Tissue-specific transcripomes of *Mytilus galloprovincialis* reveal new functions



Moreira R, Canchaya C, Novoa B, Posada D, Figueras A









Economic importance

• Worldwide mussel production (FAO):

Worldwide mussel production



- nº1 China > 60%
- nº2 Spain >20%
 - Galicia ≈100%

Objectives



Sampling

10 um



PG

CpG

V.

anguillarum 10⁶ CFU/mL SCHIL

华大基因

Assembly and annotation statistics

Reads sta	tistics	All	Hemocytes	Mantle	Muscle	Gill
Millions of raw reads			112.706	111.322	113.045	56.244
Millions of clean reads			107.386	106.060	107.127	53.335
Total Megabases			9,665	9,545	9,641	4,800
% GC content			38.99%	38.32%	38.28%	37.54%
Assembly sta	tistics					
Number of contigs			261,332	428,939	313,554	238,650
Number of unigenes		151,320	107,045	131,935	120,572	120,254
Average unigene length		570				
N50 unigene length		774				
Range unigene length		200 - 17,690				
Number of unigenes <500pb		104,757				
Number of unigenes >500pb		46,563				
Annotation sta	tistics	Unigenes	Percentage			
Annotated unigenes by nt		14,207	9.4%			
Annotated unigenes by nr		45,182	29.8%			
Annotated unigenes by SwissProt		36,656	24.2%			
Annotated unigenes by KEGG		31,144	20.6%			
Annotated unigenes by COG		14,503	9.6%			
TOTAL annotated unigenes		50,998	33.7%			
Unigenes with GO terms		18,899	12.5%			

KEGG pathways

256 pathways

- Defense
- Signal transduction
- Cancer

			70 ľ	(EQ	u a	nne	otat	ea	uni	ger	ies	
		%0	2%	4%	%9	88	10%	12%	14%	16%	2	90T
	Amino acid metabolism									_		
	Lipid metabolism	1										
	Carbohydrate metabolism											
	Nucleotide metabolism	-										
	Glycan biosynthesis and metabolism	-										
Metabolism	Metabolism of cofactors and vitamins	-										
	1Xenobiotics biodegradation and metabolism											
	Energy metabolism											
	Metabolism of terpenoids and polyketides											
	Biosynthesis of other secondary metabolites	1										
	Translation	-										
	Folding, sorting and degradation	-										
Genetic Information Processing	Transcription	-										
	Replication and repair	-		Т								
	Signal transduction											
vironmental Information Processing	Signaling molecules and interaction	-										
	Membrane transport											
	Cell communication	-										
	Transport and catabolism	-						1				
Cellular Processes	Cell growth and death	-										
	Cell motility	-			1							
		-		Τ								
	Directive system											
	Nervey system	-										
	Endocrine system						-					
Organismal Systems	Circulatery system	-										
	Circulatory system	-										
	Development	-										
	Sensory system	-										
	Excretory system	-										
	Environmental adaptation	-										
	Infectious diseases: Bacterial	-									_	1
	Cancers	-									•	
	infectious diseases: Viral	-						-	•			
	Intectious diseases: Parasitic	-								•		
Human diseases	Neurodegenerative diseases	-										
	Cardiovascular diseases	-										
	Substance dependence	-										
	Immune diseases	-										
	Endocrine and metabolic diseases				_					_		
		%	2%	4%	89	88	10%	12%	14%	16%	8	ę

Mussel vs oyster



Transcriptomes general description



Exclusive transcriptomes

Reads	Hemocytes top exclusive	Reads	Mantle top exclusive
466	C1q domain containing protein MgC1q28	305	von Willebrand factor D and EGF domain-containing protein
300	defensin	293	Fibroin heavy chain
268	mytilin B	283.5	C1q domain containing protein MgC1q95
159	fibrinogen-related protein	245	nacrein B3
146	mytilin-6	215	ADAM family mig-17
141	Serine protease inhibitor Cvsi-2	165.5	fibrocystin L
112.5	complement component 4	160.5	Gigasin-6
73.5	C1q domain containing protein MgC1q56	152.5	C1q domain containing protein MgC1q69
69.5	myticin C gene	149	ATP-dependent RNA helicase A
56.5	fibrinogen-related protein (FREP_G1)	115.5	peroxidasin homolog (Drosophila)-like
Reads	Muscle top exclusive	Reads	Gill top exclusive
107	ribosomal RNA	6644	Perlucin-like protein
62	hsp90-2 gene for heat shock protein 90	3223	C1q domain containing protein MgC1q71
31	myticin C	997	Fucolectin
30.5	Angiopoietin-4	989	Collagen alpha-1(XII) chain
27	stress-70 protein, mitochondrial-like	916	C1q domain containing protein MgC1q17
25.5	ficolin-2-like, partial	897	C1q domain containing protein MgC1q52
12.5	C1q domain containing protein MgC1q22	889	C1q domain containing protein MgC1q36
11.5	Small heat shock protein hspl, mitochondrial	858	C1q domain containing protein MgC1q81
11	DnaJ homolog dnj-10	795	fibrinogen-related protein
9.5	Macrophage receptor MARCO	772	Fibrinogen C domain-containing protein 1

