GWAS analysis for resistance against enteric septicemia of catfish using the first-generation interspecific backcrosses

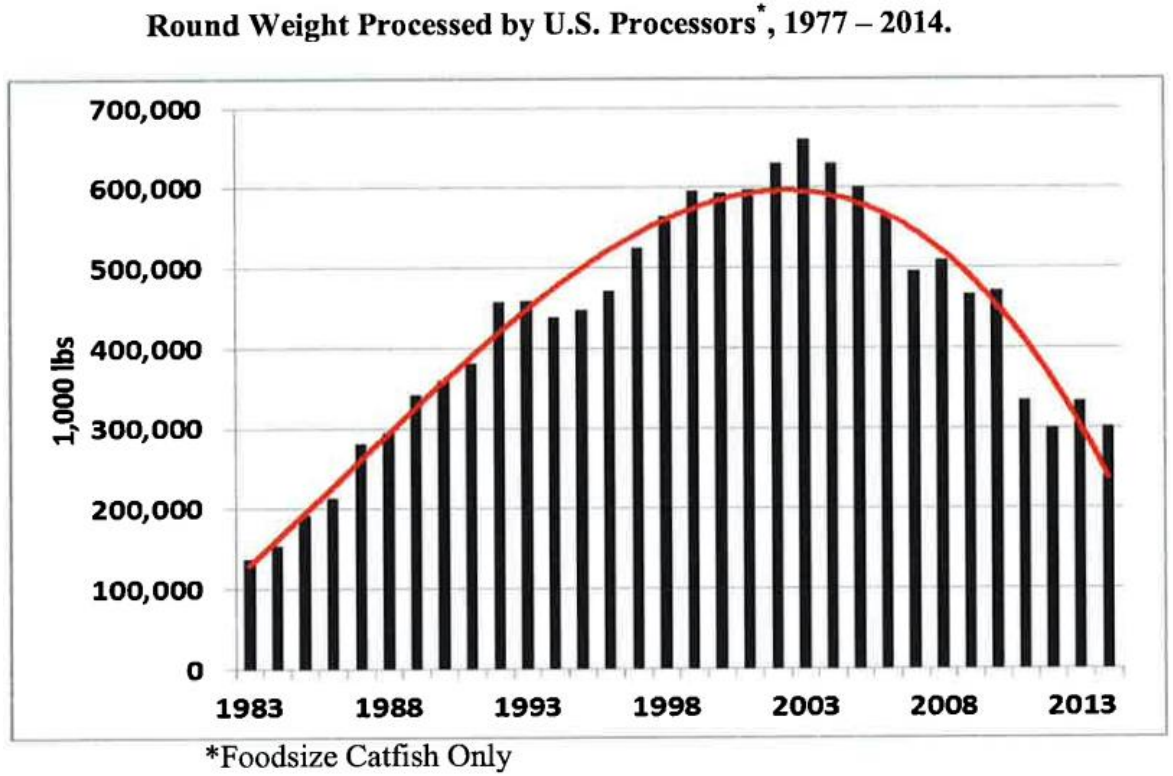
Suxu Tan
Auburn University
szt0038@auburn.edu

AUBURN
SCHOOL OF FISHERIES,
AQUACULTURE AND AQUATIC SCIENCES

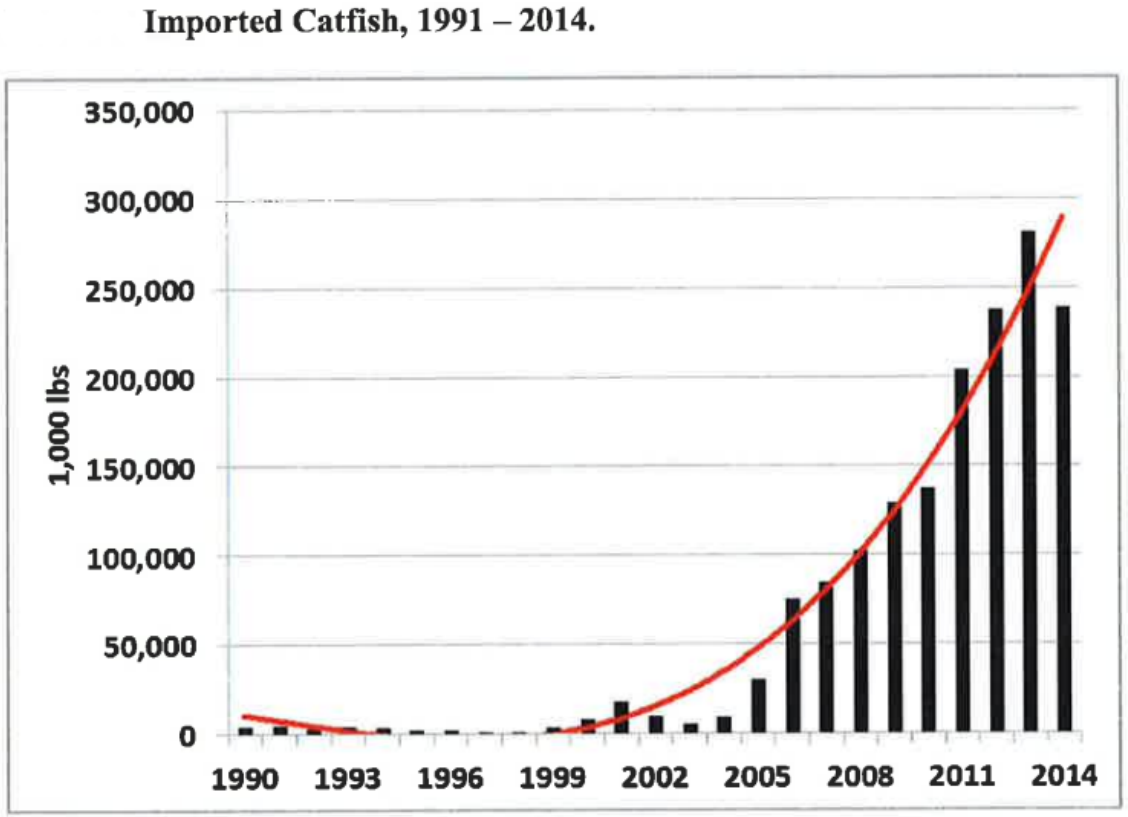
Introduction - Catfish

- Catfish can survive in a wide range of freshwater habitats such as lakes, rivers, and streams.
- Channel catfish, blue catfish, black bullhead, brown bullhead, flathead catfish, white catfish, yellow bullhead
- Catfish industry is the largest aquaculture industry in the United States, accounting for over 50% of all US aquaculture production.

