific acyl-CoA dehydrogenase, mitochondrial	6.66	4.89E-03
XP_688983.3	PREDICTED: xanthine dehydrogenase/oxidase	6.59	4.29E-02
XP_002662923.1	galactoside-binding soluble lectin 9	6.48	2.19E-02
NP_001018510.1	guanine deaminase	6.46	2.92E-04
NP_001017821.1	regulator of G-protein signaling 16	6.44	3.59E-02
NP_001077304.1	vascular cell adhesion protein 1 precursor	6.44	3.59E-02
NP_996939.1	phospholipase A1 member A precursor	6.38	4.81E-02
NP_001007448.1	carnitine O-palmitoyltransferase 2, mitochondrial	6.37	4.22E-03
NP_001004682.1	fatty acid binding protein 11a	6.33	2.50E-05
NP_001038808.1	abhydrolase domain containing 12B	6.28	1.73E-06
XP_001344951.3	PREDICTED: supervillin	6.27	4.08E-02
XP_003197736.1	PREDICTED: sterile alpha motif domain-containing protein 9-like	6.24	1.40E-02
NP_694498.2	homogentisate 1,2-dioxygenase	6.23	0.00E+00
NP_001035452.1	elongation of very long chain fatty acids protein 2	6.10	7.48E-06
NP_956838.1	nucleolar protein of 40 kDa	6.09	0.00E+00

NP_957001.1	retinol dehydrogenase 8a	6.09	3.67E-02
NP_001082927.1	1,25-dihydroxyvitamin D(3) 24-hydroxylase, mitochondrial	6.07	5.60E-06
NP_001008639.1	very long-chain acyl-CoA synthetase	6.07	3.74E-05
NP_001002686.1	flavin reductase	6.05	4.63E-04
XP_001345541.4	PREDICTED: alpha-2-macroglobulin	6.03	0.00E+00
NP_956640.1	diablo, IAP-binding mitochondrial protein a	6.02	5.36E-12
NP_001073330.1	apolipoprotein A-IV precursor	5.97	2.98E-03
NP_001025370.1	diamine N-acetyltransferase 1	5.97	6.76E-05
NP_997787.1	alpha-2-HS-glycoprotein	5.97	0.00E+00
NP_694494.1	fibroblast growth factor receptor 1-A	5.83	4.27E-03
NP_001038735.2	finTRIM family, member 14	5.83	4.33E-02
XP_003199195.1	PREDICTED: small serum protein 2-like	5.79	0.00E+00
XP_696711.2	PREDICTED: acyl-coenzyme A thioesterase 2, mitochondrial	5.66	1.56E-02
NP_001018330.1	solute carrier family 2 (facilitated glucose transporter), member 15b	5.63	2.81E-02
NP_001187789.1	fam195a	5.62	5.53E-03
XP_003448977.1	PREDICTED: coagulation factor XI-like	5.59	3.50E-11
XP_003199171.1	PREDICTED: retrotransposable element Tf2 155 kDa protein type 1-like	5.51	3.68E-02
NP_001025328.2	catechol-O-methyltransferase a	5.51	6.92E-13
XP_002662790.2	lanosterol 14-alpha demethylase	5.42	4.80E-03
NP_001122240.1	acyl-CoA-binding domain-containing protein 7	5.34	4.74E-03
XP_001919038.3	PREDICTED: activated CDC42 kinase 1-like	5.32	9.27E-03
NP_958456.1	trans-2,3-enoyl-CoA reductase	5.18	3.14E-07
NP_998086.1	methionine-R-sulfoxide reductase B2, mitochondrial precursor	5.16	1.69E-06
XP_003445571.1	PREDICTED: molybdenum cofactor biosynthesis protein 1-like	5.15	4.84E-02
NP_956727.1	autophagy-related protein 13	5.15	6.12E-03
XP_699016.3	PREDICTED: si:dkey-5n18.1	5.11	4.18E-07
NP_001008593.1	protein NDRG2	5.07	9.57E-05
NP_571676.1	solute carrier family 3, member 2a	4.99	1.14E-03
NP_955823.1	C-1-tetrahydrofolate synthase, cytoplasmic	4.97	2.43E-03
XP_700558.3	PREDICTED: transmembrane protein 56-B	4.95	4.09E-02
XP_002667926.1	tubulin polyglutamylase complex subunit 2	4.95	2.05E-03
NP_001187139.1	granzyme-like III precursor	4.93	2.19E-02
NP_998089.1	insulin-like growth factor-binding protein 7 precursor	4.90	1.13E-02
NP_001018161.1	peroxisomal carnitine O-octanoyltransferase	4.88	1.55E-02
XP_698776.4	bcl2-associated X protein, b	4.85	3.74E-02
NP_957106.1	peptidyl-prolyl cis-trans isomerase FKBP1B	4.82	4.56E-02
NP_998058.1	dihydropyrimidine dehydrogenase	4.82	3.59E-04
XP_003198307.1	ATP-citrate synthase	4.82	2.26E-03
NP_001002636.1	ketohexokinase	4.80	4.23E-03

XP_001336460.1	PREDICTED: wu:fc03a11	4.77	1.06E-06
NP_957475.1	long-chain specific acyl-CoA dehydrogenase, mitochondrial	4.77	3.88E-05
NP_001006068.1	delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial	4.76	1.02E-02
NP_001191453.1	elongation of very long chain fatty acids protein 4-like	4.75	3.73E-06
XP_686922.3	PREDICTED: peroxisomal sarcosine oxidase	4.73	7.06E-03
AET85182.1	TC1-like transposase	4.72	4.30E-02
XP_001922687.3	PREDICTED: antihemorragic factor cHLP-B	4.68	0.00E+00
NP_001093614.2	apolipoprotein A-Ib precursor	4.64	0.00E+00
XP_003198228.1	retinol dehydrogenase 12	4.64	5.93E-03
NP_998231.1	hexokinase-2	4.64	5.93E-03
NP_001167257.1	Coagulation factor X precursor	4.63	1.92E-14
NP_999938.1	tartrate-resistant acid phosphatase type 5 precursor	4.58	3.26E-02
NP_001157278.1	im:6895556	4.56	5.21E-07
AAN41455.1	interleukin-8 variant 3	4.55	3.51E-13
NP_001073508.1	methyltransferase like 7A precursor	4.47	1.70E-02
XP_002665851.1	ubiquitin-conjugating enzyme E2 L3	4.47	4.71E-03
NP_991165.1	small integral membrane protein 14	4.46	2.65E-06
NP_001181918.1	fibrinogen alpha chain precursor	4.46	0.00E+00
NP_001082963.1	2-amino-3-carboxymuconate-6-semialdehyde decarboxylase	4.45	2.44E-04
XP_688042.3	glutathione peroxidase 4 precursor	4.44	1.91E-02
NP_001018541.1	stearoyl-CoA desaturase b	4.41	1.91E-03
NP_991242.1	cytidine deaminase	4.40	3.68E-05
NP_001007391.1	3-hydroxyanthranilate 3,4-dioxygenase	4.39	1.11E-07
NP_001002331.1	alanine-glyoxylate aminotransferase a	4.37	5.30E-12
NP_001077298.1	proprotein convertase subtilisin/kexin type 5b preproprotein	4.31	9.88E-03
NP_001077041.1	histidine triad nucleotide-binding protein 2, mitochondrial	4.30	9.35E-03
NP_001071226.1	serine proteinase inhibitor, clade A, member 1 precursor	4.29	0.00E+00
NP_998168.1	S100 calcium binding protein A10b	4.29	1.58E-07
NP_001122176.1	lipoamide kinase isozyme 4	4.22	7.07E-04
XP_003199037.1	glycerol-3-phosphate acyltransferase 3	4.22	2.00E-02
NP_001032485.1	solute carrier family 16, member 7 (monocarboxylic acid transporter 2)	4.20	2.10E-04
NP_001099059.2	serpin peptidase inhibitor-like precursor	4.18	1.74E-05
NP_631919.1	lysozyme-like protein 2 precursor	4.17	0.00E+00
NP_694499.1	DNA-binding protein inhibitor ID-3	4.16	3.24E-03
AFC88124.1	immunoglobulin light chain variable region	4.16	5.25E-14
NP_001001949.2	granulin-a precursor	4.16	1.57E-03
NP_956892.1	interferon regulatory factor 6	4.12	3.33E-02
NP_998559.1	glutathione S-transferase, alpha-like	4.11	1.11E-16
NP_001018598.1	centrosomal protein of 55 kDa	4.10	0.00E+00

NP_001017806.1	single-stranded DNA-binding protein, mitochondrial	4.06	9.06E-03
NP_001007446.1	phytanoyl-CoA dioxygenase domain-containing protein 1	4.04	2.67E-09
NP_571924.2	alcohol dehydrogenase class-3	4.04	8.56E-03
NP_571185.1	carbonic anhydrase	4.00	3.11E-02
NP_998466.2	aldehyde dehydrogenase 2.2	3.99	3.16E-10
NP_001003981.2	pleiotrophin precursor	3.98	4.85E-02
NP_998223.1	78 kDa glucose-regulated protein precursor	3.97	1.11E-16
NP_001005938.1	protein DJ-1	3.97	5.23E-07
NP_840073.2	methionine-R-sulfoxide reductase B1-A	3.95	6.29E-04
NP_001070796.1	3-keto-steroid reductase	3.94	2.91E-02
NP_001070141.1	transmembrane protein 229b	3.93	3.08E-02
NP_955910.1	beta-ureidopropionase	3.92	2.77E-03
NP_919360.1	carbonyl reductase NADP(+)	3.90	2.60E-02
NP_001107204.1	leukemia inhibitory factor receptor alpha b precursor	3.90	3.05E-02
NP_001018539.2	carboxypeptidase B2 precursor	3.90	1.59E-08
NP_001012249.1	peptidyl-prolyl cis-trans isomerase FKBP11 precursor	3.87	7.88E-04
NP_956878.1	glutathione S-transferase theta 1b	3.86	6.97E-10
NP_001104707.1	sorting nexin-15	3.86	2.17E-02
NP_571508.1	insulin-like growth factor 2a precursor	3.83	1.06E-04
NP_991206.1	purine nucleoside phosphorylase 4b	3.83	1.48E-02
NP_955904.2	dnaJ homolog subfamily C member 3	3.81	1.16E-04
NP_001073502.1	nonhistone chromosomal protein HMG-14A-like	3.76	3.86E-04
NP_001025442.1	maleylacetoacetate isomerase isoform 1	3.75	1.93E-11
NP_001093483.1	complement component 3-like	3.75	0.00E+00
XP_003449657.1	epithelial membrane protein 2	3.75	1.61E-02
NP_001070105.1	phosphoethanolamine methyltransferase	3.75	3.38E-09
XP_002665648.1	PREDICTED: somatostatin receptor type 2-like	3.74	1.26E-03
XP_691265.5	diphosphomevalonate decarboxylase	3.72	3.85E-02
NP_001103993.1	preimplantation protein 4-like	3.72	1.00E-02
NP_001165890.1	sorbitol dehydrogenase	3.71	9.08E-04
NP_001121716.1	ras-associated and pleckstrin homology domains-containing protein 1	3.68	0.00E+00
NP_998327.1	alanine-glyoxylate aminotransferase b	3.67	1.78E-02
XP_700128.4	PREDICTED: zinc transporter ZIP9-B-like	3.64	3.45E-02
NP_998094.1	protein ETHE1, mitochondrial	3.63	1.52E-02
NP_001006088.1	N-acylsphingosine amidohydrolase 1 precursor	3.62	4.58E-02
NP_997895.1	sideroflexin-2	3.60	1.84E-02
NP_001104551.1	growth hormone receptor b precursor	3.58	4.05E-02
NP_001245154.1	interferon responsive immune	3.57	1.01E-03
XP_002160050.1	PREDICTED: similar to transposase-like	3.56	1.96E-03

NP_919392.2	receptor-interacting serine/threonine-protein kinase 2	3.55	4.86E-02
NP_001071049.1	platelet receptor Gi24 precursor	3.52	7.81E-06
NP_001007361.1	cholesteryl ester transfer protein precursor	3.52	6.25E-05
NP_001119936.1	insulin-like growth factor-binding protein 2-B precursor	3.51	2.16E-06
NP_878275.2	TGF-beta receptor type-2 precursor	3.50	2.66E-02
XP_001921185.1	MHC class I antigen	3.49	9.65E-03
NP_001258237.1	acetyl-Coenzyme A carboxylase alpha	3.49	8.15E-03
XP_701334.3	PREDICTED: ornithine aminotransferase, mitochondrial-like isoform 2	3.49	7.95E-07
NP_001017824.1	cytokeratin-like	3.47	1.97E-02
NP_001243761.1	solute carrier family 16 (monocarboxylic acid transporters), member 6a	3.46	4.40E-02
AAA82596.1	immunoglobulin light chain F class	3.45	1.22E-04
NP_998398.1	cyclic AMP-dependent transcription factor ATF-4	3.43	2.32E-10
NP_956661.1	abhydrolase domain-containing protein 14B	3.41	7.83E-03
XP_003197710.1	nuclear factor, interleukin 3 regulated, member 6	3.40	3.70E-02
XP_001332864.2	PREDICTED: vitelline membrane outer layer protein 1 homolog	3.39	0.00E+00
NP_001093446.1	crystallin, zeta (quinone reductase)	3.39	7.73E-03
XP_693392.5	alpha-1,3-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase C	3.37	8.11E-03
NP_001096599.1	cystatin B	3.36	1.84E-02
NP_001028900.1	protein kinase C and casein kinase substrate in neurons protein 1	3.34	4.49E-02
NP_001243535.1	vesicle-associated membrane protein 5-like isoform 1	3.34	4.05E-02
XP_001340422.4	PREDICTED: transmembrane protease serine 3	3.33	5.97E-09
NP_001018407.1	cystinosin precursor	3.33	8.39E-03
NP_998415.1	casein kinase I isoform delta-B	3.33	3.30E-02
NP_956239.1	FK506 binding protein 1A, 12kDa	3.31	3.13E-02
NP_998488.1	phosphatidylethanolamine binding protein	3.31	4.71E-03
NP_956430.1	peroxisomal multifunctional enzyme type 2	3.29	1.41E-04
NP_001122255.1	neural proliferation, differentiation and control, 1 precursor	3.29	2.52E-03
XP_698073.1	PREDICTED: tumor suppressor candidate 5 homolog	3.27	4.79E-13
NP_878283.1	antithrombin-III	3.27	8.33E-15
NP_001003515.1	hydroxyacyl-coenzyme A dehydrogenase, mitochondrial	3.26	6.99E-05
NP_919369.2	tubulin alpha-1C chain	3.26	7.09E-03
NP_001013509.1	hexosaminidase D	3.24	5.30E-03
NP_001098598.2	selenoprotein X, 1b	3.24	8.98E-03
NP_955826.1	elongation of very long chain fatty acids protein 6	3.24	1.51E-06
NP_932328.1	transcription factor PU.1	3.24	4.45E-03
NP_571770.1	MHC class II integral membrane protein alpha chain 3 precursor	3.24	4.15E-12
NP_001007393.1	membrane-associated progesterone receptor component 1	3.23	5.94E-11
NP_001098596.2	tubulin alpha-1B chain	3.20	4.08E-03
NP_001071025.1	multiple inositol polyphosphate histidine phosphatase, 1a	3.19	4.01E-11

XP_003965907.1	moesin	3.09	9.10E-07
NP_001004680.1	cytochrome c oxidase subunit VIa-like	3.09	4.09E-04
NP_001038622.2	sorting nexin-17	3.08	3.40E-02
NP_001098746.1	hydroxyacyl-Coenzyme A dehydrogenase/3-ketoacyl-Coenzyme A thiolase/enoyl-Coenzyme A hydratase, alpha subunit a	3.08	1.68E-03
NP_001018499.1	BTB and CNC homology 1 transcription factor	3.05	4.85E-02
NP_001007382.1	PCI domain-containing protein 2	3.05	1.24E-06
XP_002663445.1	PREDICTED: sterol 26-hydroxylase, mitochondrial isoform 1	3.05	7.62E-04
NP_958916.1	aldehyde dehydrogenase family 9 member A1-B	3.04	1.05E-02
NP_001108152.1	complement component 3-like isoform 2 precursor	3.03	7.88E-15
NP_001082802.1	solute carrier organic anion transporter family, member 1D1	3.01	3.61E-05
NP_853537.1	glutamine synthetase 1	3.01	8.05E-03
NP_001012498.1	betaine--homocysteine S-methyltransferase 1	3.01	5.88E-09
NP_001003743.1	short-chain specific acyl-CoA dehydrogenase, mitochondrial	2.99	4.47E-02
NP_001124060.1	protein tyrosine phosphatase, receptor type, Na precursor	2.99	2.52E-02
NP_001187096.1	interferon induced protein 2	2.98	4.39E-06
XP_699339.5	dipeptidyl peptidase 2 precursor	2.97	9.81E-04
XP_686012.3	PREDICTED: integrin beta-2-like	2.97	4.13E-02
NP_001077310.1	pantothenate kinase 1	2.95	2.29E-02
AAB67871.1	MHC class II beta chain	2.95	1.14E-04
NP_001188375.1	reticulon-4	2.95	4.05E-02
NP_998590.1	1-acyl-sn-glycerol-3-phosphate acyltransferase gamma	2.94	4.59E-02
NP_001002169.1	spermidine/spermine N1-acetyltransferase	2.94	6.49E-09
NP_001006057.1	interferon gamma inducible protein 30 precursor	2.93	7.64E-07
NP_001001400.1	actin-related protein 8	2.88	1.98E-02
NP_570987.1	catalase	2.88	1.80E-03
NP_001116535.1	dihydropyrimidinase	2.85	1.30E-05
NP_001025233.1	apolipoprotein Bb	2.84	6.55E-07
NP_878300.1	heparin cofactor 2 precursor	2.83	0.00E+00
NP_001076345.1	sulfotransferase family 3, cytosolic sulfotransferase 3	2.83	7.55E-07
NP_998217.1	3-ketoacyl-CoA thiolase, mitochondrial	2.81	4.05E-02
NP_001093626.1	Fc receptor, IgE, high affinity I, gamma polypeptide like precursor	2.81	5.52E-03
NP_001017670.1	membrane-spanning 4-domains, subfamily A, member 17A.11	2.81	6.77E-04
XP_003199737.1	PREDICTED: oncoprotein-induced transcript 3 protein-like	2.79	6.00E-03
NP_998668.1	splicing factor 3B subunit 3	2.78	3.20E-02
XP_001335910.2	interferon-inducible protein Gig1	2.78	1.28E-02
NP_001005939.1	multiple coagulation factor deficiency 2 precursor	2.78	6.92E-06
NP_001005587.1	carnitine O-acetyltransferase	2.77	7.26E-04
NP_001103192.1	cytochrome P450, family 2, subfamily AA, polypeptide 7	2.76	1.10E-02
ADO28130.1	microsomal glutathione S-transferase 1	2.75	2.69E-06

NP_957340.1	serine hydroxymethyltransferase, cytosolic	2.74	1.21E-02
NP_001007292.1	3-oxoacid CoA transferase 1a	2.74	2.13E-03
NP_998662.2	cysteine-rich protein 2	2.74	2.57E-02
NP_775366.1	lysosome membrane protein 2	2.73	2.73E-02
NP_001071013.1	ribosome-releasing factor 2, mitochondrial	2.73	1.87E-02
XP_696560.3	PREDICTED: la-related protein 1	2.73	1.10E-02
NP_001004546.1	organic solute transporter subunit alpha	2.73	2.31E-03
NP_991129.1	ADP-ribosylation factor-like protein 8B-A	2.73	2.19E-02
NP_001007389.1	tetratricopeptide repeat protein 36	2.72	9.57E-04
NP_001012315.1	heat-responsive protein 12-like isoform 1	2.72	1.14E-05
NP_775361.1	V-type proton ATPase subunit E 1	2.71	1.25E-02
NP_998297.1	fructose-1,6-bisphosphatase 1b	2.71	0.00E+00
NP_955996.1	lysosome-associated membrane glycoprotein 1 precursor	2.71	2.79E-03
XP_001332180.4	TRAF2 binding protein	2.71	2.68E-04
XP_003199241.1	cytochrome c oxidase subunit VIb-like	2.71	3.27E-11
NP_001017687.1	protein archease	2.71	3.34E-02
NP_958917.1	actin related protein 2/3 complex, subunit 5B	2.71	3.34E-02
NP_001103873.1	heat shock cognate 71 kDa protein	2.70	2.46E-05
NP_001073442.1	TATA-binding protein-associated factor 2N	2.69	4.61E-02
NP_001116744.1	sterol O-acyltransferase 1	2.68	3.87E-02
NP_955897.2	acidic chitinase-like precursor	2.68	3.95E-03
AAI55834.1	Si:dkey-10p5.7	2.68	4.47E-02
NP_001121757.1	UDP glucuronosyltransferase 5 family, polypeptide E1 precursor	2.68	3.57E-02
NP_001070621.1	zinc finger protein 36, C3H type-like 1a	2.68	1.71E-02
NP_001026846.1	poly A binding protein, cytoplasmic 1 a	2.66	2.22E-02
NP_001134813.1	Membrane-spanning 4-domains subfamily A member 4A	2.66	2.09E-03
NP_571637.1	bcl2-associated X protein, a	2.65	9.16E-03
ACY69995.1	heat shock cognate 70	2.64	1.09E-03
NP_001025343.1	peroxisomal membrane protein 11C	2.62	2.02E-02
NP_001029354.2	secreted immunoglobulin domain 4 precursor	2.62	0.00E+00
NP_955829.1	NHP2 non-histone chromosome protein 2-like 1	2.61	2.01E-02
NP_001071016.1	cytochrome P450, family 3, subfamily C, polypeptide 4	2.61	4.29E-04
NP_997921.1	granulin 2 precursor	2.60	4.84E-03
NP_694484.1	cytoglobin-1	2.59	5.52E-03
XP_693522.1	delta(14)-sterol reductase	2.58	2.50E-03
NP_942570.1	acylamino-acid-releasing enzyme	2.58	8.11E-03
NP_001017619.1	L-xylulose reductase	2.58	4.17E-03
NP_571251.1	mitochondrial uncoupling protein 2	2.57	8.86E-03
NP_001074159.1	protein CREG1 precursor	2.57	1.51E-03

NP_956373.1	pancreatic progenitor cell differentiation and proliferation factor A	2.55	4.03E-06
NP_001002574.1	bile acid receptor	2.55	1.67E-02
NP_956113.1	phosphoserine aminotransferase isoform 1	2.55	4.22E-02
NP_001028265.1	alpha globin-like	2.55	0.00E+00
NP_571625.1	pyruvate carboxylase, mitochondrial	2.52	2.53E-05
NP_571954.1	cytochrome P450, family 1, subfamily A	2.51	2.63E-11
NP_954686.1	cytochrome P450, family 4, subfamily T, polypeptide 8	2.50	1.39E-03
NP_001013045.1	ba1 globin, like	2.50	0.00E+00
NP_571447.1	CD74 molecule, major histocompatibility complex, class II invariant chain b	2.49	2.95E-12
NP_956185.1	ecdysoneless	2.49	3.33E-02
XP_695816.5	transferrin receptor protein 2	2.47	1.44E-02
XP_002664937.1	26S protease regulatory subunit 6A	2.47	2.62E-03
NP_001076265.2	bifunctional purine biosynthesis protein PURH	2.47	2.84E-02
NP_955914.1	V-type proton ATPase subunit d 1	2.46	2.16E-02
XP_003199922.1	fucosidase, alpha-L-1, tissue precursor	2.46	3.58E-02
NP_957159.1	non-specific lipid-transfer protein	2.45	4.76E-06
NP_001002117.1	ribokinase	2.43	2.89E-02
XP_003200452.1	PREDICTED: protein Z-dependent protease inhibitor	2.42	1.38E-11
NP_001070069.1	biliverdin reductase A	2.41	3.74E-02
NP_001073131.1	enoyl-CoA hydratase domain-containing protein 3, mitochondrial	2.41	3.91E-02
NP_001018199.1	glycogen starch synthase, liver	2.40	2.34E-04
NP_001008584.1	low affinity cationic amino acid transporter 2	2.39	2.41E-03
NP_001019582.1	MAWD binding protein like	2.39	1.69E-03
NP_001116330.1	mannan-binding lectin serine protease 2 precursor	2.39	0.00E+00
XP_685178.1	grancalcin	2.38	3.15E-04
NP_991258.2	eukaryotic translation initiation factor 4H isoform 2	2.37	1.69E-02
NP_957333.1	WD repeat-containing protein 1	2.37	3.89E-03
NP_571320.1	DNA-binding protein inhibitor ID-1	2.37	2.74E-03
NP_998255.1	scinderin like b	2.37	3.97E-02
AAH66622.1	Zgc:56585 protein	2.37	8.91E-04
NP_001035136.1	glia maturation factor gamma	2.36	3.00E-02
NP_571665.1	CD74 molecule, major histocompatibility complex, class II invariant chain a	2.34	5.86E-10
XP_004078254.1	PREDICTED: complement C1s subcomponent-like	2.34	1.12E-02
NP_571439.2	E3 ubiquitin-protein ligase Mdm2	2.33	4.23E-02
NP_001020335.1	inter-alpha-trypsin inhibitor heavy chain H4 precursor	2.32	0.00E+00
XP_002662791.1	PREDICTED: 5-formyltetrahydrofolate cyclo-ligase-like	2.31	2.74E-03
NP_954969.1	ADP-ribosylation factor 5	2.31	4.03E-02
XP_699028.2	PREDICTED: 25-hydroxycholesterol 7-alpha-hydroxylase-like	2.31	4.54E-02
NP_997935.1	carboxypeptidase N catalytic chain precursor	2.31	2.79E-03

NP_001038917.2	5-hydroxyisourate hydrolase	2.30	2.96E-02
AAH71456.2	Zgc:110339	2.30	2.52E-02
NP_571958.1	prosaposin precursor	2.30	2.43E-09
NP_956007.1	calreticulin precursor	2.29	1.20E-04
NP_998219.1	fibrinogen gamma polypeptide precursor	2.28	0.00E+00
NP_958482.1	major vault protein	2.28	2.85E-04
XP_700727.5	prefoldin subunit 1	2.28	5.09E-03
NP_998587.2	UDP glucuronosyltransferase 1 family polypeptide a3 precursor	2.28	4.80E-02
NP_001137225.1	thrombopoietin	2.27	4.09E-03
NP_997981.1	glycine N-methyltransferase	2.27	1.08E-04
NP_991121.1	AP-1 complex subunit sigma-2	2.27	4.06E-02
NP_999903.1	glucose-6-phosphate translocase	2.26	6.10E-03
NP_001116095.1	2-acylglycerol O-acyltransferase 1	2.25	7.70E-04
NP_001025299.1	dynamin-2	2.25	8.93E-03
NP_937853.1	endoplasmin precursor	2.25	1.60E-08
NP_001076267.1	14-3-3 protein beta/alpha-A	2.25	7.53E-09
NP_571893.1	angiopoietin-related protein 3	2.24	4.97E-11
NP_957362.1	epoxide hydrolase 1	2.24	9.72E-06
XP_003198717.1	PREDICTED: zinc finger protein 729-like	2.24	1.79E-07
NP_001005598.2	transthyretin (prealbumin, amyloidosis type I) precursor	2.24	2.53E-08
NP_955816.1	rho-related GTP-binding protein RhoE	2.24	2.77E-04
NP_956862.1	keratin-like	2.23	5.99E-03
NP_998807.1	MHC class I alpha chain, partial	2.23	4.01E-03
NP_956647.1	uncoupling protein 3	2.22	4.23E-02
NP_001002706.1	lysozyme g-like 1	2.22	4.69E-07
NP_570996.2	F-box only protein 50	2.21	3.99E-02
NP_001124169.1	beta thymosin-like	2.21	0.00E+00
NP_957161.1	serine/arginine-rich splicing factor 5a	2.20	3.46E-02
NP_001020668.1	novel immune-type receptor 11b precursor	2.20	2.86E-02
NP_998803.1	myosin light chain alkali, smooth-muscle isoform	2.20	2.42E-04
NP_775331.1	epididymal secretory protein E1 precursor	2.20	4.58E-02
NP_571601.1	10 kDa heat shock protein, mitochondrial	2.19	1.80E-04
XP_003450068.1	selenophosphate synthetase 2	2.19	1.14E-03
NP_001098578.1	THO complex subunit 4	2.19	8.11E-03
NP_001098576.1	regulator of microtubule dynamics protein 1	2.19	1.62E-02
NP_957395.1	H3 histone, family 3D	2.19	9.93E-04
NP_001159390.1	beta thymosin-like protein 2	2.17	2.76E-03
NP_998581.1	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 1b	2.17	4.70E-02
NP_957099.1	peroxiredoxin-6	2.17	2.49E-07

