COCKLES

Co-Operation for Restoring CocKle SheLlfisheries and its Ecosystem-Services in the Atlantic Area
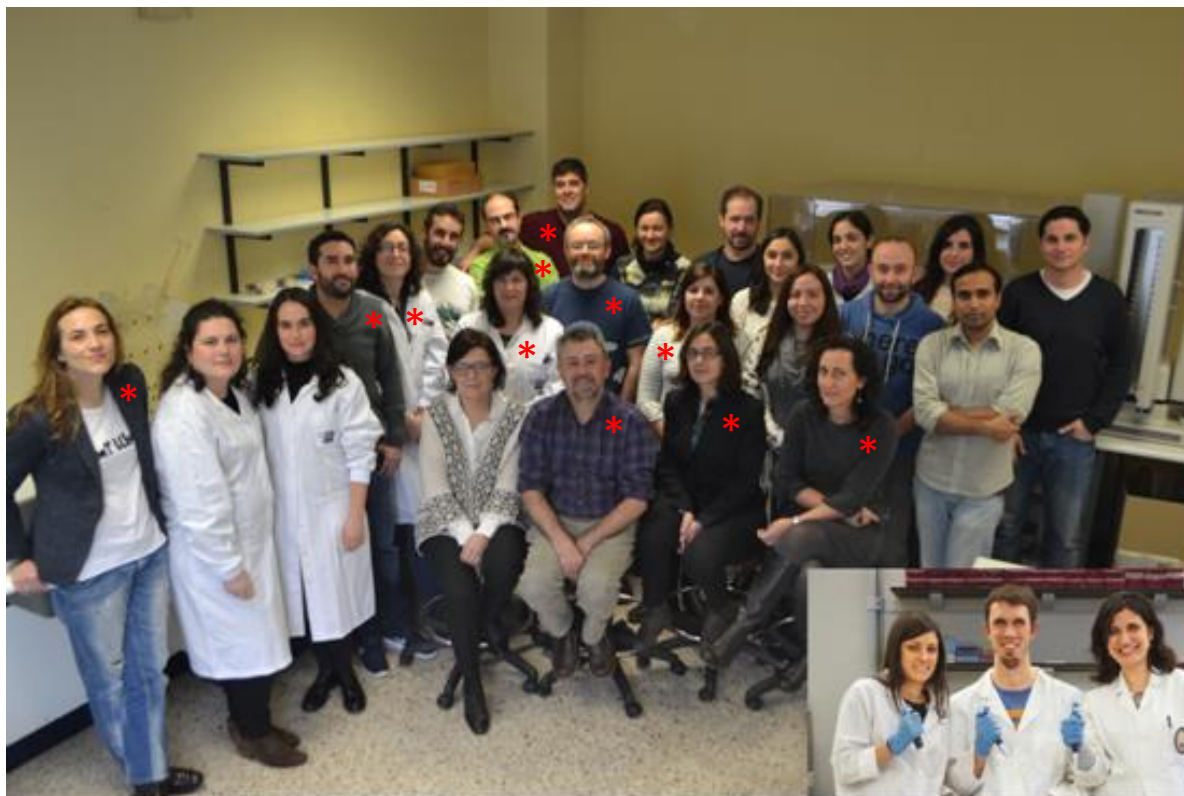
Genetic diversity of COCKLES in the Atlantic Area (AA)

Paulino Martínez, Manuel Vera
Universidade de Santiago de Compostela

FINAL VIRTUAL CONFERENCE
March 2021

ACUIGEN GENETICS GROUP

UNIVERSITY OF SANTIAGO DE COMPOSTELA (USC)

➢ Development of genomic resources (all WPs)
  • Whole genome assembly
  • Highly dense genetic map
  • Large SNP resources
  • Molecular tool for *Cerastoderma edule* and *C. glaucum* identification
 • Transcriptome (DG and other tissues)

➢ Population Genomics Approach
  • Genetic sructure in the Atlantic Area (WP4-WP6)
  • Vilagarcía Ría (marteiliosis outbreak 2012 vs 2018): genetic makers for resistance (WP4-WP5-WP7)

