

# COCKLES

Co-Operation for Restoring Cockle Shellfisheries  
and its Ecosystem-Services in the Atlantic Area

## OPPORTUNITIES FROM SELECTION TO DISEASE RESISTANCE

Paulino Martínez, Belén G Pardo  
Universidade de Santiago de Compostela

**FINAL VIRTUAL CONFERENCE**  
**March 2021**

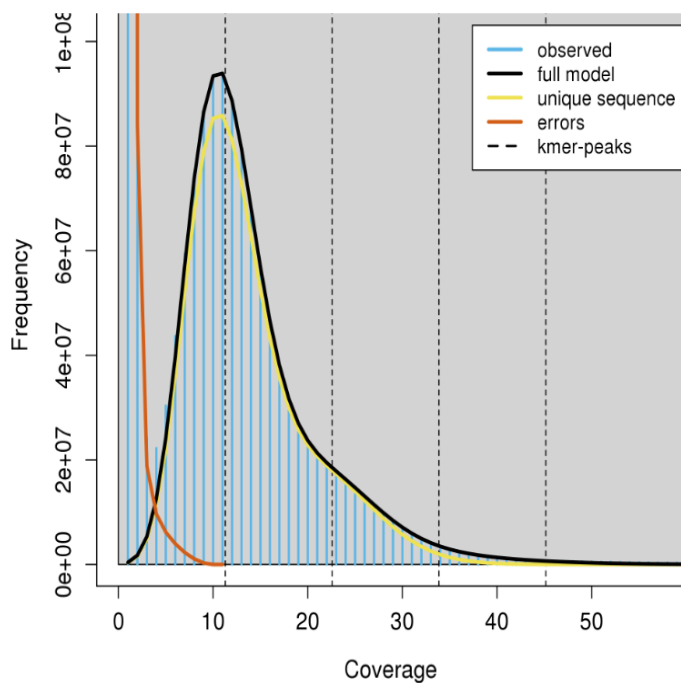


## OVERVIEW OF GENETICS-USC WORK AND INTERACTIONS BETWEEN WPS

- Development of genomic resources (all WPs)
  - Whole genome assembly
  - Highly dense genetic map
  - Large SNP resources
  - Transcriptome (DG and other tissues) (WP7-WP5)
  
- Population Genomics Atlantic Area (WP4-WP6)
  
- Differentially expressed genes across time and the infection process: candidate genes and associated SNPs for selection (WP7)
  
- Population Genomics of Vilagarcía Ría (martellosis outbreak 2012 vs 2018): genetic makers of resistance (WP4-WP5-WP7)
  
- Validation of genetic markers in a new ongoing experiment in the field

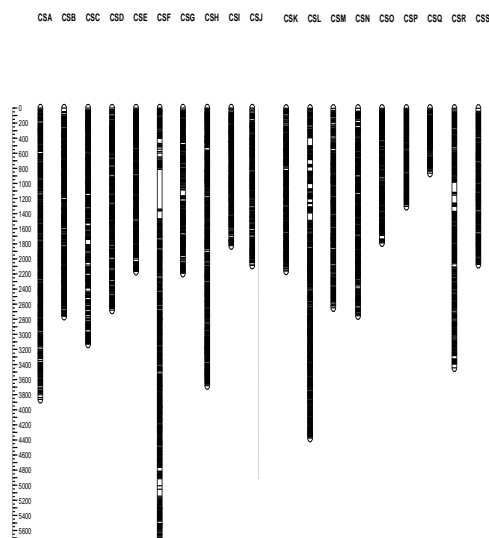
## NEW GENOMIC RESOURCES

### SCUBA CANCER ERC PROJECT (IP: Dr José Tubío-USC)



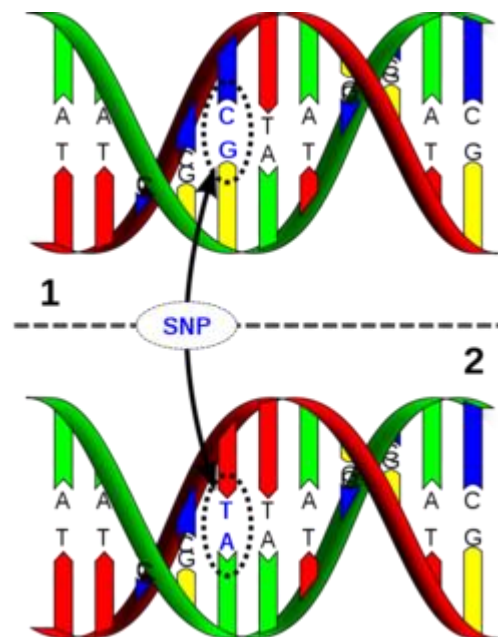
The size of the cockle's genome is ~0.8 Gb; N50 ~40 Mb; 19 major scaffolds corresponding 19 chromosomes of the haploid karyotype

### COCKLE'S GENETIC MAP



Consensus map: 19 Linkage groups; 13,868 SNPs, total length 51,778.7 cM

### LARGE SNPs RESOURCESs



- ~ 14,000 anonymous SNPs
- > 50,000 candidate genes associated SNP

## WP7: Procedures for developing cockle culture, supporting genetic breeding programmes for resistance to marteiliosis and restoring exhausted cockle beds

### Action 7.3. Identification of candidate genes and genetic markers of resistance to marteiliosis

**General objective:** Identification of candidate genes and markers for their use in marteiliosis resistance breeding programs to recover cockle's production.

**Main participants:**

USC: RNA extraction, gene expression and marker association analyses.