NP_954970.1	H1 histone family, member X	2.17	4.62E-02
NP_001187725.1	c-type lectin domain family 4 member f	2.14	1.89E-15
NP_001082894.1	peroxiredoxin-4 precursor	2.14	1.44E-06
XP_003199766.1	proteasome subunit alpha type-1	2.13	6.40E-04
NP_956985.1	tubulin alpha-4A chain	2.13	6.57E-06
NP_001106980.1	interleukin-27 subunit beta precursor	2.13	9.78E-03
XP_693183.3	PREDICTED: inter-alpha-trypsin inhibitor heavy chain H3-like	2.13	0.00E+00
XP_694331.5	24-dehydrocholesterol reductase	2.12	3.59E-02
NP_571870.2	proteasome alpha 6b subunit	2.11	1.74E-02
ADV35592.1	lily-type lectin-2	2.11	1.76E-02
XP_694827.4	PREDICTED: apolipoprotein B-100	2.10	1.12E-02
NP_998503.1	aminoacyl tRNA synthase complex-interacting multifunctional protein 2	2.09	2.80E-02
NP_001073512.1	transmembrane protein with metallophosphoesterase domain	2.09	2.41E-02
NP_001032515.1	cytochrome P450, family 3, subfamily A, polypeptide 65	2.09	1.71E-05
XP_001332828.3	PREDICTED: phosphatidylcholine-sterol acyltransferase-like	2.08	6.91E-03
NP_571317.1	complement component c3a precursor	2.08	0.00E+00
NP_997937.1	tubulin, alpha 8 like	2.07	8.02E-04
NP_878288.1	coagulation factor IX precursor	2.07	2.73E-02
NP_001039029.2	proteasome subunit beta type-7	2.07	1.66E-02
NP_997914.2	small GTPase RhoA	2.07	4.89E-03
NP_997805.1	solute carrier family 1 (glial high affinity glutamate transporter), member 3a	2.06	8.15E-03
NP_998569.1	lysosomal acid lipase/cholesteryl ester hydrolase precursor	2.06	1.16E-04
NP_001039316.1	aminoacyl tRNA synthase complex-interacting multifunctional protein 1	2.05	1.47E-02
NP_001018360.1	transmembrane protein 127	2.05	4.61E-02
NP_956747.1	elongation of very long chain fatty acids protein 5	2.04	4.87E-04
NP_957412.2	protein AMBP precursor	2.04	0.00E+00
NP_932329.1	angiotensinogen precursor	2.04	0.00E+00
NP_001004540.1	aldehyde dehydrogenase family 8 member A1	2.04	1.40E-03
NP_998508.1	signal recognition particle 19 kDa protein	2.04	5.50E-03
NP_954688.1	adenosylhomocysteinase	2.03	1.53E-14
XP_685817.2	V-type proton ATPase subunit e 1	2.03	6.63E-03
NP_001103939.1	quinoid dihydropteridine reductase a	2.02	1.33E-05
NP_997931.1	ran-specific GTPase-activating protein	2.02	8.02E-03
NP_571639.1	amyloid beta A4 protein precursor	2.02	4.07E-02
NP_957222.1	ras-related protein Rab-7a	2.01	1.68E-02
NP_001038536.2	protein Z-dependent protease inhibitor precursor	2.01	1.61E-06
NP_998298.1	MARCKS-like 1a	-2.01	3.55E-02
NP_001014827.1	stanniocalcin-2 precursor	2.01	1.95E-02
NP_957060.1	polyadenylate-binding protein-interacting protein 2B	2.01	4.26E-04

NP_955827.1	ras-related protein Rap-1b precursor	2.00	4.78E-03
NP_001077323.1	indoleamine 2,3-dioxygenase 2	-2.01	4.24E-02
NP_694512.1	hepatocyte growth factor-like protein	-2.01	2.05E-06
XP_003978992.1	pre-mRNA-processing factor 39	-2.02	2.70E-02
NP_001153818.1	zinc finger, FYVE domain containing 21	-2.02	4.59E-03
NP_001003993.1	adenylate kinase isoenzyme 1	-2.02	2.24E-03
NP_001014354.1	lysosomal alpha-mannosidase precursor	-2.02	2.47E-02
NP_956314.1	protein BTG1	-2.02	1.20E-02
NP_001018160.1	tetraspanin-2	-2.02	1.19E-03
NP_001123295.1	proteasome subunit beta type-3	-2.03	3.61E-04
NP_571702.1	C-X-C motif chemokine 14 precursor	-2.03	2.94E-05
NP_998694.1	transmembrane protein 50A	-2.03	1.98E-02
NP_001017767.1	ubiquitin-conjugating enzyme E2 G2	-2.04	1.30E-02
NP_001120841.1	neuropeptide B precursor	-2.04	7.09E-03
NP_571959.2	CCAAT/enhancer-binding protein beta	-2.05	8.88E-16
NP_001014311.1	eukaryotic translation initiation factor 4, gamma 2a	-2.05	4.00E-02
NP_956843.1	MACRO domain-containing protein 2	-2.05	2.77E-02
ACX35602.1	VHSV-induced protein	-2.05	1.40E-04
NP_958864.1	flotillin 1b	-2.06	2.97E-02
NP_001002609.1	proteasome subunit beta type-2	-2.06	6.99E-05
NP_001264046.1	acetyl-coenzyme A synthetase, cytoplasmic	-2.06	1.01E-04
NP_571445.2	acetyl-CoA acetyltransferase, cytosolic	-2.06	3.09E-03
XP_001921030.3	PREDICTED: RING finger protein 213	-2.06	4.88E-02
NP_001025316.1	unconventional prefoldin RPB5 interactor	-2.07	4.47E-02
XP_003200136.1	PREDICTED: NACHT, LRR and PYD domains-containing protein 3-like, partial	-2.07	1.94E-02
NP_956312.1	bleomycin hydrolase	-2.08	4.80E-02
NP_001136054.1	interleukin-1 receptor accessory protein-like 1-B precursor	-2.08	1.49E-10
Q75VN4.3	H2B_RHASC RecName: Full=Histone H2B	-2.08	4.98E-02
NP_955880.1	alpha-actinin-4	-2.09	3.17E-03
NP_997768.2	acidic (leucine-rich) nuclear phosphoprotein 32 family, member B	-2.10	1.89E-02
NP_571582.1	coatomer subunit zeta-2	-2.10	2.46E-02
CBN81124.1	Retinol dehydrogenase 11 (All-trans/9-cis/11-cis)	-2.10	1.77E-04
NP_919240.2	ornithine decarboxylase antizyme 1	-2.10	1.89E-02
XP_001336197.4	PREDICTED: NACHT, LRR and PYD domains-containing protein 12	-2.11	6.81E-03
NP_955846.1	H1 histone family, member 0	-2.12	1.96E-02
NP_997911.2	Krueppel-like factor 15	-2.12	3.64E-02
NP_571369.1	superoxide dismutase	-2.12	1.66E-09
NP_001107790.1	interleukin-6 receptor subunit alpha	-2.12	1.37E-07
NP_001093608.2	RNA-binding protein 5	-2.14	1.58E-02

NP_997829.1	tax1-binding protein 1 homolog B	-2.14	1.34E-03
XP_003199901.1	phospholipid transfer protein precursor	-2.15	2.21E-05
NP_957371.2	formimidoyltransferase-cyclodeaminase	-2.15	3.82E-02
NP_998302.1	ras homolog gene family, member A	-2.15	2.54E-02
NP_861431.1	annexin A11b	-2.16	7.78E-04
NP_956103.1	3 beta-hydroxysteroid dehydrogenase type 7	-2.16	5.65E-04
XP_003200078.1	solute carrier family 35 (UDP-glucuronic acid/UDP-N-acetylgalactosamine dual transporter), member D1a	-2.16	4.50E-02
NP_998100.1	acyl-CoA dehydrogenase-like	-2.18	2.49E-02
NP_571779.1	major histocompatibility complex class I UDA precursor	-2.18	1.08E-05
NP_001002569.2	probable thiopurine S-methyltransferase	-2.19	4.77E-02
NP_958480.1	splicing factor, arginine/serine-rich 3b	-2.19	6.23E-04
NP_954973.2	influenza virus NS1A-binding protein homolog A	-2.20	2.11E-04
NP_571593.2	CD81 antigen	-2.20	1.53E-03
NP_998355.1	tumor necrosis factor receptor superfamily member 1A precursor	-2.20	2.11E-02
NP_991125.1	barrier-to-autointegration factor	-2.20	6.01E-03
XP_001339829.4	apolipoprotein C-I	-2.20	0.00E+00
NP_001243136.1	p53 apoptosis effector related to PMP-22	-2.20	7.33E-03
NP_001103588.1	major histocompatibility complex class I ZE like precursor	-2.20	1.06E-04
NP_001076548.1	proline-rich nuclear receptor coactivator 1	-2.21	5.99E-05
NP_956486.1	serine carboxypeptidase 1 precursor	-2.21	1.69E-05
NP_001003424.1	proteasome maturation protein	-2.22	1.07E-07
NP_997769.2	cystathionase (cystathionine gamma-lyase)	-2.22	4.85E-02
NP_957224.1	PDZ domain-containing protein GIPC2	-2.23	1.00E-02
NP_001032501.1	high mobility group protein B2	-2.23	6.37E-08
NP_998288.1	matrix metalloproteinase-9 precursor	-2.23	5.11E-05
NP_956660.1	PRELI domain-containing protein 1, mitochondrial	-2.23	4.61E-03
NP_001018874.1	protein C-ets-2	-2.23	3.02E-03
XP_001919541.2	oxidation resistance protein 2 variant A, partial	-2.23	3.68E-02
NP_001104638.1	serine/threonine-protein phosphatase 4 catalytic subunit B	-2.23	4.60E-02
NP_001018311.1	bloodthirsty	-2.23	1.40E-02
NP_998349.1	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha a	-2.24	5.28E-08
NP_957489.1	nucleoside diphosphate kinase, mitochondrial	-2.24	6.84E-04
NP_956311.1	cold inducible RNA binding protein isoform 2	-2.24	2.07E-04
NP_955991.1	transcription initiation factor IIB	-2.25	2.00E-02
NP_571648.1	death associated protein 1b	-2.25	1.61E-02
NP_001018440.1	katanin p60 ATPase-containing subunit A1	-2.26	3.02E-02
NP_956308.1	charged multivesicular body protein 1b	-2.26	3.10E-03
NP_956694.1	DnaJ (Hsp40) homolog, subfamily C, member 5 gamma b	-2.26	4.84E-02
XP_002666912.2	D-dopachrome decarboxylase	-2.27	1.08E-04

NP_955923.1	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha b	-2.27	3.01E-06
NP_956650.1	protein C precursor	-2.27	7.38E-03
NP_998667.1	eukaryotic translation initiation factor 3 subunit J-B	-2.27	3.77E-03
NP_571546.1	major histocompatibility complex class I UBA precursor	-2.27	3.11E-05
NP_956010.1	mps one binder kinase activator-like 2A	-2.27	3.01E-02
NP_001128601.1	urocanate hydratase 1	-2.28	3.45E-09
XP_687660.2	coagulation factor V precursor	-2.29	7.84E-09
NP_957251.1	solute carrier family 15 member 4	-2.29	4.12E-02
NP_001003781.1	transport and golgi organization 2 homolog	-2.29	1.61E-04
XP_003198245.1	sulfotransferase 1C1	-2.30	6.12E-03
NP_956027.1	tsukushin precursor	-2.32	1.53E-02
NP_991134.2	BCL2/adenovirus E1B interacting protein 3-like b	-2.32	3.63E-02
NP_957108.1	pterin-4-alpha-carbinolamine dehydratase	-2.33	2.37E-04
NP_001191177.1	bifunctional aminoacyl-tRNA synthetase	-2.33	7.20E-03
XP_003199582.1	PREDICTED: synaptogyrin-2-like	-2.34	1.10E-04
NP_001071206.1	transmembrane protease, serine 4a	-2.34	1.25E-02
NP_001071019.1	agmatinase, mitochondrial	-2.34	3.67E-02
NP_001077337.1	interleukin-10 receptor subunit beta precursor	-2.34	4.83E-03
NP_001018570.1	tumor protein D52-like 1	-2.34	3.24E-02
NP_571746.1	ATPase, Na+/K+ transporting, beta 1b polypeptide	-2.35	2.23E-05
NP_998388.1	myristoylated alanine-rich C kinase substrate 2	-2.35	9.06E-03
NP_956628.1	putative hydroxypyruvate isomerase	-2.35	2.75E-02
NP_001018389.1	peptidyl-prolyl cis-trans isomerase NIMA-interacting 4	-2.36	8.24E-04
NP_001004674.1	high-mobility group box 2b	-2.36	3.61E-11
NP_998622.1	charged multivesicular body protein 4c	-2.36	3.41E-03
NP_001025358.1	proteasome subunit alpha type-3	-2.36	2.23E-04
NP_001002342.1	nuclear prelamin A recognition factor	-2.36	2.64E-02
NP_954968.1	regulator of G-protein signaling 4	-2.36	8.54E-04
NP_997915.1	jun B proto-oncogene b	-2.37	2.22E-16
NP_001007775.1	pre-mRNA-splicing factor SPF27	-2.37	3.67E-02
NP_956328.1	stress-associated endoplasmic reticulum protein 2	-2.37	3.49E-11
NP_956207.1	mitogen-activated protein kinase kinase kinase kinase 5	-2.38	3.02E-02
NP_957507.1	glutaminyl-tRNA synthetase	-2.38	4.77E-02
XP_003198602.1	PREDICTED: NACHT, LRR and PYD domains-containing protein 3	-2.38	0.00E+00
NP_956500.2	zgc:56194 precursor	-2.38	2.96E-03
NP_956558.2	protein NipSnap homolog 3A	-2.38	3.70E-02
NP_571505.1	fibroblast growth factor receptor 4 precursor	-2.39	2.91E-02
NP_571570.2	apoptosis-associated speck-like protein containing a CARD	-2.39	3.65E-02
NP_997749.1	cathepsin L, 1 a precursor	-2.39	0.00E+00

NP_997853.1	cathepsin H precursor	-2.39	1.38E-06
NP_957461.1	4-hydroxyphenylpyruvate dioxygenase	-2.39	0.00E+00
NP_955958.1	ubiquitin-conjugating enzyme E2D 1	-2.40	6.31E-03
XP_001342084.3	PREDICTED: phosphoinositide 3-kinase adapter protein 1	-2.40	3.42E-03
NP_998069.1	charged multivesicular body protein 2b	-2.41	4.75E-02
XP_002664001.1	annexin A6	-2.41	4.40E-04
NP_001019602.2	HMG box-containing protein 1	-2.41	2.70E-02
NP_001013551.1	lysosomal membrane glycoprotein 2 precursor	-2.43	1.24E-02
NP_998593.1	CD9 antigen, b	-2.43	3.04E-02
NP_001013548.1	transmembrane protein 248	-2.43	1.63E-02
NP_997939.1	fibrinogen beta chain precursor	-2.43	0.00E+00
NP_963870.1	transgelin 2	-2.44	3.95E-08
NP_001103872.1	UDP-glucose 6-dehydrogenase	-2.44	4.90E-02
NP_001122201.1	Kruppel-like factor 9	-2.44	3.88E-02
NP_957422.1	sarcosine dehydrogenase, mitochondrial	-2.45	3.65E-02
NP_001139039.1	3'-5' exoribonuclease CSL4 homolog	-2.45	1.33E-03
NP_001035389.1	UDP-glucose 4-epimerase	-2.45	4.23E-02
ABQ59678.1	vacuolar protein sorting 52	-2.45	2.46E-02
NP_956864.1	selenium-binding protein 1	-2.46	5.97E-06
NP_919376.1	baculoviral IAP repeat-containing 2	-2.46	2.67E-02
NP_955898.1	microtubule-associated proteins 1A/1B light chain 3B	-2.47	6.48E-03
NP_001229699.1	glucagon b precursor	-2.48	4.77E-02
NP_001132933.1	vitronectin b precursor	-2.48	0.00E+00
NP_956617.2	hydroxysteroid 11-beta-dehydrogenase 1-like protein precursor	-2.48	1.49E-02
NP_001107891.1	hemK methyltransferase family member 1	-2.48	4.85E-02
NP_998504.1	palmitoyl-protein thioesterase 1 precursor	-2.49	1.65E-03
NP_001092231.1	histone cluster 1, H4-like	-2.50	1.03E-03
NP_942104.2	tubulin, beta 2c	-2.50	2.94E-02
NP_955365.1	pyruvate kinase isozymes M1/M2	-2.51	7.48E-04
XP_684448.2	oxysterol binding protein-like 3a	-2.51	2.64E-02
NP_956887.2	serine/arginine-rich splicing factor 1B	-2.51	3.17E-02
NP_997777.1	probable ATP-dependent RNA helicase DDX5	-2.53	7.24E-05
NP_001107061.1	heat shock cognate 70-kd protein, like	-2.53	2.92E-02
NP_954682.1	serine/threonine-protein kinase Sgk1	-2.54	1.07E-08
NP_998512.1	NEDD4 family-interacting protein 2	-2.54	3.73E-07
XP_693301.4	PREDICTED: poly ADP-ribose polymerase 10-like	-2.54	4.38E-02
NP_932336.3	transketolase-like protein 2	-2.55	2.30E-08
NP_001070762.1	DENN/MADD domain containing 2D	-2.56	1.11E-02
XP_002661254.1	PREDICTED: mitogen-activated protein kinase kinase kinase 14-like	-2.56	1.20E-02

NP_001017787.1	leptin receptor gene-related protein	-2.56	3.39E-04
NP_956065.1	ras-related C3 botulinum toxin substrate 1	-2.56	9.42E-03
NP_001005964.1	AP-3 complex subunit sigma-1	-2.57	3.52E-02
XP_002667651.2	microtubule-associated protein, RP/EB family, member 3b	-2.57	1.63E-03
NP_001002184.1	lipopolysaccharide-induced tumor necrosis factor-alpha factor homolog	-2.57	7.22E-04
NP_956667.1	sorcin	-2.58	1.61E-04
NP_998519.1	conserved oligomeric Golgi complex subunit 2	-2.59	2.73E-04
NP_658909.1	glucose phosphate isomerase a	-2.59	4.44E-09
NP_999924.1	legumain precursor	-2.59	1.69E-07
NP_998240.2	flotillin-2a	-2.60	2.72E-02
NP_001073542.1	RING finger protein 115	-2.60	2.28E-02
NP_998196.1	growth arrest and DNA-damage-inducible, beta a	-2.61	3.30E-02
NP_957320.1	AP-2 complex subunit mu-B	-2.61	2.05E-02
NP_957459.2	protein Mpv17	-2.61	6.25E-03
XP_684081.4	mannose-P-dolichol utilization defect 1a	-2.62	3.74E-10
NP_001020727.1	radical S-adenosyl methionine domain-containing protein 2	-2.63	2.72E-03
NP_001038213.2	serine/threonine-protein kinase TBK1	-2.66	7.66E-03
NP_001020708.1	homocysteine-responsive endoplasmic reticulum-resident ubiquitin-like domain member 1 protein	-2.66	3.19E-03
NP_956438.1	uridine phosphorylase 2	-2.67	0.00E+00
NP_956932.1	complement component C6	-2.67	0.00E+00
XP_699436.4	PREDICTED: peptidyl-glycine alpha-amidating monooxygenase A	-2.67	3.20E-03
NP_956780.1	solute carrier family 25 member 39	-2.68	3.61E-02
NP_001159387.1	pendrin	-2.68	3.63E-02
NP_001073497.1	monocarboxylate transporter 10	-2.68	9.18E-05
NP_001038775.1	glutamate NMDA receptor-associated protein 1	-2.69	5.38E-06
NP_954681.1	CTP synthase 1	-2.70	4.16E-02
NP_001117013.1	CD9 antigen	-2.70	9.64E-12
NP_001007162.1	ras-related protein Rab-1A	-2.70	3.93E-02
NP_001038936.1	protein FAM195B	-2.70	2.07E-02
NP_956994.1	PPPDE peptidase domain-containing protein 2	-2.73	4.05E-02
CAP19417.1	novel protein similar to vertebrate RAB33B, member RAS oncogene family	-2.74	1.29E-02
NP_956612.1	ADP-ribosylation factor-like protein 4A	-2.74	3.96E-02
NP_001003879.1	DNA polymerase beta	-2.75	1.65E-02
NP_571019.1	delta-like protein C precursor	-2.75	2.61E-02
NP_001018154.1	component of gems 4	-2.77	4.02E-03
NP_991248.1	glyoxylate reductase/hydroxypyruvate reductase b	-2.77	0.00E+00
XP_687605.4	PREDICTED: NACHT, LRR and PYD domains-containing protein 12-like	-2.77	3.19E-05
NP_571838.1	claudin b	-2.79	3.20E-02
NP_956434.1	ras-like protein family member 11B	-2.80	1.24E-02

NP_001002334.1	tetraspanin 13a	-2.81	1.40E-03
NP_001038418.1	glutaminyl-peptide cyclotransferase-like	-2.81	0.00E+00
NP_001038447.2	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 4a	-2.81	1.83E-05
NP_958464.1	nuclear distribution protein nudE-like 1-B	-2.81	1.13E-02
XP_002660664.2	PREDICTED: microfibril-associated glycoprotein 4	-2.81	0.00E+00
XP_002661250.2	Putative transposase	-2.81	6.20E-03
XP_003201439.1	PREDICTED: von Willebrand factor A domain-containing protein 5A-like	-2.81	6.78E-04
NP_001070070.1	solute carrier family 25 member 38-A	-2.82	6.72E-05
AAI15338.1	LOC555678 protein	-2.83	6.95E-04
NP_991273.1	interferon regulatory factor 9	-2.83	2.01E-06
NP_001035126.1	bone morphogenetic protein 1a precursor	-2.83	7.14E-03
NP_956081.1	NEDD4-binding protein 1	-2.84	1.91E-02
NP_956300.1	heme transporter hrg1-A	-2.84	3.53E-02
NP_001074088.1	solute carrier family 39 (zinc transporter), member 3	-2.84	4.24E-04
XP_696748.3	transferrin receptor 1a	-2.84	5.74E-04
NP_956366.1	lectin, galactoside-binding, soluble, 9 (galectin 9)-like 1	-2.84	1.09E-09
NP_571531.1	catenin alpha-1	-2.85	2.95E-03
NP_705930.1	histone H2A.V	-2.85	1.60E-07
NP_001019576.1	ubiquitin carboxyl-terminal hydrolase isozyme L3	-2.85	8.59E-03
NP_957039.1	endothelial differentiation-related factor 1 homolog	-2.86	1.36E-03
NP_001038672.1	arf-GAP with SH3 domain, ANK repeat and PH domain-containing protein 2	-2.86	2.87E-02
NP_937785.1	basigin precursor	-2.87	0.00E+00
XP_002660683.1	novel protein similar to vertebrate solute carrier family 23 (nucleobase transporters), member 2 (SLC23A2, zgc:110789)	-2.87	3.11E-02
NP_001018380.1	WW domain-binding protein 2	-2.88	1.52E-02
NP_001140147.1	CDP-diacylglycerol--serine O-phosphatidyltransferase	-2.89	4.82E-02
XP_692125.5	PREDICTED: interleukin-4 receptor subunit alpha	-2.89	3.88E-02
XP_700015.4	PREDICTED: myosin-X	-2.90	3.11E-02
NP_957132.2	Golgin subfamily A member 1	-2.90	2.79E-02
NP_956677.1	CDGSH iron-sulfur domain-containing protein 2	-2.90	2.13E-02
NP_571876.1	ornithine decarboxylase	-2.90	9.64E-04
NP_991271.1	proteasome subunit alpha type-5	-2.91	4.41E-08
NP_958922.2	solute carrier family 3, member 2b	-2.92	6.91E-04
ACS16044.1	C-type lectin	-2.92	0.00E+00
NP_956136.1	guanine nucleotide-binding protein G(i) subunit alpha-2	-2.93	5.77E-03
NP_956630.1	v-maf musculoaponeurotic fibrosarcoma oncogene homolog f	-2.93	1.03E-02
NP_705949.1	protein spinster homolog 1	-2.93	1.49E-02
NP_571895.1	cadherin 1, epithelial precursor	-2.94	1.29E-06
NP_001002651.1	LIM domain kinase 2	-2.95	3.01E-03
NP_956417.1	RAB11a, member RAS oncogene family, like	-2.95	1.60E-02

XP_687430.2	PREDICTED: BTB/POZ domain-containing protein 10	-2.95	4.65E-04
XP_001336816.1	transposase	-2.95	2.44E-02
NP_001129720.2	C-C motif chemokine 19-like precursor	-2.96	0.00E+00
NP_001070085.1	cytokine-inducible SH2-containing protein	-2.96	3.04E-03
NP_001038800.1	ribosomal protein L22-like 1	-2.96	1.29E-04
NP_998252.1	45 kDa calcium-binding protein precursor	-2.97	3.51E-04
NP_001032767.2	solute carrier organic anion transporter family, member 2b1	-2.97	3.89E-04
NP_001017660.1	histone H1-like	-2.97	3.53E-08
XP_693568.5	epoxide hydrolase 2	-2.97	8.78E-03
NP_956381.1	thioredoxin interacting protein a	-3.00	0.00E+00
XP_003197916.1	PREDICTED: RING finger protein 31-like, partial	-3.00	1.28E-03
NP_001071005.1	H(+)/Cl(-) exchange transporter 7	-3.00	1.66E-02
NP_001038402.1	atrial natriuretic peptide receptor 1 precursor	-3.02	2.01E-02
NP_001073305.1	glycoprotein integral membrane protein 1 precursor	-3.02	1.41E-04
XP_001335593.2	PREDICTED: putative transcription factor p65 homolog	-3.04	1.70E-02
NP_001108211.2	MAP kinase-interacting serine/threonine-protein kinase 1	-3.04	1.79E-02
NP_001187403.1	ras-related protein rab-7a	-3.04	4.11E-05
NP_001002683.1	dipeptidyl peptidase 3	-3.04	1.25E-02
NP_001119855.1	serine/threonine-protein kinase SIK1	-3.05	2.10E-02
NP_956879.1	major histocompatibility complex class I UXA2 precursor	-3.05	9.09E-04
XP_002663230.2	PREDICTED: Niemann-Pick C1-like protein 1-like	-3.06	2.76E-02
NP_001083051.1	solute carrier organic anion transporter family, member 2A1	-3.07	9.27E-03
NP_999977.1	erythrocyte membrane protein band 4.1-like 3b	-3.07	3.08E-02
NP_001017833.1	inorganic pyrophosphatase	-3.08	0.00E+00
NP_957126.1	mid1-interacting protein 1-B	-3.09	1.13E-12
NP_998523.1	phosphatidylinositol 4-kinase type 2-alpha	-3.10	4.54E-02
NP_998377.1	myosin, light chain 9b, regulatory	-3.11	5.73E-03
XP_693668.1	PREDICTED: LAG1 longevity assurance homolog 2-like	-3.11	3.48E-03
NP_956004.1	WW domain binding protein 2	-3.12	0.00E+00
NP_840092.1	chemokine (C-X-C motif) ligand 12a (stromal cell-derived factor 1) precursor	-3.12	7.59E-06
NP_942110.1	stearoyl-CoA desaturase 5	-3.15	1.40E-03
NP_998676.1	N-acetylserotonin O-methyltransferase-like protein	-3.16	3.76E-02
NP_001077039.1	ERBB receptor feedback inhibitor 1	-3.16	4.86E-02
NP_999905.1	P3 protein precursor	-3.17	2.20E-05
NP_001183982.1	musculoskeletal, embryonic nuclear protein 1	-3.18	2.48E-02
NP_001071210.2	procollagen-lysine,2-oxoglutarate 5-dioxygenase 1 precursor	-3.18	2.48E-02
NP_001037807.2	putative helicase mov-10-B.1	-3.19	2.62E-02
NP_571466.1	proteasome subunit beta type-9	-3.19	2.85E-05
NP_991286.1	gamma-aminobutyric acid receptor-associated protein-like 2	-3.20	2.21E-07

XP_694922.4	PREDICTED: ATP-dependent DNA helicase Q5-like	-3.20	1.82E-06
NP_571861.2	cryptochrome-2	-3.20	1.94E-02
NP_001002543.1	proteasome subunit beta type-10	-3.21	2.23E-04
XP_684527.3	PREDICTED: prominin-1-A	-3.21	1.10E-04
NP_571660.1	ferritin heavy chain	-3.22	0.00E+00
NP_998362.1	zinc finger protein 207	-3.24	3.12E-02
NP_956871.1	N-acylsphingosine amidohydrolase precursor	-3.24	9.14E-07
NP_001004635.1	AP-1 complex subunit sigma-3	-3.25	1.18E-02
NP_001139082.1	transmembrane protein 176	-3.25	1.10E-02
XP_001337112.2	PREDICTED: phosphatidylserine synthase 2	-3.26	3.48E-02
NP_957043.1	acyl-protein thioesterase 2	-3.26	4.26E-05
NP_001116524.1	endoplasmic reticulum aminopeptidase 2 precursor	-3.27	1.84E-03
NP_001108532.1	chemokine (C-X-C motif) ligand C1c	-3.27	9.78E-03
NP_956857.1	charged multivesicular body protein 1a	-3.27	2.24E-05
NP_001116523.1	bloodthirsty-related gene family, member 22	-3.27	2.56E-02
NP_571450.1	proteasome activator complex subunit 1	-3.28	7.77E-15
NP_001018325.1	transmembrane protein 179B	-3.29	3.51E-02
NP_956385.2	signal transducer and activator of transcription 1b	-3.30	3.35E-05
NP_001070133.1	ras-related protein Rab-21	-3.31	1.41E-03
NP_001002461.1	thioredoxin	-3.31	0.00E+00
NP_001119851.1	long-chain-fatty-acid--CoA ligase ACSBG2	-3.32	6.91E-04
XP_687184.3	PREDICTED: leucine-rich repeat-containing G-protein coupled receptor 4	-3.33	4.14E-02
NP_001013554.2	receptor accessory protein 6-like	-3.33	3.21E-02
XP_695257.5	PREDICTED: lymphocyte antigen 75, partial	-3.33	2.13E-02
XP_001339327.1	PREDICTED: ras-related protein Rab-35-like	-3.34	1.66E-03
XP_003200235.1	PREDICTED: protein BTG3-like	-3.34	2.32E-03
NP_001155073.1	S-adenosylmethionine decarboxylase proenzyme	-3.34	2.32E-03
NP_001018464.1	glycogen phosphorylase, muscle form	-3.35	2.57E-03
NP_955837.1	CD63 antigen	-3.35	6.79E-14
XP_692717.4	PREDICTED: TNFAIP3-interacting protein 2-like	-3.38	1.44E-02
NP_001013321.2	PQ-loop repeat-containing protein 2-like	-3.39	1.47E-03
XP_699900.5	glycoprotein M6Bb	-3.40	1.02E-02
NP_001035412.1	coiled-coil domain-containing protein 104	-3.40	4.00E-02
NP_571168.1	Janus kinase 2a	-3.40	5.69E-03
XP_001340114.1	PREDICTED: stAR-related lipid transfer protein 13	-3.40	5.69E-03
NP_001038787.1	crumbs homolog 3a precursor	-3.42	3.62E-02
NP_001124067.1	toll-like receptor 5b precursor	-3.43	0.00E+00
XP_001345947.4	PREDICTED: microfibril-associated glycoprotein 4-like	-3.46	0.00E+00
NP_955969.1	tropomyosin alpha-4 chain	-3.46	3.65E-02

