

Article

Genome-Wide Marker Analysis for Traits of Economic Importance in Asian Seabass *Lates Calcarifer*

Nguyen Hong Nguyen ^{1,2,*} and Pham Van Khang ^{1,2}

¹ University of the Sunshine Coast, GenCology Research Centre, Locked Bag 4, Maroochydore DC QLD 4558, Australia; Khang.Pham@research.usc.edu.au

² University of the Sunshine Coast, School of Science, Technology and Engineering, Maroochydore, QLD 4558, Australia

* Correspondence: to nnguyen@usc.edu.au

Table S1. Basic statistics of single nucleotide polymorphisms (SNPs).

Marker Information.	Mean	Min	Max
Call Rate	0.937	0.203	1.0
Minor Allele Frequency	0.260	0.003	0.5
Major Allele Frequency	0.740	0.500	1.0
Expected HWE P	0.023	0.000	1.0
Minor Allele Count	36.415	0.000	92.0
Major Allele Count	18.592	0.000	186.0
Heterozygote count	120.240	1.000	186.0
Missing Allele Count	11.752	0.000	149.0

Table S2. Number of SNPs (N) and false discovery rates (FDR) for traits studied ($p < 5 \times 10^{-8}$ or $-\log_{10}p > 5$).

Trait	Correlation Trend Test		Correlation with PCA Corrections		Regression with PCA Corrections		Single Locus Mixed Model	
	N	FDR	N	FDR	N	FDR	N	FDR
Weight	5	0.01	2	0.01	3	0.05	2	0.01
Length	2	0.03	1	0.05	1	0.02	0	>0.05
Survival	0	>0.05	1	0.10	1	0.05	0	>0.05
Cannibalism	1	0.29	0	>0.05	0	>0.05	0	>0.05
Disease resistance	0	>0.05	1	0.12	1	0.05	0	>0.05

Table S3. A list of significant SNPs ($p < 0.01$) linked to candidate genes with known functions for body weight.

Scheme .	Chromosome	GO term	Description
32918924 F 0-23:G>T-23	12	F:GO:0005131; F:GO:0005179; C:GO:0005576;mone activity; C:extracellular region; C:extra- C:GO:0005615; P:GO:0010469; F:GO:0046872	F:growth hormone receptor binding; F:hor- cellular space; P:regulation of signaling re- ceptor activity; F:metal ion binding
32923452 F 0-63:C>G-63	1	F:GO:0005131; F:GO:0005179; C:GO:0005576; C:GO:0005615; P:GO:0010469; F:GO:0046872	As above
32923308 F 0-16:C>T-16	22	F:GO:0005131; F:GO:0005179; C:GO:0005576; C:GO:0005615; P:GO:0010469; F:GO:0046872	As above
32923534 F 0-34:A>T-34	16	F:GO:0005131; F:GO:0005179; C:GO:0005576; C:GO:0005615; P:GO:0010469; F:GO:0046872	As above

32918658 F 0-31:A>G-31	3	F:GO:0005131; F:GO:0005179; C:GO:0005576; C:GO:0005615; P:GO:0010469; F:GO:0046872	As above
32919179 F 0-31:C>T-31	1	F:GO:0005131; F:GO:0005179; C:GO:0005576; C:GO:0005615; P:GO:0010469; F:GO:0046872	As above
32920021 F 0-54:A>G-54	2	P:GO:0015944; C:GO:0016020; C:GO:0016021; P:GO:0045333; F:GO:0046872; F:GO:0051536; F:GO:0051539	P:formate oxidation; C:membrane; C:integral component of membrane; P:cellular respira- tion; F:metal ion binding; F:iron-sulfur clus- ter binding; F:4 iron, 4 sulfur cluster binding
32923298 F 0-68:C>A-68	7	F:GO:0003677; F:GO:0003700; P:GO:0006355	F:DNA binding; F:DNA-binding transcrip- tion factor activity; P:regulation of transcrip- tion, DNA-templated
12505620 F 0-53:A>G-53	1	C:GO:0016020; C:GO:0016021	C:membrane; C:integral component of mem- brane
32920034 F 0-10:A>G-10	3	P:GO:0000160; F:GO:0003677; P:GO:0006355	P:phosphorelay signal transduction system; F:DNA binding; P:regulation of transcrip- tion, DNA-templated
32923723 F 0-44:G>A-44	7	F:GO:0003904; P:GO:0006281	F:deoxyribodipyrimidine photo-lyase activ- ity; P:DNA repair
32920021 F 0-54:A>G-54	2	P:GO:0015944; C:GO:0016020; C:GO:0016021; P:GO:0045333; F:GO:0046872; F:GO:0051536; F:GO:0051539	P:formate oxidation; C:membrane; C:integral component of membrane; P:cellular respira- tion; F:metal ion binding; F:iron-sulfur clus- ter binding; F:4 iron, 4 sulfur cluster binding

*sequences of SNPs provided upon request.