Introduction - Catfish



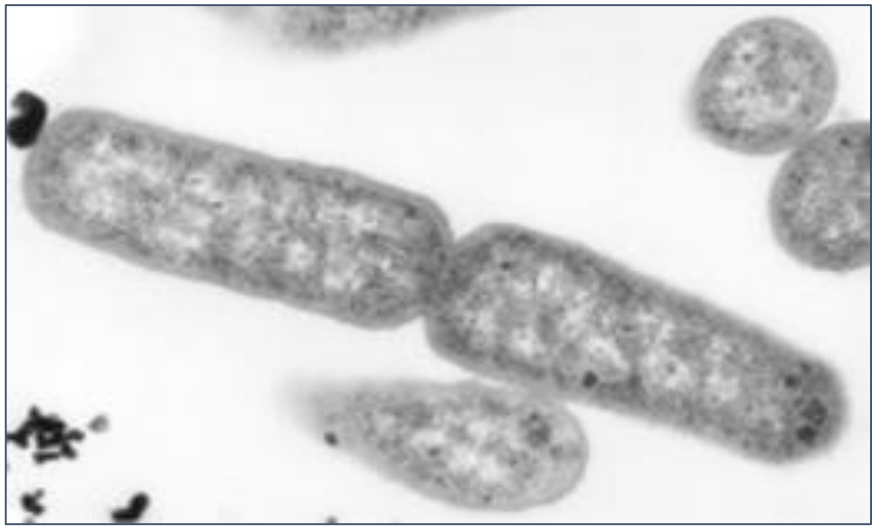
- increased feed and fuel costs
- international competition
- devastating diseases



Source: http://www.agecon.msstate.edu/whatwedo/budgets/docs/catfish2014.pdf

Introduction - ESC

- **Wide distribution**: in all catfish producing areas in the world, and in the US, mostly Mississippi, Alabama, Arkansas, and Louisiana
- **Huge losses**: $40-60 million annually



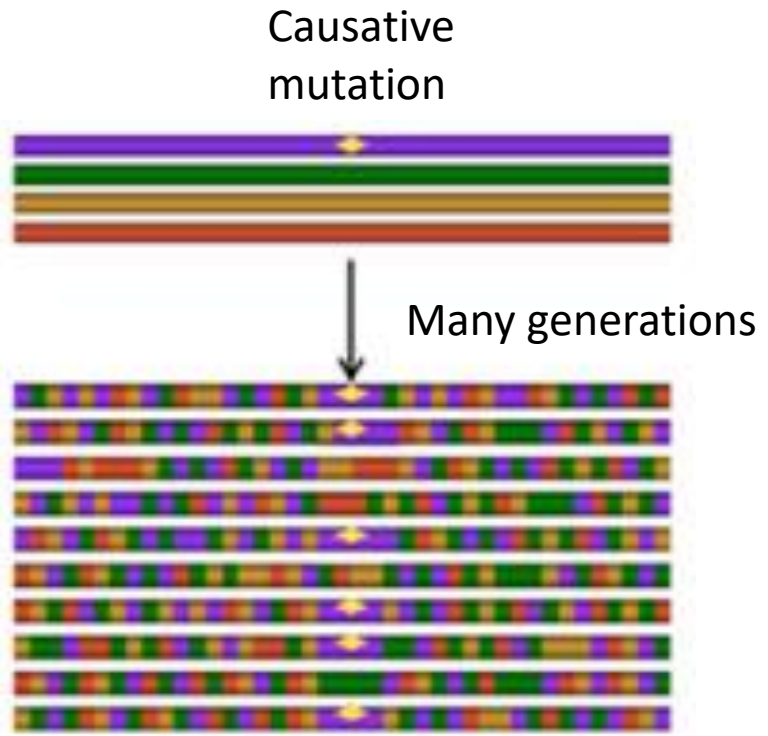
Transmission electron micrograph
Edwardsiella ictaluri

Symptoms including hole in the head

Introduction - GWAS

- GWAS can link genotype and phenotype, which examines a genome-wide set of genetic variants in individuals to find variants that associate with a phenotype of interest.

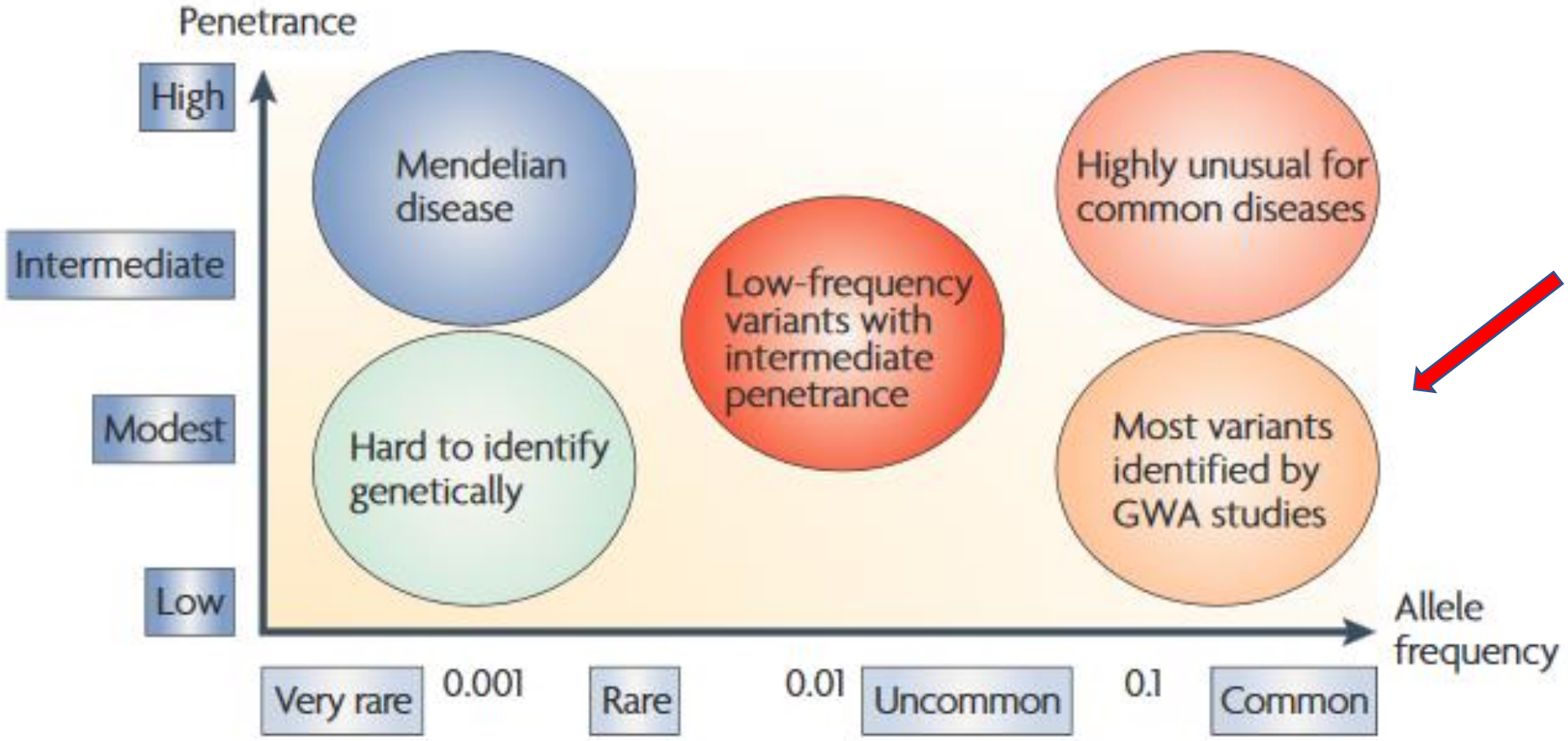
- GWAS is based upon the principle of linkage disequilibrium (LD) which is the nonrandom association between alleles at different loci.