CIMA: Experimental design and sampling

**Major Outputs:**

- ✓ transcriptome database enriched in digestive gland- and immune-related genes;
- ✓ list of candidate genes and markers of marteiliosis-resistance to be used in selective breeding programmes to produce marteiliosis resistant cockles strains.

**Other goals not initially foreseen**

- ✓ *Marteilia cochillia* transcriptome
- ✓ Validation of genetic markers by crossing population genomics and transcriptomic approaches in a new field experiment in Vilagarcía Ría

## ACTION 7.3. IDENTIFICATION OF CANDIDATE GENES AND GENETIC MARKERS OF RESISTANCE TO MARTEILIOSIS

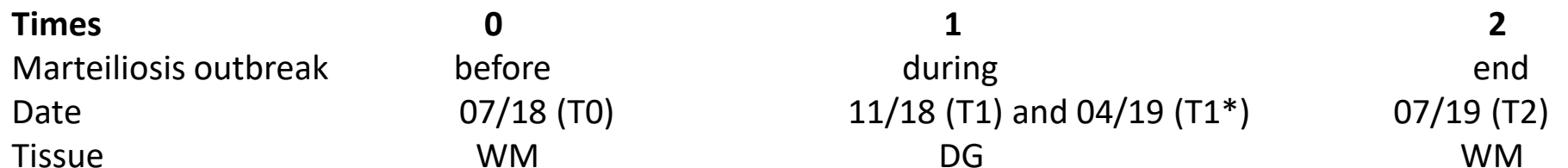
### Sampling

B Cockles from a marteiliosis outbreak (07/18 to 07/19)



Digestive gland (DG)

Whole meat (WM)

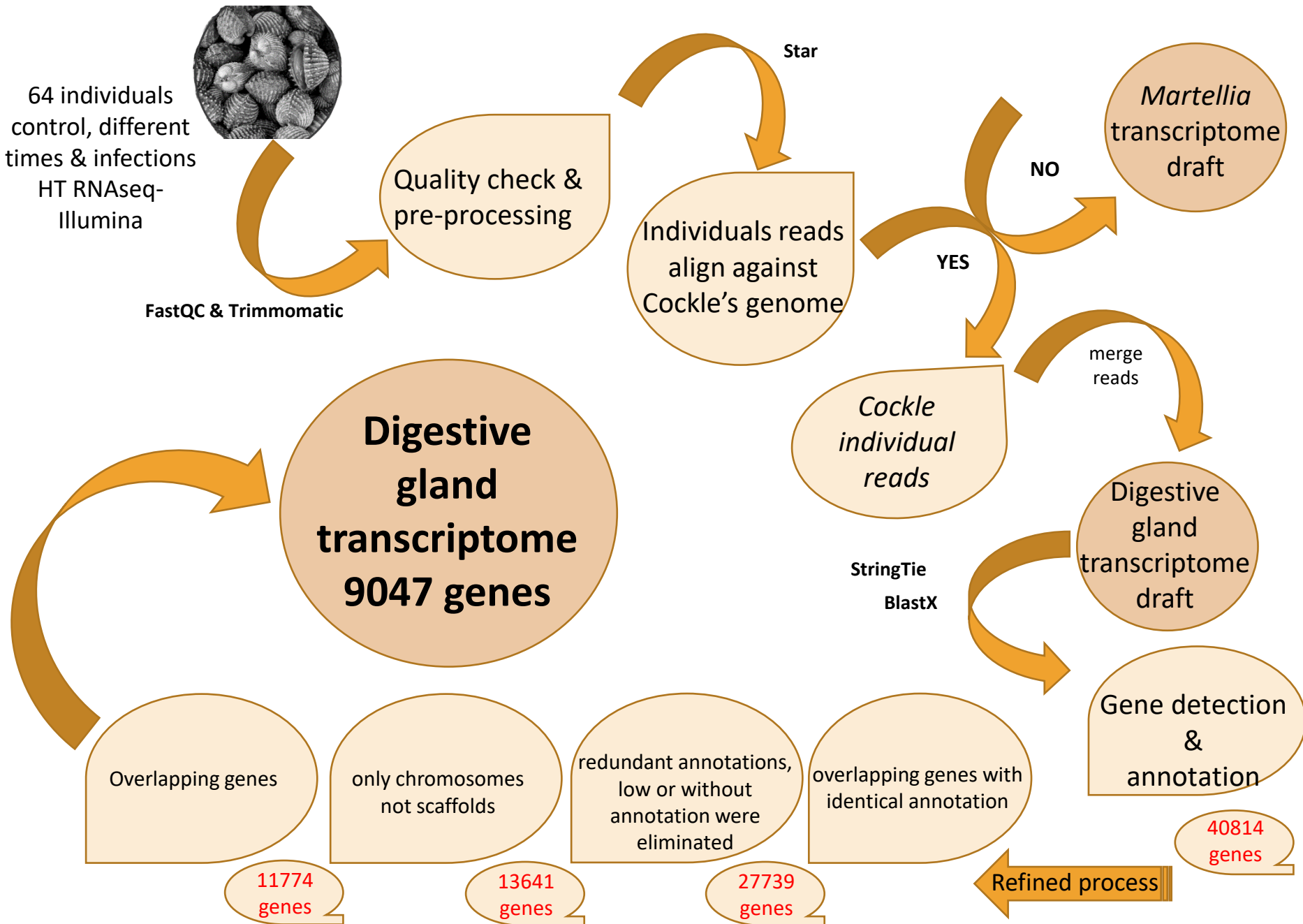


### Outbreak samples

T0 (WM)		T1 (DG)		T1* (DG)		T2 (WM)	
Degree	Nº	Degree	Nº	Degree	Nº	Degree	Nº
0	50	0	101	0	31	0	38
		1	13	1	9	1	5
		2	10	2	5	2	5
NOIA (DG)		3	17	3	3	3	1
Degree	Nº	4	4	4	1	4	1
0	50						