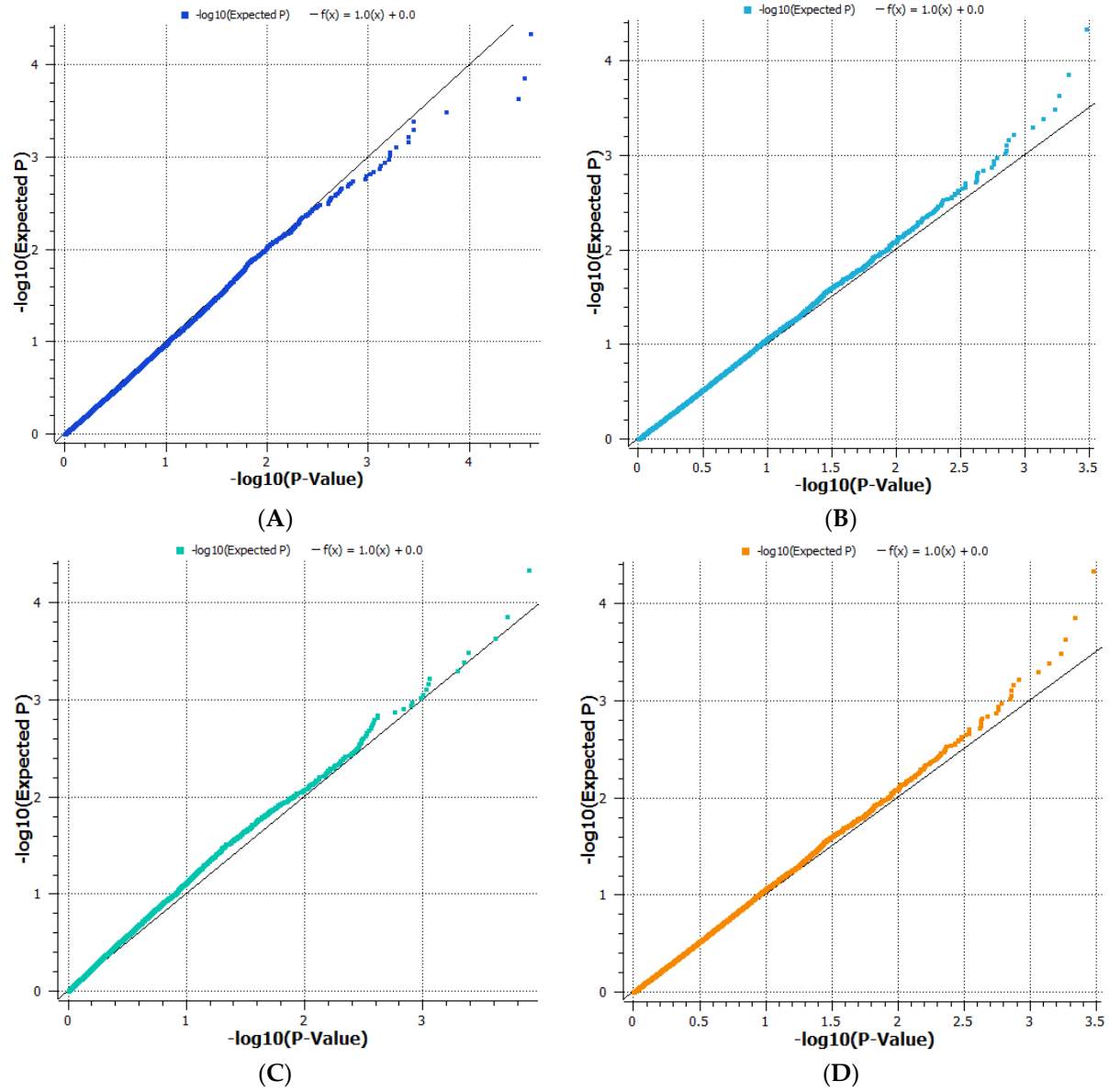
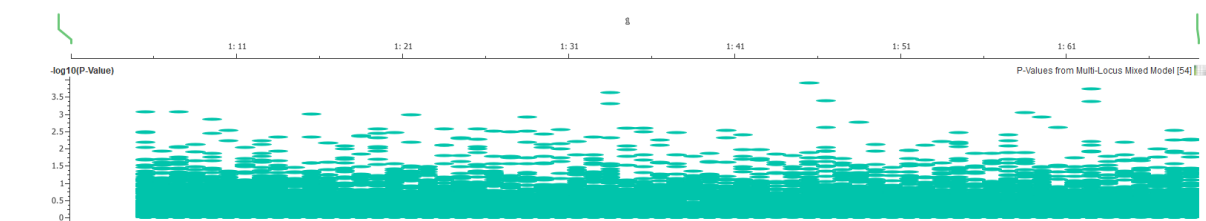
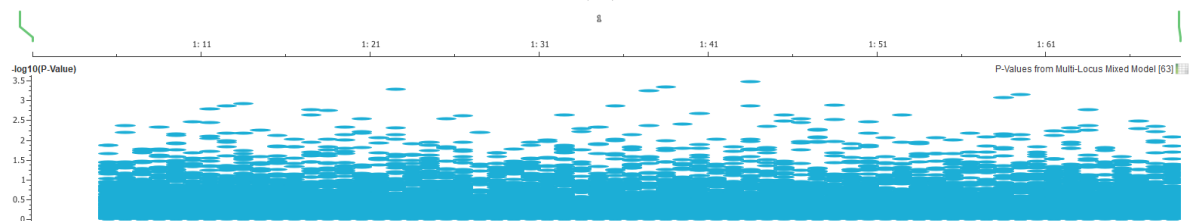