➢ Differentially expressed genes across time and the infection process: candidate genes and associated SNPs for selection (WP7)

➢ Validation of genetic markers in a new ongoing experiment in the field (WP7)
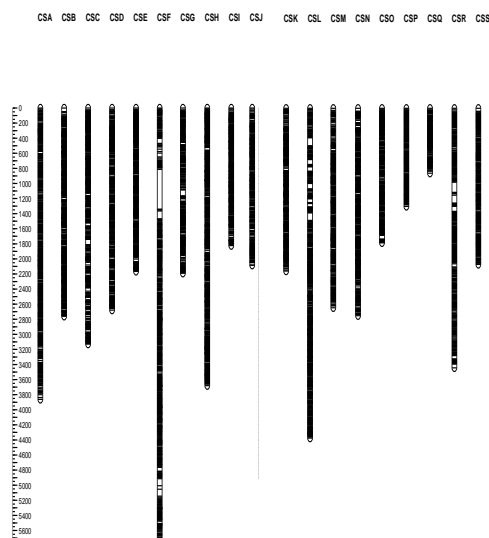
NEW GENOMIC RESOURCES

SCUBA CANCER ERC PROJECT (IP: Dr José Tubío-USC)
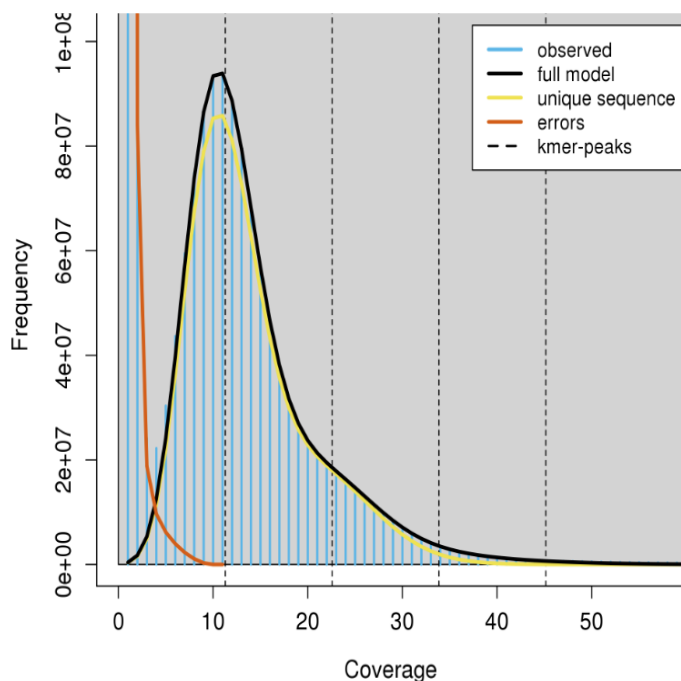
The size of the cockle's genome is ~0.8 Gb; N50 ~40 Mb; 19 major scaffolds corresponding 19 chromosomes of the haploid karyotype
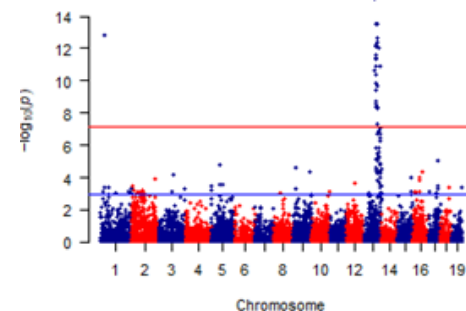
COCKLE'S GENETIC MAP

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

LARGE SNPs RESOURCES
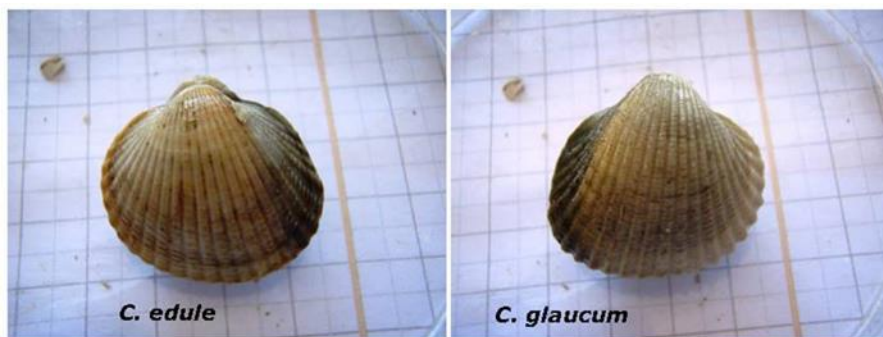- ~ 10,000 anonymous SNPs
- > 50,000 candidate genes associated SNPs