NP_998528.1	SUMO-activating enzyme subunit 2	-3.46	5.62E-04
XP_001340167.2	PREDICTED: poly ADP-ribose polymerase 9	-3.50	1.81E-02
XP_003198892.1	PREDICTED: NFX1-type zinc finger-containing protein 1-like	-3.50	3.15E-02
NP_001074091.1	CD276 antigen precursor	-3.51	7.49E-03
NP_001013490.1	thyroid cancer-1	-3.51	1.73E-11
NP_571882.1	bcl-2-like protein 1	-3.52	2.87E-03
NP_001107050.1	FAD-dependent oxidoreductase domain-containing protein 1	-3.53	2.30E-02
NP_001091647.2	myosin-9	-3.53	1.04E-04
NP_001002166.1	lysine-specific demethylase 5B-B	-3.54	2.34E-02
XP_002667876.1	PREDICTED: peptidyl-prolyl cis-trans isomerase C-like	-3.54	1.56E-03
NP_958899.1	isovaleryl-CoA dehydrogenase, mitochondrial	-3.55	2.60E-02
NP_958904.1	hypoxia-inducible factor 1-alpha inhibitor	-3.57	4.33E-02
NP_998468.1	DNA polymerase delta subunit 2	-3.58	4.47E-02
NP_001073498.1	arrestin domain-containing protein 3	-3.59	3.92E-02
NP_997933.1	methyl-CpG-binding domain protein 2	-3.59	9.00E-04
ACI66294.1	C19orf12 homolog	-3.59	1.56E-07
XP_695742.5	PREDICTED: sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1-like	-3.60	1.51E-02
XP_003200020.1	PREDICTED: probable ATP-dependent RNA helicase DDX41-like, partial	-3.60	8.43E-03
NP_955927.1	WD repeat and SOCS box-containing protein 1	-3.62	3.42E-04
XP_693535.5	PREDICTED: A disintegrin and metalloproteinase with thrombospondin motifs 1	-3.62	2.56E-03
XP_002660860.1	PREDICTED: leucine-rich repeat-containing protein 8D	-3.62	1.85E-02
XP_001338257.2	PREDICTED: poly(ADP-ribose) glycohydrolase	-3.63	3.38E-02
NP_001106976.1	interleukin-6 receptor subunit beta precursor	-3.63	1.25E-04
XP_001331985.3	PREDICTED: nuclear receptor coactivator 7-like	-3.63	1.19E-02
NP_571126.1	T-box transcription factor TBX2b	-3.64	3.80E-02
NP_571752.1	proteasome beta 10 subunit	-3.64	4.57E-05
XP_691115.5	PREDICTED: poly ADP-ribose polymerase 14, partial	-3.64	1.38E-02
AAT52141.1	CC chemokine SCYA107	-3.66	9.18E-08
NP_957212.1	basic leucine zipper and W2 domain-containing protein 2	-3.66	4.80E-03
NP_001134823.1	Programmed cell death 1 ligand 1 precursor	-3.66	2.06E-02
NP_001018462.1	alpha-methylacyl-CoA racemase	-3.67	2.55E-02
NP_919396.1	protein phosphatase 2, regulatory subunit B', epsilon isoform a	-3.67	4.17E-02
NP_001038487.1	transmembrane 4 L six family member 18	-3.68	1.69E-02
XP_682981.1	PREDICTED: antigen peptide transporter 2-like isoform 1	-3.68	1.55E-03
ABA54949.1	CC chemokine SCYA101	-3.68	2.90E-04
NP_998204.1	AN1-type zinc finger protein 5	-3.68	5.55E-04
NP_001018470.1	homer protein homolog 2	-3.69	1.11E-04
NP_001070838.2	protein kinase, cAMP-dependent, regulatory, type II, alpha A	-3.71	1.83E-02
NP_001004672.1	protein lin-7 homolog C	-3.72	4.20E-03

NP_001001840.2	nuclear factor NF-kappa-B p100 subunit	-3.72	1.25E-11
NP_997887.1	alpha-enolase	-3.74	0.00E+00
NP_956047.1	calponin 3, acidic a	-3.74	3.79E-02
NP_998323.1	tropomyosin alpha-4 chain isoform 2	-3.74	3.79E-02
NP_956971.1	interferon regulatory factor 7	-3.77	2.30E-02
NP_001187232.1	NK-lysin type 3 precursor	-3.77	1.88E-04
XP_693604.4	PREDICTED: receptor-transporting protein 2	-3.77	4.00E-02
NP_001034710.1	follistatin-like 1b precursor	-3.80	2.01E-03
NP_957376.1	calumenin-B precursor	-3.80	9.59E-10
NP_001018368.1	complement factor D precursor	-3.83	1.32E-07
NP_957370.1	SH2 domain containing 3Cb	-3.83	3.22E-03
XP_003446639.1	PREDICTED: NACHT, LRR and PYD domains-containing protein 3-like	-3.83	2.93E-03
NP_942571.1	allograft inflammatory factor 1-like	-3.84	6.13E-05
NP_001038932.1	transgelin	-3.84	2.52E-04
NP_001017637.1	transmembrane gamma-carboxyglutamic acid protein 4 precursor	-3.84	1.39E-03
NP_997753.1	sodium-dependent phosphate transporter 1-B	-3.84	7.52E-07
XP_001920595.2	PREDICTED: n-myc-interactor	-3.84	3.23E-09
NP_571614.1	serine/threonine-protein kinase pim-2	-3.84	1.15E-03
NP_001191383.1	P2Y purinoceptor 11	-3.85	2.98E-02
NP_001017722.1	RNA/RNP complex-1-interacting phosphatase	-3.86	1.49E-02
NP_001025408.1	mitochondrial carnitine/acylcarnitine carrier protein CACL	-3.87	2.25E-02
NP_001093536.2	optineurin	-3.88	1.29E-08
NP_955967.1	a disintegrin and metalloproteinase domain 17a precursor	-3.89	4.83E-04
NP_991189.1	BCL-6 corepressor	-3.89	3.84E-02
NP_942099.1	bisphosphoglycerate mutase 1a	-3.89	0.00E+00
NP_997730.1	dual specificity protein phosphatase 5	-3.94	8.04E-03
XP_003198109.1	PREDICTED: interferon-induced very large GTPase 1-like	-3.95	2.34E-02
NP_958466.1	suppressor of fused homolog	-3.96	4.04E-02
NP_956573.2	Golgi membrane protein 1-like	-3.96	4.39E-05
NP_998274.2	UMP-CMP kinase	-3.98	4.72E-02
NP_571238.1	beta-2-microglobulin precursor	-3.98	0.00E+00
XP_002665164.1	PREDICTED: stonustoxin subunit alpha-like	-3.98	1.84E-02
NP_001124074.1	cytochrome c oxidase subunit VIIb	-3.99	5.36E-06
XP_001341822.1	PREDICTED: transmembrane protein 20-like	-3.99	1.08E-02
NP_957406.1	ADP-ribosylation factor-like protein 6-interacting protein 1	-4.00	2.37E-03
NP_998646.1	guanine nucleotide binding protein (G protein), beta polypeptide 1, like	-4.01	7.72E-03
NP_956527.1	hypoxia-inducible factor 1-alpha	-4.02	4.44E-16
NP_878286.2	glutamine synthetase	-4.02	4.58E-03
NP_997826.1	CD82 antigen	-4.03	5.54E-03

XP_001346157.4	PREDICTED: ribonuclease inhibitor-like	-4.03	3.30E-03
NP_571677.1	DNA-binding death effector domain-containing protein 2	-4.05	3.65E-02
NP_956489.1	charged multivesicular body protein 4b	-4.06	3.87E-02
NP_001008644.2	BMP-2-inducible protein kinase isoform 2	-4.07	8.53E-03
XP_686013.4	ORM1-like protein 3	-4.09	3.88E-02
NP_001002061.1	ras-related C3 botulinum toxin substrate 2	-4.09	4.83E-03
NP_998331.1	proteasome subunit alpha type-7-like	-4.10	1.67E-08
XP_003198142.1	PREDICTED: interleukin-12 receptor subunit beta-2	-4.12	1.51E-06
NP_999898.1	5-azacytidine-induced protein 2	-4.12	2.18E-03
NP_997868.1	hypoxia up-regulated protein 1 precursor	-4.12	6.87E-04
NP_956634.1	pleckstrin homology domain containing, family F (with FYVE domain) member 1	-4.14	2.25E-02
NP_001002216.1	growth arrest and DNA-damage-inducible, alpha, b	-4.14	6.75E-03
XP_692034.2	PREDICTED: TBC1 domain family member 9 isoform 2	-4.16	3.81E-04
XP_002666484.2	PREDICTED: leucine-rich repeat-containing G-protein coupled receptor 6-like	-4.16	2.75E-02
XP_002759284.1	PREDICTED: metastasis suppressor protein 1 isoform 2	-4.17	1.39E-02
NP_999915.1	E3 ubiquitin-protein ligase RNF170	-4.18	3.35E-02
NP_954965.1	carboxypeptidase A precursor	-4.18	1.07E-06
XP_684461.4	PREDICTED: beta-1,4-galactosyltransferase 3	-4.20	1.03E-09
NP_001008578.1	mitochondrial fission regulator 1	-4.21	7.45E-03
NP_001013533.1	lipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex, mitochondrial	-4.21	1.53E-02
NP_001019612.1	proteasome alpha 2 subunit	-4.22	1.24E-09
NP_957303.1	acetoacetyl-CoA synthetase	-4.22	7.24E-07
XP_003276069.1	PREDICTED: cAMP-responsive element modulator isoform 1	-4.22	6.89E-13
XP_687719.1	glutathione peroxidase 1	-4.22	3.72E-02
NP_001157280.1	catechol-O-methyltransferase domain containing 1 precursor	-4.25	6.65E-12
XP_002667267.2	PREDICTED: polypeptide N-acetylgalactosaminyltransferase 5-like, partial	-4.27	3.04E-02
ACM08269.1	Tetraspanin-1	-4.27	4.74E-08
NP_956342.1	arginine--tRNA ligase, cytoplasmic	-4.28	1.23E-04
NP_001032648.1	Y+L amino acid transporter 1	-4.28	1.86E-02
NP_001124061.1	myosin, light chain 12, genome duplicate 2	-4.28	1.35E-04
XP_003199326.1	protein dpy-19 homolog 1	-4.29	4.98E-02
NP_001076277.1	PX domain-containing protein kinase-like protein	-4.29	2.92E-03
NP_956394.1	HIG1 domain family member 1A	-4.33	8.46E-08
NP_919375.2	myeloid cell leukemia sequence 1b	-4.34	0.00E+00
NP_001002707.1	gamma-aminobutyric acid receptor-associated protein-like	-4.34	4.29E-09
NP_001025120.1	leucine rich repeat containing 8 family, member A	-4.35	2.13E-04
NP_999851.1	sulfotransferase 6B1	-4.36	2.97E-12
NP_001071197.1	serine/threonine-protein kinase 17A	-4.36	1.27E-03
NP_001038484.1	TCDD-inducible poly ADP-ribose polymerase	-4.37	7.14E-03

NP_001104702.1	cystathionine-beta-synthase a	-4.38	0.00E+00
NP_957229.1	phosphatidylinositol transfer protein alpha isoform	-4.41	2.27E-02
XP_690797.3	eukaryotic translation initiation factor 4E-binding protein 3	-4.41	2.18E-03
XP_699217.2	PREDICTED: interferon-induced guanylate-binding protein 1	-4.42	4.06E-03
NP_001020351.1	six-cysteine containing astacin protease 4 precursor	-4.44	3.21E-02
NP_956233.1	arrestin domain-containing protein 2	-4.44	1.33E-02
XP_688060.3	PREDICTED: interferon-induced 35 kDa protein homolog	-4.44	4.52E-09
NP_001013544.1	six-cysteine containing astacin protease 3 precursor	-4.46	7.55E-10
XP_685769.5	PREDICTED: mucin-5AC-like	-4.49	5.92E-03
XP_001921323.1	PREDICTED: protein phosphatase Slingshot homolog 2	-4.49	3.05E-02
NP_001133520.1	K0141 protein	-4.50	1.26E-06
NP_001003494.1	ubiquitin-conjugating enzyme E2 E2	-4.51	4.47E-03
NP_001076467.1	fos-related antigen 2	-4.52	6.50E-06
XP_689309.4	PREDICTED: poliovirus receptor-related protein 2-like	-4.56	1.47E-02
NP_001002648.1	solute carrier family 38, member 5b	-4.56	9.10E-15
XP_695496.5	PREDICTED: probable ATP-dependent RNA helicase YTHDC2-like	-4.58	1.87E-02
XP_003198001.1	PREDICTED: CD59 glycoprotein-like	-4.59	0.00E+00
NP_998109.1	transmembrane 6 superfamily member 1	-4.59	4.12E-02
NP_998461.1	tissue inhibitor of metalloproteinase 2b precursor	-4.60	4.85E-05
NP_001245246.1	im:7137886	-4.65	5.31E-05
NP_957457.2	asparagine synthetase	-4.66	2.22E-16
NP_001038857.1	basic leucine zipper transcriptional factor ATF-like 3	-4.67	1.27E-02
XP_002666630.1	PREDICTED: putative protein phosphatase 1 regulatory inhibitor subunit 3G-like	-4.67	5.55E-05
NP_957474.1	glycogen starch synthase, muscle	-4.70	1.09E-04
NP_001191098.1	si:rp71-1c23.2	-4.72	0.00E+00
NP_001018432.1	pleckstrin homology-like domain family A member 2	-4.72	2.39E-02
XP_700051.4	PREDICTED: si:ch211-197g15.7	-4.75	1.16E-03
XP_003199078.1	PREDICTED: RNA-binding protein with multiple splicing-like	-4.78	1.27E-02
NP_001007444.1	rho-related GTP-binding protein RhoU	-4.78	1.27E-02
NP_955932.2	mitochondrial import receptor subunit TOM34	-4.78	1.10E-02
XP_698749.5	PREDICTED: protein NLRC3	-4.81	2.50E-02
CAC28060.1	putative transposase	-4.81	4.34E-06
NP_001139067.1	L-2-hydroxyglutarate dehydrogenase, mitochondrial	-4.84	2.92E-02
NP_001025129.1	acyl-CoA:lysophosphatidylglycerol acyltransferase 1	-4.86	1.06E-03
NP_956125.1	DIRAS family, GTP-binding RAS-like 1a	-4.90	1.27E-02
NP_001004617.1	cytosolic acyl coenzyme A thioester hydrolase	-4.91	1.27E-03
NP_001035430.1	sulfatase-modifying factor 1 precursor	-4.92	6.60E-05
NP_998176.1	amylase-3 protein precursor	-4.93	1.26E-05
NP_001017769.1	high mobility group protein B3	-4.94	2.43E-03

NP_999862.1	reverse transcriptase-like protein	-4.96	6.85E-03
NP_001177684.1	neutral amino acid transporter B(0)	-4.97	8.63E-03
NP_001071007.1	sodium channel subunit beta-1 precursor	-4.98	1.47E-11
NP_001002168.1	ranBP-type and C3HC4-type zinc finger-containing protein 1	-5.04	1.46E-02
NP_571555.1	signal transducer and activator of transcription 1-alpha/beta	-5.05	4.24E-03
NP_001002176.1	threonine aldolase 1	-5.07	3.05E-03
NP_001122257.1	ubiquitin-like	-5.08	5.04E-04
NP_001035183.2	coproporphyrinogen-III oxidase, mitochondrial	-5.08	2.79E-04
NP_001018577.1	calcium-binding and coiled-coil domain-containing protein 2	-5.09	5.63E-08
XP_001343219.1	PREDICTED: sphingosine-1-phosphate phosphatase 1	-5.09	3.04E-02
NP_991247.1	heterogeneous nuclear ribonucleoprotein H1	-5.09	1.61E-02
XP_001922665.2	PREDICTED: serine/threonine-protein kinase TAO3-like isoform 1	-5.11	4.27E-06
NP_001070859.1	proto-oncogene serine/threonine-protein kinase pim-1	-5.14	1.24E-05
NP_998410.1	L-threonine 3-dehydrogenase, mitochondrial	-5.15	1.77E-02
XP_692922.3	PREDICTED: tumor necrosis factor alpha-induced protein 3	-5.16	1.45E-03
NP_571751.1	proteasome beta 11 subunit	-5.18	1.90E-09
XP_688551.2	PREDICTED: lipoamide kinase isozyme 3	-5.18	3.89E-02
NP_001041614.1	syndecan-4 precursor	-5.19	1.10E-11
NP_571148.1	tyrosine-protein kinase JAK1	-5.19	7.46E-05
NP_001035459.1	tumor necrosis factor receptor superfamily member 9 precursor	-5.20	7.23E-04
NP_001002217.1	carboxypeptidase A4 precursor	-5.22	4.63E-06
XP_003218673.1	PREDICTED: epithelial-stromal interaction protein 1-like	-5.22	8.58E-03
XP_002660417.1	PREDICTED: interferon-induced protein 44	-5.24	4.42E-03
XP_694014.4	PREDICTED: ankyrin repeat domain-containing protein 34B	-5.25	2.29E-03
NP_001138285.1	C-X-C chemokine receptor type 3-like	-5.30	6.00E-04
NP_571049.1	TAP binding protein precursor	-5.31	2.80E-12
NP_957258.1	cyclic AMP-dependent transcription factor ATF-3	-5.36	2.28E-03
NP_001001828.2	RING finger protein 114	-5.37	6.51E-08
XP_696367.4	PREDICTED: NACHT, LRR and PYD domains-containing protein 14	-5.37	1.37E-03
NP_001025437.1	inositol oxygenase	-5.40	0.00E+00
NP_998476.1	purine nucleoside phosphorylase 5a	-5.40	1.05E-07
NP_001041531.1	caspase 3, apoptosis-related cysteine protease b	-5.40	1.79E-06
NP_001002754.1	ras-related C3 botulinum toxin substrate 3a (rho family, small GTP binding protein Rac3)	-5.41	1.81E-03
NP_001073421.1	TNFAIP3-interacting protein 1	-5.48	4.73E-13
NP_878292.1	tissue factor pathway inhibitor precursor	-5.49	4.38E-03
NP_571449.1	proteasome activator complex subunit 2	-5.50	0.00E+00
NP_001108170.1	immunity-related GTPase family, f1	-5.51	6.09E-03
NP_955930.1	60S ribosomal protein L10a	-5.52	3.16E-03
NP_001037818.1	phosphatidate phosphatase LPIN1	-5.53	2.47E-02

XP_686458.2	PREDICTED: inositol-trisphosphate 3-kinase C isoform 1	-5.57	8.09E-03
XP_001921668.2	PREDICTED: aquaporin-11, partial	-5.57	1.50E-04
NP_001074030.1	ubiquitin-associated domain-containing protein 2	-5.58	3.29E-03
NP_001006594.1	ATP-binding cassette, sub-family B (MDR/TAP), member 3 like 1	-5.67	1.51E-02
NP_001001818.1	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 2a	-5.69	3.44E-14
XP_002666105.2	PREDICTED: protein POF1B	-5.69	3.15E-02
NP_956421.1	adenylyltransferase and sulfurtransferase MOCS3	-5.69	8.44E-03
NP_998469.1	suppressor of cytokine signaling 3b	-5.70	1.06E-10
NP_001092893.1	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase delta-3-A	-5.70	2.58E-02
NP_999886.1	guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-5	-5.70	2.58E-02
XP_001923366.3	PREDICTED: protein NLRC3-like	-5.75	4.48E-02
NP_001155979.1	ras-related protein Rab-33B	-5.77	8.44E-04
NP_998358.1	2-aminoethanethiol (cysteamine) dioxygenase b	-5.81	5.20E-05
NP_919379.1	anti-apoptotic protein NR13	-5.83	0.00E+00
XP_002663242.2	PREDICTED: spermatogenesis-associated protein 5-like, partial	-5.83	5.94E-05
NP_571467.2	proteasome subunit beta type-8	-5.88	0.00E+00
NP_001004120.1	nuclear factor interleukin-3-regulated protein	-5.89	1.29E-07
NP_956281.1	transcription factor AP-1	-5.92	1.99E-11
XP_700448.2	PREDICTED: phospholipase A2, major isoenzyme	-5.95	3.12E-05
NP_998262.1	85 kDa calcium-independent phospholipase A2	-5.96	5.19E-03
XP_002661398.1	PREDICTED: ubl carboxyl-terminal hydrolase 18-like	-5.97	6.27E-03
XP_001338123.1	PREDICTED: junctional adhesion molecule C	-5.99	3.38E-03
NP_001025262.1	protein kinase C zeta type	-6.01	4.68E-02
NP_001076331.1	UPF0249 protein ydjC homolog	-6.01	4.68E-02
XP_690266.2	PREDICTED: monocarboxylate transporter 7-like	-6.04	3.89E-03
NP_955905.1	arginase-2, mitochondrial	-6.08	4.24E-11
NP_997743.1	mini-chromosome maintenance complex-binding protein	-6.11	2.31E-02
NP_001004582.1	chymotrypsin-like precursor	-6.24	0.00E+00
NP_001132910.1	acyl-CoA synthetase family member 2, mitochondrial precursor	-6.26	4.11E-02
XP_002664795.2	aldose reductase	-6.26	3.15E-03
XP_001335167.1	PREDICTED: syncollin-like	-6.27	0.00E+00
NP_001036185.1	copper-transporting ATPase 1	-6.29	1.05E-02
NP_001189369.1	E3 ubiquitin-protein ligase RNF19B	-6.30	5.39E-04
XP_001340407.1	PREDICTED: thiosulfate sulfurtransferase/rhodanese-like domain-containing protein 1-like	-6.32	4.89E-02
XP_002665053.1	PREDICTED: antigen peptide transporter 1-like	-6.35	2.75E-05
NP_919365.1	fructose-bisphosphate aldolase C-B	-6.39	0.00E+00
NP_570993.2	frizzled homolog 8a	-6.40	1.53E-02
NP_001082991.1	nucleoside diphosphate-linked moiety X motif 6	-6.41	5.18E-12
NP_001038687.1	peptidoglycan recognition protein 6	-6.42	7.45E-03

NP_001070115.1	XIAP-associated factor 1	-6.47	6.05E-08
XP_001920864.3	transposase	-6.47	4.59E-02
NP_956874.1	solute carrier family 25 member 34	-6.52	1.70E-02
XP_002667032.2	PREDICTED: myotubularin-related protein 4	-6.54	1.33E-02
NP_957217.2	very low-density lipoprotein receptor precursor	-6.54	1.33E-02
ADN37680.1	hyopthetical protein	-6.55	1.43E-05
NP_998259.1	glyceraldehyde 3-phosphate dehydrogenase 2	-6.55	0.00E+00
AAT52140.1	CC chemokine SCYA106	-6.57	0.00E+00
NP_001076493.1	transcription factor ETV6	-6.61	4.17E-02
NP_956931.1	disintegrin and metalloproteinase domain-containing protein 8 precursor	-6.66	2.98E-03
XP_694358.2	PREDICTED: pentraxin-related protein PTX3	-6.70	5.76E-04
NP_001104716.2	nicotinamide N-methyltransferase	-6.73	1.86E-02
NP_001018589.1	prolyl endopeptidase	-6.74	8.85E-12
NP_955809.2	PR domain zinc finger protein 1	-6.76	2.70E-03
NP_571475.1	thyrotrophic embryonic factor	-6.79	1.44E-02
NP_001076519.1	regulator of microtubule dynamics protein 3	-6.80	9.90E-05
XP_002662830.1	PREDICTED: lymphocyte antigen 75	-6.80	2.26E-02
XP_688915.1	PREDICTED: tripartite motif-containing protein 16	-6.80	1.08E-02
NP_963882.1	paxillin	-6.86	8.06E-03
NP_001014328.1	leukemia inhibitory factor receptor alpha a precursor	-6.91	4.58E-02
NP_998513.2	protein NDRG1 isoform 2	-6.92	7.66E-03
XP_690447.3	PREDICTED: tumor necrosis factor receptor superfamily member 5	-7.03	1.55E-14
AAG27458.1	AF305694_1 latent nuclear antigen	-7.04	1.93E-02
NP_998044.1	interferon regulatory factor 10	-7.10	1.44E-02
XP_002663310.2	PREDICTED: periphilin-1-like	-7.13	4.11E-02
XP_001920559.3	PREDICTED: poly ADP-ribose polymerase 15	-7.15	1.07E-02
XP_001919284.3	PREDICTED: MTSS1-like protein-like	-7.21	1.74E-02
NP_999856.1	prolyl 4-hydroxylase, alpha polypeptide I b precursor	-7.24	2.23E-02
NP_998009.1	interleukin 1, beta	-7.30	1.52E-05
NP_001003426.1	elastase-like precursor	-7.30	0.00E+00
NP_957140.1	ADP-ribosylation factor-like 5C	-7.31	2.14E-07
NP_001038320.2	vascular endothelial growth factor A precursor	-7.42	2.01E-02
NP_001017853.1	chromobox homolog 7a	-7.45	5.28E-09
NP_001076500.1	autophagy-related protein 9A	-7.46	1.88E-05
NP_571783.1	cationic trypsin-3 precursor	-7.50	0.00E+00
NP_001070770.2	serum/glucocorticoid regulated kinase 1-like	-7.50	0.00E+00
NP_899190.2	cytosolic sulfotransferase 2	-7.54	3.42E-02
NP_001138718.1	CD40 antigen precursor	-7.57	4.40E-10
NP_001116329.1	DAB2 interacting protein	-7.67	3.14E-02

XP_002666567.2	PREDICTED: pleckstrin homology domain-containing family G member 5-like	-7.67	3.14E-02
NP_001001826.1	protein CYR61 precursor	-7.68	2.32E-02
NP_001075159.1	chymotrypsinogen 2-like precursor	-7.75	0.00E+00
NP_001185677.1	normal mucosa of esophagus-specific gene 1	-7.79	2.33E-13
NP_571321.1	L-lactate dehydrogenase A chain	-7.83	0.00E+00
NP_956370.1	methionine--tRNA ligase, cytoplasmic	-7.96	0.00E+00
NP_958488.1	retinol dehydrogenase 10-B	-7.99	2.68E-02
NP_001073470.1	tuftelin	-8.08	4.17E-14
AAZ99069.1	CC chemokine SCYA114B	-8.09	1.85E-07
NP_998475.1	egl nine homolog 3	-8.18	2.00E-07
NP_570997.1	protein BTG2	-8.19	0.00E+00
NP_001007429.1	cell adhesion molecule 4 precursor	-8.26	3.26E-02
AAX25877.2	SJCHGC09533 protein	-8.28	0.00E+00
NP_956180.1	elastase 2 like precursor	-8.29	0.00E+00
XP_688254.3	collectin-11 precursor	-8.35	1.19E-02
NP_001188367.1	arrestin domain containing 3-like	-8.36	5.51E-05
NP_001171406.2	type III iodothyronine deiodinase	-8.42	0.00E+00
XP_700028.3	PREDICTED: AMP deaminase 2	-8.46	1.89E-05
NP_001124100.1	bloodthirsty-related gene family, member 9	-8.47	4.20E-02
XP_689435.4	PREDICTED: 6-phosphofructokinase type C isoform 5	-8.47	4.20E-02
NP_571543.1	peroxisome proliferator-activated receptor delta b	-8.50	4.22E-02
NP_571370.1	cytosolic phospholipase A2	-8.50	2.93E-02
XP_003441202.1	PREDICTED: protein DEPP-like	-8.54	3.83E-03
NP_991310.1	interferon regulatory factor 1	-8.54	5.30E-09
XP_002663077.1	PREDICTED: alkaline ceramidase 2-like	-8.54	3.86E-09
XP_003452034.1	PREDICTED: transmembrane protein 150A-like	-8.65	4.88E-04
XP_002666175.2	PREDICTED: AMP deaminase 2-like	-8.72	1.04E-07
NP_001004578.1	EH domain-containing protein 1	-8.75	2.63E-02
NP_001012371.1	hypoxia-inducible factor 1, alpha subunit, like 2	-8.75	1.84E-02
NP_001239559.1	suppressor of cytokine signaling 1b	-8.80	1.14E-03
NP_571739.1	LIM domain and actin-binding protein 1	-8.84	2.87E-04
NP_001076405.1	cytochrome P450, family 2, subfamily AD, polypeptide 6	-8.95	3.54E-02
NP_001018443.1	caspase-7	-9.02	3.37E-02
NP_997954.2	type II iodothyronine deiodinase	-9.05	4.97E-02
NP_001103491.1	carboxypeptidase B	-9.05	0.00E+00
NP_957145.1	protein N-terminal asparagine amidohydrolase	-9.07	4.86E-02
NP_991230.1	small nuclear ribonucleoprotein polypeptides B and B1	-9.09	9.36E-07
NP_955977.1	eukaryotic translation initiation factor 4E-binding protein 3-like	-9.13	0.00E+00
NP_001038662.1	arginase-1	-9.25	3.15E-02