(Zhu et al. 2008)

Introduction - GWAS

Common disease/common variant hypothesis



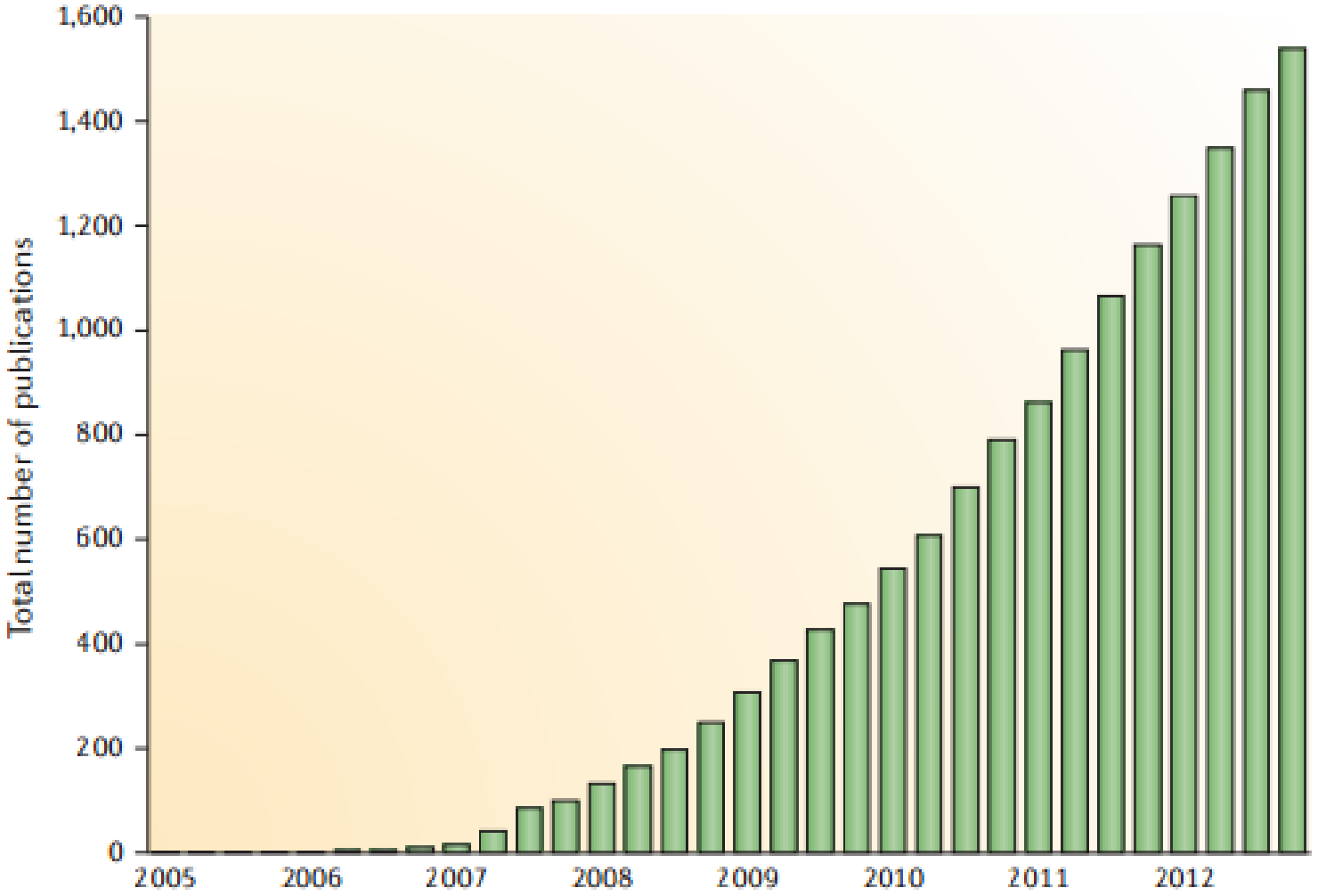
McCarthy et al. 2008

Introduction - GWAS

The first successful GWAS was reported in 2005. (Klein et al.)

96 cases
50 controls

Found a common intron variant associated with age-related macular degeneration



Pace of GWAS publications since 2005
(Manolio 2013)

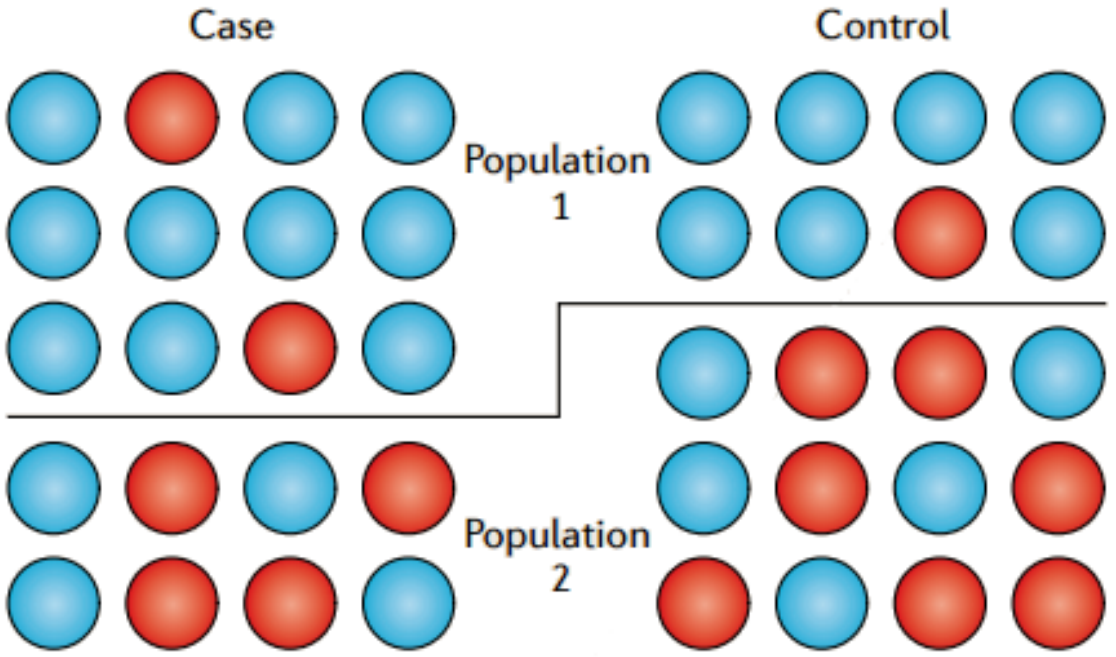
Introduction - GWAS

Advantages:

- The association mapping has high resolution.
- No pedigree information required

Disadvantages:

- Expensive
- Genotyping error
- Susceptible to population stratification



Balding 2006

Objectives

- Identification of quantitative trait locus (QTL) and SNPs associated with ESC resistance at species level

- Identification of potential candidate genes and pathways controlling ESC resistance