MLM Analysis of Trait STRIPE in both families

A major QTL associated with cockle's color

Development of a molecular tool for discrimination between *C. edule* and *C. glaucum* and their putative hybrids

C. edule

C. glaucum

Taking advantage of genomic resources (RADseq + *C. edule* genome), development of a **molecular tool (SNaPshot technique) composed by seven SNPs**, with different variants fixed for each species.

GCAT
TACG
GCAT *genes*   MDPI

Type of the Paper (Short Communication)

A useful SNP panel to distinguish two cockle species, *Cerastoderma edule* and *C. glaucum*, co-occurring in some European beds, and their putative hybrids.

Francesco Maroso [1a], Celia Pérez de Gracia [1], David Iglesias [2], Asunción Cao [2], Seila Díaz [3], Antonio Villalba [2,4,5], Manuel Vera [1,6*] and Paulino Martínez [1,6]

Output: Genetic diversity and structure of Atlantic cockle beds

- Genetic diversity in Atlantic cockle beds.

- Temporal and geographic genetic structure: demographic parameters.

- Signals of natural selection regarding environmental and biological factors.

- Definition of operational conservation units for cockles.

Expected results: Genetic diversity, structure and effect of biotic and abiotic variables on cockle distribution

- NGS methodologies allow population genetics analysis at genomic level, increasing the power to detect the effect of evolutionary forces and environmental variables depicting the structure of the species, useful for its appropriate management.
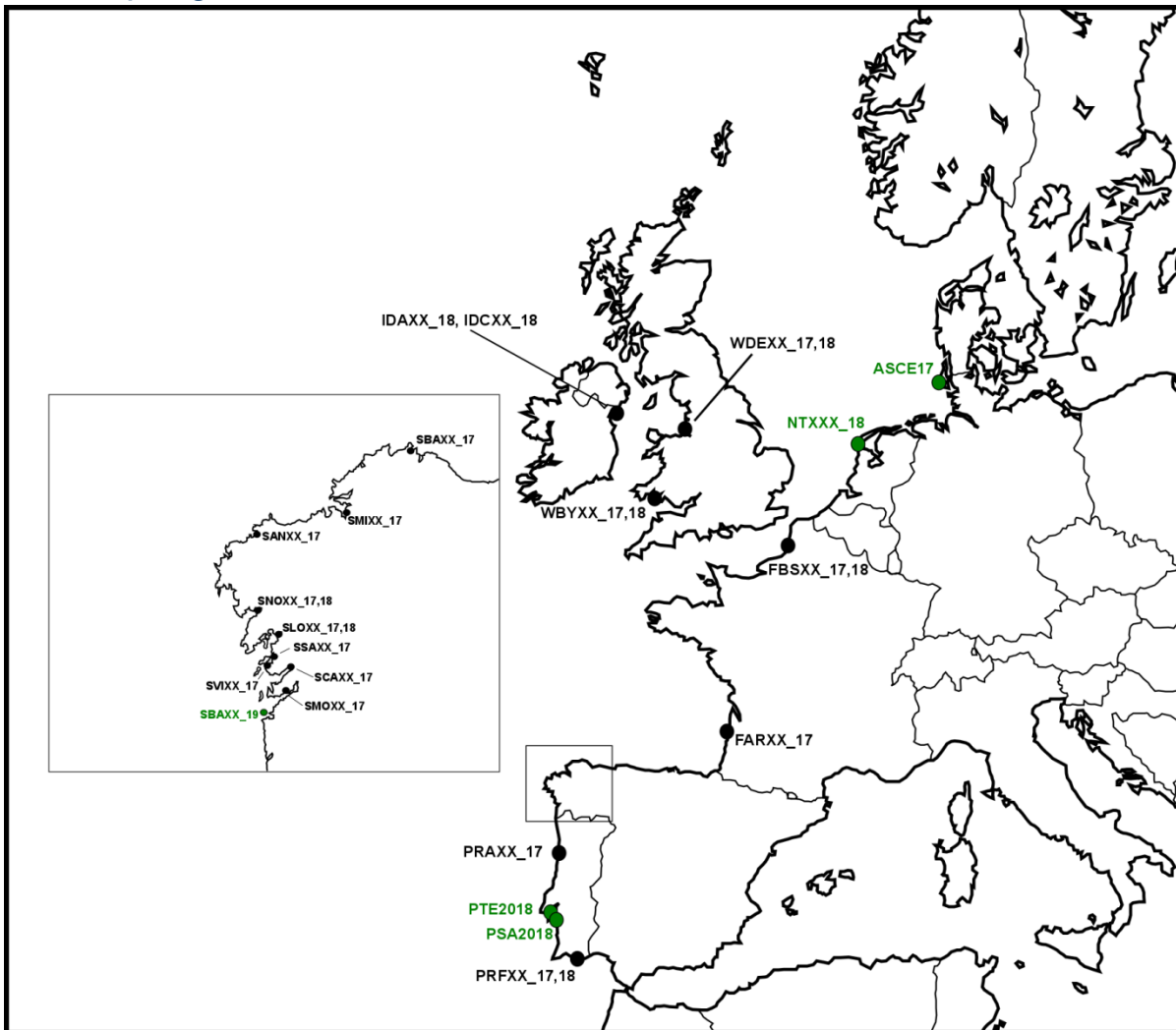
22 beds (746 individuals):