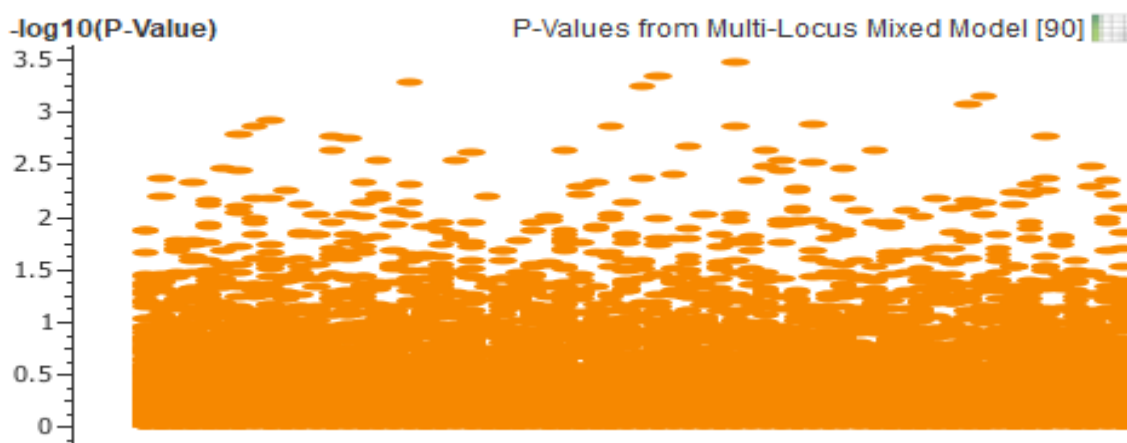
Figure S1. A–D: Q-Q plots for body weight (A), survival (B), cannibalism (C) and disease resistance (D). Lambda values close to one for all traits.



(A)



(B)



(C)

Figure S2. A–C: The Manhattan plot showing the $-\log_{10}(p\text{-values})$ of SNPs on cannibalism (A)(top), survival (B)(middle) and disease resistance (C)(bottom) using a multi-locus mixed.model.