XP_698635.3	PREDICTED: 6-phosphofructokinase, liver type-like	-9.60	1.80E-02
NP_001132936.1	elastase 2 precursor	-9.65	0.00E+00
NP_991122.2	lysophospholipid acyltransferase LPCAT4	-9.66	7.83E-04
XP_003197852.1	PREDICTED: NAD-dependent deacetylase sirtuin-5-like	-9.88	3.68E-02
NP_001092722.1	testis-expressed sequence 2 protein	-9.90	0.00E+00
NP_001074052.1	retinol dehydrogenase 10-A	-9.99	1.31E-03
NP_956163.1	insulin-induced gene 1 protein	-10.11	2.98E-09
NP_001038939.1	phorbol-12-myristate-13-acetate-induced protein 1	-10.25	9.41E-05
NP_991172.1	thromboxane-A synthase	-10.44	3.09E-02
NP_001003467.1	suppressor of cytokine signaling 1	-10.47	0.00E+00
NP_001012372.1	hepatoma-derived growth factor-related protein 3	-10.49	4.74E-02
XP_684417.1	PREDICTED: ADM-like	-10.51	9.34E-12
XP_690395.4	PREDICTED: thrombospondin-1	-10.60	3.08E-03
NP_001182713.2	fep15 selenoprotein precursor	-10.64	6.31E-05
XP_001341936.1	Transposable element Tcb2 transposase	-10.67	7.35E-03
P41386.1	TBB_HALDI RecName: Full=Tubulin beta chain; AltName: Full=Beta-tubulin	-10.69	7.87E-04
AAH95653.1	Si:busm1-6a2.1 protein, partial	-10.90	6.40E-03
NP_001070091.1	cAMP-dependent protein kinase type I-beta regulatory subunit	-11.07	4.08E-02
NP_001122149.1	prolactin receptor precursor	-11.17	3.51E-03
NP_001108058.2	ATP-binding cassette, sub-family A (ABC1), member 1B	-11.20	0.00E+00
NP_001070771.1	filamin-binding LIM protein 1	-11.50	2.13E-02
NP_001002340.1	claudin-7-A	-11.52	1.28E-02
NP_001073171.1	tyrosyl-DNA phosphodiesterase 2	-11.52	1.28E-02
NP_998562.1	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3	-11.54	0.00E+00
NP_001243322.1	low-density lipoprotein receptor-related protein 8 precursor	-11.63	3.42E-07
XP_002663487.2	PREDICTED: ankyrin repeat domain-containing protein 57	-12.05	3.04E-02
NP_956401.1	DNA damage-inducible transcript 4 protein	-12.08	4.32E-12
XP_003199913.1	PREDICTED: transmembrane emp24 domain-containing protein 11-like	-12.15	1.71E-02
NP_956244.1	suppressor of cytokine signaling 3a	-12.21	0.00E+00
XP_003992662.1	PREDICTED: hephaestin-like protein 1	-12.27	5.61E-03
NP_001020344.1	carboxyl ester lipase, like precursor	-12.29	0.00E+00
NP_001093209.1	acyl-CoA synthetase long-chain family member 4, like	-12.40	0.00E+00
NP_001003552.1	solute carrier family 16 (monocarboxylic acid transporters), member 9	-12.44	1.79E-03
NP_001161936.1	tumor necrosis factor ligand superfamily member 13	-12.63	3.92E-05
NP_571483.1	vascular endothelial growth factor A-A isoform 1 precursor	-12.63	4.61E-02
NP_001122048.1	leptin precursor	-12.96	0.00E+00
NP_001074151.1	leucine-rich repeat-containing protein 15 precursor	-13.02	2.31E-02
XP_002660612.1	PREDICTED: SERTA domain-containing protein 2	-13.08	2.38E-04
XP_003199953.1	replication factor C subunit 5	-13.31	3.99E-02

NP_001153492.1	calcium-binding mitochondrial carrier protein SCaMC-2-B	-13.40	3.52E-04
NP_001014828.1	prostaglandin E synthase	-13.73	2.68E-05
XP_001339993.2	PREDICTED: interferon-induced protein 44-like	-13.75	6.66E-16
XP_001341326.3	PREDICTED: ubiquitin-like modifier-activating enzyme 1-like	-13.81	1.73E-14
NP_001091654.1	A-kinase anchor protein 12	-14.04	1.41E-03
NP_001018529.1	cytosolic beta-glucosidase	-14.09	3.33E-02
NP_001002343.1	guanylate binding protein 1	-14.14	0.00E+00
NP_958458.1	fragile X mental retardation syndrome-related protein 1	-14.19	1.66E-02
XP_002661386.1	PREDICTED: nicotinamide phosphoribosyltransferase-like	-14.25	5.20E-06
NP_001159702.1	WNT1 inducible signaling pathway protein 1a precursor	-14.27	8.63E-03
NP_955899.2	trypsinogen precursor	-14.52	0.00E+00
NP_001123607.1	adrenomedullin 2a precursor	-14.98	6.08E-14
NP_001018318.1	carboxypeptidase A1 (pancreatic) precursor	-16.04	0.00E+00
NP_001018333.1	caspase-6	-16.21	1.38E-14
NP_001002595.2	egl nine homolog 1	-16.38	2.01E-03
XP_003200135.1	ST3 beta-galactoside alpha-2,3-sialyltransferase 5, like	-16.70	4.12E-02
XP_683729.3	PREDICTED: protein FAM13A isoform 1	-16.84	3.98E-02
NP_955913.1	cysteine/serine-rich nuclear protein 1	-17.04	2.55E-10
NP_991254.1	growth arrest and DNA damage-inducible protein GADD45 gamma	-17.13	6.98E-04
CCD11006.1	chemokine CXCL-F5	-17.63	1.26E-03
NP_001188418.1	EF-hand domain-containing family member B	-17.75	1.45E-02
NP_001165872.1	plasminogen activator, urokinase a precursor	-17.76	4.89E-05
XP_002662574.1	PREDICTED: solute carrier family 2, facilitated glucose transporter member 1	-18.62	4.44E-16
XP_003444689.1	PREDICTED: protocadherin-10-like	-19.68	9.34E-03
NP_001017559.1	MORN repeat-containing protein 4	-20.08	2.27E-02
NP_956165.1	S-adenosylmethionine synthase isoform type-1	-20.71	0.00E+00
NP_001070802.1	hepatic leukemia factor	-20.73	5.87E-05
NP_956372.1	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 4b	-21.05	1.93E-02
NP_001073523.1	zinc finger protein 395a	-21.96	1.97E-03
NP_998655.1	tubulin, beta 2B	-22.35	2.17E-04
NP_001129726.1	chemokine (C-C motif) ligand 20-like	-22.35	1.55E-02
NP_001015041.1	connective tissue growth factor precursor	-22.70	2.90E-06
NP_001120988.1	heme oxygenase 1	-23.48	0.00E+00
NP_998721.2	transcription factor jun-B	-23.63	1.31E-11
NP_001108031.1	plasminogen activator inhibitor 1 precursor	-23.81	4.61E-06
NP_001018629.1	interferon, gamma 1-1 precursor	-23.81	4.05E-02
XP_003457928.1	PREDICTED: nuclear factor 7, ovary-like	-23.81	4.05E-02
NP_998439.1	claudin 5 precursor	-24.62	1.06E-02
NP_997783.1	chymotrypsin B1 precursor	-25.12	0.00E+00

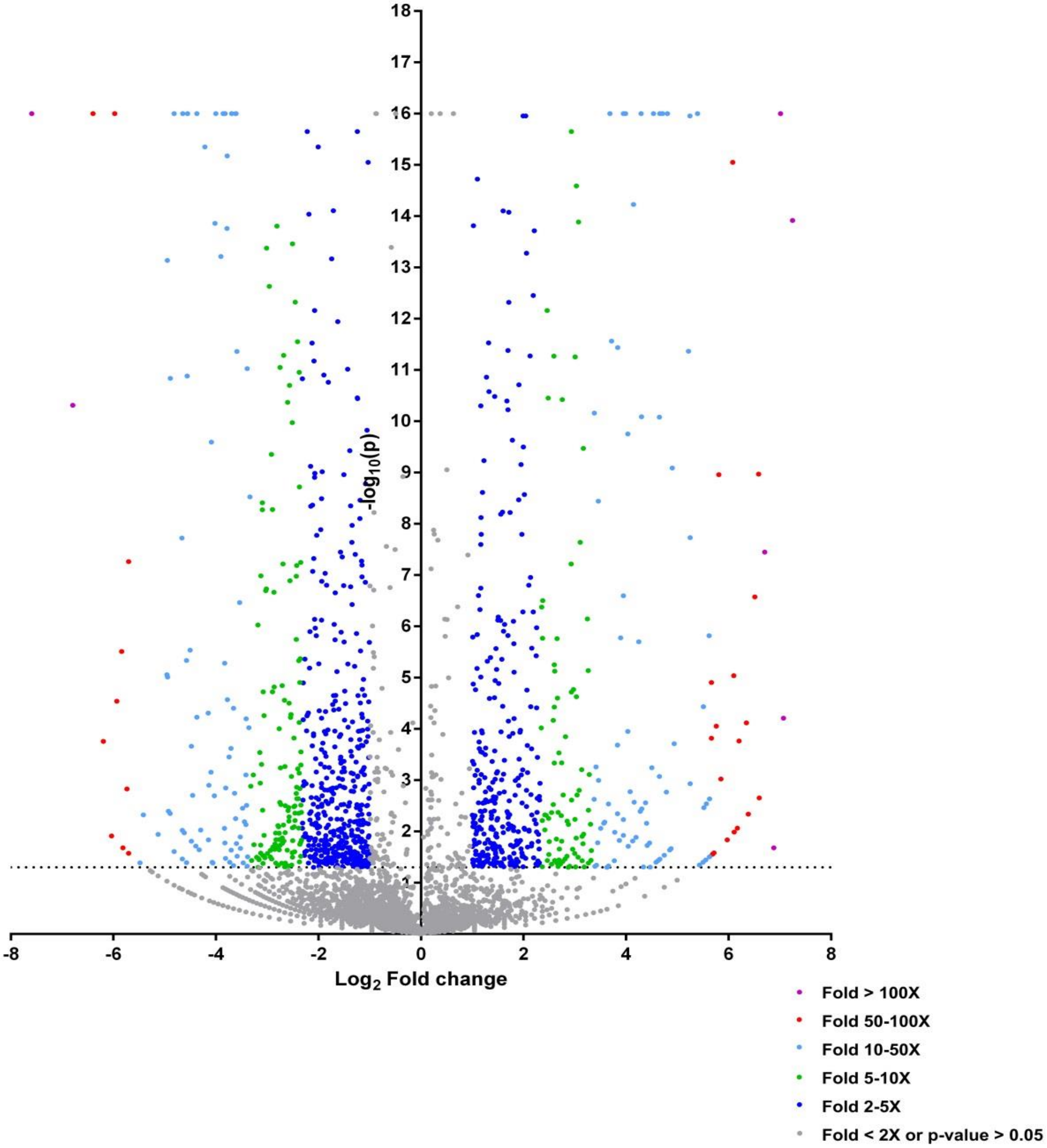
NP_001002190.1	response gene to complement 32 protein	-25.26	9.47E-03
XP_689109.3	PREDICTED: lymphatic vessel endothelial hyaluronic acid receptor 1	-25.26	3.41E-02
XP_002665479.1	PREDICTED: tumor necrosis factor receptor superfamily member 11B	-25.40	1.88E-08
NP_998105.2	ligand of numb-protein X 2b	-28.19	2.45E-02
NP_001116443.1	myxovirus (influenza virus) resistance G	-28.22	0.00E+00
XP_001923114.3	PREDICTED: tumor necrosis factor receptor superfamily member 5-like	-29.67	1.45E-11
XP_003197671.1	PREDICTED: poliovirus receptor-related protein 3-like	-29.80	4.46E-03
XP_001342072.1	PREDICTED: RIKEN cDNA 9130211I03-like isoform 1	-30.45	4.01E-03
XP_003201233.1	PREDICTED: protein FAM110A isoform 3	-30.82	9.66E-06
AAP49829.1	interferon-inducible protein Gig2	-30.85	7.24E-14
CBN81178.1	Pol polyprotein	-31.09	8.73E-06
NP_001035079.1	DNA-binding protein inhibitor ID-4	-34.97	1.14E-02
NP_001155024.1	fos-related antigen 1	-42.75	4.74E-03
NP_001029093.1	glucosamine--fructose-6-phosphate aminotransferase isomerizing 2	-44.70	4.09E-02
NP_705943.1	prostaglandin-endoperoxide synthase 2 precursor	-52.15	5.45E-08
NP_001038207.1	immune-related, lectin-like receptor 3 isoform b	-52.20	2.65E-02
NP_001107915.1	bloodthirsty-related gene family, member 18	-53.44	1.48E-03
CAD60788.1	novel protein similar to DNA polymerases	-56.36	2.08E-02
NP_001104630.1	prostacyclin synthase	-57.33	3.07E-06
CAG14936.1	interleukin-11	-61.22	2.88E-05
NP_997873.1	monocarboxylate transporter 4	-62.94	0.00E+00
NP_001243120.1	nuclear receptor subfamily 0 group B member 2	-65.77	1.22E-02
NP_878294.1	metalloproteinase inhibitor 2 precursor	-73.36	1.74E-04
NP_001076375.2	chemokine CCL-C5a precursor	-84.35	0.00E+00
ABA54956.1	CC chemokine SCYA109	-111.02	4.86E-11
NP_998115.1	parathyroid hormone 1a	-193.37	0.00E+00

Table 5. Comparison of gene expression between resistant fish and susceptible fish after infection with *Edwardsiella ictaluri*.

Differentially expressed genes between resistant and susceptible fish	1,255
Gene expressed higher in resistant fish	528
>100 fold	5

50-100 fold	18
10-50 fold	86
5-10 fold	94
2-5 fold	325
Genes expressed lower in resistant fish	727
>100 fold	2
50-100 fold	10
10-50 fold	86
5-10 fold	159
2-5 fold	470

Figure 1. Volcano plot of genes differentially expressed between resistant and susceptible fish. The dots located in the positive area stand for genes expressed higher in resistant fish, and dots located in the negative area stand for genes expressed higher in susceptible fish. As shown in graphic symbol, different color were used to scale different expression fold changes; purple stands for expression fold changes higher than 100-fold; red stands for expression fold changes from 50-100 fold; light blue stands for expression fold changes from 10-50 fold; green stands for expression fold changes from 5-10 fold; blue stands for expression fold changes from 2-5 fold; and ray stands for gene expressed insignificantly (p-value>0.05 or fold change smaller than 2)



III.4 Identification of SNPs and significant SNPs

SNPs were identified by alignment of short reads to the reference assembly of the RNA-seq. A quality control was performed for SNPs, at least 6 reads were required for each group (resistant, susceptible, or control) and a total of minor allele reads count must be greater than 3 among three groups. A total of 513,371 SNPs were identified. These SNPs were located within 31,646 contigs. In order to determine SNPs with significant difference in allele frequencies between the resistant and susceptible fish (significant SNPs), Fisher's Exact test was performed. As shown in Table 6, 56,419 SNPs were identified as significant SNPs. These significant SNPs were located in 11,249 contigs. Of the 11,249 contigs, 5,480 had significant hits to known genes, and the remaining probably represented untranslated regions. The 5,480 contigs with hits to known genes represented 4,304 unique genes.

Table 6. Identification of SNPs and significant SNPs (allele frequencies statistically different between the resistant and susceptible groups) from the assembled catfish liver transcriptome

Total number of SNPs	513,371
Number of contigs containing SNPs	31,646
Number of significant SNPs	56,419 (10.99%)
Number of contigs containing significant SNPs	11,249
Number of contigs with significant hits to genes	5,480
Number of genes containing significant SNPs	4,304
Number of genomic scaffolds containing significant SNPs	2,096

III.5 Bulk frequency ratios

Although significant SNPs identified through RNA-seq analysis reflect the final ratios of different alleles at the RNA level in the two bulked samples, statistical analysis for significant SNPs using Fisher's Exact test was only the first step for screening SNPs that may be significantly associated with the trait. In order to compare the SNP allele frequencies more directly, bulk frequency ratios (BFR) were generated from the RNA-seq data between the two bulks, the resistant fish and the susceptible fish. The BFR of genes were determined by the maximum BFR of the significant SNPs generated from the Fisher's Exact test located in this gene. As shown in Figure 2, large proportion of genes containing the significant SNPs had a BFR of at least 2. A total of 359 genes had a BFR equal or greater than 4. Among these genes, 337 (93.9 %) genes had a BFR of 4-16; 23 genes had a BFR over 16; and 4 genes had a BFR over 32 (Figure 2). The four genes had the highest BFR are multidrug resistance-associated protein 5, tumor suppressor candidate 5 homolog (Interferon-induced transmembrane protein), uncharacterized protein LOC101157921, and DnaJ subfamily A member 2. The additional 19 genes with the largest BFR of over 16 are myoglobin, cytosolic phospholipase A2, metalloreductase STEAP2, suppression of tumorigenicity 5 protein, protein G7c-like, Beta-tubulin, purine nucleoside phosphorylase 4b, EF-hand domain-containing protein D2, Serum amyloid P-component precursor, plasminogen activator inhibitor 1 precursor, si:dkey-269d20.3, transmembrane protein C9orf125, MHC class II beta, bone morphogenetic protein 1a, interferon-induced very large GTPase 1-like, aldehyde dehydrogenase family 9 member A1-A, CC chemokine SCYA108, RNA-binding protein 8A and 40S ribosomal protein S3a (Figure 2). A list of the genes with various BFR values was listed in Table 7.

Table 7. BFR (⩾4) and combined allele ratios of genes containing significant SNPs.

Max BFR	Combined Allel Ratio	ID	Description	Rank of RPKM
4.28	1.23	NP_001014356.2	CASP8 associated protein 2 [Danio rerio]	11
5.06	1.44	NP_571474.1	retinoic acid receptor alpha-B [Danio rerio]	21
4.44	1.80	XP_003201798.1	PREDICTED: hypothetical protein LOC100535050 [Danio rerio]	26
4.81	1.92	XP_003442816.1	PREDICTED: hypothetical protein LOC100709414 [Oreochromis niloticus]	27
4.38	1.44	NP_001003543.1	uncharacterized protein LOC445149 [Danio rerio]	38
8.84	1.22	NP_001002679.1	ras-related protein Rab-3B [Danio rerio]	39
4.76	2.38	NP_001077282.1	brevican core protein precursor [Danio rerio]	43
6.67	2.20	NP_001177949.1	proteasome (prosome, macropain) activator subunit 4b [Danio rerio]	49
4.29	2.91	XP_003460254.1	PREDICTED: zinc finger protein 729-like [Oreochromis niloticus]	62
4.00	1.14	XP_692362.2	PREDICTED: anthrax toxin receptor 1 [Danio rerio]	66
4.55	1.56	NP_001071068.1	protein phosphatase 1L [Danio rerio]	69
4.44	1.80	NP_998133.1	protein HEG precursor [Danio rerio]	70
4.05	1.11	NP_001091661.1	uncharacterized protein LOC100000539 [Danio rerio]	72
4.50	2.17	XP_003978669.1	PREDICTED: uncharacterized protein LOC101071928 [Takifugu rubripes]	78
4.41	1.75	XP_696087.1	PREDICTED: transcription elongation factor A N-terminal and central domain-containing protein 2 [Danio rerio]	80
4.90	1.67	NP_001018499.1	BTB and CNC homology 1 transcription factor [Danio rerio]	87
5.25	2.21	NP_001120984.1	HMG domain-containing protein 4 [Danio rerio]	93
5.06	1.13	NP_001104700.1	probable tubulin polyglutamylase TTLL2 [Danio rerio]	98
5.87	1.09	F8S114.1	VSPCR_CROAD RecName: Full=Thrombin-like enzyme crotalase; Short=SVTLE; AltName: Full=Fibrinogen-clotting enzyme; AltName: Full=Snake venom serine protease 2; Short=SVSP; Flags: Precursor	99
9.00	3.23	NP_001116775.1	uncharacterized protein LOC791704 precursor [Danio rerio]	113
5.10	1.91	XP_700731.5	PREDICTED: neurabin-1-like [Danio rerio]	114
6.00	2.71	XP_001923040.3	PREDICTED: poliovirus receptor [Danio rerio]	119
6.23	1.46	XP_003454912.1	PREDICTED: uncharacterized protein C3orf23-like [Oreochromis niloticus]	124
5.37	1.94	NP_001103940.1	protein WWC3 [Danio rerio]	129
4.57	1.44	NP_991118.1	probable prolyl-tRNA synthetase, mitochondrial [Danio rerio]	132
7.78	1.15	XP_003200274.1	PREDICTED: protein AMBP-like [Danio rerio]	147
4.58	2.20	NP_997830.1	DnaJ subfamily A member 2-like [Danio rerio]	177
6.57	4.00	XP_003445224.1	PREDICTED: hypothetical protein LOC100709243 [Oreochromis niloticus]	183
5.00	2.00	XP_684227.1	PREDICTED: sarcoplasmic/endoplasmic reticulum calcium ATPase 2 isoform 1 [Danio rerio]	190
7.90	1.93	ZP_04645459.1	pG1 protein [Lactobacillus jensenii 269-3]	191
7.84	2.94	NP_878277.1	period circadian protein homolog 2 [Danio rerio]	194
4.25	1.54	NP_998228.1	WD repeat-containing protein 24 [Danio rerio]	198
6.60	2.00	XP_688194.1	PREDICTED: polypeptide N-acetylgalactosaminyltransferase 12 [Danio rerio]	210
6.25	2.40	NP_001006031.1	retinol dehydrogenase 14a (all-trans/9-cis/11-cis) [Danio rerio]	227
4.90	1.10	XP_692676.1	PREDICTED: polyamine-modulated factor 1 [Danio rerio]	236

14.44	2.07	XP_003200093.1	PREDICTED: sacsin-like [Danio rerio]	237
8.18	2.42	NP_001091660.1	N-acetylated-alpha-linked acidic dipeptidase-like protein [Danio rerio]	240
4.33	2.54	XP_003443838.1	PREDICTED: uncharacterized protein C1orf172 homolog [Oreochromis niloticus]	245
4.06	2.00	NP_001091715.1	B-cell receptor CD22 precursor [Danio rerio]	246
11.18	5.50	NP_001038650.1	endothelin 2 precursor [Danio rerio]	249
8.73	2.58	NP_001073558.1	NF-kappa-B inhibitor epsilon [Danio rerio]	258
7.88	3.33	NP_001019243.1	integrin beta-like protein 1 precursor [Danio rerio]	266
4.21	2.67	XP_003442194.1	PREDICTED: cyclin-dependent kinase 12-like [Oreochromis niloticus]	277
5.70	3.33	XP_001345254.4	PREDICTED: titin [Danio rerio]	285
4.96	1.50	AEJ32000.1	serine proteinase 6 [Crotalus adamanteus]	295
5.86	2.17	NP_956579.1	proton-coupled folate transporter [Danio rerio]	297
5.06	1.32	NP_001070849.1	NXPE family member 3 precursor [Danio rerio]	302
5.87	2.71	XP_001344847.2	PREDICTED: mitochondrial intermediate peptidase [Danio rerio]	346
5.09	1.29	NP_001025407.1	coronin-7 [Danio rerio]	357
5.34	1.55	NP_001075100.1	Rho guanine nucleotide exchange factor (GEF) 1a [Danio rerio]	374
10.15	1.35	NP_001005934.1	glycerol-3-phosphate dehydrogenase 1-like protein [Danio rerio]	380
6.07	2.26	NP_956140.1	tyrosine-protein phosphatase non-receptor type 11 [Danio rerio]	385
4.70	1.81	ADO28274.1	glyoxal reductase [Ictalurus furcatus]	402
6.16	1.63	NP_001004543.1	tyrosine-protein kinase Lyn [Danio rerio]	408
4.47	1.07	NP_997815.2	zinc finger and BTB domain-containing protein 11 [Danio rerio]	418
5.88	2.06	XP_001344126.2	PREDICTED: TNFAIP3-interacting protein 1-like [Danio rerio]	424
4.60	1.38	NP_001007344.1	sorting and assembly machinery component 50 homolog B [Danio rerio]	428
6.27	1.15	XP_002662144.1	PREDICTED: tetratricopeptide repeat protein 37-like [Danio rerio]	430
8.85	1.58	NP_001005293.1	synembryn-B [Danio rerio]	431
5.46	1.82	XP_003200204.1	PREDICTED: integrin alpha-10-like [Danio rerio]	434
4.07	1.75	XP_002666958.1	PREDICTED: beta-2-syntrophin-like isoform 1 [Danio rerio]	435
4.05	1.09	NP_001082837.2	protein RRP5 homolog [Danio rerio]	439
4.60	1.90	NP_001116100.1	uncharacterized protein LOC100142653 [Danio rerio]	440
5.75	2.36	NP_956549.1	DNA methyltransferase 1-associated protein 1 [Danio rerio]	441
5.93	1.81	XP_692914.1	PREDICTED: lysine-specific demethylase 7A [Danio rerio]	448
7.82	3.77	NP_001035469.1	methionine adenosyltransferase II, alpha-like [Danio rerio]	459
8.84	2.36	NP_001191199.1	poly (ADP-ribose) polymerase family, member 2 [Danio rerio]	463
20.00	6.08	NP_956880.1	myoglobin [Danio rerio]	504
5.67	1.10	NP_956913.2	UPF0668 protein C10orf76 homolog [Danio rerio]	530
6.39	1.75	XP_700224.3	PREDICTED: hypothetical protein LOC571530 [Danio rerio]	551
6.91	1.92	NP_956947.2	transcriptional regulator ATRX [Danio rerio]	573
6.01	1.76	NP_001002536.1	uncharacterized protein LOC436809 [Danio rerio]	593
6.71	1.76	XP_003198205.1	PREDICTED: BAI1-associated protein 3-like [Danio rerio]	594
4.24	1.07	NP_001005993.1	BRCA1-A complex subunit Abraxas [Danio rerio]	601

4.32	1.08	XP_688866.2	PREDICTED: TIM21-like protein, mitochondrial-like [Danio rerio]	602
7.50	2.43	NP_956468.2	uridine 5'-monophosphate synthase [Danio rerio]	613
4.85	2.31	NP_001002616.1	protein pellino homolog 2 [Danio rerio]	623
4.92	2.26	NP_001005600.1	trans-1,2-dihydrobenzene-1,2-diol dehydrogenase [Danio rerio]	634
4.60	1.38	NP_001107107.1	MAP7 domain containing 1 [Danio rerio]	645
6.00	1.05	XP_709654.2	PREDICTED: u3 small nucleolar ribonucleoprotein protein MPP10 isoform 2 [Danio rerio]	655
4.88	1.21	ACQ59049.1	THAP domain-containing protein 4 [Anoplopoma fimbria]	658
5.06	1.06	NP_001008585.1	large neutral amino acids transporter small subunit 4 [Danio rerio]	676
4.91	2.62	NP_001107268.1	exocyst complex component 5 [Danio rerio]	681
4.21	1.64	NP_001154919.2	nipped-B-like protein B [Danio rerio]	712
7.22	2.50	NP_998550.1	serine/threonine-protein kinase VRK1 [Danio rerio]	717
7.50	1.86	XP_002667251.2	PREDICTED: sugar phosphate exchanger 3 [Danio rerio]	726
6.60	1.24	NP_001004597.1	disrupted in renal carcinoma 2 [Danio rerio]	730
5.12	1.37	NP_001189381.1	chromodomain-helicase-DNA-binding protein 8 [Danio rerio]	740
4.27	1.78	NP_001038492.1	C-C chemokine receptor family-like [Danio rerio]	742
5.01	2.22	NP_001028889.2	adenosine deaminase CECR1-A precursor [Danio rerio]	743
13.06	3.00	NP_001073660.2	spermatogenesis-associated protein 2 [Danio rerio]	751
4.15	1.43	NP_001028893.1	integrin alpha-V precursor [Danio rerio]	760
5.63	1.88	XP_002662798.2	PREDICTED: hypothetical protein LOC100334463 [Danio rerio]	769
7.20	1.61	NP_956091.1	N-myc downstream regulated family member 3b [Danio rerio]	774
4.91	2.27	XP_001338803.3	PREDICTED: tumor protein p53-inducible protein 13-like [Danio rerio]	783
7.19	3.43	XP_001336902.1	PREDICTED: baculoviral IAP repeat-containing protein 6 [Danio rerio]	788
4.98	2.93	NP_001038584.1	INO80 complex homolog 1 [Danio rerio]	795
7.53	1.90	NP_001038745.1	plakophilin 3 [Danio rerio]	796
4.44	1.53	NP_001002612.1	uncharacterized protein LOC436885 [Danio rerio]	797
5.40	2.50	NP_001002646.1	adenosine deaminase [Danio rerio]	807
6.23	1.13	NP_001139062.1	si:ch211-206k20.5 [Danio rerio]	810
16.80	7.25	NP_571370.1	cytosolic phospholipase A2 [Danio rerio]	825
4.38	1.36	XP_002666500.1	PREDICTED: UPF0640 protein C3orf78 homolog [Danio rerio]	829
5.11	1.53	NP_001003983.1	cadherin-5 precursor [Danio rerio]	851
4.57	1.88	NP_955888.1	transmembrane protein 177 [Danio rerio]	859
4.14	1.25	NP_775350.1	telomeric repeat-binding factor 2 [Danio rerio]	869
4.26	1.31	XP_695742.5	PREDICTED: sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1-like [Danio rerio]	888
6.15	1.01	XP_001920363.2	PREDICTED: b-cell receptor CD22-like [Danio rerio]	890
5.37	1.13	XP_001921194.2	PREDICTED: inositol 1,4,5-trisphosphate receptor type 1 [Danio rerio]	897
8.64	2.36	XP_001922696.1	PREDICTED: stonustoxin subunit beta [Danio rerio]	911
4.13	2.27	XP_694124.3	PREDICTED: interferon-induced helicase C domain-containing protein 1 [Danio rerio]	926
5.93	1.22	NP_955857.1	deoxyhypusine hydroxylase [Danio rerio]	942
4.18	1.63	NP_001001823.1	tudor domain-containing protein 3 [Danio rerio]	952