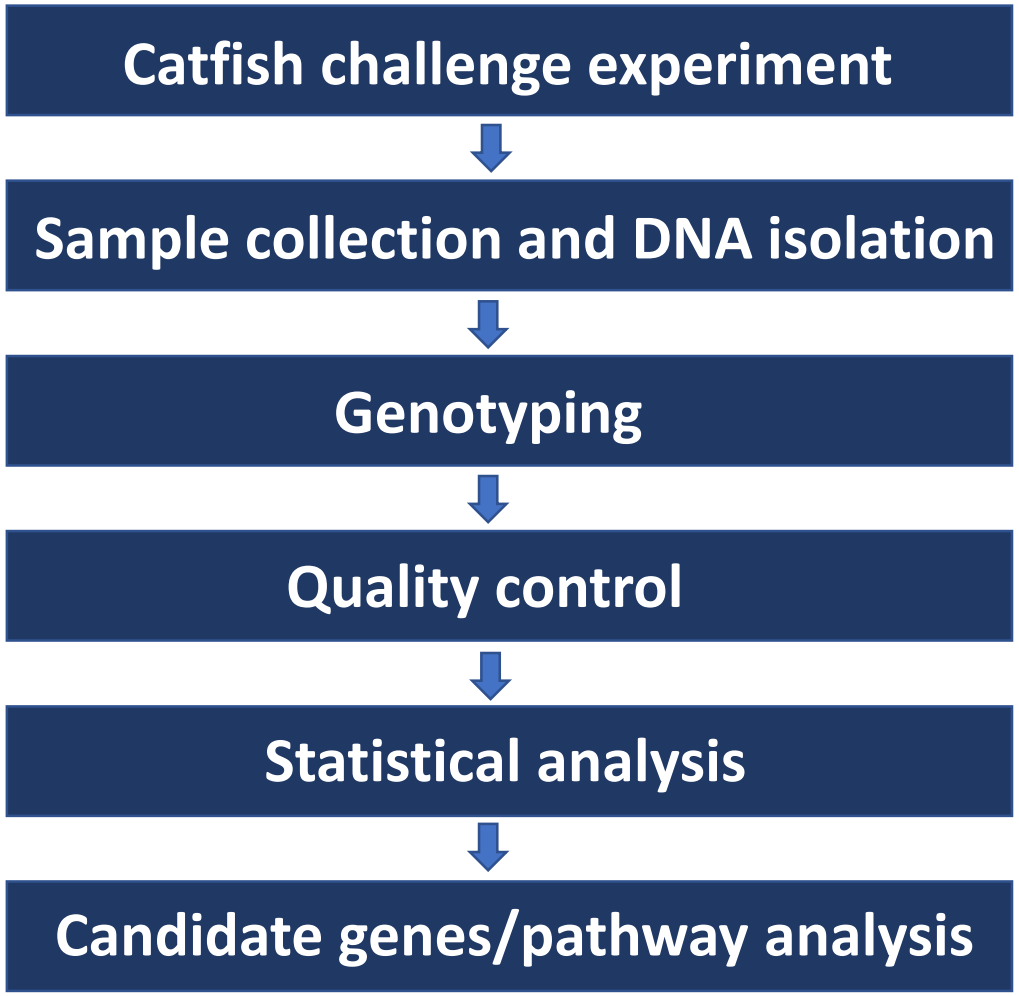
Resources Required for GWAS in this study

- ✓ Catfish population
- ✓ Development of the catfish 690K SNP array (Zeng et al., 2017)
- ✓ Powerful statistical tools: SVS software packages, PLINK, etc.

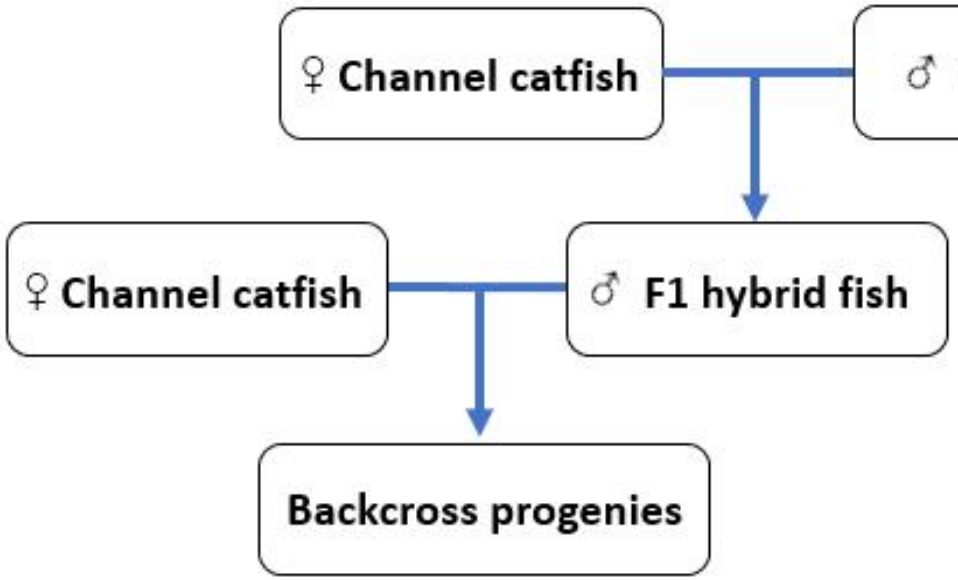
GOLDEN HELIX SNP & VARIATION SUITE

plink...

Flowchart


Catfish challenge experiment
Sample collection and DNA isolation
Genotyping
Quality control
Statistical analysis
Candidate genes/pathway analysis

Experiment design



Selective genotyping
288 phenotypic extremes in three 690k catfish SNP arrays

Family ID	Dam	Sire	Sample number	Susceptible sample number	Resistant sample number
1	Channel 1	Hybrid 1	71	36	35
2	Channel 2	Hybrid 2	70	34	36
3	Channel 3	Hybrid 3	70	36	34
4	Channel 4	Hybrid 4	77	36	41

Quality control and LD pruning

- Sample quality control

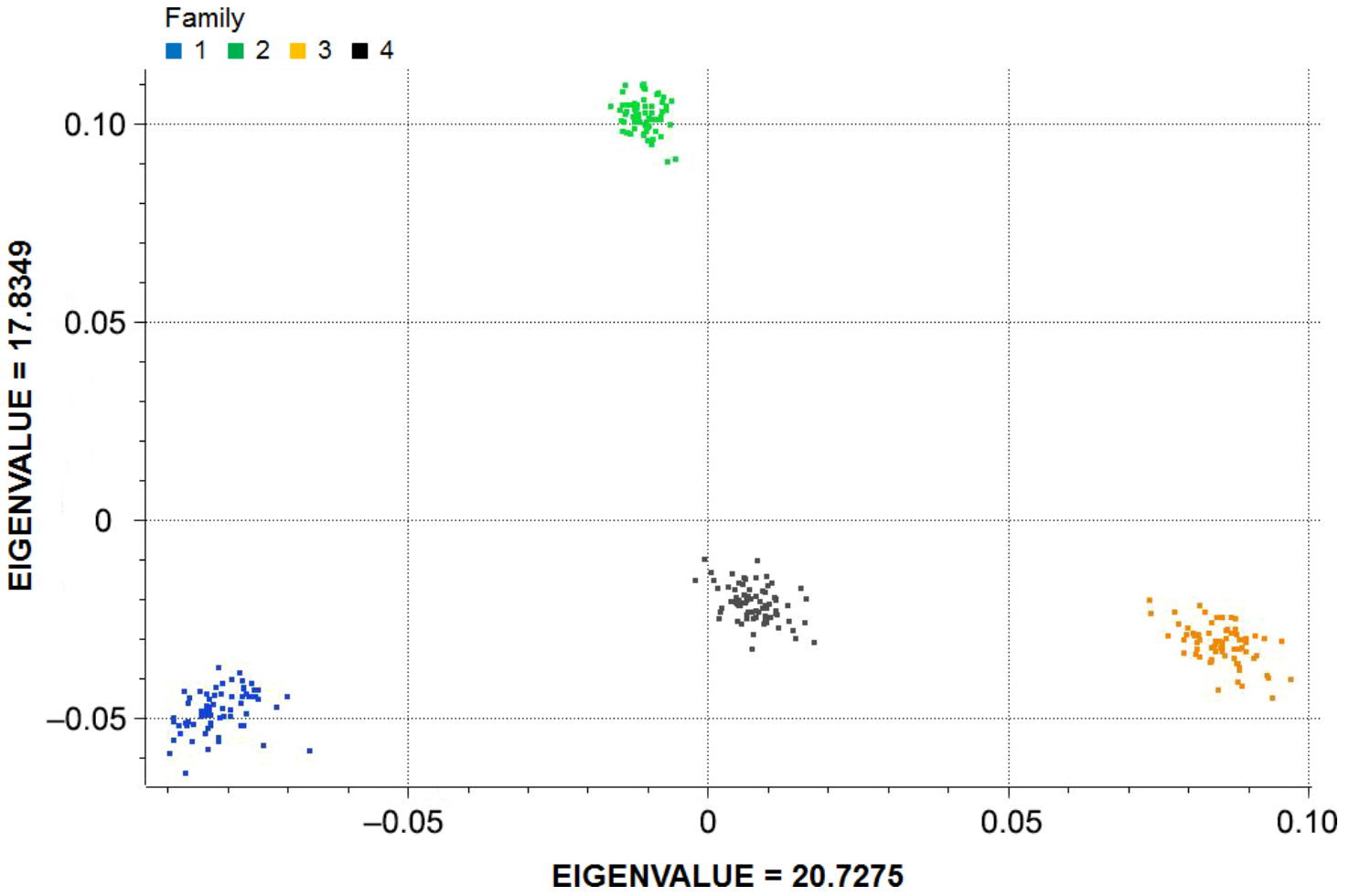
  No sample with genotype missingness > 5%
- SNP quality control