64 samples for RNA-Seq  
(Samples for RNA extraction > 64)

# COCKLE'S TRANSCRIPTOME RECONSTRUCTION



## COCKLE'S DIGESTIVE GLAND TRANSCRIPTOME: 9047 ANNOTATED AND MAPPED GENES

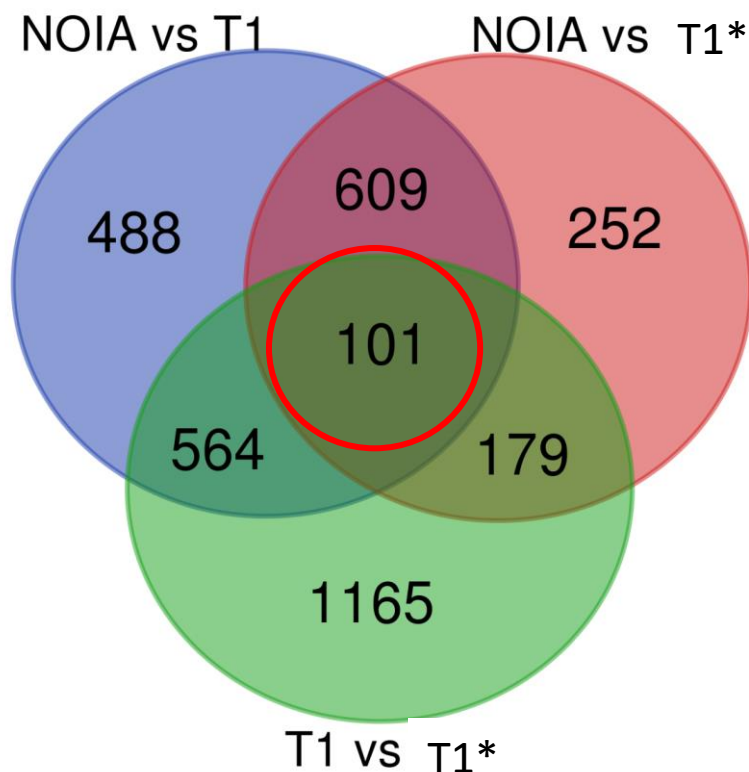
chromosome	start	end	CODE	ANNOTATION	EVALUE
1	53119542	53119772	GENCO1723	Exodeoxyribonuclease III	2,00E-41
1	47476673	47476974	GENCO1508	DUF4817 domain-containing protein	1,00E-111
1	23443410	23443713	GENCO755	Pol-like protein	5,00E-10
1	3259734	3260044	GENCO99	2OG-Fe(II) oxygenase	0
1	64192454	64192813	GENCO2109	Integrase catalytic domain-containing protein	1,00E-164
1	52846367	52846777	GENCO1707	reverse transcriptase family protein	0
1	15018326	15018807	GENCO482	Exodeoxyribonuclease III	4,00E-07
1	58135193	58135754	GENCO1894	Exodeoxyribonuclease III	0
1	47502407	47503037	GENCO1509	52 kDa repressor of the inhibitor of the protein kinase-like	7,00E-160
1	31718611	31719278	GENCO1067	Endo/exonuclease/phosphatase domain-containing protein	2,00E-15
1	2120836	2121536	GENCO64	Reverse transcriptase domain-containing protein (Fragment)	3,00E-23
1	9403090	9403812	GENCO301	Reverse transcriptase domain-containing protein	4,00E-60
1	5378754	5379485	GENCO178	multiple epidermal growth factor-like domains protein 10	5,00E-23
1	2162056	2162793	GENCO67	Exodeoxyribonuclease III	3,00E-17
1	19180601	19181384	GENCO610	putativeuclease HARBI1 isoform X2	1,00E-61
1	2204813	2205664	GENCO71	DUF4817 domain-containing protein	0
1	30724064	30724936	GENCO1027	Reverse transcriptase domain-containing protein (Fragment)	1,00E-29
1	54580166	54581108	GENCO1780	Ribosomal protein S6 kinase delta-1	3,00E-11
1	11793838	11794875	GENCO385	lysosomal-trafficking regulator isoform X7	1,00E-30
1	25037643	25038711	GENCO803	Reverse transcriptase domain-containing protein	5,00E-169
1	3456050	3457146	GENCO106	transmembrane emp24 domain-containing protein 7-like	0
1	51671673	51672826	GENCO1667	THAP-type domain-containing protein	5,00E-16