7.84	2.50	NP_001006096.1	ADP-ribosylation factor-like 6 interacting protein 5a [Danio rerio]	953
8.58	1.80	NP_957198.1	fermitin family homolog 3 [Danio rerio]	956
4.11	1.32	NP_001071043.2	thyroid hormone receptor interactor 4 [Danio rerio]	958
4.47	1.77	NP_001013316.1	prion protein, related sequence 3 precursor [Danio rerio]	984
6.88	1.77	NP_001002375.1	negative elongation factor E [Danio rerio]	985
4.56	3.29	XP_003442188.1	PREDICTED: hypothetical protein LOC100698035 [Oreochromis niloticus]	987
5.82	1.70	NP_001004600.1	RING finger protein 103 [Danio rerio]	988
8.00	2.68	NP_997750.1	IST1 homolog [Danio rerio]	1023
11.75	3.15	XP_706414.3	PREDICTED: hypothetical protein LOC567610 [Danio rerio]	1027
5.00	2.00	NP_001004596.1	general transcription factor IIH subunit 1 [Danio rerio]	1028
5.92	1.30	XP_003448760.1	PREDICTED: elongation of very long chain fatty acids protein 6-like [Oreochromis niloticus]	1029
9.64	1.61	NP_991283.2	testin [Danio rerio]	1035
10.94	1.64	XP_003198617.1	PREDICTED: beta-glucuronidase-like [Danio rerio]	1052
7.05	1.40	NP_001032486.1	integrin beta-1-binding protein 1 [Danio rerio]	1055
5.07	1.21	NP_956758.2	intron-binding protein aquarius [Danio rerio]	1066
9.70	2.52	NP_958477.2	importin subunit alpha-3 [Danio rerio]	1096
4.66	1.73	NP_001018387.1	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily E member 1-related isoform 1 [Danio rerio]	1109
4.29	2.69	NP_001073132.2	RILP-like protein 1 [Danio rerio]	1127
5.75	1.57	NP_001073430.1	gametogenetin-binding protein 2 [Danio rerio]	1142
5.14	1.94	NP_001070727.2	beta-1,4-galactosyltransferase 1 [Danio rerio]	1143
6.34	1.57	XP_002662576.2	PREDICTED: putative helicase mov-10-B.1 [Danio rerio]	1144
4.42	2.83	XP_001337829.1	PREDICTED: tyrosine-protein kinase ABL1-like isoform 1 [Danio rerio]	1145
4.92	1.32	NP_001017620.1	transmembrane protein 256 precursor [Danio rerio]	1149
7.71	1.20	NP_001002157.1	isocitrate dehydrogenase [NAD] subunit beta, mitochondrial [Danio rerio]	1166
11.22	2.03	NP_001038542.1	SH2 domain-containing adapter protein D [Danio rerio]	1174
4.97	1.12	NP_001018311.1	bloodthirsty [Danio rerio]	1185
4.88	1.42	NP_996970.1	N(6)-adenine-specific DNA methyltransferase 2 [Danio rerio]	1197
7.53	1.26	NP_956980.1	mitochondrial fission process protein 1 [Danio rerio]	1212
5.77	4.76	NP_001092893.1	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase delta-3-A [Danio rerio]	1228
9.65	1.41	NP_001008619.1	probable 2-oxoglutarate dehydrogenase E1 component DHKTD1, mitochondrial [Danio rerio]	1229
11.75	3.15	XP_002938031.1	PREDICTED: hypothetical protein LOC100497494 [Xenopus (Silurana) tropicalis]	1231
4.17	1.56	NP_001013554.2	receptor accessory protein 6-like [Danio rerio]	1236
5.00	1.45	XP_700128.4	PREDICTED: zinc transporter ZIP9-B-like [Danio rerio]	1249
4.53	2.14	NP_001018515.1	PRELI domain containing 1b [Danio rerio]	1252
4.04	2.31	NP_571389.1	LIM domain-binding protein 2 [Danio rerio]	1255
5.00	5.00	BAE46429.1	ORF2-encoded protein [Danio rerio]	1265
4.33	1.00	NP_001139048.1	glycogenin-2 [Danio rerio]	1269
5.34	2.10	XP_690443.3	PREDICTED: SAM and SH3 domain-containing protein 3-like [Danio rerio]	1287
4.56	1.46	NP_001035021.1	S-acyl fatty acid synthase thioesterase, medium chain [Danio rerio]	1297

6.03	1.61	NP_001003751.1	uncharacterized protein LOC445294 [Danio rerio]	1298
4.57	1.36	NP_001017589.1	zinc finger CCHC-type and RNA-binding motif-containing protein 1 [Danio rerio]	1309
4.52	1.60	NP_001119928.1	immunoresponsive gene 1 [Danio rerio]	1310
4.57	3.00	XP_699956.3	PREDICTED: semaphorin-4A [Danio rerio]	1334
5.56	2.85	NP_001103502.1	stomatin-like protein 1 [Danio rerio]	1349
4.40	1.37	NP_998624.1	calpain 1, large subunit [Danio rerio]	1354
4.00	1.25	NP_998605.1	pleiotropic regulator 1 [Danio rerio]	1397
5.21	2.49	NP_998685.2	transmembrane protein 184A [Danio rerio]	1398
6.32	2.22	XP_002666807.2	PREDICTED: protein MON2 homolog, partial [Danio rerio]	1402
11.63	1.72	NP_997780.1	E3 ubiquitin-protein ligase RNF128 precursor [Danio rerio]	1419
4.38	2.40	NP_001070037.1	myelin-associated neurite-outgrowth inhibitor [Danio rerio]	1423
9.58	5.75	XP_001344050.1	PREDICTED: hypothetical protein LOC100004848 [Danio rerio]	1440
4.41	1.70	NP_958445.1	large proline-rich protein BAT3 [Danio rerio]	1443
5.66	1.44	NP_001038382.1	segment polarity protein dishevelled homolog DVL-1 [Danio rerio]	1451
5.42	2.60	XP_001921323.1	PREDICTED: protein phosphatase Slingshot homolog 2 [Danio rerio]	1462
6.72	1.04	NP_001093465.1	ran guanine nucleotide release factor [Danio rerio]	1465
6.88	2.65	XP_690924.5	PREDICTED: zinc finger protein Rlf [Danio rerio]	1495
6.66	1.49	NP_001082989.1	sodium-dependent glucose transporter 1 [Danio rerio]	1515
5.05	2.00	NP_001017718.2	transmembrane protein 184B [Danio rerio]	1534
7.87	1.00	NP_001038850.2	glucokinase [Danio rerio]	1540
7.56	1.64	CAE17634.2	novel protein similar to human telomerase-associated protein 1 (TEP1) [Danio rerio]	1546
15.63	1.80	NP_001076324.1	39S ribosomal protein L37, mitochondrial [Danio rerio]	1553
7.33	3.76	NP_001017593.1	epithelial cell adhesion molecule precursor [Danio rerio]	1567
4.14	2.78	XP_003444806.1	PREDICTED: hypothetical protein LOC100710310 [Oreochromis niloticus]	1586
5.50	6.40	XP_693604.4	PREDICTED: receptor-transporting protein 2 [Danio rerio]	1589
15.12	11.36	XP_003458913.1	PREDICTED: hypothetical protein LOC100699620 [Oreochromis niloticus]	1598
6.93	2.41	NP_001007304.1	ADP-ribosylation factor GTPase-activating protein 1 [Danio rerio]	1614
4.28	1.36	NP_956588.1	UDP-N-acetylhexosamine pyrophosphorylase-like protein 1 [Danio rerio]	1615
5.56	2.36	NP_001185678.1	uncharacterized LOC563012 [Danio rerio]	1644
8.75	8.15	NP_942104.2	tubulin, beta 2c [Danio rerio]	1685
6.10	1.33	NP_956848.1	S-methyl-5'-thioadenosine phosphorylase [Danio rerio]	1695
21.21	3.00	XP_001921185.1	PREDICTED: metalloreductase STEAP2 [Danio rerio]	1700
5.54	1.64	ACM09793.1	Transposable element Tc1 transposase [Salmo salar]	1748
15.53	1.03	NP_571422.1	uroporphyrinogen decarboxylase [Danio rerio]	1766
6.02	1.86	NP_001009987.2	transmembrane protein 111-like [Danio rerio]	1786
4.38	1.72	NP_956174.1	acylglycerol kinase, mitochondrial precursor [Danio rerio]	1814
4.64	1.53	NP_956236.1	fructose-1,6-bisphosphatase 1a [Danio rerio]	1822
6.24	1.92	NP_955878.1	hippocampus abundant transcript 1 protein [Danio rerio]	1852
5.42	1.85	NP_001019619.1	myosin-11 [Danio rerio]	1860

4.60	1.53	NP_998703.1	crk-like protein [Danio rerio]	1898
16.87	3.69	NP_001002327.1	suppression of tumorigenicity 5 protein [Danio rerio]	1899
4.73	1.92	XP_002665164.1	PREDICTED: stonustoxin subunit alpha-like [Danio rerio]	1920
4.35	1.14	NP_001013506.1	coiled-coil domain containing 174 [Danio rerio]	1922
6.27	2.77	XP_003460241.1	PREDICTED: hypothetical protein LOC100705656 [Oreochromis niloticus]	1934
17.79	3.22	XP_003197747.1	PREDICTED: protein G7c-like [Danio rerio]	1952
4.51	2.47	NP_001116713.1	nuclear receptor subfamily 1, group H, member 5 [Danio rerio]	1956
16.58	6.75	P41386.1	TBB_HALDI RecName: Full=Tubulin beta chain; AltName: Full=Beta-tubulin	1958
7.13	1.68	NP_001073422.2	isochorismatase domain-containing protein 2, mitochondrial [Danio rerio]	1963
4.51	1.68	NP_001002539.1	AP-3 complex subunit sigma-2 [Danio rerio]	1991
4.44	2.07	XP_691291.2	PREDICTED: solute carrier family 12 member 4 [Danio rerio]	1994
5.78	1.14	NP_001018093.2	NFU1 iron-sulfur cluster scaffold homolog [Danio rerio]	1997
4.34	1.64	XP_690269.4	PREDICTED: uncharacterized protein C7orf57 homolog isoform 2 [Danio rerio]	1999
5.63	1.45	NP_001002445.1	ras homolog gene family, member Ad [Danio rerio]	2011
9.04	1.85	NP_998268.1	tRNA-splicing ligase RtcB homolog [Danio rerio]	2016
4.33	2.78	XP_003198249.1	PREDICTED: TOM1-like protein 2-like [Danio rerio]	2018
12.67	4.88	NP_001093459.1	spectrin beta chain, brain 1 [Danio rerio]	2040
10.82	3.26	NP_775365.1	ATP-dependent RNA helicase DDX19B [Danio rerio]	2154
5.73	1.24	NP_956332.1	developmentally-regulated GTP-binding protein 1 [Danio rerio]	2172
11.22	2.24	NP_001025355.1	uncharacterized protein LOC563956 [Danio rerio]	2184
4.64	2.80	NP_956828.1	methylosome protein 50 [Danio rerio]	2198
4.33	2.82	XP_001336816.1	PREDICTED: quinone oxidoreductase PIG3-like [Danio rerio]	2206
6.68	1.31	NP_997944.1	eosinophil peroxidase precursor [Danio rerio]	2264
5.93	2.92	XP_686922.3	PREDICTED: peroxisomal sarcosine oxidase [Danio rerio]	2269
20.97	2.67	NP_991206.1	purine nucleoside phosphorylase 4b [Danio rerio]	2276
5.08	2.44	NP_956151.1	heat shock protein 4b [Danio rerio]	2282
4.78	1.62	NP_001007435.1	nitric oxide synthase-interacting protein [Danio rerio]	2293
22.59	3.52	NP_001038676.1	EF-hand domain-containing protein D2 [Danio rerio]	2296
4.78	1.63	NP_998485.1	charged multivesicular body protein 3 [Danio rerio]	2304
5.15	1.33	NP_958463.1	alpha-2-macroglobulin receptor-associated protein [Danio rerio]	2307
4.47	1.51	NP_001017824.1	cytokeratin-like [Danio rerio]	2315
20.83	1.07	ACQ58309.1	Serum amyloid P-component precursor [Anoplopoma fimbria]	2319
4.83	2.13	NP_775337.1	protein SDA1 homolog [Danio rerio]	2340
4.34	1.35	NP_956900.1	E3 UFM1-protein ligase 1 [Danio rerio]	2349
4.55	1.43	NP_957422.1	sarcosine dehydrogenase, mitochondrial [Danio rerio]	2395
4.30	1.81	NP_999897.1	D-amino acid oxidase [Danio rerio]	2399
4.93	3.52	NP_571924.2	alcohol dehydrogenase class-3 [Danio rerio]	2410
6.09	2.17	NP_956166.1	oxysterol-binding protein-related protein 2 [Danio rerio]	2439
4.84	1.80	NP_001017898.1	etoposide induced 2.4 [Danio rerio]	2447

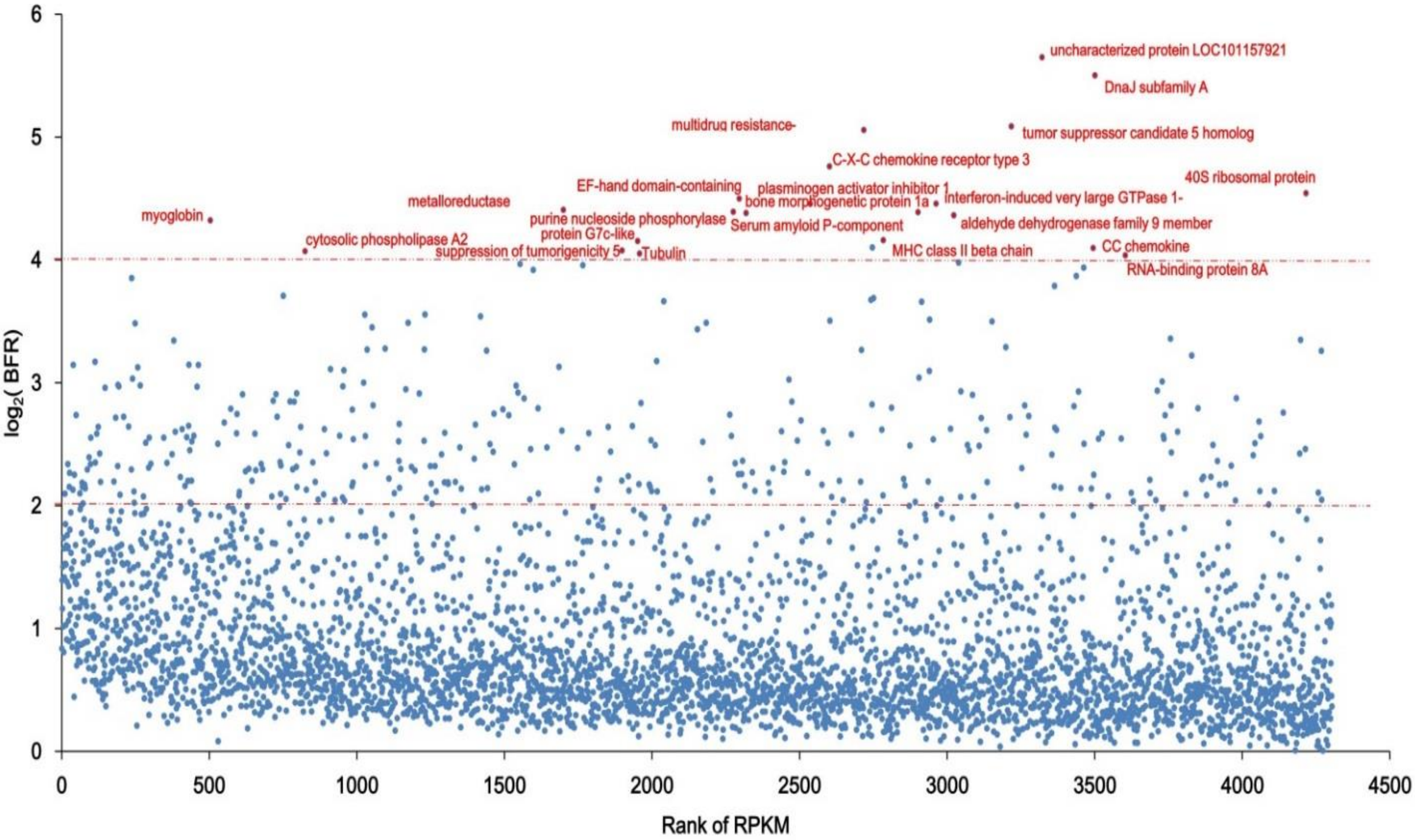
5.11	2.13	NP_001038662.1	arginase-1 [Danio rerio]	2450
8.14	5.38	NP_001189369.1	E3 ubiquitin-protein ligase RNF19B [Danio rerio]	2464
7.19	2.07	NP_001003738.1	anamorsin [Danio rerio]	2474
5.77	1.39	NP_956682.1	coiled-coil domain-containing protein 25 [Danio rerio]	2493
6.46	2.76	XP_001335910.2	PREDICTED: intelectin-1a-like [Danio rerio]	2505
4.81	1.82	NP_001035389.1	UDP-glucose 4-epimerase [Danio rerio]	2528
21.94	17.50	NP_001108031.1	plasminogen activator inhibitor 1 precursor [Danio rerio]	2537
4.50	1.36	XP_693354.3	PREDICTED: leucine rich repeat and Ig domain containing 1a, partial [Danio rerio]	2576
6.10	4.48	NP_001122201.1	Kruppel-like factor 9 [Danio rerio]	2581
5.69	1.71	NP_001038823.1	28S ribosomal protein S14, mitochondrial [Danio rerio]	2596
27.11	4.36	NP_001138285.1	si:dkey-269d20.3 [Danio rerio]	2602
11.35	1.74	XP_003201450.1	PREDICTED: aminopeptidase N-like [Danio rerio]	2603
4.20	1.27	NP_001038391.1	MTSS1-like protein [Danio rerio]	2609
4.60	1.92	NP_998662.2	cysteine-rich protein 2 [Danio rerio]	2635
4.12	2.38	NP_571439.2	E3 ubiquitin-protein ligase Mdm2 [Danio rerio]	2648
5.97	1.72	NP_001006093.1	sterol carrier protein 2b [Danio rerio]	2676
4.58	3.27	NP_998593.1	CD9 antigen, b [Danio rerio]	2707
9.63	1.00	NP_001002535.1	RWD domain-containing protein 1 [Danio rerio]	2710
33.28	3.13	NP_001182542.1	multidrug resistance-associated protein 5 [Danio rerio]	2718
4.08	1.04	BAE46430.1	reverse transcriptase [Danio rerio]	2726
12.78	5.49	NP_001034710.1	follistatin-like 1b precursor [Danio rerio]	2742
7.08	1.01	NP_694494.1	fibroblast growth factor receptor 1-A precursor [Danio rerio]	2746
17.16	1.14	XP_003459899.1	PREDICTED: transmembrane protein C9orf125 homolog [Oreochromis niloticus]	2747
12.89	10.43	NP_001155073.1	S-adenosylmethionine decarboxylase proenzyme [Danio rerio]	2751
6.14	1.52	NP_998314.1	peptidyl-prolyl cis-trans isomerase FKBP5 [Danio rerio]	2779
4.23	1.63	NP_001018460.1	prefoldin subunit 3 [Danio rerio]	2783
17.88	2.39	AAB67871.1	MHC class II beta chain [Ictalurus punctatus]	2784
6.94	2.58	NP_001025235.1	ribonuclease T2 precursor [Danio rerio]	2812
4.65	3.65	XP_003458380.1	PREDICTED: hypothetical protein LOC100703668 [Oreochromis niloticus]	2853
4.48	1.30	ACQ58579.1	C12orf62 homolog [Anoplopoma fimbria]	2857
5.61	1.50	XP_002933173.1	PREDICTED: hypothetical protein LOC100488659 [Xenopus (Silurana) tropicalis]	2874
4.09	1.29	AFK08733.1	oxidation resistance protein 2 variant A, partial [Danio rerio]	2892
20.93	2.83	NP_001035126.1	bone morphogenetic protein 1a precursor [Danio rerio]	2902
8.23	1.27	NP_001094432.1	vacuolar protein sorting 4 homolog b-like [Danio rerio]	2905
12.63	1.32	NP_001170924.2	elongation factor-1, delta, a [Danio rerio]	2914
4.58	3.69	XP_004084042.1	PREDICTED: uncharacterized protein LOC101164525 [Oryzias latipes]	2936
8.54	1.53	XP_003201117.1	PREDICTED: ribosome-binding protein 1 [Danio rerio]	2940
11.41	2.48	NP_001038442.2	tubulointerstitial nephritis antigen-like precursor [Danio rerio]	2941
5.81	4.00	XP_692034.2	PREDICTED: TBC1 domain family member 9 isoform 2 [Danio rerio]	2953

21.96	16.55	XP_682842.1	PREDICTED: interferon-induced very large GTPase 1-like [Danio rerio]	2963
4.00	2.25	NP_957423.1	nuclear receptor coactivator 4 [Danio rerio]	2965
4.15	1.56	NP_001103197.1	uncharacterized protein LOC796128 [Danio rerio]	2980
6.17	4.93	XP_004070564.1	PREDICTED: uncharacterized protein C3orf32-like [Oryzias latipes]	3012
20.56	1.74	NP_958879.1	aldehyde dehydrogenase family 9 member A1-A [Danio rerio]	3023
15.76	2.32	NP_956997.1	Golgi reassembly-stacking protein 2 [Danio rerio]	3039
4.20	1.40	XP_003451558.1	PREDICTED: hypothetical protein LOC100693979 [Oreochromis niloticus]	3046
7.62	1.53	NP_001001815.1	insulin-like growth factor 2b precursor [Danio rerio]	3047
5.63	1.40	XP_002661059.2	PREDICTED: UPF0562 protein C7orf55 homolog [Danio rerio]	3069
5.46	2.69	NP_001008639.1	very long-chain acyl-CoA synthetase [Danio rerio]	3075
7.47	7.27	XP_001342589.3	PREDICTED: hypothetical protein LOC100002920 [Danio rerio]	3086
4.21	2.58	XP_002666630.1	PREDICTED: putative protein phosphatase 1 regulatory inhibitor subunit 3G-like [Danio rerio]	3090
5.60	4.82	NP_001122225.1	uncharacterized protein LOC569781 [Danio rerio]	3109
6.55	2.03	NP_705939.1	P2X purinoceptor 4 [Danio rerio]	3116
6.12	1.50	NP_958864.1	flotillin 1b [Danio rerio]	3134
4.57	1.62	NP_938177.1	protein Tob1 [Danio rerio]	3136
11.31	3.97	XP_003197916.1	PREDICTED: RING finger protein 31-like, partial [Danio rerio]	3152
9.77	1.37	NP_955897.2	acidic chitinase-like precursor [Danio rerio]	3199
6.59	2.31	NP_957257.1	fumarate hydratase, mitochondrial precursor [Danio rerio]	3213
33.96	2.34	XP_003200303.1	PREDICTED: tumor suppressor candidate 5 homolog [Danio rerio]	3218
4.00	1.31	NP_958459.1	vesicle-trafficking protein SEC22b-B precursor [Danio rerio]	3237
4.94	1.34	NP_001116727.1	RING finger protein 10 [Danio rerio]	3253
7.04	1.30	NP_001007292.1	3-oxoacid CoA transferase 1a [Danio rerio]	3263
5.97	1.53	NP_001091722.1	uncharacterized protein LOC569842 precursor [Danio rerio]	3268
6.63	1.15	NP_001104551.1	growth hormone receptor b precursor [Danio rerio]	3277
50.19	19.32	XP_004086099.1	PREDICTED: uncharacterized protein LOC101157921 [Oryzias latipes]	3322
5.33	4.71	XP_003454888.1	PREDICTED: protein jagged-1b-like [Oreochromis niloticus]	3355
4.46	1.18	XP_694827.4	PREDICTED: apolipoprotein B-100 [Danio rerio]	3363
13.82	5.60	NP_956314.1	protein BTG1 [Danio rerio]	3364
6.20	1.51	NP_001077036.1	lanosterol synthase [Danio rerio]	3365
6.13	1.14	XP_001343562.1	PREDICTED: hypothetical protein LOC100004199 [Danio rerio]	3371
4.42	2.55	XP_003199342.1	PREDICTED: vitamin K-dependent gamma-carboxylase-like, partial [Danio rerio]	3385
7.01	3.60	NP_001038800.1	ribosomal protein L22-like 1 [Danio rerio]	3430
14.61	2.20	ACI68661.1	NADH dehydrogenase 1 subunit C1, mitochondrial precursor [Salmo salar]	3438
7.61	1.94	XP_003198717.1	PREDICTED: zinc finger protein 729-like [Danio rerio]	3445
4.40	1.45	NP_001008642.1	epoxide hydrolase 2 [Danio rerio]	3453
15.31	1.67	NP_001017727.1	peroxisomal trans-2-enoyl-CoA reductase [Danio rerio]	3463
5.67	1.07	NP_001003515.1	hydroxyacyl-coenzyme A dehydrogenase, mitochondrial [Danio rerio]	3464
4.40	1.38	NP_001019582.1	MAWD binding protein like [Danio rerio]	3489

17.11	17.78	AAT52142.1	CC chemokine SCYA108 [Ictalurus furcatus]	3495
4.76	4.15	XP_003198696.1	PREDICTED: hypothetical protein LOC100536019 [Danio rerio]	3497
45.28	1.95	NP_998658.1	DnaJ subfamily A member 2 [Danio rerio]	3501
5.83	2.38	NP_957511.1	nucleolar protein 56 [Danio rerio]	3513
6.01	1.42	NP_956063.1	hsc70-interacting protein [Danio rerio]	3525
4.22	1.41	NP_001184218.1	dimethylanaline monooxygenase-like [Danio rerio]	3546
5.84	2.39	NP_998188.2	ATPase, H+ transporting, lysosomal accessory protein 2 [Danio rerio]	3590
16.42	1.74	NP_001013363.1	RNA-binding protein 8A [Danio rerio]	3604
4.30	3.98	NP_955905.1	arginase-2, mitochondrial [Danio rerio]	3624
4.10	2.74	NP_001003944.1	actin-related protein 3 [Danio rerio]	3631
4.63	2.67	XP_002661250.2	PREDICTED: multidrug resistance-associated protein 1-like [Danio rerio]	3687
4.30	1.46	NP_955870.1	splicing factor, arginine/serine-rich 11 [Danio rerio]	3688
4.13	1.20	NP_957091.1	reactive oxygen species modulator 1 [Danio rerio]	3707
7.65	1.22	NP_001004540.1	aldehyde dehydrogenase family 8 member A1 [Danio rerio]	3712
8.06	2.01	NP_956863.1	TSC22 domain family protein 3 [Danio rerio]	3729
5.92	4.68	XP_002660417.1	PREDICTED: interferon-induced protein 44 [Danio rerio]	3733
5.83	1.78	NP_955879.1	voltage-dependent anion-selective channel protein 2 [Danio rerio]	3735
6.66	2.49	ACX35602.1	VHSV-induced protein [Salmo salar]	3739
10.25	11.44	XP_003211698.1	PREDICTED: caspase-1-like [Meleagris gallopavo]	3757
7.04	12.47	NP_001007446.1	phytanoyl-CoA dioxygenase domain-containing protein 1 [Danio rerio]	3759
5.40	1.26	NP_997981.1	glycine N-methyltransferase [Danio rerio]	3760
6.07	1.41	NP_001035396.2	kynurenine 3-monooxygenase [Danio rerio]	3781
4.27	11.40	NP_001038401.1	carboxylesterase 3 precursor [Danio rerio]	3809
9.33	2.48	NP_998158.1	probable ribosome biogenesis protein RLP24 [Danio rerio]	3829
6.93	1.41	NP_955949.1	rab GDP dissociation inhibitor beta [Danio rerio]	3850
4.64	2.67	NP_001038735.2	finTRIM family, member 14 [Danio rerio]	3864
4.70	1.49	NP_570987.1	catalase [Danio rerio]	3869
4.24	2.31	NP_955817.1	mitochondrial uncoupling protein 3 [Danio rerio]	3879
4.71	1.00	NP_001002169.1	spermidine/spermine N1-acetyltransferase [Danio rerio]	3894
5.63	1.81	NP_956358.1	branched-chain-amino-acid aminotransferase, cytosolic [Danio rerio]	3901
5.14	1.73	NP_001091951.1	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit E [Danio rerio]	3916
4.50	2.67	XP_698776.4	PREDICTED: protein NLRC3 [Danio rerio]	3923
4.52	1.22	AAP49009.1	transposase [Rana pipiens]	3944
5.01	1.75	NP_001007209.2	coagulation factor V precursor [Danio rerio]	3963
4.12	1.35	XP_001339274.1	PREDICTED: hypothetical protein LOC798849 [Danio rerio]	3977
7.33	5.26	XP_002661412.1	PREDICTED: hypothetical protein LOC100332348 [Danio rerio]	3980
5.32	1.92	NP_001017833.1	inorganic pyrophosphatase [Danio rerio]	4037
5.69	1.34	NP_001006081.1	DNA replication complex GINS protein PSF1 [Danio rerio]	4043
6.43	4.92	NP_955899.2	trypsinogen precursor [Danio rerio]	4057

4.34	1.13	XP_002666578.2	PREDICTED: sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1-like, partial [Danio rerio]	4060
5.92	1.90	NP_999924.1	legumain precursor [Danio rerio]	4063
4.03	3.51	XP_003198084.1	PREDICTED: hypothetical protein LOC100535628 [Danio rerio]	4090
4.34	2.47	NP_955930.1	60S ribosomal protein L10a [Danio rerio]	4109
6.76	2.48	NP_001153598.1	complement factor properdin precursor [Danio rerio]	4139
5.37	1.58	NP_001132933.1	vitronectin b precursor [Danio rerio]	4195
10.19	1.72	XP_003199195.1	PREDICTED: small serum protein 2-like [Danio rerio]	4198
5.51	1.90	NP_571385.2	heat shock protein HSP 90-beta [Danio rerio]	4214
23.29	2.23	NP_956353.1	40S ribosomal protein S3a [Danio rerio]	4216
4.30	1.33	NP_001020335.1	inter-alpha-trypsin inhibitor heavy chain H4 precursor [Danio rerio]	4258
9.58	2.58	NP_001003728.1	40S ribosomal protein S6 [Danio rerio]	4268
4.13	2.17	NP_001116330.1	mannan-binding lectin serine protease 2 precursor [Danio rerio]	4269

Figure 2. Genes harbouring significant SNPs, plotted by the $\log_2$(BFR) versus rank of RPKM. Genes which have BFR larger than 16 ($\log_2 = 4$) were highlighted by red and their gene names were labelled.