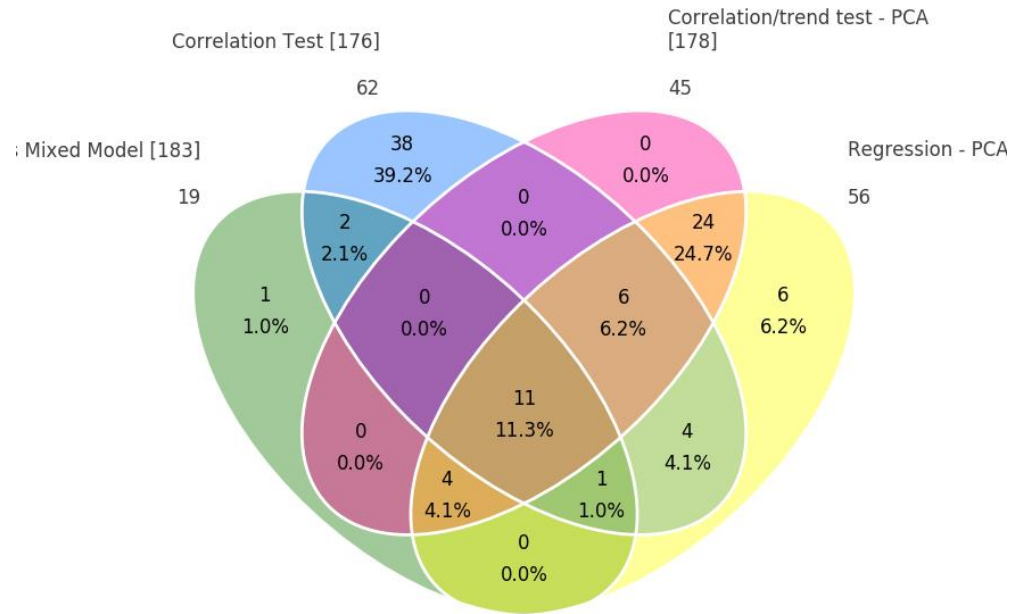


Figure S3. Venn diagram to show common SNPs involved with body weight, using four different statistical methods: 1) Correlation/trend test, 2) Correlation with PCA corrections, 3) Regression with PCA corrections and 4) Multi-locus mixed model at a significant level of $p < 0.001$ or $-\log_{10}(p) > 3$.

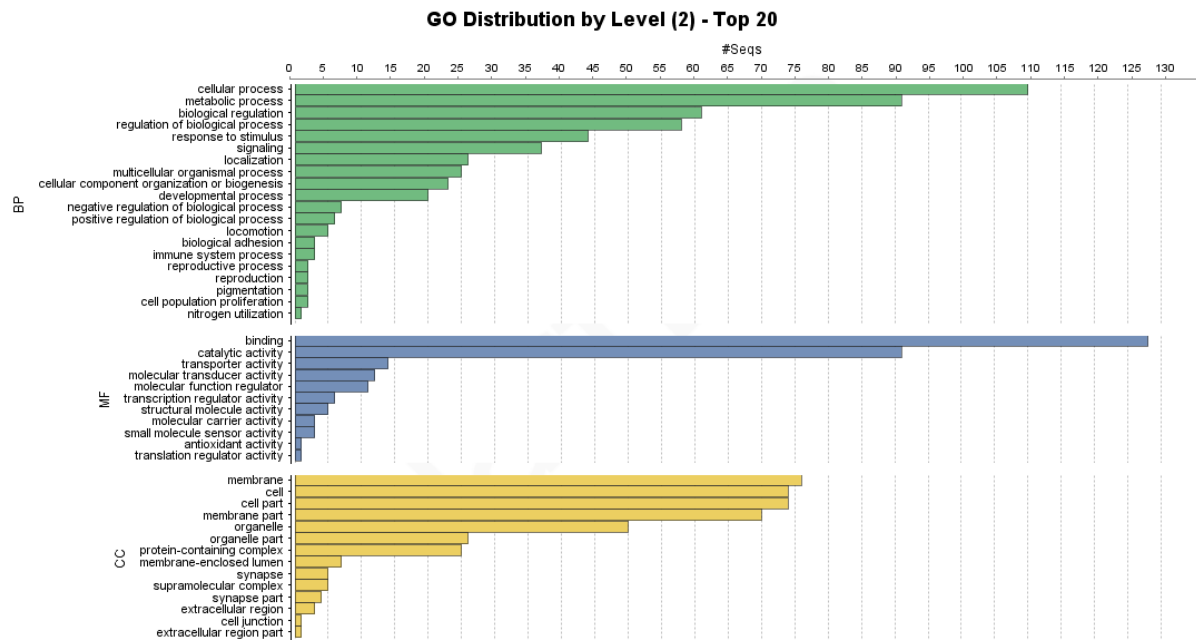


Figure S4. Gene Ontology distribution of 11,084 SNPs used in the study.