  Excluded SNPs with a minor allele frequency (MAF) < 0.05 or a call rate < 95%

- LD pruning was conducted to achieve a set of independent SNPs and LD blocks.
- LD pruning is a good practice prior to IBS analysis and PCA analysis which may be biased by large blocks of redundant SNPs.

PCA analysis

- Each dot represents one individual.
- Each family was grouped into a separate cluster.



Statistical analysis (EMMAX and QFAM)

- EMMAX (Efficient Mixed-Model Association eXpedited) analyses

$$Y = X\beta + Zu + e$$

Where **Y** is the vector of phenotype; **β** is the coefficient vector of fixed effects including first three principle components and fish body weight; **u** is the vector of the random effect, *Var*(u) = $G\sigma_g^2$, where σ_g^2 is the additive genetic variance and G is the genomic kinship matrix using the IBS; **e** is the vector of random residuals; **X** is the matrix of fixed effects and **Z** is the matrix of random additive genetic effects.

- QFAM (Family based association test for quantitative traits)

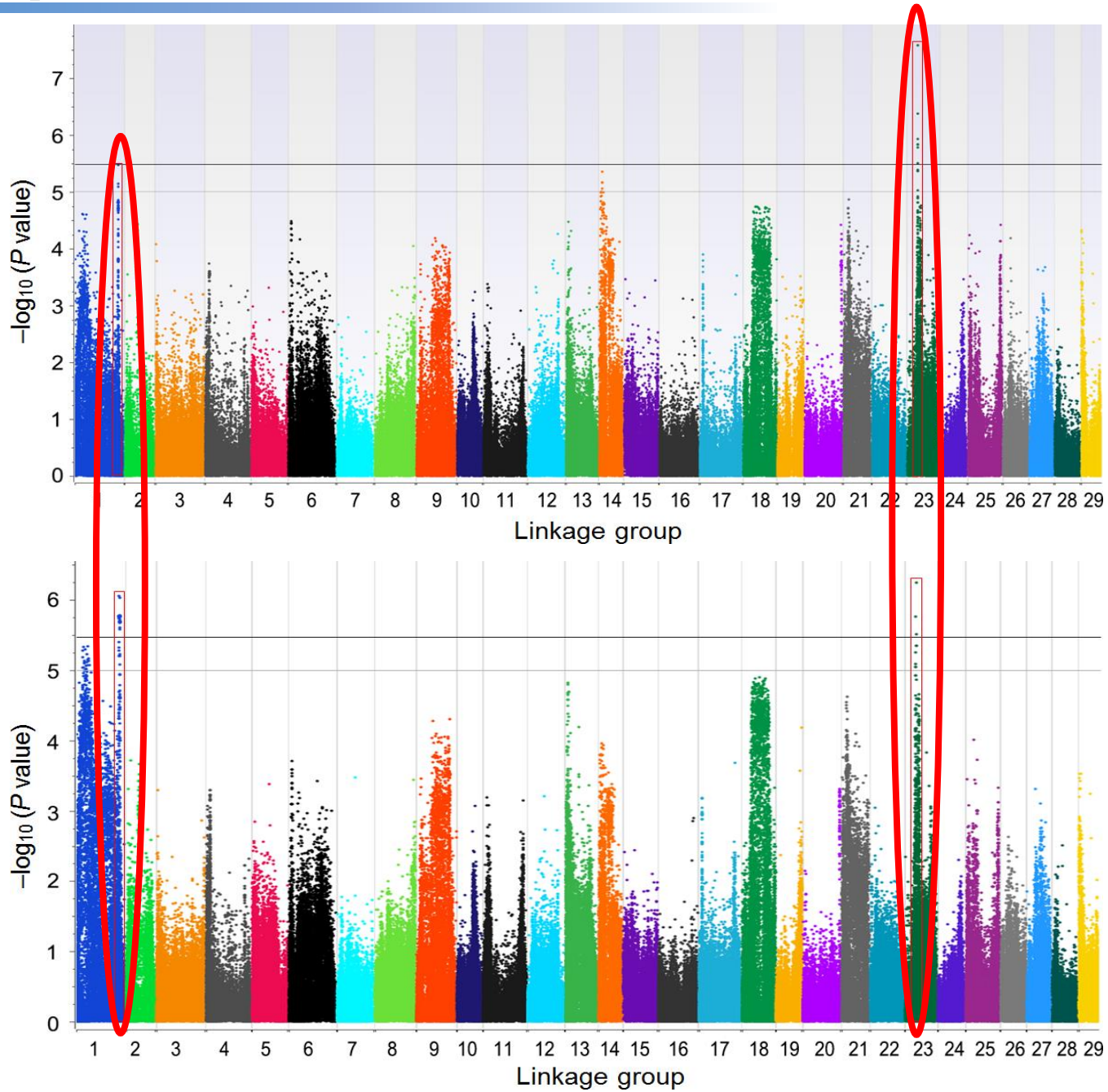
$$\hat{y}_{ij} = \mu + \beta_b b_i + \beta_w w_{ij}$$

QFAM partitions the genotypes into between-family (b) and within-family (w) components. The within-family analysis used in this study is robust to population stratification, which assesses transmission of alleles within a family, but without making use of allelic association observed across families.

Results - Manhattan plot

EMMAX result

QFAM result



Results - SNP

Examination of the associated SNPs reveals superior blue catfish alleles responsible for strong resistance against ESC.