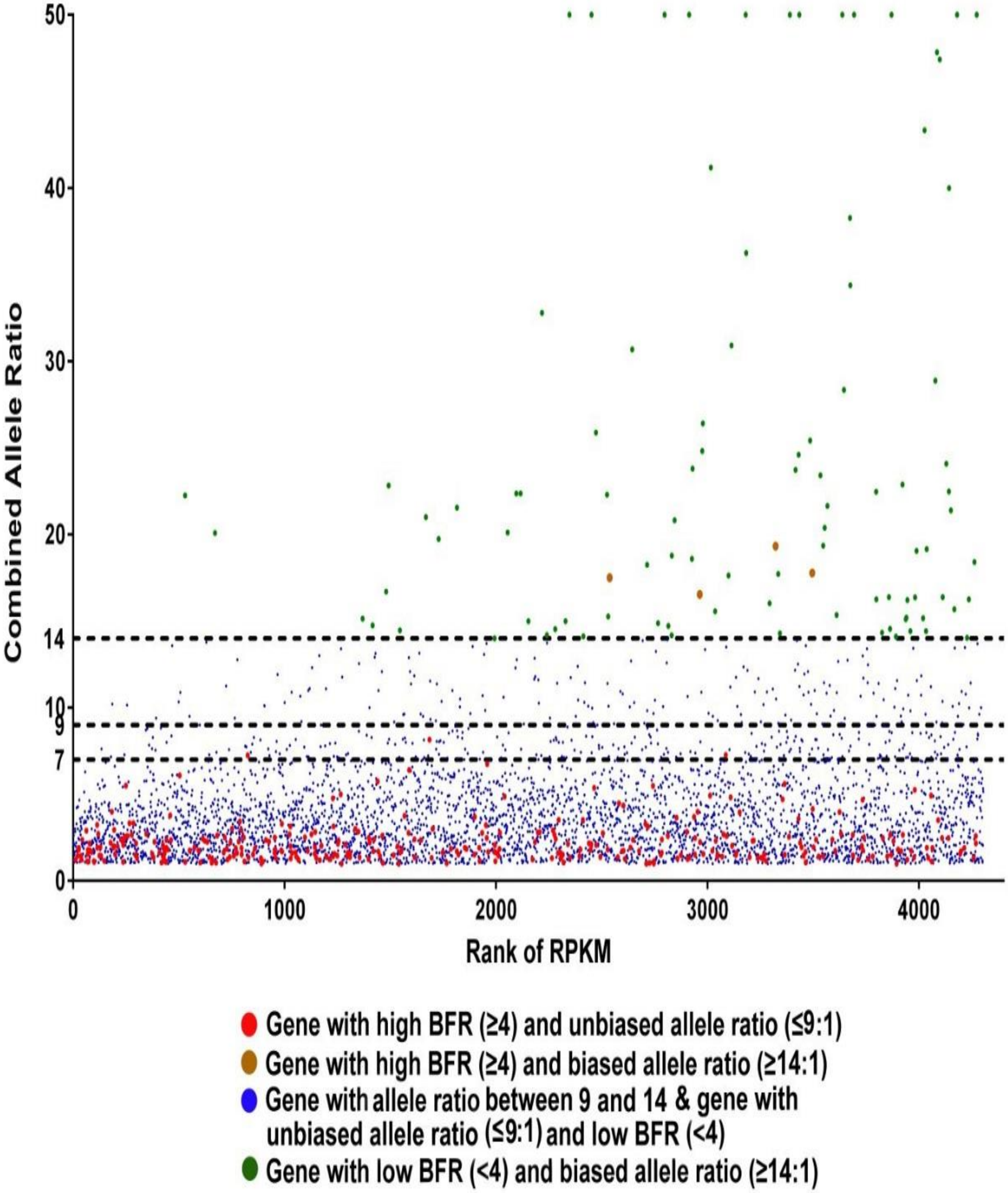


III.6 Genes with large BFR caused by genetic segregation

As RNA-seq data is analyzed in terms of RPKM at the RNA level, the allele ratios obtained by RNA-seq are compounded by two factors: the genotype allele frequencies at the DNA level and the relative expression levels of the two alleles at the RNA level. For instance, the two alleles may have very different genotype allele frequencies in the two bulked samples, and in these cases, even if the expression is not regulated at the transcriptional level, the final ratio of the two alleles between the bulked samples are expected to be different. However, if one of the two alleles is differentially regulated, the final allele ratio at the RNA level would be different from the allele ratio at the DNA level.

In order to differentiate SNPs with large BFR caused by genotype allele frequency difference from those with large BFR caused by allele-specific expression, the ratio of the two alleles was analyzed with combined bulk of both resistance and susceptible bulks. Theoretically, now the combined bulk should include alleles at expected segregation ratios without any connection with the phenotypes.

Figure 3. Genes harbouring significant SNPs, plotted by their combined allele ratios versus rank of RPKM. Red dots represent genes with BFR≥4 and combined allele ratio≤9; brown dots represent genes with BFR≥4 and allele ratio≥14; green dots represent genes with BFR<4 and allele ratio≥14, blue stands for genes with allele ratio from 9 to 14, or genes with BFR<4 and allele ratio≤9. The three threshold lines of combined allele ratio of 1:7 (maximal possible ratio at the DNA level for any polymorphic SNP with two families), 1:9 (3X maximal possible ratio at the DNA level for SNPs polymorphic in both families); and 1:14 (2X maximal possible ratio at the DNA level for any polymorphic SNP with two families) were drawn for references.


Combined Allele Ratio
50
40
30
20
14
10
9
7
0
0
1000
2000
3000
4000
Rank of RPKM
● Gene with high BFR (≥4) and unbiased allele ratio (≤9:1)
● Gene with high BFR (≥4) and biased allele ratio (≥14:1)
● Gene with allele ratio between 9 and 14 & gene with unbiased allele ratio (≤9:1) and low BFR (<4)
● Gene with low BFR (<4) and biased allele ratio (≥14:1)

If the large BFR is caused by different allele frequencies at the DNA level in the two bulks but not by allele-specific expression, the ratio of the two alleles in the combined bulk (resistant bulk plus susceptible bulk) should be relatively small and predictable with the family structure. For instance, at an AA x AG SNP site in a single family, the progenies should have a 3A:1G allele frequency at the genomic level. However, the situation is more complex if more than one family is used. Nevertheless, the largest allele ratio at the DNA level can still be predicted. For instance, if two families are used as in this study, at an SNP site, the largest possible allele ratio at the DNA level is 7:1, i.e., AA x AG in one family, and AA x AA in the second family, where the largest possible allele ratio at the DNA level is 7A:1G. Any other combinations would result in a smaller combined allele ratio at the DNA level (Table 8). Therefore, we differentiated the large BFR caused by genetic segregation and those caused by allele-specific expression based on the combined allele ratio of the resistant and susceptible bulks. While those large BFRs with very large combined allele ratios are likely caused, at least in part, by allele-specific expression, and those large BFRs with small combined allele ratios are likely caused by genetic segregation, the BFRs in the transitional zone could be caused by both genetic segregation and allele-specific expression (Figure 3).

Of the 359 genes with large BFR (≥4), 347 had a combined bulk allele ratio of 7 or less. The vast majority of these had a combined bulk allele ratio of 1-3, suggesting that the large BFR of these genes were mostly caused by genetic segregation.

Table 8. Allele ratio combination in two families

1st Family/ 2nd Family	AA xAA	AA x AG	AG x AG
AA x AG	7:1	3:1	5:3

AG x AG	3:1	5:3	1:1
AA x AA	-	7:1	3:1

III.7 Genes with large BFR caused by allele-specific expression

As shown in Figure 3, a large number of genes harboring significant SNPs had a significantly higher combined allele ratio, with 286 genes had an allele ratio of greater than 9:1 (genes above the threshold of 9 in Figure 3). Considering that most of the genes with BFR >4 had a combined allele ratio of 1-3 (see above and Figure 3), many of the genes with combined allele ratio of greater than 9 could be caused by allele-specific expression. On the cautious side, even in the extreme cases of 7:1 allele ratios at the DNA levels, twice the largest possible allele ratio at the DNA level should be 14:1. As shown in Figure 3, 98 genes had a combined allele ratio of greater than 14, indicating that these large allele ratios are caused by, at least in part, allele-specific expression. A list of these allele-specific expressed genes was provided in Table 9. Of the 98 genes with high combined allele ratios, 4 genes were with BFR higher than 4 and allele ratio higher than 14. They are plasminogen activator inhibitor 1, interferon-induced very large GTPase 1-like, uncharacterized protein LOC101157921 and CC chemokine SCYA108. Apparently, these large combined allele ratios were caused by both genetic segregation and by allele-specific expression.

Table 9. Allele-specific expressed genes.

Max BFR	Combined Allel_ratio	ID	Description
1.06	22.25	AAI52592.1	LOC569167 protein [Danio rerio]

1.23	20.09	XP_002943664.1	PREDICTED: hypothetical protein LOC100486110 [Xenopus (Silurana) tropicalis]
1.30	15.14	NP_956613.2	macoilin-1 [Danio rerio]
1.50	14.75	XP_002667032.2	PREDICTED: myotubularin-related protein 4 [Danio rerio]
1.10	16.70	NP_001032787.1	sorting nexin-2 [Danio rerio]
1.15	22.83	NP_001032777.1	Hermansky-Pudlak syndrome 1 protein [Danio rerio]
1.12	14.47	XP_003937480.1	PREDICTED: zinc finger protein 585B [Saimiri boliviensis boliviensis]
1.67	21.00	NP_001091654.1	A-kinase anchor protein 12 [Danio rerio]
1.15	19.75	NP_001191383.1	P2Y purinoceptor 11 [Danio rerio]
1.21	21.56	NP_958886.1	collagen, type I, alpha 1b precursor [Danio rerio]
1.15	14.00	NP_001133536.1	Abl interactor 1 [Salmo salar]
1.10	20.11	NP_001017908.2	integrin alpha FG-GAP repeat containing 2 [Danio rerio]
1.11	22.38	NP_001191060.1	lecithin retinol acyltransferase a [Danio rerio]
1.36	22.38	NP_001007448.1	carnitine O-palmitoyltransferase 2, mitochondrial [Danio rerio]
1.39	15.00	NP_999977.1	erythrocyte membrane protein band 4.1-like 3b [Danio rerio]
1.14	32.80	XP_003452034.1	PREDICTED: transmembrane protein 150A-like [Oreochromis niloticus]
1.07	14.19	XP_691812.1	PREDICTED: uncharacterized protein DKFZp762I1415-like [Danio rerio]
1.41	14.54	NP_997895.1	sideroflexin-2 [Danio rerio]
1.19	15.00	XP_692717.4	PREDICTED: TNFAIP3-interacting protein 2-like [Danio rerio]
1.74	50.00	NP_001092735.1	uncharacterized protein LOC100093712 precursor [Danio rerio]
1.39	14.12	AAI55834.1	Si:dkey-10p5.7 [Danio rerio]
1.08	50.00	XP_690141.2	PREDICTED: hypothetical protein LOC561636 [Danio rerio]
1.21	25.88	XP_002195523.1	PREDICTED: uncharacterized protein LOC100231554 [Taeniopygia guttata]
1.09	22.30	AFU76492.1	short chain alpha neurotoxin A1 [Micrurus laticollaris]
1.43	15.27	NP_955823.1	C-1-tetrahydrofolate synthase, cytoplasmic [Danio rerio]
21.94	17.50	NP_001108031.1	plasminogen activator inhibitor 1 precursor [Danio rerio]
1.14	30.69	NP_001001400.1	actin-related protein 8 [Danio rerio]
1.09	18.25	NP_001243120.1	nuclear receptor subfamily 0 group B member 2 [Danio rerio]
1.33	14.88	NP_955910.1	beta-ureidopropionase [Danio rerio]
2.32	50.00	XP_003453579.1	PREDICTED: fetuin-B-like [Oreochromis niloticus]
1.10	14.73	NP_998403.1	transmembrane protein 43 [Danio rerio]
1.16	14.17	NP_001004529.1	enoyl-CoA hydratase, mitochondrial [Danio rerio]
1.11	18.79	NP_958473.1	tyrosine--tRNA ligase, cytoplasmic [Danio rerio]
1.25	20.82	XP_682981.1	PREDICTED: antigen peptide transporter 2-like isoform 1 [Danio rerio]
1.42	50.00	NP_001004661.1	aquaporin 8a, tandem duplicate 1 [Danio rerio]
1.21	18.60	NP_570996.2	F-box only protein 50 [Danio rerio]
1.07	23.80	NP_956144.1	c20orf24 homolog [Danio rerio]
21.96	16.55	XP_682842.1	PREDICTED: interferon-induced very large GTPase 1-like [Danio rerio]
1.19	24.82	XP_002160050.2	PREDICTED: uncharacterized protein LOC100206654 [Hydra magnipapillata]
1.06	26.41	NP_001077323.1	indoleamine 2,3-dioxygenase 2 [Danio rerio]

1.21	41.18	NP_001002579.1	mitochondrial glycine cleavage system H protein [Danio rerio]
1.07	15.57	XP_695816.5	PREDICTED: protein zyg-11 homolog [Danio rerio]
2.57	17.63	XP_688314.4	PREDICTED: hypothetical protein LOC559835 [Danio rerio]
1.06	30.92	NP_001025388.1	sorting nexin-6 [Danio rerio]
1.03	50.00	NP_998475.1	egl nine homolog 3 [Danio rerio]
1.31	36.25	NP_001002372.1	transmembrane 4 L6 family member 5 [Danio rerio]
1.17	16.03	NP_998232.1	dual specificity protein phosphatase 1 [Danio rerio]
50.19	19.32	XP_004086099.1	PREDICTED: uncharacterized protein LOC101157921 [Oryzias latipes]
1.05	17.71	NP_956155.1	protein-tyrosine phosphatase-like member B [Danio rerio]
1.09	14.29	NP_956849.2	programmed cell death protein 10-A [Danio rerio]
1.17	50.00	NP_001008651.1	prostaglandin reductase 1 [Danio rerio]
2.37	23.73	NP_001091731.1	uncharacterized protein LOC100002730 precursor [Danio rerio]
1.20	24.60	NP_571320.1	DNA-binding protein inhibitor ID-1 [Danio rerio]
1.04	50.00	NP_001008584.1	low affinity cationic amino acid transporter 2 [Danio rerio]
1.22	25.43	NP_958456.1	trans-2,3-enoyl-CoA reductase [Danio rerio]
17.11	17.78	AAT52142.1	CC chemokine SCYA108 [Ictalurus furcatus]
1.08	23.41	NP_997956.1	adenosine kinase isoform 1 [Danio rerio]
1.13	19.35	NP_001159390.1	beta thymosin-like protein 2 [Danio rerio]
1.07	20.39	NP_999939.1	tetraspanin-6 [Danio rerio]
1.05	21.65	NP_001070105.1	phosphoethanolamine methyltransferase [Danio rerio]
1.59	15.36	NP_001119936.1	insulin-like growth factor-binding protein 2-B precursor [Danio rerio]
1.05	50.00	NP_001139083.1	uncharacterized protein LOC571915 [Danio rerio]
1.05	28.35	NP_001002621.1	glutathione S-transferase omega 1 [Danio rerio]
1.84	38.27	NP_001038953.1	type IV antifreeze protein precursor [Danio rerio]
3.76	34.38	NP_001001409.2	actin, alpha, cardiac muscle 1a [Danio rerio]
1.09	50.00	NP_001002217.1	carboxypeptidase A4 precursor [Danio rerio]
1.07	22.47	NP_998035.1	probable 4-hydroxy-2-oxoglutarate aldolase, mitochondrial precursor [Danio rerio]
1.55	16.26	NP_001018363.1	complement C1q subcomponent subunit A precursor [Danio rerio]
1.09	14.34	NP_957180.1	glutaryl-Coenzyme A dehydrogenase [Danio rerio]
1.24	16.39	NP_571395.1	plastin-2 [Danio rerio]
1.12	14.56	NP_001002495.2	cytochrome b-c1 complex subunit 8 [Danio rerio]
1.05	50.00	NP_001004582.1	chymotrypsin-like precursor [Danio rerio]
1.14	14.10	NP_956842.2	RNA-binding protein PNO1 [Danio rerio]
1.12	22.88	AFC88124.1	immunoglobulin light chain variable region [Hemibagrus macropterus]
1.25	15.11	NP_001071025.1	multiple inositol polyphosphate histidine phosphatase, 1a [Danio rerio]
1.16	15.19	NP_998544.1	aspartate aminotransferase 2a [Danio rerio]
1.41	16.21	NP_001007152.1	sex hormone binding globulin precursor [Danio rerio]
1.05	14.42	NP_001002521.1	signal peptidase complex catalytic subunit SEC11A [Danio rerio]
1.31	16.38	NP_001018539.2	carboxypeptidase B2 precursor [Danio rerio]

1.22	19.06	NP_001019822.1	fatty acid binding protein 1-B.1 [Danio rerio]
1.39	15.16	NP_001076267.1	14-3-3 protein beta/alpha-A [Danio rerio]
1.03	43.33	XP_685374.1	PREDICTED: ATP synthase subunit f, mitochondrial isoform 1 [Danio rerio]
1.10	14.43	NP_997887.1	alpha-enolase [Danio rerio]
1.14	19.16	NP_001071070.1	cytochrome P450, family 4, subfamily V, polypeptide 8 [Danio rerio]
1.18	28.88	NP_631919.1	lysozyme-like protein 2 precursor [Danio rerio]
1.03	47.85	XP_002665065.1	PREDICTED: transmembrane protein 176B-like [Danio rerio]
1.41	47.43	NP_571322.1	L-lactate dehydrogenase B-A chain [Danio rerio]
1.08	16.39	AAT52140.1	CC chemokine SCYA106 [Ictalurus furcatus]
1.09	24.09	XP_001922616.3	PREDICTED: saxitoxin and tetrodotoxin-binding protein 2-like [Danio rerio]
1.09	22.48	NP_001120807.1	nucleolar RNA helicase 2 [Danio rerio]
1.08	40.00	AFM86947.1	hypothetical protein, partial [Callorhinchus milii]
1.05	21.39	XP_003198150.1	PREDICTED: hypothetical protein LOC100535957 [Danio rerio]
1.22	15.69	NP_956640.1	diablo, IAP-binding mitochondrial protein a [Danio rerio]
1.01	50.00	JC1348	hypothetical 18K protein - goldfish mitochondrion
1.15	14.03	NP_001002378.1	ferritin heavy chain [Danio rerio]
1.12	16.26	NP_001070200.1	40S ribosomal protein S16 [Danio rerio]
1.02	18.40	NP_956340.1	40S ribosomal protein S12 [Danio rerio]
1.00	50.00	XP_002805692.1	PREDICTED: hypothetical protein LOC701528 [Macaca mulatta]

III.8 Location of genes with high bulk frequency ratio (BFR)

To determine the genomic location of SNPs with high levels of BFR, genes containing SNPs with high BFR (BRF≥4) were initially used as query for BLAST searches against the draft catfish genome sequence scaffolds in relation to the linkage map. A total of 354 genes were identified to contain significant SNPs with high BFR (Table 7). BLASTN searches were conducted to determine the locations of the 354 genes on the scaffolds of the catfish genome draft sequence, and they were found to be within 201 genomic sequence scaffolds. Of the 201 scaffolds, 134 can be located in linkage groups, and the remaining cannot be located on linkage groups because there were no markers from these scaffolds were mapped.

In order to further analyze the linkage disequilibrium (LD) and to reduce the possibility of false positives, we placed additional stringent parameters including SNPs with BFR ≥5, along with combined bulk allele ratio <4 (to assure that the high BFR is likely caused by genetic segregation) and the probability of significant SNP <0.001. As shown in Figure 4, eleven LG potentially harbor genes with significant SNPs. However, with only two families used in the infection group (resistant and suseptible), if a gene is truly involved in resistance, long stretches of genomic segments are expected to be in LD because of genetic linkage. Therefore, the locations and distributions of genes containing significant SNPs with the following characteristics were further analyzed: 1) at least 5 genes were involved in the LG with significant SNPs; 2) or if the genes with significant SNPs were fewer than 5, at least one SNPs have a BFR equal or greater than 10. Using this set of criteria, eight LG appeared to harbor QTLs involved in ESC disease resistance. These LG were LG 1, 3, 6, 9, 15, 17, 18, and 25. Of these eight LG, six contained genes with SNPs having BFR>10 (Figure 4). Of the eight LG, LG6, 15, and 17 had the largest numbers of genes with significant SNPs.

Figure 4. The distribution of high BFR genes (BFR≥4) on linkage groups.

Blue stands for the number of genes with 5≤BFR<10, red stands for the number of genes with BFR≥10.

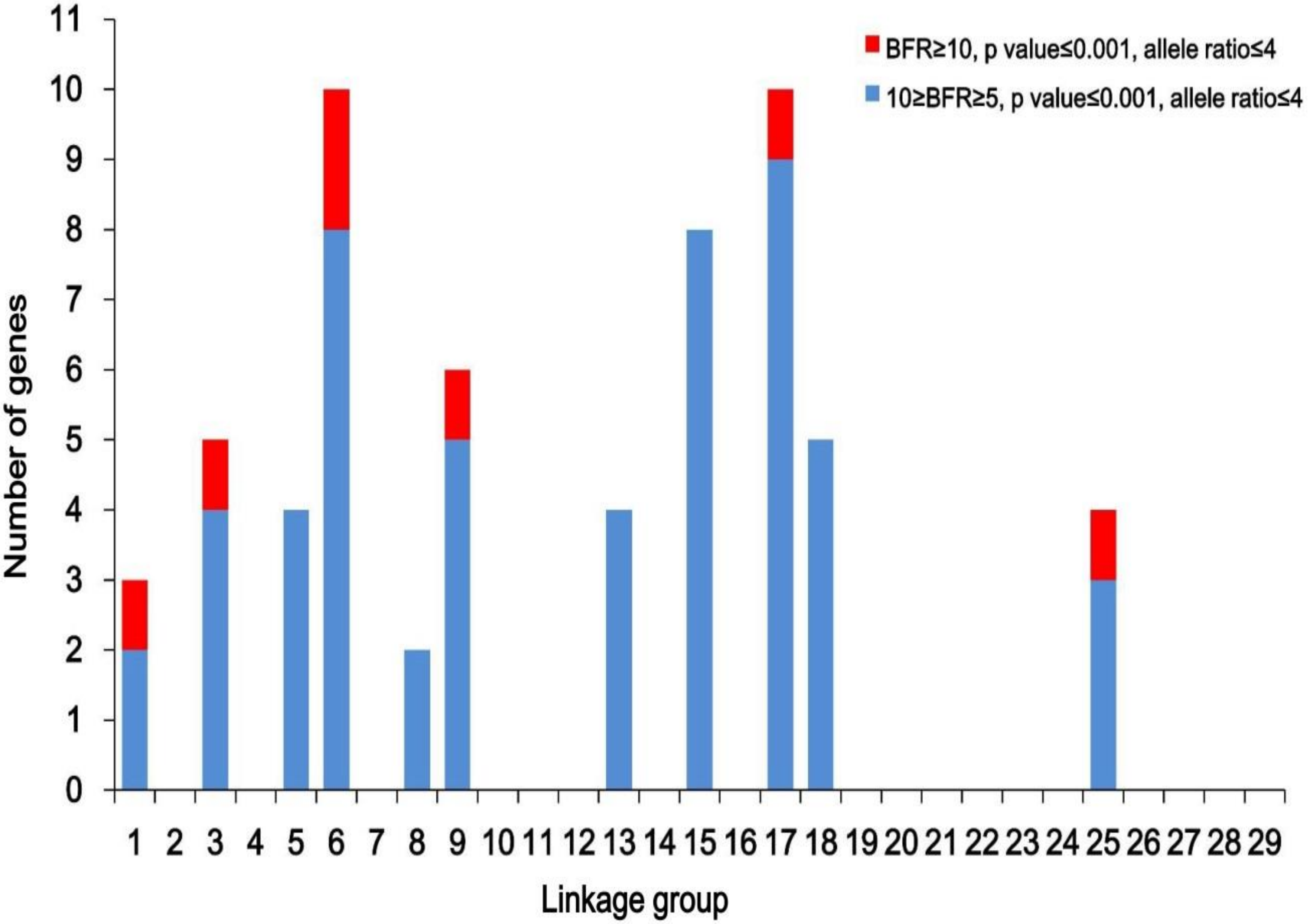

BFR≥10, p value≤0.001, allele ratio≤4
10≥BFR≥5, p value≤0.001, allele ratio≤4
Number of genes
0
1
2
3
4
5
6
7
8
9
10
11
Linkage group
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29

Figure 5. Distribution of genes containing SNPs of high BFR in linkage group 6

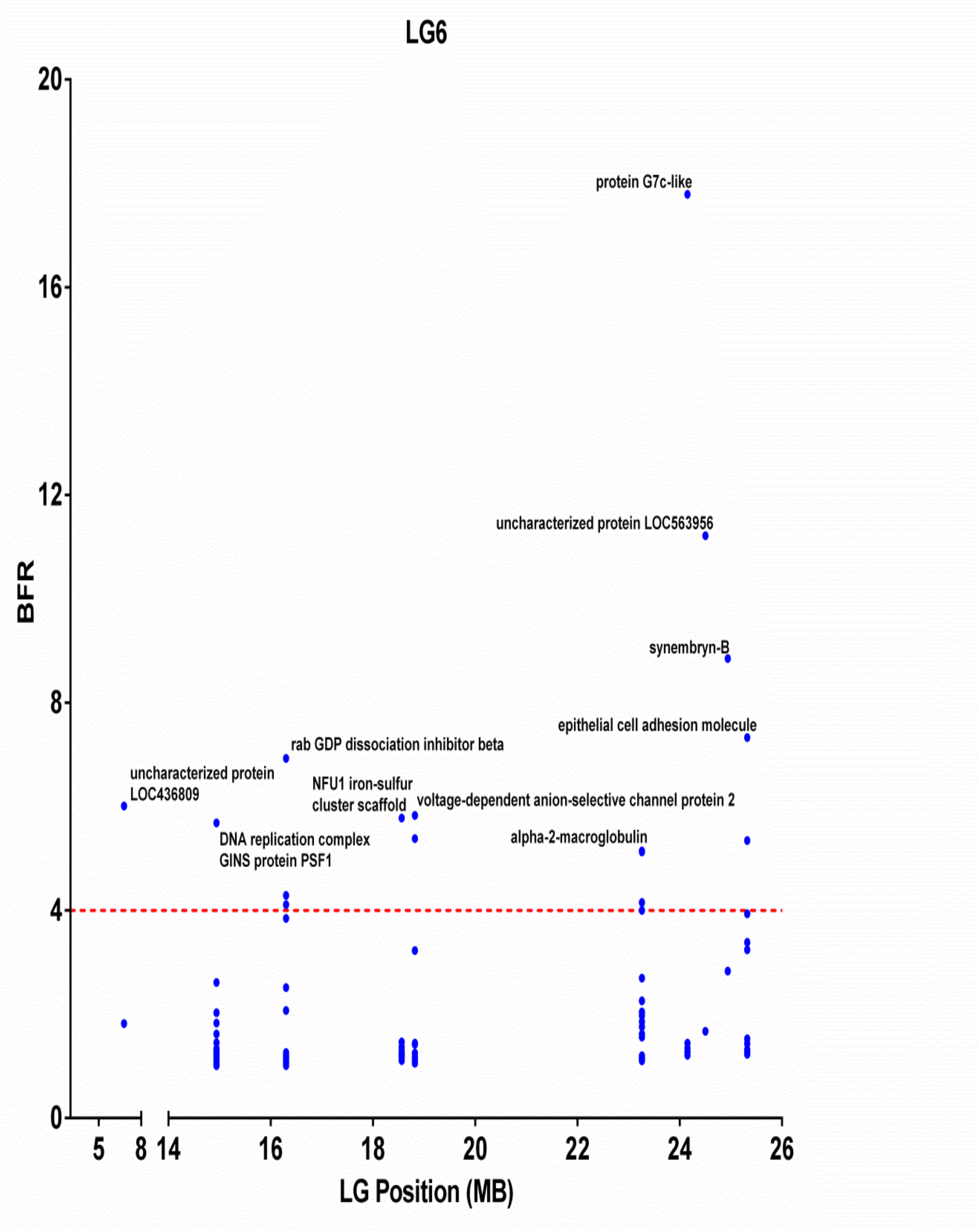


Figure 6. Distribution of genes containing SNPs of high BFR in linkage group 15

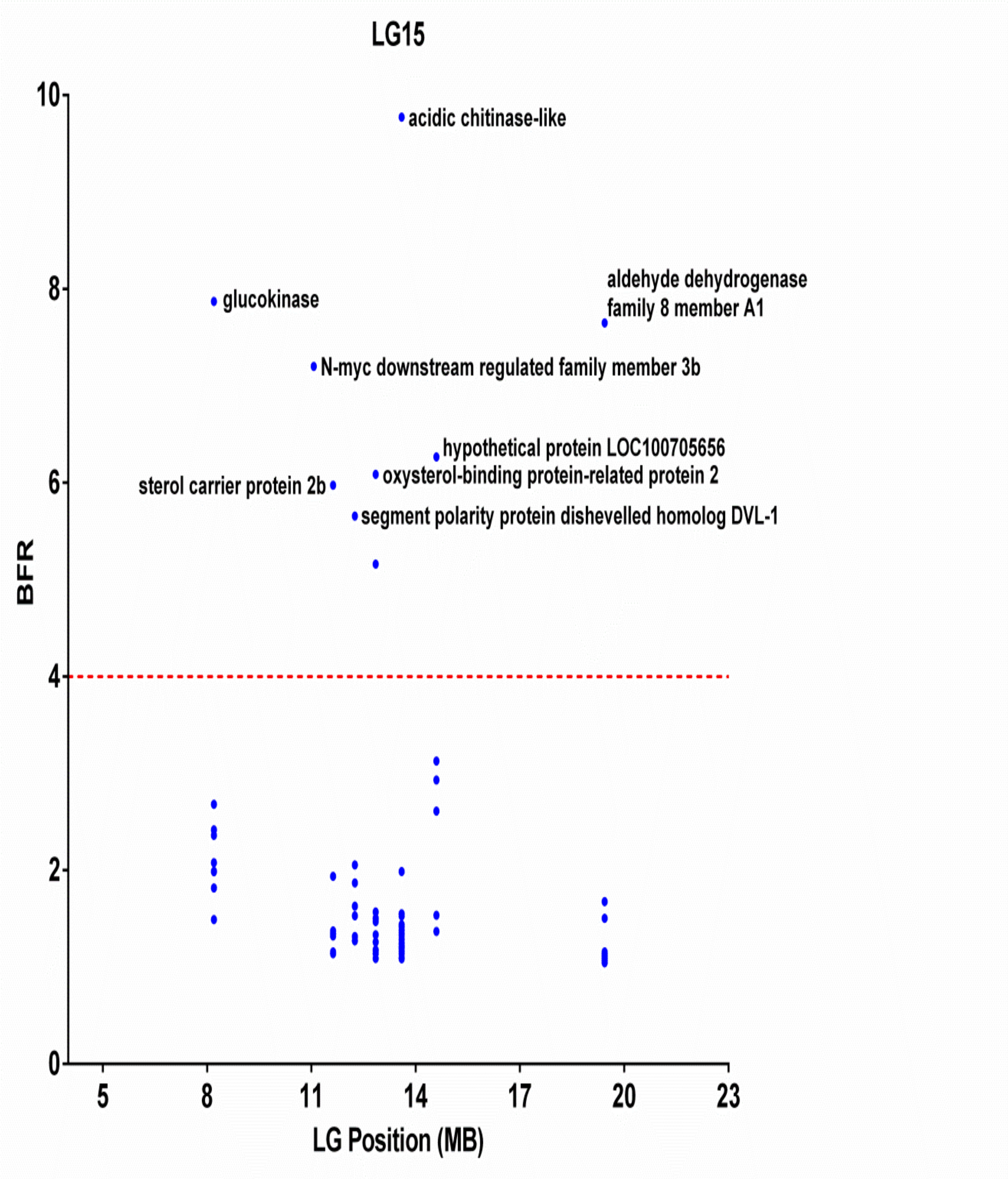
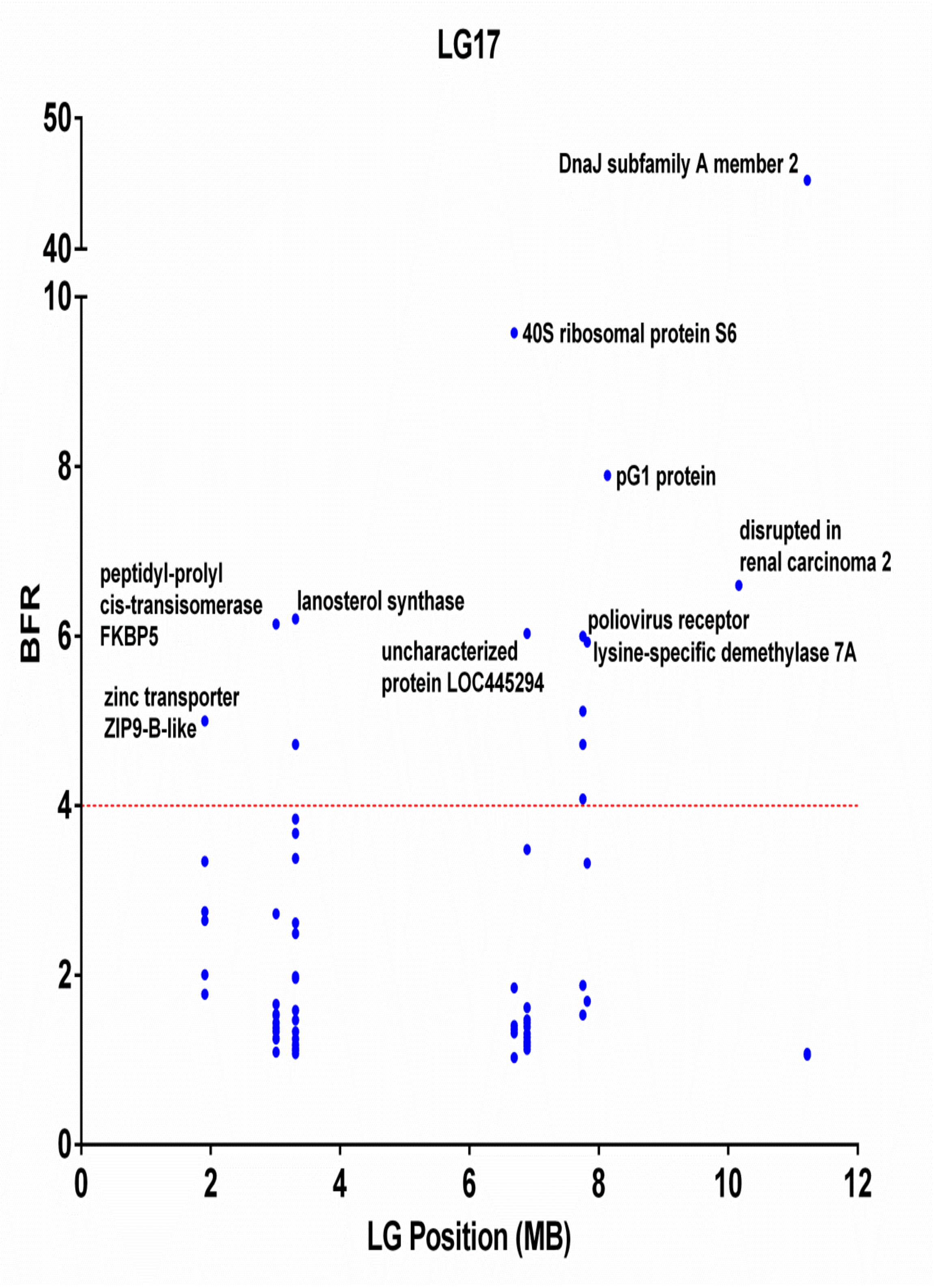