Macrogeographical analysis:
14 beds (including six temporal replicates)

Microgeographical analysis:
10 beds from Galician coasts (two of them included on macrogeographical analysis

Workflow:

Tissue arrival (fixed in absolut ethanol) → DNA extraction → Genomic library construction (2b-RAD libaries, 90 ind/library)

↓

Population genomics analyses ← Bioinformatic work: Genotyping ← Sequencing in NGS platform
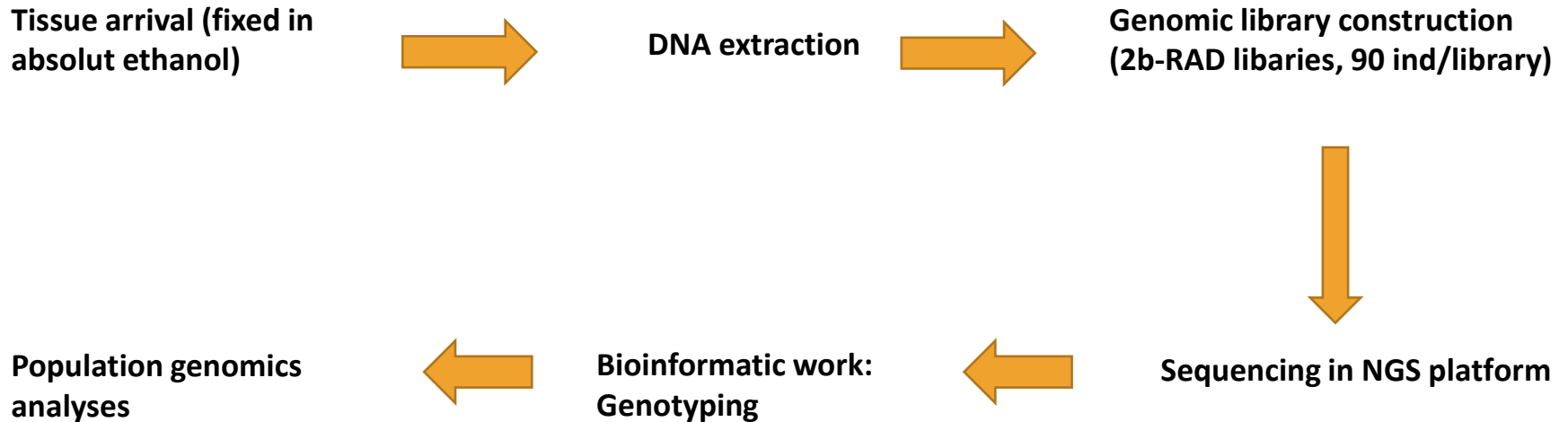
9309 markers after quality and population filtering