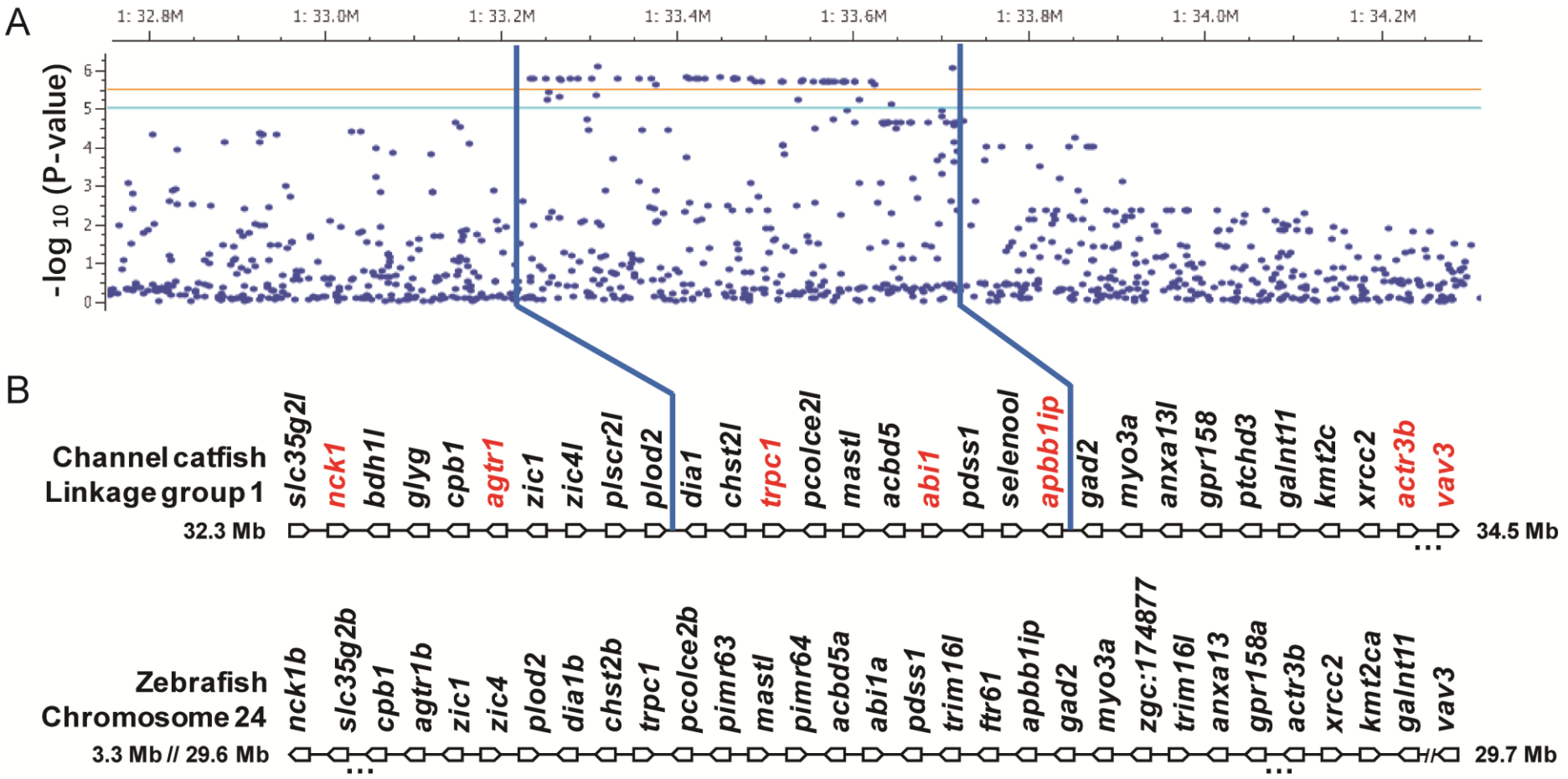
1. the interspecific SNPs on LG1

Example SNP
Channel catfish:
ATATTTATGCAGAAAACAACAAAGCAGAAGTCCTGCCCAAGATGACATTCAGCTTTACTTCTCACTAACCA
Blue catfish:
ATATTTATGCAGAAAACAACAAAGCAGAAGTCCTGACCAAGATGACATTCAGCTTTACTTCTCACTAACCA

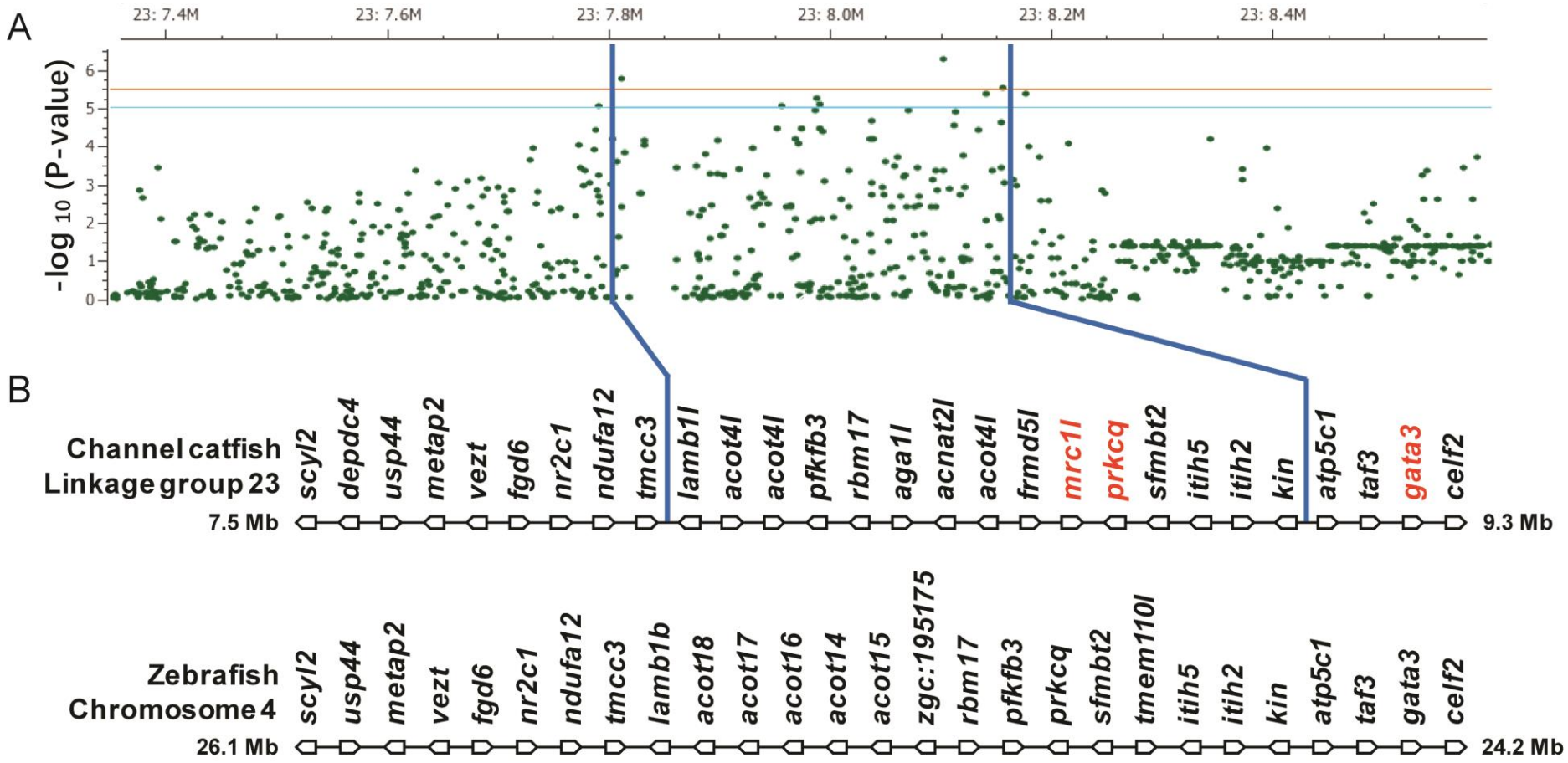
2. channel catfish-specific SNPs on LG23