## COCKLE'S TRANSCRIPTOME: 50 HIGHEST EXPRESSED GENES

ID new	Annotation	Código uniref90	tpm	ID new	Annotation	Código uniref90	tpm
GENCO27210	Cytochrome b	A0A343F4F0	44790,3	GENCO3327	Apolipophorin	A0A210PJU9	3786,29
GENCO5275	Serine protease	D2KKK2	24152,2	GENCO23986	SCO-spondin-like isoform X2	UPI000B4586FB	3364,23
GENCO16088	Metalloendopeptidase	A0A210QK76	9614,26	GENCO1762	Galectin	A0A3R5VPR7	3361,57
GENCO18457	acyl carrier protein, mitochondrial-like isoform X4	UPI000D725F68	8603,36	GENCO8126	V-type proton ATPase 16 kDa proteolipid subunit	UPI0005C3C2FC	3279,24
GENCO23593	Glutathione S-transferase sigma-class	G9IBV2	7781,38	GENCO20294	Peptidylglycine monooxygenase	A0A1S3H7I0	3237,67
GENCO23749	Elongation factor 1-alpha	A0A194ALQ1	6868,93	GENCO4680	S-adenosylmethionine synthase	A0A2C9KBX6	3218,53
GENCO22890	Cartilage intermediate layer protein 2	A0A210QK88	6236,65	GENCO4853	Ovochymase-1	A0A210QS89	3210,94
GENCO14869	B-cell translocation gene 1	B5TRS0	5991,46	GENCO21689	insoluble matrix shell protein 5-like	UPI000BAFC570	3155,26
GENCO11070	Adenosylhomocysteinase	K1RW85	5928,51	GENCO9889	ATP synthase subunit g, mitochondrial	UPI00096B4B7A	3055,8
GENCO22498	leucine-rich repeat-containing protein 40	UPI00148A8D8E	5855,67	GENCO21138	Microsomal glutathione S-transferase	C8CBM6	3022,24
GENCO2501	selenoprotein Pb-like	A0A1S3JI53	5822,22	GENCO17828	Phenylalanine--tRNA ligase	A0A1S3JL40	2931,22
GENCO3046	BHLH domain-containing protein	A0A0K0ENM6	5331,72	GENCO4156	FHA domain-containing protein	W4YQC5	2928,66
GENCO5390	Histone domain-containing protein	A0A3B3YT84	5090,57	GENCO15539	Metallothionein 1	A5A2I2	2852,7
GENCO1935	Cystatin	A0A2R3SJZ3	5056,43	GENCO9197	Collagen alpha-6(VI) chain	A0A210PI34	2807,03
GENCO16166	Serine/threonine-protein kinase 16	A0A210PWQ4	4890,35	GENCO21720	V-type proton ATPase subunit G 1-like	UPI0014587A18	2804,85
GENCO8206	transforming growth factor-beta-induced protein ig-h3-like	UPI001457FC9C	4769,99	GENCO3516	deleted in malignant brain tumors 1 protein-like	UPI0000E47227	2676,73
GENCO15787	ATP synthase peripheral stalk subunit OSCP	A0A286Q4W7	4659,94	GENCO4470	zinc metalloproteinase nas-4-like	UPI000BAF5B52	2665,3
GENCO18555	Heat shock protein 22 isoform 1	C8CBN4	4646,4	GENCO13514	Histone H1-delta	K1P7W5	2507,65
GENCO7909	ganglioside GM2 activator	UPI0005C39317	4484,06	GENCO13587	Profilin	F4XXT7	2457,99
GENCO25037	Sialic acid binding lectin 8	A0A451ET91	4381,71	GENCO23173	serine protease inhibitor Cvs1-2-like	UPI00148AD200	2448,42
GENCO17546	Ferritin	A0A067XI00	4297,62	GENCO9807	Heat shock protein 70	B6CQR3	2444,91
GENCO5278	Putative chitinase 3	K1QYL9	4251,7	GENCO5861	Brain protein I3	A0A6A4VHF0	2419,42
GENCO22524	transforming growth factor-beta-induced protein ig-h3-like	UPI001458CB45	4248,82	GENCO13463	Histone H1-delta	K1P7W5	2415,23
GENCO10377	saxiphilin-like	UPI000BB08257	4089,15	GENCO12254	Putative thaumatin-like protein 1a	A0A194AQG2	2407,45
GENCO20306	endoglucanase	UPI00148A5ECF	3918,85	GENCO22889	Cartilage intermediate layer protein 2	A0A210QK88	2392,35



## GENE EXPRESSION ANALYSIS IN THE WILD: TEMPORAL COMPARISON vs NOIA (DG)

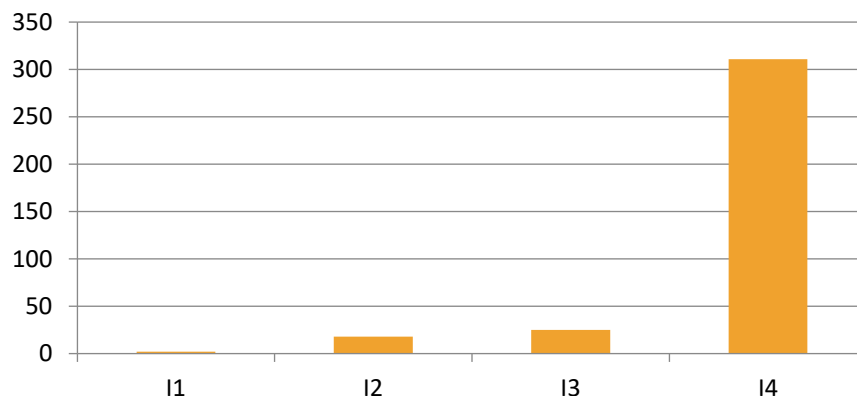


Gene code	ANNOTATION	CHR	start	end
GENCO45	Neuronal PAS domain-containing protein 4	1	1403140	1425633
GENCO538	protein numb-like isoform X1	1	16762704	16816473
GENCO707	dehydrogenase/reductase SDR family member 1-like	1	21421187	21443538
GENCO815	Kazal-like domain-containing protein	1	25496572	25599526
GENCO1081	RBR-type E3 ubiquitin transferase	1	32259773	32270203
GENCO1374	NUC153 domain-containing protein	1	42706173	42724973
GENCO1685	WD repeat-containing protein 27	1	52210593	52236120
GENCO2145	Arylsulfatase B	2	838248	839539
GENCO2861	ATP-binding cassette sub-family D member 3	2	22015395	22040660
GENCO2964	Peroxin-13	2	25766303	25772544
GENCO3342	RING finger protein 11	2	35678399	35698233
GENCO3518	non-specific lipid-transfer protein-like	2	40037905	40071507
GENCO3548	zinc finger protein 260-like	2	40662196	40676685