Figure 7. Distribution of genes containing SNPs of high BFR in linkage group 17



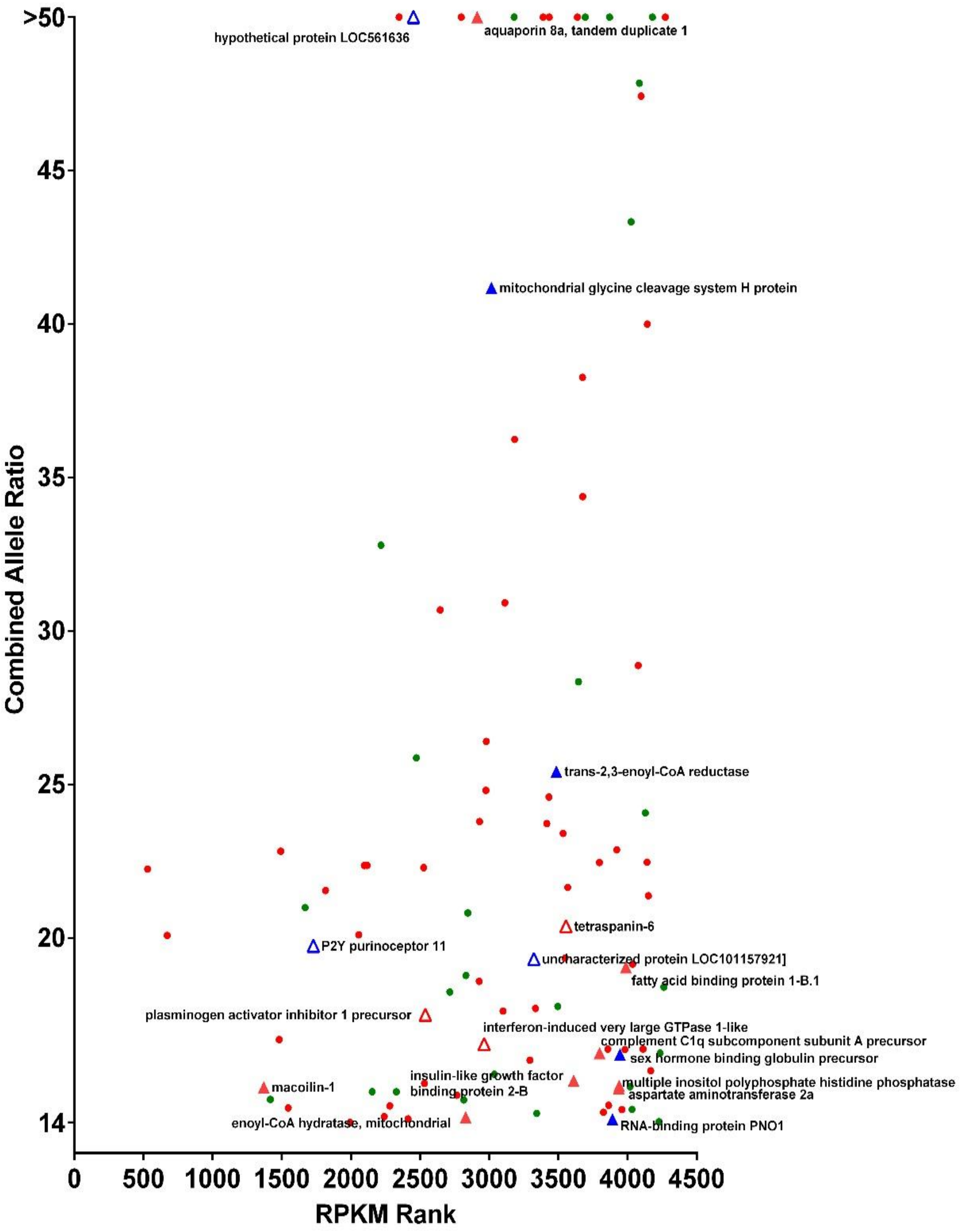
Detailed distributions of genes within LG were determined by locating the genes on the scaffolds along the linkage group. As shown in Figure 5, in LG6, SNPs within the von Willebrand factor A domain-containing protein 7-like (G7c-like) gene had a very high BFR (17.8), and the nearby genes had a much lower BFR, suggesting that the resistant gene(s) were near the G7c-like gene. In LG 15, the SNP with the highest BFR was found in the middle of LG15 within the acidic chitinase-like gene, and BFR values in surrounding genes were gradually lower along both sides of the LG, suggesting tight linkage of the resistance gene(s) near the acidic chitinase gene (Figure 6). In LG17, one SNP was found to have extremely high BFR (45.3) within the DnaJ subfamily A member 2 gene (homologue of HSP40). BFR values of SNPs within genes left of the DnaJ subfamily A member 2 gene dropped sharply, but no genes with high BFR values were found on the right side of the DnaJ subfamily A member 2 gene (Figure 7).

III.9 Parental origin of highly expressed alleles

As blue catfish is generally more resistant to ESC disease than channel catfish, an interesting question to ask is which allele was preferentially expressed after ESC infection within the backcross progenies used in this study. We therefore attempted to analyze the parental origin of the alleles for the genes with high combined allele ratios. As shown in Figure 8, a total of 98 genes harboring SNPs with combined allele ratio of 14 or greater were identified. Of these, the parental origins could be determined for 18 genes with existing genome information, while the parental origin of the remaining 80 could not be determined. Of the 18 genes, 11 were of channel catfish origin and 7 were of blue catfish origin. Of the 11 genes preferentially expressed with channel catfish alleles, six were expressed high in resistant fish and five were expressed highly in susceptible fish. Similarly, of the 7 genes preferentially expressed with the blue catfish alleles, four were expressed highly in resistant fish while 3 were expressed highly in susceptible fish (Figure 8).

Figure 8. Genes harbouring significant SNPs, plotted by their combined allele ratios versus the rank of RPKM.

Red dots stand for genes with the preferentially expressed allele expressed higher in resistant group and their parental origin unknown; Solid red triangles stand for genes with the preferentially expressed allele expressed higher in resistant group and their parental origin being channel catfish; Unfilled red triangles stand for genes with the preferentially expressed allele expressed higher in susceptible group and their parental origin being channel catfish; Solid blue triangles stand for genes with the preferentially expressed allele expressed higher in resistant group and their parental origin being blue catfish; Unfilled blue triangles stand for genes with the preferentially expressed allele expressed higher in susceptible group and their parental origin being blue catfish; and green dots stand for genes with the preferentially expressed allele expressed higher in susceptible group and their parental origin unknown.

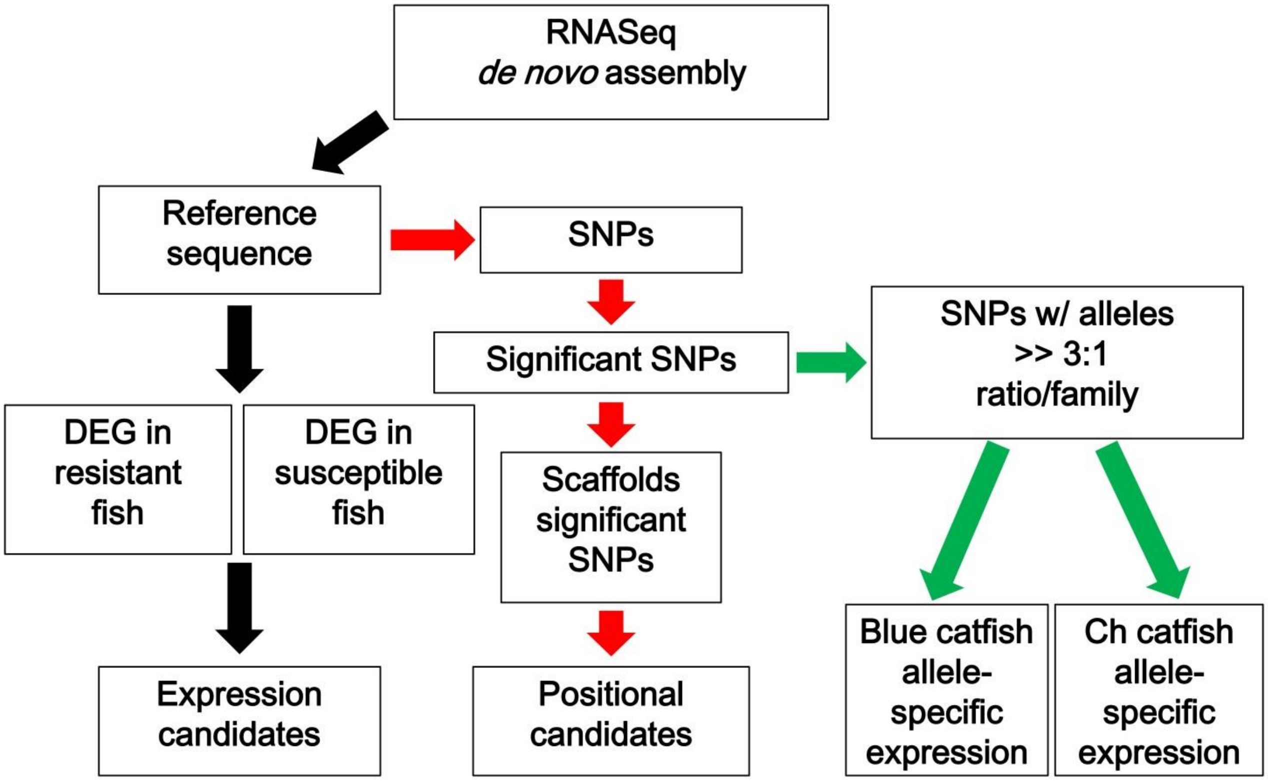


● Gene with low BFR (<4) and biased allele ratio (≥14:1), their major alleles were mainly contributed from susceptible to group
● Gene with low BFR (<4) and biased allele ratio (≥14:1), their major alleles were mainly contributed from resistent group
▲ Blue catfish allele specific expressed gene, their major alleles were mainly contributed from resistent group
▲ Channel catfish allele specific expressed gene, their major alleles were mainly contributed from resistent group
△ Channle catfish allele specific expressed gene, their major alleles were mainly contributed from susceptible group
△ Blue catfish allele specific expressed gene, their major alleles were mainly contributed from susceptible group

IV Discussion

In this study, we conducted BSR-Seq by combining the NGS-based RNA-seq with bulk segregant analysis for the analysis of genes involved in disease response and disease resistance against enteric septicemia of catfish (ESC). Such a simple combination of RNA-seq and BSA analysis allowed identification of differentially expressed genes, localization of disease resistance-related genes in linkage groups, and analysis of allele-specific expression (Figure 9).

Figure 9. Flow through of BSR-seq.



BSR-Seq carried the full capability of RNA-seq that allowed identification of differentially expressed genes. Comparison of expression in resistant fish pool and susceptible fish pool with the control allowed the identification of differentially expressed genes after infection. A total of 1,240 and 224 genes were identified to be differentially expressed after ESC infection in susceptible fish and resistant fish, respectively. In the susceptible fish, many of the up-regulated genes represent the acute phase response protein genes, as previously reported [106]. Apparently, microarray studies were limited to the gene probes existing on the array, while RNA-seq analysis has the ability to detect all induced genes, depending on their expression levels. Clearly, the greater numbers of genes identified from this study after infection indicated that RNA-seq is more sensitive than the microarray analysis. Although RNA-seq analysis was also previously conducted after ESC infection [107], the tissues were different in these studies. In the work of Li et al. [107], intestine tissue was used, while liver was used in this study. Nevertheless, many of the differentially expressed genes identified here in the liver were among the differentially expressed genes in the intestine, as well as those identified in the microarray studies. For instance, the acute phase response (APR) genes such as CC chemokines, Toll-like receptors, complement component proteins, catechol-O-methyltransferase domain containing 1, apolipoprotein proteins, fibrinogens, angiotensinogen, MHC class I and II, Tumor necrosis factors were all found to be up-regulated, as found previously [106].

The use of phenotypic extremes, resistant and susceptible fish, allowed comparison of expression patterns of genes involved in immune responses, although the time point was quite different. For instance, a number of immune-related genes were found to be expressed higher in resistant fish than in susceptible fish including apolipoprotein A IV, apolipoprotein Ab, apolipoprotein Eb, apolipoprotein Bb, apolipoprotein B100, apolipoprotein M and complement component 1q (C1q), complement component 1s (C1s), complement component 3, complement component 3a, fibrinogen alpha, fibrinogen gamma, MHC class I, and MHC II. These genes expressed at higher levels in resistant fish could indicate their importance in the related disease resistance.

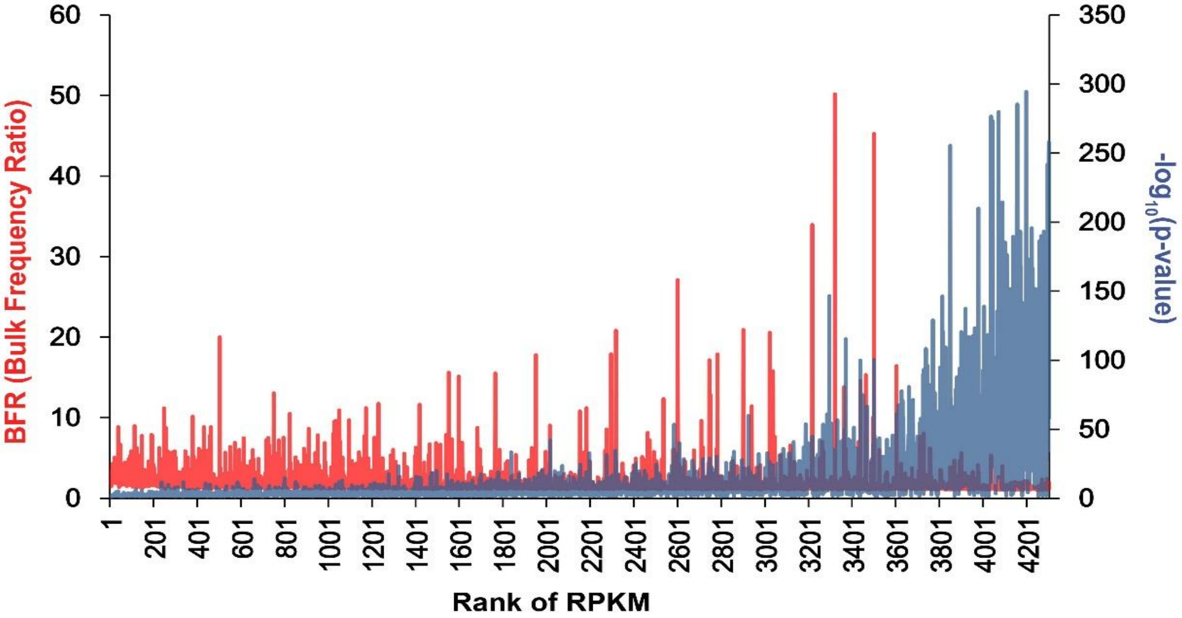
Apolipoproteins have been shown to be important for disease resistance in mice [108-110] and chickens [111, 112].

Apparently, much larger number of differentially expressed genes were identified in susceptible fish than in resistant fish as compared with the control. We believe that much of this difference was caused by the different time point (3-5 days versus 2 weeks) between these two groups. As heavy mortalities occurred 4-6 days after infection, the vast majority of differentially expressed genes in susceptible fish represented the massive responses of the host against the infection including the acute phase proteins and genes involved in inflammation and immune responses [106]. In contrast, the resistant fish samples were collected two weeks after infection from the survivors. As such, these resistant fish were either "resistant" or not infected, or may have cleared the bacteria from their system. The massive host responses to infection may have been over by the time of two weeks after infection.

In addition to its ability to identify differentially expressed genes, BSR-Seq carried the ability to identify significant SNPs between the pools of very strong phenotype contrast. Localization of the significant SNPs along the chromosomes should allow identification of positional candidate genes responsible for the trait. In this study, use of F2-generation backcross progenies, when coupled to pooling of samples from phenotypic extremes, allowed analysis of significant SNPs between the resistant and susceptible fish. We initially identified a set of significant SNPs between the two bulks using Fisher's Exact test. Although Fisher's Exact test is technically suitable for the binary SNP markers, it is compromised by the expression levels. Because the "allele frequencies" were called at the RNA level, highly expressed genes had much lower p-values (Figure 10). Therefore, p-values were used only as the initial step for the identification of positional candidate SNPs linked with resistance QTLs.

Figure 10. Relationship between RPKM, BFR and p-value.

The BFR is shown on the left-Y axis (red) and P-values are shown on right-Y axis (blue), both against the rank of RPKM on X-axis. Note there is correlation between p-values and RPKM, but not the BFR.



Bulk frequency ratio (BFR) was previously used as an effective parameter for genetic analysis in BSA or BSR-Seq [95]. However, in those cases, genotypes at the DNA level were known. Here in this study, the “allele frequency” was calculated from the mapped reads of RNA samples, and thus the calculated BFR could be compromised by allele-specific expression. In order to identify positional candidates for resistance using transcriptome datasets generated from BSR-Seq, we need to differentiate allele frequencies caused by genetic segregation and those caused by allele-specific induction/suppression: Significant SNPs with large BFR and small combined allele ratio are caused by genetic segregation; significant SNPs with small BFR but large combined allele ratio are caused by allele-specific expression, while significant SNPs with large BFR and large combined allele ratios are caused by more than just genetic segregation including allele-specific expression. For instance, if the allele ratio at an A/G SNP site is 10 to 1 in resistance pool, and 1 to 10 in susceptible pool, the BFR should be 10. When the two bulks were combined, now the allele ratio of A/G is 11:11 = 1. This SNP, with a high BFR and a low combined allele ratio, should be a SNP with allele frequency difference between the bulks caused by genetic segregation. In contrast, when one of the two alleles is differentially up-regulated, the combined allele ratio will stay large. For instance, at an A/G SNP site, if A is significantly up-regulated in the resistant fish, say 100A:5G, and in susceptible fish, A and G are roughly expressed equally, both at low levels, say 5A:5G. In this case, the BFR = (100/105)/(5/10) = 1.91; when the two bulks are combined, the combined allele ratio would be 10.5. Apparently in this case, the large allele ratio is caused by allele-specific expression.

In this study, each bulk was made of 24 fish with 12 fish from one family and 12 fish from a second family. As the exact allele ratio at each SNP site is unknown in the two families, we made several assumptions for the analysis of SNPs due to genetic segregation and those due to allele-specific expression. At an A/G SNP site, the parent in one family could be AA x GG, AA x AG, or AG x AG. In these cases, the largest allele ratio can be 3:1 (in the case of AA x AG) at the DNA level. When two families were used, as in this study, the largest allele ratio at the DNA level could be 7:1, i.e., AA x AG in one family, and homozygous AA x AA in another family. Any SNPs with significantly larger combined allele ratio than 7:1 would suggest allele-specific expression. We identified SNPs with combined allele ratio of greater than 14 (twice the largest possible allele ratio at the DNA level) as being allele-specifically expressed. Apparently, many SNPs fell between the two possibilities, and to obtain reasonable assessment of those caused by genetic segregation and those caused by allele-specific expression, we neglected those ones between the two possibilities (Figure 3).

Location of significant SNPs caused by genetic segregation along a well-assembled genome sequence should allow identification of genomic segments controlling the trait. In catfish, the whole genome sequence assembly is still under way. We therefore, mapped the significant SNPs to scaffolds and then examined the patterns of the SNP distribution. Within a QTL region, statistical significance should be the highest with the gene underlining the performance trait, and LD should decay gradually on both sides of the chromosome surrounding the gene [113]. In our study, quite many significant SNPs were located on unmapped scaffolds, but many were also mapped to genetic linkage groups including LG6, LG15, and LG17 that included at least 10 genes with high BFR (≥5) and low combined allele ratio (≤4). As shown in Figures 5. 6, and 7, the LD appeared to be decaying around the most significant SNPs, suggesting that these genomic regions indeed contain resistance-related genes. For instance, in LG6, the gene containing the most significant SNP was protein G7c-like gene located at the 24Mb position, and the BFR values on both sides of this gene were significant, but lower than BFR within the G7c-like gene (Figure 5). Similarly, the gene containing the highest BFR was acidic chitinase-like gene that located in the middle of a 23 Mb DNA in LG15, and the BFRs were lower on both sides along this LG (Figure 6). In LG17, the SNPs with highest BFR was located within the DnaJ subfamily A member 2 gene close to the end of the 12 Mb DNA in LG17, and the BFR on both sides were lower (Figure 7). It was unknown if the detected genes with the highest BFR were themselves involved in disease resistance. This was because some linked genes with even greater BFR were not yet mapped to the linkage group, staying as isolated scaffolds, and therefore cannot be viewed under the same “Manhattan plot”, or because the expression level of the real disease resistance gene was so low that it was not detected under the BSR-Seq analysis.

One logical thought is that if one gene is seriously involved in disease resistance, it should be correlated with positional candidate genes as well as expression candidate genes. In other words, it should be differentially expressed either between resistant and susceptible fish or differentially expressed after infection, and located at a genomic location the resistance phenotype is mapped. A cross examination of differentially expressed genes and genes with high BFR but low combined allele ratio allowed identification of 17 genes (Table 10). Only four of these 17 genes have been mapped to linkage groups, whereas the remaining 13 were just mapped to isolated scaffolds. Further mapping studies are required to determine how many QTLs are involved. These genes included some important immune function related genes such as the NLR, MHC-related genes, and Mannan-binding lectin serine peptidase. Regardless of their direct involvement in the disease resistance, the importance of these genes in resistance and immune responses to ESC should not be neglected. Additional research is warranted to determine if these genes are responsible for the resistance against ESC.

Table 10. Differentially expressed genes after infection with high bulk frequency ratios and low combined allele ratio

Gene	BFR	Combined allele ratio	Known function
Metalloreductase STEAP2	21.2	3	Regulates iron ion homeostasis and involved in endocytosis
Tumor suppressor candidate 5 homolog	34.0	2.3	Inhibits breast tumor formation in *vivo*
Protein G7c-like	17.8	3.2	Effect the susceptibility to lung tumors
Acidic chitinase-like	9.8	1.4	Participates in the defense against nematodes, fungi and

			bacteria. Plays a role in T-helper cell type 2 (Th2) immune response.
Mannan-binding lectin serine peptidase 2	4.1	2.2	Lectin complement pathway actication
Inter-alpha-trypsin inhibitor heavy chain H4	4.3	1.3	Type II acute-phase protein (APP) involves in inflammatory responses to trauma. May also play a role in liver development or regeneration.
D-amino acid oxidase	4.3	1.8	Regulation of the glutamatergic neurotransmission; may play a role in the glutamatergic mechanisms of schizophrenia
MAWD binding protein like	4.4	1.4	Inhibits proliferation and invasion in gastric cancer
Cytokeratin-like	4.5	1.5	Mediate epithelial innate defense through their antimicrobial properties
Spermidine/spermine N1-acetyltransferase	4.7	1	Involves in polyamine homeostasis
Stonustoxin subunit alpha-like	4.7	1.9	Induces hemolytic activities, displays edema-inducing activities, increases vascular permeability and interferes irreversibly with neuromuscular function.
UDP-glucose 4-epimerase	4.8	1.8	Catalyzes the epimerization of UDP-glucose to UDP-galactose and the epimerization of UDP-N-acetylglucosamine to UDP-N-acetylgalactosamine.
Major histocompatibility complex class I UDA	5.1	3.1	Play an important role in immune response and antigen processing and presentation
NLR-C8	5.9	1.4	Involves in the gram negative bacteria recognition. Against the intracellular pathogen.
MARCKS-like 1a	6.1	2.5	Most prominent cellular substrate for protein kinase C. It can bind

			calmodulin, actin, and synapsin.
3-oxoacid CoA transferase 1a	7.0	1.3	Key enzyme for ketone body catabolism. Also plays and important roles in the energy metabolism of spermatozoa.
MHC class II beta chain	17.9	2.4	Involves in antigen processing and presentation of peptide or polysaccharide antigen via MHC class II

A total of 98 genes were identified as genes with allele-specific expression (Figure 3). One obvious question is what causes allele-specific expression (ASE). Two hypotheses were previously proposed to account for ASE [114, 115]. In the first hypothesis, mutations in cis-acting DNA elements can cause differences in binding of trans-acting factors, especially when such mutations are located in the promoter or enhancer regions. Inversely, mutations in the trans-acting factors would also cause their differences in binding to their target sequence. In the second hypothesis, epigenetic factors such as differential methylation of the two alleles can cause differences in expression levels. It has long been known that mutations in non-coding regions which affect gene expression can cause human genetic disease [114, 116]. A differentially expressed gene exhibits *cis*-acting variation when the differential expression is caused by factors linked to the differentially expressed alleles, such as differences in promoter sequences or chromatin state. The list of examples in which *cis*-acting regulatory variation plays a key role in phenotypic variation are increasing [114]. *In vitro* experiments prove that variants in gene promoter regions effect rates of transcription and these variants may also lead to a significant proportion of differential allelic expression [115]. In additions, expressed genes contain *trans*-acting variation when the differential expression is caused by factors unlinked to the differentially expressed alleles, such as differences caused by genetic background and regulatory networks.

The importance of DNA methylation as a driving factor in allele specific expression has been claimed and proved by a number of studies [115, 117]. In these studies, a directly correlation between allele specific expression and methylation was observed. Clearly, the different epigenetic state of each haploid genome is a major factor in the expression of the two alleles. Although X-chromosome inactivation and silencing are usually considered to be mainly related to epigenetic effect [118], some studies also suggested that the change of gene regulation caused by epigenetic modification of sequence variation might be a common pathogenic mechanism in mammals [119]. One hypothetical role for epigenetics is genetic imprinting leading to mono-allelic expression [120]. Both *in vivo* and *in vitro* experiments indicate that allele-specific differences in the rate of transcription are commonly existed, if not all genes are likely to show differential allelic expression in different individuals [115]. However, the role of ASE genes in complex traits is still not clear.