Example SNP
Channel catfish:
CACATACAACAGAGATAAAACAAATGAGCTTTTACAGATGGGTATATAACACAGGCATGGGCTATGGAGCC
Channel catfish:
CACATACAACAGAGATAAAACAAATGAGCTTTTACGGATGGGTATATAACACAGGCATGGGCTATGGAGCC

Results - Gene



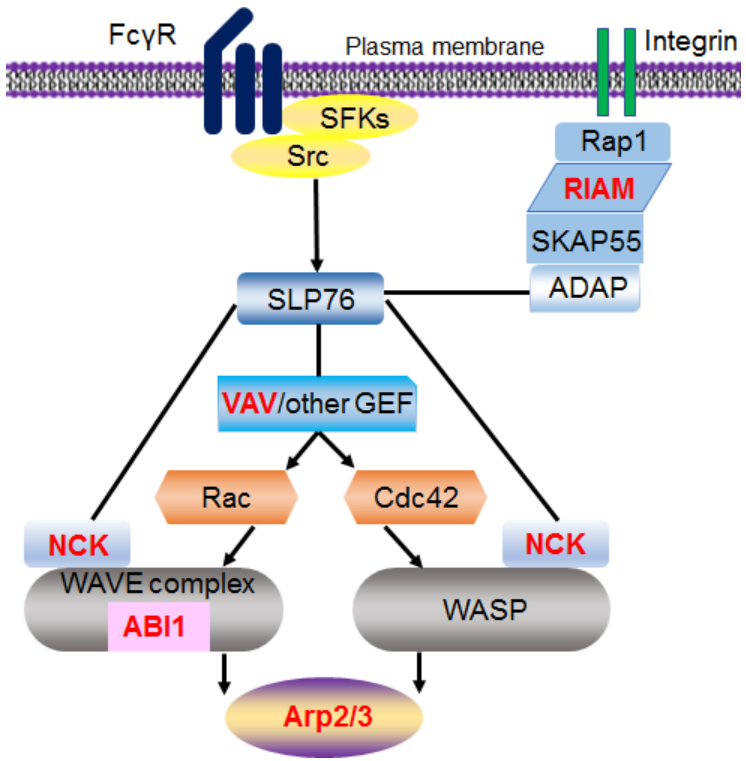
Results - Gene



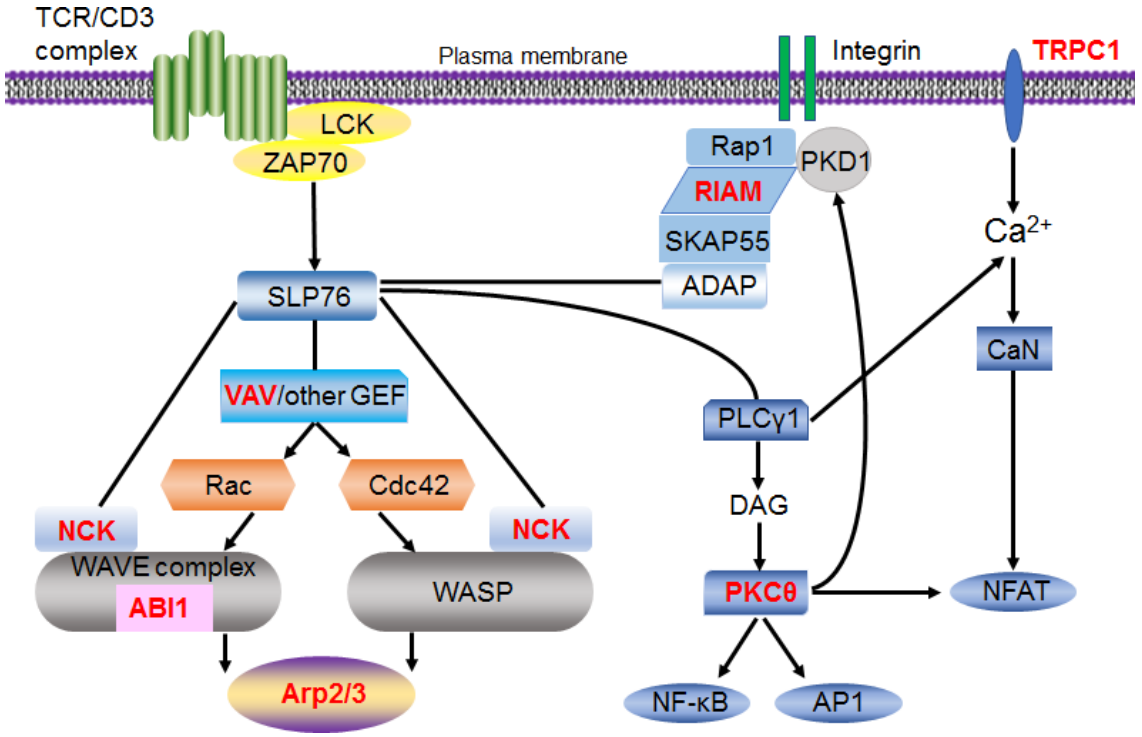
Results - Function

Linkage group	Gene	Location (bp)	Function
1	*nck1*	32,352,283-32,413,183	actin filament organization Phagocytosis T cell activation B cell receptor signaling
	agtr1	32,480,471-32,494,698	regulation of inflammatory response
	trpc1	33,472,942-33,504,891	calcium ion transport B cell receptor signaling
	abi1	33,560,300-33,609,297	actin polymerization or depolymerization Phagocytosis
	apbb1ip	33,632,904-33,677,852	T cell activation
	actr3b	34,277,661-34,297,129	actin nucleation Phagocytosis
	vav3	34,542,282-34,629,025	Phagocytosis B cell receptor signaling
23	*mrc1l*	7,943,812-7,946,175	cellular response to lipopolysaccharide endocytosis T cell activation
	prkcq	7,954,997-7,964,953	inflammatory response T cell activation
	gata3	8,239,408-8,259,340	inflammatory response T cell differentiation humoral immune response

Results - Involved pathway



Phagocytosis



T-cell activation

Conclusions

1. Two significantly associated QTL for ESC resistance were identified on LG1 and LG23.
2. The significant QTL on LG1 is consistent with the finding of previous studies (Zhou et al. 2017), reflecting the power of GWAS.
3. Examination of the associated SNPs revealed superior blue catfish alleles responsible for strong resistance against ESC.
4. The candidate genes were found to be involved in the pathways of phagocytosis and T-cell activation.
5. The positionally related immune genes were functionally related in similar pathways.