101 consistent genes differentially expressed across time

## GENE EXPRESSION ANALYSIS IN THE WILD: INFECTION LEVEL

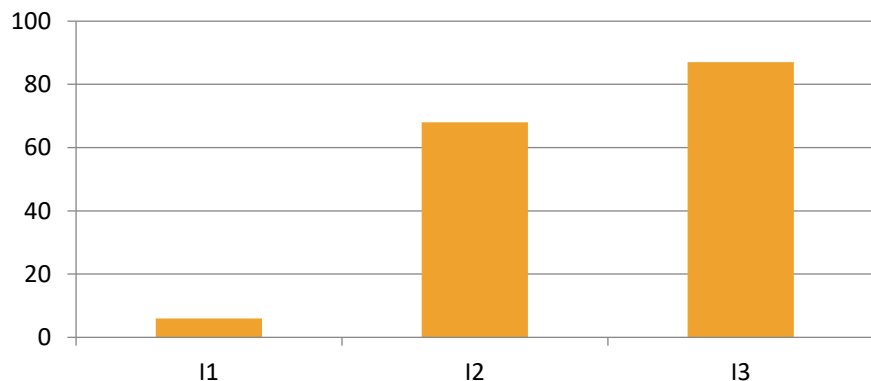
**T1 infection level**



**357 DE genes**

17-beta-hydroxysteroid dehydrogenase 14  
 26S proteasome non-ATPase regulatory subunit 2  
 26S proteasome non-ATPase regulatory subunit 6-like  
 26S proteasome non-ATPase regulatory subunit 9-like  
 3 beta-hydroxysteroid dehydrogenase/Delta 5-->4-isomerase type 4  
 3-hydroxyacyl-[acyl-carrier-protein] dehydratase  
 7-dehydrocholesterol reductase-like  
 A disintegrin and metalloproteinase with thrombospondin motifs  
 gon-1-like isoform X1  
 AAC-rich mRNA clone AAC4 protein-like  
 Adenylate cyclase

**T1\* infection level**



**152 DE genes**

LAM\_G\_DOMAIN domain-containing protein  
 Ferric uptake regulation protein FUR  
 mucin-4 isoform X1  
 Heat shock protein 70  
 Calpain-5  
 G\_PROTEIN\_RECEP\_F1\_2 domain-containing protein  
 Calmodulin  
 ganglioside GM2 activator  
 Cathepsin L  
 glutamine synthetase-like  
 protein PML-like

## GENE EXPRESSION ANALYSIS IN THE WILD: T0 VS T2 (WHOLE MEAT)

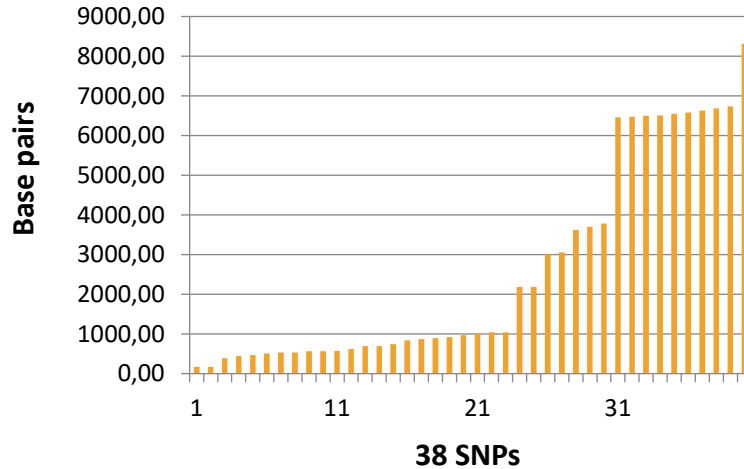


target_id	Anotacion	pval	FC
GENCO15891	tryptase-2-like n=1	2,73E-07	2,4656149
GENCO1081	RBR-type E3 ubiquitin transferase	2,05E-06	-2,53182085
GENCO4651	Serine-enriched protein	1,38E-06	3,62065859
GENCO4710	Interferon regulatory factor 2	1,64E-06	1,56588876
GENCO22383	ABC transporter G family member 14-like	1,55E-06	4,33694101
GENCO11232	Eukaryotic translation initiation factor 6	2,47E-06	-1,30511677
GENCO12585	Protein yippee-like	4,50E-06	1,38901597
GENCO20294	Peptidylglycine monooxygenase	5,24E-06	6,20300616
GENCO2635	Transmembrane protein 214	8,56E-06	-1,23842729
GENCO20373	ATP-binding cassette sub-family F member 3	8,01E-06	-1,57961643
GENCO20688	AAA domain-containing protein	7,61E-06	-1,39398645
GENCO2299	Elongation of very long chain fatty acids protein	1,16E-05	-1,63616225
GENCO11472	E3 ubiquitin-protein ligase	1,10E-05	1,37366202
GENCO12702	peroxisomal membrane protein 2-like	1,24E-05	-1,23699665
GENCO974	Tripartite motif-containing protein 45	1,38E-05	1,79247125
GENCO3058	Zinc transporter 7	1,71E-05	-1,33161805