Most ASE studies have been conducted in humans, and no studies have been conducted in fish. In addition to the above discussed possibilities, it is noteworthy that we used an interspecific hybrid system in this study. In the channel catfish backcrossed progenies, overall 50% of chromosomes are “homozygous” from channel catfish while 50% chromosomes are heterozygous with one chromosome being from channel catfish and the other from blue catfish. If a *trans*-acting factor is transcribed from channel catfish genes, it would bind the *cis*-acting elements from channel catfish with greater affinity, causing allelic expression. This would be an additional logical explanation for the observed ASE.

The significance of the observed allele-specific expression in relation to phenotype is unknown at present. To date, the majority of expression analysis focuse on the total amount of the transcripts. However, emerging evidence underlies the importance of understanding the allele specific transcript in cases of disease [114, 121]. In our studies here, among the 98 allele-specifically expressed genes, parental origin of alleles can be determine for only 18 genes. Among these 18 genes, the channel catfish allele was expressed higher in 11 genes (six high in resistant fish and five high in susceptible fish), and blue catfish allele was expressed higher in 7 genes (four high in resistant fish and three high in susceptible fish). There was no correlation of resistance with a specific parent. However, it is possible that certain combinations of alleles would warrant resistance and certain combinations of alleles would lead to susceptibility. This clearly warrants future studies.

V Conclusions

In this study, we demonstrated the application of BSR-Seq to study disease resistance by combining RNA-seq with bulk segregant analysis. It has the full capacity for the identification of differentially expressed genes, the capacity to identify significant SNPs between phenotypic bulks, the capacity to potentially identify the positional candidate genes, and the ability to identify allelic expressed genes. Among many differentially expressed genes, 17 genes were also among the genes that had high BFR and low combined allele ratio, suggesting that these genes could be potentially involved in disease resistance.

Reference

1. Chatakondi N, Yant D, Dunham R: **Commercial production and performance evaluation of channel catfish, Ictalurus punctatus female × blue catfish, Ictalurus furcatus male F-1 hybrids.** *Aquaculture* 2005, **247**.
2. Miller N, Wilson M, Bfufltén E, Stuge T, Warr G, Ciem W: **Functional and molecular characterization of teleost leukocytes.** *Immunological Reviews* 1998, **166:**187-197.
3. Vallejo AN, Miller NW, William Clem L: **Antigen processing and presentation in teleost immune responses.** *Annual Review of Fish Diseases* 1992, **2:**73-89.
4. Barker K, Khayat M, Miller N, Wilson M, Clem LW, Bengtén E: **Immortal and mortal clonal lymphocyte lines from channel catfish: comparison of telomere length, telomerase activity, tumor suppressor and heat shock protein expression.** *Developmental & Comparative Immunology* 2002, **26:**45-51.
5. Zhou H, Bengtén E, Miller NW, Clem LW, Wilson M: **The T Cell Receptor β Locus of the Channel Catfish, Ictalurus punctatus, Reveals Unique Features.** *The Journal of Immunology* 2003, **170:**2573-2581.
6. Shen L, Stuge TB, Evenhuis JP, Bengtén E, Wilson M, Chinchar VG, Clem LW, Miller NW: **Channel catfish NK-like cells are armed with IgM via a putative FcμR.** *Developmental & Comparative Immunology* 2003, **27:**699-714.
7. Antao AB, Wilson M, Wang J, Bengtén E, Miller NW, Clem LW, Chinchar VG: **Genomic organization and differential expression of channel catfish MHC class I genes.** *Developmental & Comparative Immunology* 2001, **25:**579-595.
8. Hawke NA, Yoder JA, Haire RN, Mueller MG, Litman RT, Miracle AL, Stuge T, Shen L, Miller N, Litman GW: **Extraordinary variation in a diversified family of immune-type receptor genes.** *Proceedings of the National Academy of Sciences* 2001, **98:**13832-13837.
9. Bao B, Peatman E, Peng X, Baoprasertkul P, Wang G, Liu Z: **Characterization of 23 CC chemokine genes and analysis of their expression in channel catfish (Ictalurus punctatus).** *Developmental and Comparative Immunology* 2006, **30:**783.
10. Peatman E, Liu Z: **Evolution of CC chemokines in teleost fish: a case study in gene duplication and implications for immune diversity.** *Immunogenetics* 2007, **59:**613-623.
11. Bao B, Peatman E, Xu P, Li P, Zeng H, He C, Liu Z: **The catfish liver-expressed antimicrobial peptide 2 (LEAP-2) gene is expressed in a wide range of tissues and developmentally regulated.** *Molecular Immunology* 2006, **43:**367-377.
12. Xu P, Bao B, He Q, Peatman E, He C, Liu Z: **Characterization and expression analysis of bactericidal permeability-increasing protein (BPI) antimicrobial peptide gene from channel catfish< i> Ictalurus punctatus</i>.** *Developmental & Comparative Immunology* 2005, **29:**865-878.
13. Wang Q, Wang Y, Xu P, Liu Z: **NK-lysin of channel catfish: gene triplication, sequence variation, and expression analysis.** *Molecular Immunology* 2006, **43:**1676-1686.
14. Wang Q, Bao B, Wang Y, Peatman E, Liu Z: **Characterization of a NK-lysin antimicrobial peptide gene from channel catfish.** *Fish and Shellfish Immunology* 2006, **20:**419-426.

15. Baoprasertkul P, Peatman E, Abernathy J, Liu Z: **Structural characterisation and expression analysis of Toll-like receptor 2 gene from catfish.** *Fish & Shellfish Immunology* 2007, **22:**418-426.
16. Baoprasertkul P, Peatman E, Somridhivej B, Liu Z: **Toll-like receptor 3 and TICAM genes in catfish: species-specific expression profiles following infection with Edwardsiella ictaluri.** *Immunogenetics* 2006, **58:**817-830.
17. Sha Z, Xu P, Takano T, Liu H, Terhune J, Liu Z: **The warm temperature acclimation protein Wap65 as an immune response gene: its duplicates are differentially regulated by temperature and bacterial infections.** *Molecular Immunology* 2008, **45:**1458.
18. Rajendran K, Zhang J, Liu S, Peatman E, Kucuktas H, Wang X, Liu H, Wood T, Terhune J, Liu Z: **Pathogen recognition receptors in channel catfish: II. Identification, phylogeny and expression of retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs).** *Developmental & Comparative Immunology* 2012.
19. Rajendran K, Zhang J, Liu S, Kucuktas H, Wang X, Liu H, Sha Z, Terhune J, Peatman E, Liu Z: **Pathogen recognition receptors in channel catfish: I. Identification, phylogeny and expression of NOD-like receptors.** *Developmental & Comparative Immunology* 2011.
20. Zhang J, Liu S, Rajendran KV, Sun L, Zhang Y, Sun F, Kucuktas H, Liu H, Liu Z: **Pathogen recognition receptors in channel catfish: Ш. phylogeny and expression analysis of Toll-like receptors.** *Developmental and Comparative Immunology* 2013, **in press**.
21. Takano T, Sha Z, Peatman E, Terhune J, Liu H, Kucuktas H, Li P, Edholm E, Wilson M, Liu Z: **The two channel catfish intelectin genes exhibit highly differential patterns of tissue expression and regulation after infection with Edwardsiella ictaluri.** *Developmental and Comparative Immunology* 2008, **32:**693.
22. Bao B, Peatman E, Li P, He C, Liu Z: **Catfish hepcidin gene is expressed in a wide range of tissues and exhibits tissue-specific upregulation after bacterial infection.** *Developmental & Comparative Immunology* 2005, **29:**939-950.
23. Liu H, Takano T, Abernathy J, Wang S, Sha Z, Jiang Y, Terhune J, Kucuktas H, Peatman E, Liu Z: **Structure and expression of transferrin gene of channel catfish, Ictalurus punctatus.** *Fish & Shellfish Immunology* 2010, **28:**159.
24. Jiang Y, Abernathy J, Peatman E, Liu H, Wang S, Xu D, Kucuktas H, Klesius P, Liu Z: **Identification and characterization of matrix metalloproteinase-13 sequence structure and expression during embryogenesis and infection in channel catfish (Ictalurus punctatus).** *Developmental and Comparative Immunology* 2010, **34:**590-597.
25. Niu D, Peatman E, Liu H, Lu J, Kucuktas H, Liu S, Sun F, Zhang H, Feng T, Zhou Z: **Microfibrillar-associated protein 4 (MFAP4) genes in catfish play a novel role in innate immune responses.** *Developmental & Comparative Immunology* 2011, **35:**568-579.
26. Zhou Z, Liu H, Liu S, Sun F, Peatman E, Kucuktas H, Kaltenboeck L, Feng T, Zhang H, Niu D: **Alternative complement pathway of channel catfish (Ictalurus punctatus): molecular characterization, mapping and expression analysis of factors Bf/C2 and Df.** *Fish & Shellfish Immunology* 2012, **32:**186.
27. Peatman E, Baoprasertkul P, Terhune J, Xu P, Nandi S, Kucuktas H, Li P, Wang S, Somridhivej B, Dunham R, Liu Z: **Expression analysis of the acute phase response in**

channel catfish (Ictalurus punctatus) after infection with a Gram-negative bacterium. *Developmental & Comparative Immunology* 2007, **31:**1183-1196.

28. Serapion J, Kucuktas H, Feng J, Liu Z: **Bioinformatic mining of type I microsatellites from expressed sequence tags of channel catfish (Ictalurus punctatus).** *Marine biotechnology* 2004, **6:**364-377.
29. Somridhivej B, Wang S, Sha Z, Liu H, Quilang J, Xu P, Li P, Hu Z, Liu Z: **Characterization, polymorphism assessment, and database construction for microsatellites from BAC end sequences of channel catfish (Ictalurus punctatus): A resource for integration of linkage and physical maps.** *Aquaculture* 2008, **275:**76-80.
30. Kucuktas H, Wang S, Li P, He C, Xu P, Sha Z, Liu H, Jiang Y, Baoprasertkul P, Somridhivej B: **Construction of genetic linkage maps and comparative genome analysis of catfish using gene-associated markers.** *Genetics* 2009, **181:**1649-1660.
31. Waldbieser GC, Bosworth BG, Nonneman DJ, Wolters WR: **A microsatellite-based genetic linkage map for channel catfish, Ictalurus punctatus.** *Genetics* 2001, **158:**727-734.
32. Wang S, Xu P, Thorsen J, Zhu B, de Jong PJ, Waldbieser G, Kucuktas H, Liu Z: **Characterization of a BAC library from channel catfish Ictalurus punctatus: indications of high levels of chromosomal reshuffling among teleost genomes.** *Marine biotechnology* 2007, **9:**701-711.
33. Xu P, Wang S, Liu L, Thorsen J, Kucuktas H, Liu Z: **A BAC-based physical map of the channel catfish genome.** *Genomics* 2007, **90:**380-388.
34. Liu H, Jiang Y, Wang S, Ninwichian P, Somridhivej B, Xu P, Abernathy J, Kucuktas H, Liu Z: **Comparative analysis of catfish BAC end sequences with the zebrafish genome.** *BMC genomics* 2009, **10:**592.
35. Wang S, Peatman E, Abernathy J, Waldbieser G, Lindquist E, Richardson P, Lucas S, Wang M, Li P, Thimmapuram J: **Assembly of 500,000 inter-specific catfish expressed sequence tags and large scale gene-associated marker development for whole genome association studies.** *Genome biology* 2010, **11:**R8.
36. Morin PA, Luikart G, Wayne RK, the SNPwg: **SNPs in ecology, evolution and conservation.** *Trends in Ecology & Evolution* 2004, **19:**208-216.
37. Abasht B, Lamont SJ: **Genome-wide association analysis reveals cryptic alleles as an important factor in heterosis for fatness in chicken F2 population.** *Animal Genetics* 2007, **38:**491-498.
38. Duijvesteijn N, Knol E, Merks J, Crooijmans R, Groenen M, Bovenhuis H, Harlizius B: **A genome-wide association study on androstenone levels in pigs reveals a cluster of candidate genes on chromosome 6.** *BMC Genetics* 2010, **11:**42.
39. Du Z-Q, Zhao X, Vukasinovic N, Rodriguez F, Clutter AC, Rothschild MF: **Association and Haplotype Analyses of Positional Candidate Genes in Five Genomic Regions Linked to Scrotal Hernia in Commercial Pig Lines.** *PLoS ONE* 2009, **4:**e4837.
40. Khatkar M, Nicholas F, Collins A, Zenger K, Cavanagh J, Barris W, Schnabel R, Taylor J, Raadsma H: **Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel.** *BMC Genomics* 2008, **9:**187.
41. Kim E, Berger P, Kirkpatrick B: **Genome-wide scan for bovine twinning rate QTL using linkage disequilibrium.** *Animal Genetics* 2009, **40:**300 - 307.
42. Meuwissen T, Hayes B, Goddard M: **Prediction of total genetic value using genome-wide dense marker maps.** *Genetics* 2001, **157:**1819 - 1829.

43. Brooks S, Gabreski N, Miller D, Brisbin A, Brown H, Streeter C, Mezey J, Cook D, Antczak D: **Whole-Genome SNP Association in the Horse: Identification of a Deletion in Myosin Va Responsible for Lavender Foal Syndrome.** *Plos Genet* 2010, **6:**e1000909.
44. Becker D, Tetens J, Brunner A, Burstel D, Ganter M, Kijas J, Drogemuller C, Consortium I: **Microphthalmia in Texel Sheep Is Associated with a Missense Mutation in the Paired-Like Homeodomain 3 (PITX3) Gene.** *PLoS ONE* 2010, **5:**e8689.
45. Kijas J, Townley D, Dalrymple B, Heaton M, Maddox J, McGrath A, Wilson P, Ingersoll R, McCulloch R, McWilliam S: **A genome wide survey of SNP variation reveals the genetic structure of sheep breeds.** *PLoS ONE* 2009, **4:**e4668.
46. Eck S, Benet-Pages A, Flisikowski K, Meitinger T, Fries R, Strom T: **Whole genome sequencing of a single Bos taurus animal for single nucleotide polymorphism discovery.** *Genome Biology* 2009, **10:**R82.
47. Moen T, Hayes B, Baranski M, Berg P, Kjoglum S, Koop B, Davidson W, Omholt S, Lien S: **A linkage map of the Atlantic salmon (Salmo salar) based on EST-derived SNP markers.** *BMC Genomics* 2008, **9:**223.
48. Hubert S, Higgins B, Borza T, Bowman S: **Development of a SNP resource and a genetic linkage map for Atlantic cod (Gadus morhua).** *BMC Genomics*, **11:**191.
49. He C, Chen L, Simmons M, Li P, Kim S, Liu Z: **Putative SNP discovery in interspecific hybrids of catfish by comparative EST analysis.** *Anim Genet* 2003, **34:**445 - 448.
50. Wang S, Peatman E, Abernathy J, Waldbieser G, Lindquist E, Richardson P, Lucas S, Wang M, Li P, Thimmapuram J: **Assembly of 500,000 inter-specific catfish expressed sequence tags and large scale gene-associated marker development for whole genome association studies.** *Genome Biol* 2010, **11:**R8.
51. Wang S, Sha Z, Sonstegard T, Liu H, Xu P, Somridhivej B, Peatman E, Kucuktas H, Liu Z: **Quality assessment parameters for EST-derived SNPs from catfish.** *BMC Genomics* 2008, **9:**450.
52. Mardis E: **The impact of next-generation sequencing technology on genetics.** *Trends Genet* 2008, **24:**133 - 141.
53. Novaes E, Drost D, Farmerie W, Pappas G, Grattapaglia D, Sederoff R, Kirst M: **High-throughput gene and SNP discovery in Eucalyptus grandis, an uncharacterized genome.** *BMC Genomics* 2008, **9:**312.
54. Parchman T, Geist K, Grahnen J, Benkman C, Buerkle C: **Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery.** *BMC Genomics*, **11:**180.
55. Vera J, Wheat C, Fescemyer H, Frilander M, Crawford D, Hanski I, Marden J: **Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing.** *Mol Ecol* 2008, **17:**1636 - 1647.
56. Hale M, McCormick C, Jackson J, DeWoody J: **Next-generation pyrosequencing of gonad transcriptomes in the polyploid lake sturgeon (Acipenser fulvescens): the relative merits of normalization and rarefaction in gene discovery.** *BMC Genomics* 2009, **10:**203.

57. Meyer E, Aglyamova G, Wang S, Buchanan-Carter J, Abrego D, Colbourne J, Willis B, Matz M: **Sequencing and de novo analysis of a coral larval transcriptome using 454 GSFlx.** *BMC Genomics* 2009, **10:**219.
58. Cirulli E, Singh A, Shianna K, Ge D, Smith J, Maia J, Heinzen E, Goedert J, Goldstein D, Immunology C: **Screening the human exome: a comparison of whole genome and whole transcriptome sequencing.** *Genome Biology* 2010, **11:**R57.
59. Surget-Groba Y, Montoya-Burgos J: **Optimization of de novo transcriptome assembly from next-generation sequencing data.** *Genome Res*, **20:**1432 - 1440.
60. Trick M, Long Y, Meng J, Bancroft I: **Single nucleotide polymorphism (SNP) discovery in the polyploid Brassica napus using Solexa transcriptome sequencing.** *Plant Biotechnol J* 2009, **7:**334 - 346.
61. Wold B, Myers RM: **Sequence census methods for functional genomics.** *Nat Methods* 2008, **5:**19-21.
62. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq.** *Nat Methods* 2008, **5:**621-628.
63. Simon SA, Zhai J, Nandety RS, McCormick KP, Zeng J, Mejia D, Meyers BC: **Short-read sequencing technologies for transcriptional analyses.** *Annu Rev Plant Biol* 2009, **60:**305-333.
64. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M: **The transcriptional landscape of the yeast genome defined by RNA sequencing.** *Science* 2008, **320:**1344-1349.
65. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR: **Highly integrated single-base resolution maps of the epigenome in Arabidopsis.** *Cell* 2008, **133:**523-536.
66. Cloonan N, Forrest AR, Kolle G, Gardiner BB, Faulkner GJ, Brown MK, Taylor DF, Steptoe AL, Wani S, Bethel G, et al: **Stem cell transcriptome profiling via massive-scale mRNA sequencing.** *Nat Methods* 2008, **5:**613-619.
67. Wang XW, Luan JB, Li JM, Bao YY, Zhang CX, Liu SS: **De novo characterization of a whitefly transcriptome and analysis of its gene expression during development.** *BMC Genomics* 2010, **11:**400.
68. Oleksiak MF, Roach JL, Crawford DL: **Natural variation in cardiac metabolism and gene expression in Fundulus heteroclitus.** *Nature genetics* 2004, **37:**67-72.
69. Hallman D, Boerwinkle E, Saha N, Sandholzer C, Menzel H, Csazar A, Utermann G: **The apolipoprotein E polymorphism: a comparison of allele frequencies and effects in nine populations.** *American journal of human genetics* 1991, **49:**338.
70. He G, Chen B, Wang X, Li X, Li J, He H, Yang M, Lu L, Qi Y, Wang X: **Conservation and divergence of transcriptomic and epigenomic variation in maize hybrids.** *Genome biology* 2013, **14:**R57.
71. Bevan M, Uauy C: **Genomics reveals new landscapes for crop improvement.** *Genome biology* 2013, **14:**206.
72. Young N: **QTL mapping and quantitative disease resistance in plants.** *Annual review of phytopathology* 1996, **34:**479-501.
73. Bunney WE, Bunney BG, Vawter MP, Tomita H, Li J, Evans SJ, Choudary PV, Myers RM, Jones EG, Watson SJ: **Microarray technology: a review of new strategies to discover candidate vulnerability genes in psychiatric disorders.** *American Journal of Psychiatry* 2003, **160:**657-666.

74. Wang Z, Gerstein M, Snyder M: **RNA-Seq: a revolutionary tool for transcriptomics.** *Nature Reviews Genetics* 2009, **10:**57-63.
75. Oshlack A, Robinson MD, Young MD: **From RNA-seq reads to differential expression results.** *Genome Biol* 2010, **11:**220.
76. Molinié B, Georgel P: **Genetic and epigenetic regulations of prostate cancer by genistein.** *Drug News Perspect* 2009, **22:**247-254.
77. Ng J-H, Heng J-CD, Loh Y-H, Ng H-H: **Transcriptional and epigenetic regulations of embryonic stem cells.** *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 2008, **647:**52-58.
78. He H, Hua X, Yan J: **Epigenetic regulations in hematopoietic Hox code.** *Oncogene* 2010, **30:**379-388.
79. Gururani MA, Venkatesh J, Upadhyaya CP, Nookaraju A, Pandey SK, Park SW: **Plant disease resistance genes: Current status and future directions.** *Physiological and Molecular Plant Pathology* 2012, **78:**51-65.
80. Zhao S, Zhu M, Chen H: **Immunogenomics for identification of disease resistance genes in pigs: a review focusing on Gram-negative bacilli.** *Journal of animal science and biotechnology* 2012, **3:**1-13.
81. Guangli S, Zhanjing H, Congfen H, Yinzhu S, Wang J: **Identification of the molecular markers linked to the salt-resistance locus in the wheat using RAPD-BSA technique.** *Acta Botanica Sinica* 2000, **43:**598-602.
82. Cheema KK, Grewal NK, Vikal Y, Sharma R, Lore JS, Das A, Bhatia D, Mahajan R, Gupta V, Bharaj TS: **A novel bacterial blight resistance gene from Oryza nivara mapped to 38kb region on chromosome 4L and transferred to Oryza sativa L.** *Genetical Research* 2008, **90:**397.
83. Dussle C, Quint M, Melchinger A, Xu M, Lübberstedt T: **Saturation of two chromosome regions conferring resistance to SCMV with SSR and AFLP markers by targeted BSA.** *Theoretical and Applied Genetics* 2003, **106:**485-493.
84. Asnaghi C, Roques D, Ruffel S, Kaye C, Hoarau J-Y, Telismart H, Girard J, Raboin L, Risterucci A, Grivet L: **Targeted mapping of a sugarcane rust resistance gene (Bru1) using bulked segregant analysis and AFLP markers.** *Theoretical and Applied Genetics* 2004, **108:**759-764.
85. Becker A, Chao D-Y, Zhang X, Salt DE, Baxter I: **Bulk segregant analysis using single nucleotide polymorphism microarrays.** *PLoS One* 2011, **6:**e15993.
86. Lee BY, Penman D, Kocher T: **Identification of a sex‐determining region in Nile tilapia (Oreochromis niloticus) using bulked segregant analysis.** *Animal genetics* 2003, **34:**379-383.
87. Ruyter-Spira CP, Gu Z, Van der Poel J, Groenen M: **Bulked segregant analysis using microsatellites: mapping of the dominant white locus in the chicken.** *Poultry science* 1997, **76:**386-391.
88. Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA: **Rapid SNP discovery and genetic mapping using sequenced RAD markers.** *PLoS One* 2008, **3:**e3376.
89. Wenger JW, Schwartz K, Sherlock G: **Bulk segregant analysis by high-throughput sequencing reveals a novel xylose utilization gene from Saccharomyces cerevisiae.** *PLoS genetics* 2010, **6:**e1000942.

90. Huang X, Guo Y, Shui Y, Gao S, Yu H, Cheng H, Zhou R: **Multiple alternative splicing and differential expression of dmrt1 during gonad transformation of the rice field eel.** *Biology of Reproduction* 2005, **73:**1017-1024.
91. Cloonan N, Forrest AR, Kolle G, Gardiner BB, Faulkner GJ, Brown MK, Taylor DF, Steptoe AL, Wani S, Bethel G: **Stem cell transcriptome profiling via massive-scale mRNA sequencing.** *Nature methods* 2008, **5:**613-619.
92. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y: **RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays.** *Genome research* 2008, **18:**1509-1517.
93. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq.** *Nature methods* 2008, **5:**621-628.
94. Pepke S, Wold B, Mortazavi A: **Computation for ChIP-seq and RNA-seq studies.** *Nature methods* 2009, **6:**S22-S32.
95. Trick M, Adamski NM, Mugford SG, Jiang C-C, Febrer M, Uauy C: **Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat.** *BMC Plant Biology* 2012, **12:**14.
96. Liu S, Yeh C-T, Tang HM, Nettleton D, Schnable PS: **Gene mapping via bulked segregant RNA-Seq (BSR-Seq).** *PLoS One* 2012, **7:**e36406.
97. Argue BJ, Liu Z, Dunham RA: **Dress-out and fillet yields of channel catfish, Ictalurus punctatus, blue catfish, Ictalurus furcatus, and their F1, F2 and backcross hybrids.** *Aquaculture* 2003, **228:**81-90.
98. Liu Z, Karsi A, Li P, Cao D, Dunham R: **An AFLP-based genetic linkage map of channel catfish (Ictalurus punctatus) constructed by using an interspecific hybrid resource family.** *Genetics* 2003, **165:**687-694.
99. Ninwichian P, Peatman E, Liu H, Kucuktas H, Somridhivej B, Liu S, Li P, Jiang Y, Sha Z, Kaltenboeck L: **Second-generation genetic linkage map of catfish and its integration with the BAC-based physical map.** *G3: Genes| Genomes| Genetics* 2012, **2:**1233-1241.
100. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q: **Full-length transcriptome assembly from RNA-Seq data without a reference genome.** *Nature biotechnology* 2011, **29:**644-652.
101. Huang X, Madan A: **CAP3: A DNA sequence assembly program.** *Genome research* 1999, **9:**868-877.
102. Brockman W, Alvarez P, Young S, Garber M, Giannoukos G, Lee WL, Russ C, Lander ES, Nusbaum C, Jaffe DB: **Quality scores and SNP detection in sequencing-by-synthesis systems.** *Genome research* 2008, **18:**763-770.
103. Fisher RA: **On the interpretation of χ 2 from contingency tables, and the calculation of P.** *Journal of the Royal Statistical Society* 1922, **85:**87-94.
104. Kal AJ, van Zonneveld AJ, Benes V, van den Berg M, Koerkamp MG, Albermann K, Strack N, Ruijter JM, Richter A, Dujon B: **Dynamics of gene expression revealed by comparison of serial analysis of gene expression transcript profiles from yeast grown on two different carbon sources.** *Molecular biology of the cell* 1999, **10:**1859-1872.
105. Liu S, Zhou Z, Lu J, Sun F, Wang S, Liu H, Jiang Y, Kucuktas H, Kaltenboeck L, Peatman E: **Generation of genome-scale gene-associated SNPs in catfish for the construction of a high-density SNP array.** *BMC genomics* 2011, **12:**53.

106. Peatman E, Baoprasertkul P, Terhune J, Xu P, Nandi S, Kucuktas H, Li P, Wang S, Somridhivej B, Dunham R: **Expression analysis of the acute phase response in channel catfish (Ictalurus punctatus) after infection with a Gram-negative bacterium.** *Developmental & Comparative Immunology* 2007, **31:**1183-1196.
107. Li C, Zhang Y, Wang R, Lu J, Nandi S, Mohanty S, Terhune J, Liu Z, Peatman E: **RNA-seq analysis of mucosal immune responses reveals signatures of intestinal barrier disruption and pathogen entry following Edwardsiella ictaluri</i> infection in channel catfish, Ictalurus punctatus.** *Fish & Shellfish Immunology* 2012, **32:**816-827.
108. Sigel S, Bunk S, Meergans T, Doninger B, Stich K, Stulnig T, Derfler K, Hoffmann J, Deininger S, von Aulock S: **Apolipoprotein B100 is a suppressor of Staphylococcus aureus - induced innate immune responses in humans and mice.** *European Journal of Immunology* 2012, **42:**2983-2989.
109. Kesavalu L, Lucas A, Verma R, Liu L, Dai E, Sampson E, Progulske-Fox A: **Increased atherogenesis during Streptococcus mutans infection in ApoE-null mice.** *Journal of Dental Research* 2012, **91:**255-260.
110. Peterson MM, Mack JL, Hall PR, Alsup AA, Alexander SM, Sully EK, Sawires YS, Cheung AL, Otto M, Gresham HD: **Apolipoprotein B Is an Innate Barrier against Invasive Staphylococcus aureus Infection.** *Cell host & microbe* 2008, **4:**555-566.
111. Morgan RW, Sofer L, Anderson AS, Bernberg EL, Cui J, Burnside J: **Induction of host gene expression following infection of chicken embryo fibroblasts with oncogenic Marek's disease virus.** *Journal of virology* 2001, **75:**533-539.
112. Wozniak MA, Shipley SJ, Dobson CB, Parker SP, Scott FT, Leedham-Green M, Breuer J, Itzhaki RF: **Does apolipoprotein E determine outcome of infection by varicella zoster virus and by Epstein Barr virus?** *European Journal of Human Genetics* 2007, **15:**672-678.
113. Shifman S, Kuypers J, Kokoris M, Yakir B, Darvasi A: **Linkage disequilibrium patterns of the human genome across populations.** *Human molecular genetics* 2003, **12:**771-776.
114. Buckland PR: **Allele-specific gene expression differences in humans.** *Human molecular genetics* 2004, **13:**R255-R260.
115. Bell CG, Beck S: **Advances in the identification and analysis of allele-specific expression.** *Genome Med* 2009, **1:**56.
116. Wray GA: **The evolutionary significance of cis-regulatory mutations.** *Nature Reviews Genetics* 2007, **8:**206-216.
117. Milani L, Lundmark A, Nordlund J, Kiialainen A, Flaegstad T, Jonmundsson G, Kanerva J, Schmiegelow K, Gunderson KL, Lönnerholm G: **Allele-specific gene expression patterns in primary leukemic cells reveal regulation of gene expression by CpG site methylation.** *Genome research* 2009, **19:**1-11.
118. Petronis A: **The genes for major psychosis: aberrant sequence or regulation?** *Neuropsychopharmacology* 2000, **23:**1-12.
119. Morgan HD, Sutherland HG, Martin DI, Whitelaw E: **Epigenetic inheritance at the agouti locus in the mouse.** *Nature genetics* 1999, **23:**314-318.
120. Kato MV, Shimizu T, Nagayoshi M, Kaneko A, Sasaki MS, Ikawa Y: **Genomic imprinting of the human serotonin-receptor (HTR2) gene involved in development of retinoblastoma.** *American journal of human genetics* 1996, **59:**1084.

121. Koivisto U-M, Palvimo JJ, Jänne OA, Kontula K: **A single-base substitution in the proximal Sp1 site of the human low density lipoprotein receptor promoter as a cause of heterozygous familial hypercholesterolemia.** *Proceedings of the National Academy of Sciences* 1994, **91:**10526-10530.