1. Fu Q, Yang Y, Li C, Zeng Q, Zhou T, Li N, Liu Y, Liu S, Li D, Liu ZJ. 2017. The CC and CXC chemokine receptors in channel catfish (*Ictalurus punctatus*) and their involvement in disease and hypoxia responses. Developmental and Comparative Immunology 77: 241-251.
2. Fu Q, Yang Y, Li C, Zeng Q, Zhou T, Li N, Liu Y, Li Y, Wang X, Liu S, Li D, Liu ZJ. 2017. The chemokinome superfamily: II. The 64 CC chemokines in channel catfish and their involvement in disease and hypoxia responses. Developmental and Comparative Immunology 73: 97-108.
3. Wang X, Liu S, Yang Y, Fu Q, Abebe A, Liu ZJ. 2017. Identification of NF-κB related genes in channel catfish and their expression profiles in mucosal tissues after columnaris bacterial infection.
4. Zhong X, Wang X, Zhou T, Jin Y, Tan S, Jiang C, Geng X, Li N, Shi H, Zeng Q, Yang Y, Yuan Z, Bao L, Tian C, Liu S, Li Q, Liu ZJ. 2017. Genome-wide association study reveals multiple novel QTL associated with low-oxygen tolerance in hybrid catfish. Marine Biotechnology 19: 379-390. DOI: 10.1007/s10126-017-9757-5.
5. Li Y, Geng X, Bao L, Elaswad A, Huggins KW, Dunham R, Liu ZJ. 2017. A deletion in the Hermansky-Pudlak syndrome 4 (Hps4) gene appears to be responsible for albinism in channel catfish. Molecular Genetics and Genomics, in press. DOI 10.1007/s00438-017-1302-8
6. Zhou T, Liu S, Geng X, Jin Y, Jiang C, Bao L, Yao J, Zhang Y, Zhang J, Sun L, Wang X, Li N, Tan S, Liu ZJ. 2017. GWAS analysis of QTL for enteric septicemia of catfish and their involved genes suggest evolutionary conservation of a molecular mechanism of disease resistance. Molecular Genetics and Genomics 292: 231-242. DOI 10.1007/s00438-016-1269-x
7. Gao S, and Liu ZJ. 2017. Taste receptors and gustatory associated G proteins in channel catfish, Ictalurus punctatus. Comparative Biochemistry and Physiology, part D, Genomics and Proteomics 21: 1-9. doi.org/10.1016/j.cbd.2016.10.002.
8. Gao S, Liu S, Yao J, Li N, Yuan Z, Zhou T, Li Q, and Liu ZJ. 2017. Genomic organization and evolution of olfactory receptors and trace amine-associated receptors in channel catfish, Ictalurus punctatus. Biochimica et Biophysica Acta - General Subjects 1861 (2017): 644-651. Doi 10.1016/j.bbagen.2016.10.017.
9. Zhou T, Li N, Liu S, Jin Y, Fu Q, Gao S, Wang X, Liu ZJ. 2017. The NCK and ABI adaptor genes in catfish and their involvement in ESC disease responses. Developmental and Comparative Immunology 73: 119-123.
10. Fu Q, Zeng Q, Li Y, Yang Y, Li C, Zhou T, Li N, Liu S, Yao J, Jiang C, Li D, Liu ZJ. 2017. The chemokinome superfamily in channel catfish: I. CXC subfamily and their involvement in disease defense and hypoxia responses. Fish and Shellfish Immunology 60: 380-390.
11. Tan S, Yao J, Zhou T, Liu ZJ. 2016. Identification, annotation and expression analysis of 29 Rho GTPase genes from channel catfish (Ictalurus punctatus) after bacterial infections. Developmental and Comparative Immunology 67: 445-451. DOI: 10.1016/j.dci.2016.10.005
12. Yang Y, Fu Q, Zhou T, Li Y, Liu S, Zeng Q, Wang X, Jin Y, Qin Z, Dunham R, Liu ZJ. 2016. Analysis of apolipoprotein genes and their involvement in disease response of channel catfish after bacterial infection. Developmental and Comparative Immunology 67: 464-470. doi: 10.1016/j.dci.2016.09.007.
13. Yuan Z, Liu S, Yao J, Zeng Q, Liu ZJ. 2016. Expression of Bcl-2 genes in channel catfish after bacterial infection and hypoxia stress. Developmental and Comparative Immunology 65: 79-90.
14. Wang X, Liu S, Jiang C, Geng X, Zhou T, Li N, Bao L, Li Y, Yao J, Yang Y, Jin Y, Dunham R, Liu ZJ. 2017. Multiple across-strain and within-strain QTLs suggest highly complex genetic architecture for hypoxia tolerance in channel catfish. Molecular Genetics and Genomics 292: 63–76. DOI:10.1007/s00438-016-1256-2
15. Geng X, Liu S, Yao J, Bao L, Zhang J, Li C, Wang R, Sha J, Zeng P, Zhi D, Liu ZJ. 2016. A genome wide association study identifies multiple regions associated with head size in catfish. G3:Genes Genomics Genetics 6(10):3389-3398. doi: 10.1534/g3.116.032201.
16. Li Z, Yao J, Xie Y, Geng X, Liu ZJ. 2016. Phosphoinositide 3-kinase family in channel catfish and their regulated expression after bacterial infection. Fish & Shellfish Immunology 49:364-373.
17. Fu Q, Li Y, Yang Y, Li C, Yao J, Zeng Q, Qin Z, Li Dao, Liu ZJ. 2016. Septin genes in channel catfish (Ictalurus punctatus) and their involvement in disease defense responses. Fish and Shellfish Immunology 49:110-121.
18. Chen A, Wang, R, Liu S, Peatman E, Sun L, Bao L, Jiang, C, Li C, Li Y, Zeng Q, Liu ZJ. 2016. Ribosomal protein genes are highly enriched among genes with allele-specific expression in the interspecific F1 hybrid catfish. Molecular Genetics and Genomics 291:1083-1093. DOI: 10.1007/s00438-015-1162-z
19. Jiang C, Zhang J, Yao J, Liu S, Li Y, Song L, Li C, Wang X, Liu ZJ. 2015. Complement regulatory protein genes in channel catfish and their involvement in disease defense response. Developmental and Comparative Immunology 53:33-41.
20. Xie, Y, Song L, Weng Z, Liu S, Liu ZJ. 2015. Hsp90, Hsp60 and sHsp families of heat shock protein genes in channel catfish and their expression after bacterial infections. Fish and Shellfish Immunology 44:642-651.
21. Liu S, Li Y, Qin Z, Geng X, Bao L, Kaltenboeck L, Kucuktas H, Dunham R, Liu ZJ. 2015. High-density interspecific genetic linkage mapping provides insights into genomic incompatibility between channel catfish and blue catfish. Animal Genetics 47:81-90. doi:10.1111/age.12372.
22. Geng X, Sha J, Liu S, Bao L, Zhang J, Wang R, Yao J, Li C, Feng J, Sun F, Sun L, Jiang C, Dunham R, Zhi D, Liu ZJ. 2015. A genome-wide association study in catfish reveals the presence of functional hubs of related genes within QTLs for columnaris disease resistance. BMC Genomics 16:196. DOI 10.1186/s12864-015-1409-4.
23. Sun L, Liu S, Bao L, Li Y, Feng J, Liu ZJ. 2015. Claudin multigene family in channel catfish and their expression profiles in response to bacterial infection and hypoxia as revealed by meta-analysis of RNA-Seq datasets. Comparative Biochemistry and Physiology, Part D, Genomicsand Proteomics 13:60-69.

Q and A