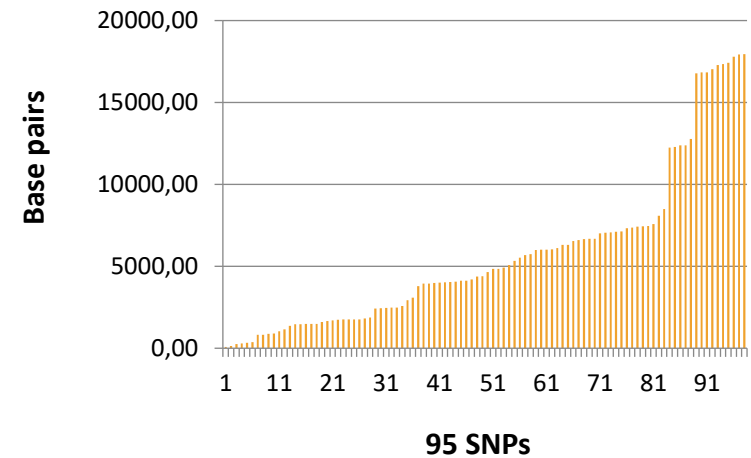
305 DEGs T0 vs T2

## GENE EXPRESSION ANALYSIS: SNPs-ASSOCIATED TO 768 DEGs

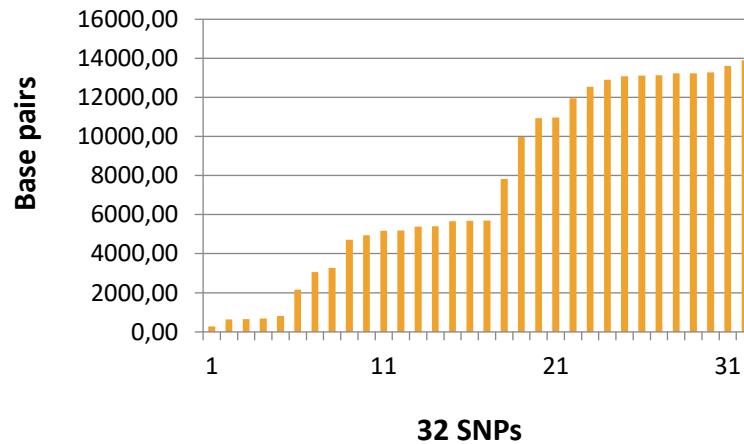
**2,4-dienoyl-CoA reductase, mitochondrial**



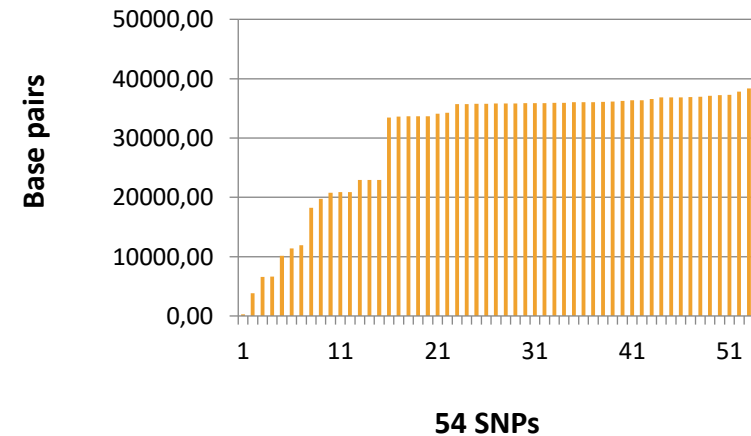
**Sulfotransferase family cytosolic 1B member 1**