Expression analysis: NOIseq



Expression charts: in red the regulated sequences p-value 0.01







Tissues differential expression



Top expressed sequences

FC	Hemocytes top expressed	FC	Mantle top expressed
1,820	Mytilin B	5,008	Usherin
709	C-type lectin	2,521	BMSP / Protein PIF
478	Defensin 2b (MGD2b)	1,489	Mytimycin
465	C1q domain containing protein	1,409	L-rhamnose-binding lectin CSL3
428	Mytilin-6	1,314	Protein diaphanous
357	Apextrin-like protein (apelP)	1,314	Vitellogenin 6
260	Serine protease inhibitor Cvs.i-2	1,261	Fibronectin 3
258	MAC/perforin- and kringle-domain containing protein	1,218	Heat shock protein 70
246	Heat shock protein 70	1,184	Hornerin / filaggrin
241	Interferon alpha-inducible protein 27 2B (IFI27 / ISG12)	1,168	Chitinase 3
FC	Muscle top expressed	FC	Gill top expressed
FC 1,226	Muscle top expressed Myosin heavy chain	FC 2,837	Gill top expressed BMSP / protein PIF
FC 1,226 1,128	Muscle top expressed Myosin heavy chain Heat shock protein 70	FC 2,837 2,817	Gill top expressed BMSP / protein PIF Perlucin
FC 1,226 1,128 787	Muscle top expressed Myosin heavy chain Heat shock protein 70 C1q domain containing protein	FC 2,837 2,817 1,833	Gill top expressed BMSP / protein PIF Perlucin Fibrinogen-related protein 7
FC 1,226 1,128 787 776	Muscle top expressed Myosin heavy chain Heat shock protein 70 C1q domain containing protein Tripartite motif-containing protein (TRIM56)	FC 2,837 2,817 1,833 1,734	Gill top expressed BMSP / protein PIF Perlucin Fibrinogen-related protein 7 Collagen alpha-1(X) chain-like
FC 1,226 1,128 787 776 481	Muscle top expressed Myosin heavy chain Heat shock protein 70 C1q domain containing protein Tripartite motif-containing protein (TRIM56) Obscurin	FC 2,837 2,817 1,833 1,734 1,479	Gill top expressed BMSP / protein PIF Perlucin Fibrinogen-related protein 7 Collagen alpha-1(X) chain-like Collagen triple helix repeat protein
FC 1,226 1,128 787 776 481 471	Muscle top expressed Myosin heavy chain Heat shock protein 70 C1q domain containing protein Tripartite motif-containing protein (TRIM56) Obscurin Calmodulin	FC 2,837 2,817 1,833 1,734 1,479 1,458	Gill top expressed BMSP / protein PIF Perlucin Fibrinogen-related protein 7 Collagen alpha-1(X) chain-like Collagen triple helix repeat protein Lipopolysaccharide binding protein
FC 1,226 1,128 787 776 481 471 471	Muscle top expressed Myosin heavy chain Heat shock protein 70 C1q domain containing protein Tripartite motif-containing protein (TRIM56) Obscurin Calmodulin Dynein light chain 1	FC 2,837 2,817 1,833 1,734 1,479 1,458 1,269	Gill top expressed BMSP / protein PIF Perlucin Fibrinogen-related protein 7 Collagen alpha-1(X) chain-like Collagen triple helix repeat protein Lipopolysaccharide binding protein Big defensin
FC 1,226 1,128 787 776 481 471 471 471	Muscle top expressed Myosin heavy chain Heat shock protein 70 C1q domain containing protein Tripartite motif-containing protein (TRIM56) Obscurin Calmodulin Dynein light chain 1 Myosin light chain	FC 2,837 2,817 1,833 1,734 1,479 1,458 1,269 1,261	Gill top expressed BMSP / protein PIF Perlucin Fibrinogen-related protein 7 Collagen alpha-1(X) chain-like Collagen triple helix repeat protein Lipopolysaccharide binding protein Big defensin Alpha 1 type V collagen
FC 1,226 1,128 787 776 481 471 471 458 452	Muscle top expressed Myosin heavy chain Heat shock protein 70 C1q domain containing protein Tripartite motif-containing protein (TRIM56) Obscurin Calmodulin Dynein light chain 1 Myosin light chain Calpain-5	FC 2,837 2,817 1,833 1,734 1,479 1,458 1,269 1,261 1,243	Gill top expressed BMSP / protein PIF Perlucin Fibrinogen-related protein 7 Collagen alpha-1(X) chain-like Collagen triple helix repeat protein Lipopolysaccharide binding protein Big defensin Alpha 1 type V collagen Apextrin-like protein

GO enrichment analyses





Hemocytes GO analysis



Mantle and muscle GO analysis



Gill GO analysis

Quantitative



Conclusions:

Almost 400 million reads obtained.

M. galloprovincialis hemocytes, mantle, muscle and gill transcriptomes and expression profile was established.

Important known and new functions were found:

- Hemocytes: antimicrobial and defense proteins.
- Mantle: hematopoietic, sensorial and antifungal function.
- Gills: recognition molecules and homeostatic functions.
- Muscle: stress and defense proteins.

The immune system of these animals seems to be a key regulator of many processes not only triggered by pathogens.

Acknowledgments:

Ministerio de Ciencia e Innovación



Mussel genome project



Thanks you for your attention!