**Fructose-bisphosphate aldolase**



**Serine/threonine-protein phosphatase**



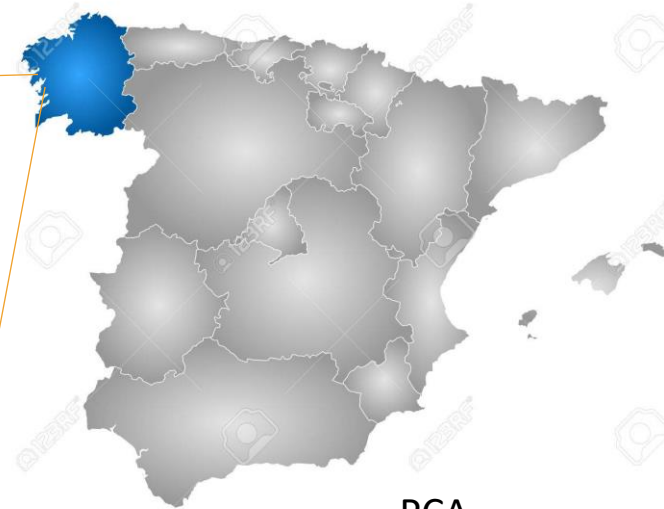
## GENE EXPRESSION ANALYSIS IN THE WILD: TRANSCRIPTOME vs PROTEOME

chromosome	start	end	CODE	ANNOTATION	EVALUE
3	46800156	46830366	GENCO5600	(S)-3-amino-2-methylpropionate transaminase n=1 Tax=Mizuhopecten yessoensis TaxID=6573 RepID=A0A210QVG7_MIZYE	0
13	11634071	11666393	GENCO18799	1,4-alpha-glucan branching enzyme n=7 Tax=Crassostrea TaxID=6564 RepID=K1QYM7_CRAGI	0
13	4902127	4918351	GENCO18577	1,5-anhydro-D-fructose reductase-like isoform X1 n=2 Tax=Pecten maximus TaxID=6579 RepID=UPI001458882A	2,00E-135
11	24748827	24768559	GENCO16855	116 kDa U5 small nuclear ribonucleoprotein component-like n=1 Tax=Pecten maximus TaxID=6579 RepID=UPI001458C48D	0
8	34345409	34356231	GENCO13144	14_3_3 domain-containing protein n=3 Tax=Pomacea canaliculata TaxID=400727 RepID=A0A2T7PD31_POMCA	0
14-3-3 protein epsilon-like					
11	16069786	16074263	GENCO16515	14-3-3 protein gamma-A-like n=1 Tax=Chanos chanos TaxID=29144 RepID=UPI0011F1AD32	4,00E-25
17	23947056	23967855	GENCO23864	14-3-3 protein zeta/delta-like protein (Fragment) n=1 Tax=Callorhinchus milii TaxID=7868 RepID=V9KDC3_CALMI	2,00E-27
7	7314836	7352077	GENCO10812	15-hydroxyprostaglandin dehydrogenase [NAD(+)]-like n=1 Tax=Crassostrea virginica TaxID=6565 RepID=UPI000B80BC4D	2,00E-111
1	54132662	54149174	GENCO1761	15-oxoprostaglandin 13-reductase n=2 Tax=Lingula unguis TaxID=7574 RepID=A0A1S3HU86_LINUN	4,00E-152
4	5393688	5412761	GENCO6034	1-acyl-sn-glycerol-3-phosphate acyltransferase delta-like n=1 Tax=Aplysia californica TaxID=6500 RepID=UPI00035A2D54	0
16	4986736	5198296	GENCO22043	26S proteasome non-ATPase regulatory subunit 1 n=1 Tax=Lottia gigantea TaxID=225164 RepID=V4AES7_LOTGI	0
3	25689139	25701163	GENCO4898	26S proteasome non-ATPase regulatory subunit 10 n=2 Tax=Crassostrea TaxID=6564 RepID=UPI00148A47CD	0
				26S proteasome non-ATPase regulatory subunit 10-like isoform X2 n=1 Tax=Mizuhopecten yessoensis TaxID=6573 RepID=UPI000B45B652	1,00E-91
11	23343701	23377239	GENCO16797	26S proteasome non-ATPase regulatory subunit 11 n=2 Tax=Pectinidae TaxID=6566 RepID=A0A210QPW7_MIZYE	0
11	9772276	9861343	GENCO16312	26S proteasome non-ATPase regulatory subunit 11-like	
4	48920365	48964015	GENCO7485	26S proteasome non-ATPase regulatory subunit 13 n=3 Tax=Pectinidae TaxID=6566 RepID=A0A210Q7K5_MIZYE	0
				26S proteasome non-ATPase regulatory subunit 1-like	
13	26957464	26976835	GENCO19315	26S proteasome non-ATPase regulatory subunit 2 n=3 Tax=Crassostrea TaxID=6564 RepID=K1PEY4_CRAGI	0
				26S proteasome non-ATPase regulatory subunit 2-like	
12	6226890	6239547	GENCO17436	26S proteasome non-ATPase regulatory subunit 4 n=3 Tax=Pectinidae TaxID=6566 RepID=A0A210R6G4_MIZYE	3,00E-168
15	16239588	16302820	GENCO21386	26S proteasome non-ATPase regulatory subunit 5 n=2 Tax=Lingula unguis TaxID=7574 RepID=A0A1S3IXM5_LINUN	3,00E-132
10	15477359	15485103	GENCO15156	26S proteasome non-ATPase regulatory subunit 6-like n=2 Tax=Pectinidae TaxID=6566 RepID=UPI000B457CAA	4,00E-13
6	11576516	11583055	GENCO9646	26S proteasome non-ATPase regulatory subunit 9-like n=1 Tax=Pecten maximus TaxID=6579 RepID=UPI001458E81	2,00E-84
				26S proteasome regulatory subunit 10B	
17	5820948	5837388	GENCO23175	26S proteasome regulatory subunit RPN11 n=13 Tax=Metazoa TaxID=33208 RepID=A0A0L8GGN2_OCTBM	0
3	51883509	51897497	GENCO5734	28 kDa heat- and acid-stable phosphoprotein n=1 Tax=Trichoplusia ni TaxID=7111 RepID=UPI000EF83B84	3,00E-28
4	12822416	12850551	GENCO6275	3-oxoacyl-[acyl-carrier-protein] reductase FabG n=2 Tax=Cryptotermes secundus TaxID=105785 RepID=A0A2J7QPM6_9NEOP	1,00E-29
3	10336682	10386459	GENCO4340	3'-phosphoadenosine-5'-phosphosulfate synthase n=11 Tax=Crassostrea TaxID=6564 RepID=K1R185_CRAGI	0
3	52263794	52420683	GENCO5743	3-phosphoinositide-dependent protein kinase 1 n=3 Tax=Pectinidae TaxID=6566 RepID=A0A210QNJ9_MIZYE	0
13	25912927	25999628	GENCO19278	40S ribosomal protein S10 n=3 Tax=Pectinidae TaxID=6566 RepID=A0A210QTB6_MIZYE	1,00E-66
13	24695961	24727549	GENCO19235	40S ribosomal protein S11 n=7 Tax=Hemiptera TaxID=7524 RepID=R4G4F1_RHOPR	9,00E-70
				40S ribosomal protein S11-like	
17	18474445	18478396	GENCO23633	40S ribosomal protein S13 n=4 Tax=Pteriomorpha TaxID=6545 RepID=K1QRZ3_CRAGI	0
				40S ribosomal protein S14	
9	12679686	12690493	GENCO13728	40S ribosomal protein S15 n=1 Tax=Phragmatopoma lapidosa TaxID=341668 RepID=A0A0A0QXJ9_9ANNE	4,00E-144

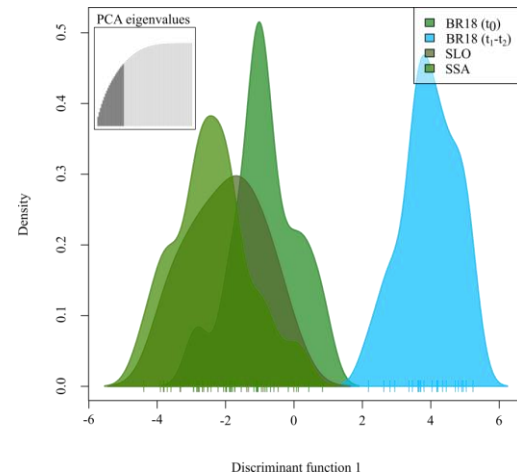
~ 50 common DE transcripts and proteins detected

# POPULATION GENOMICS MARTEILIOSIS: 121 outliers associated with resistance

## Natural outbreak 2018 vs 2012

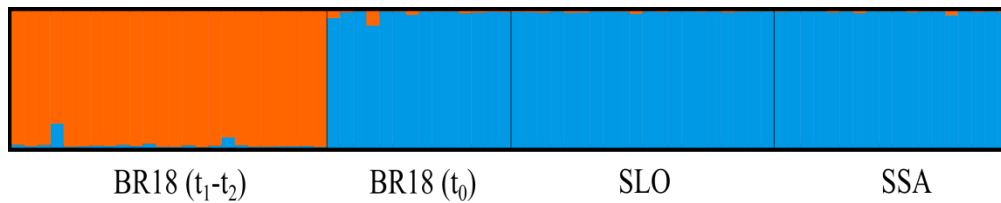


PCA



K=2

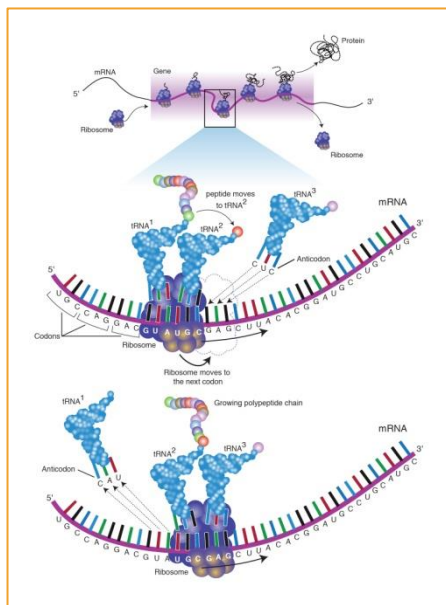
STRUCTURE (121 outliers)



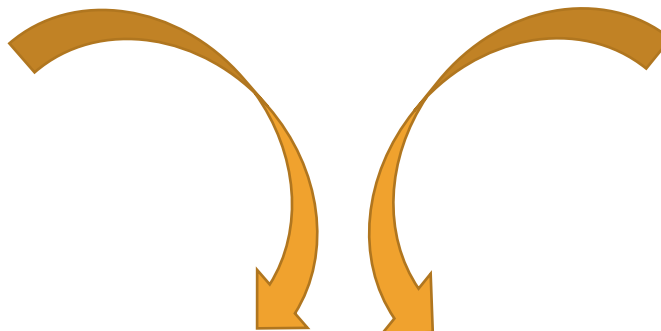


## INTEGRATING GENE EXPRESSION AND SIGNATURES OF SELECTION TO MARTEILIOSIS

768 DEGs

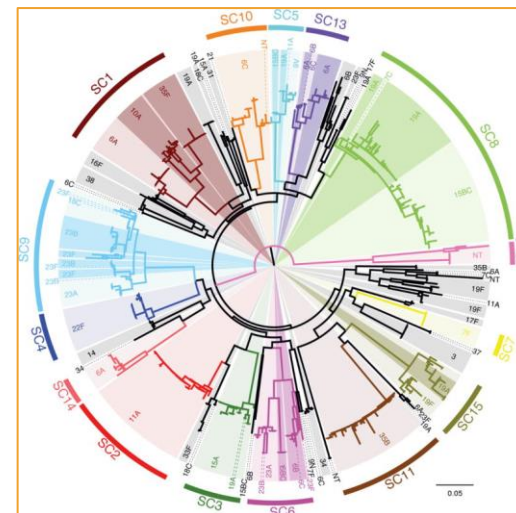


140 DEPs



56 SNPs TO BE VALIDATED  
IN THE ONGOING FIELD  
EXPERIMENT

121 OUTLIERS INF vs NI



SNPs LINKED TO GENES  
SIGNATURES SELECTION

## CONCLUSIONS:

- Broad new genomic resources developed in cockles
- Cockle's and *Marteilia* transcriptomes reconstructed for gene expression analysis of host-parasite interaction
- 121 outlier loci fully discriminate infected and non-infected cockles representing a useful tool for MAS to obtain a resistant strain
- 56 of these markers are close or within differentially expressed genes/proteins
- These markers are being validated in a new ongoing experiment in Vilagarcía Ría