Genetic diversity:

Location	N	Ho	He	F_{IS}
ASCE_17	22	0.070	0.077	0.089
NTX_18	21	0.075	0.078	0.034
WDE_17	28	0.071	0.078	0.079
WDE_18	30	0.075	0.083	0.099
WBY_17	30	0.073	0.081	0.092
WBY_18	30	0.070	0.077	0.098
IDA_18	29	0.074	0.081	0.083
IDC_18	22	0.077	0.084	0.084
FBS_17	30	0.072	0.080	0.110
FBS_18	31	0.072	0.079	0.084
FAR_17	30	0.074	0.083	0.111
SBA_17	30	0.076	0.085	0.104
SMI_17	30	0.074	0.082	0.104
SAN_17	29	0.073	0.081	0.101
SON_17	30	0.079	0.088	0.098
SNO_18	30	0.072	0.081	0.118
SLO_17	30	0.076	0.086	0.114
SLO_18	32	0.073	0.083	0.119
SSA_17	30	0.075	0.083	0.101
SVI_17	25	0.072	0.081	0.113
SCA_17	30	0.073	0.083	0.117
SMO_17	19	0.071	0.078	0.085
SBA_19	17	0.075	0.083	0.091
PRA_17	30	0.072	0.081	0.113
PTE_18	8	0.073	0.078	0.063
PSA_19	20	0.071	0.080	0.108
PRF_17	30	0.079	0.084	0.066
PRF_18	23	0.079	0.086	0.079

Temporal stability (2017-2018 cohorts):

F_{ST} WDE$_{17-18}$ = -0.0128; P = 0.9961 ± 0.0016

F_{ST} WBY$_{17-18}$ = -0.0134; P = 0.9932 ± 0.0022

F_{ST} FBS$_{17-18}$ = -0.0009; P = 0.6914 ± 0.0153

F_{ST} SNO$_{17-18}$ = -0.0049; P = 0.9990 ± 0.0002

F_{ST} SLO$_{17-18}$ = 0.0012; P = 0.2471 ± 0.0136

F_{ST} PRF$_{17-18}$ = 0.0003; P = 0. 2012 ± 0.0102

Temporal stability between cohorts

Selection footprints (outliers, BAYESCAN):

Outliers FDR (q value) < 0.05 = 460 markers

- Divergent Selection = 441

- Balancing Selection = 19

Haploid number (n) cockle:
19 chromosomes
(SCUBA-CANCER collaboration)
Outlier genomic regions associated to
(GO terms):

- Neural function and development

- Immune response and defence

- Metabolism and growth

|	Bayescan Outliers (FDR < 0.05)	
Genomic Location	Balancing selection	Divergent Selection
Mega-scaffold 1	2	40
Mega-scaffold 2	1	33
Mega-scaffold 3		39
Mega-scaffold 4	1	31
Mega-scaffold 5	2	28
Mega-scaffold 6		24
Mega-scaffold 7	2	13
Mega-scaffold 8	3	24
Mega-scaffold 9	2	18
Mega-scaffold 10	1	23
Mega-scaffold 11	1	15
Mega-scaffold 12		22
Mega-scaffold 13	1	25
Mega-scaffold 14	2	14
Mega-scaffold 15	1	16
Mega-scaffold 16		32
Mega-scaffold 17		13
Mega-scaffold 18		4
Mega-scaffold 19		11
Other scaffolds (15)		15
Not found in genome		1
Total	19	441

Population structure (pairwise F_{ST} values):

9309 SNPs (below diagonal)
441 divergent SNPs (above diagonal)

Global F_{ST} = 0.0240 (*P* < 0.001)

Population structure (fastSTRUCTURE):
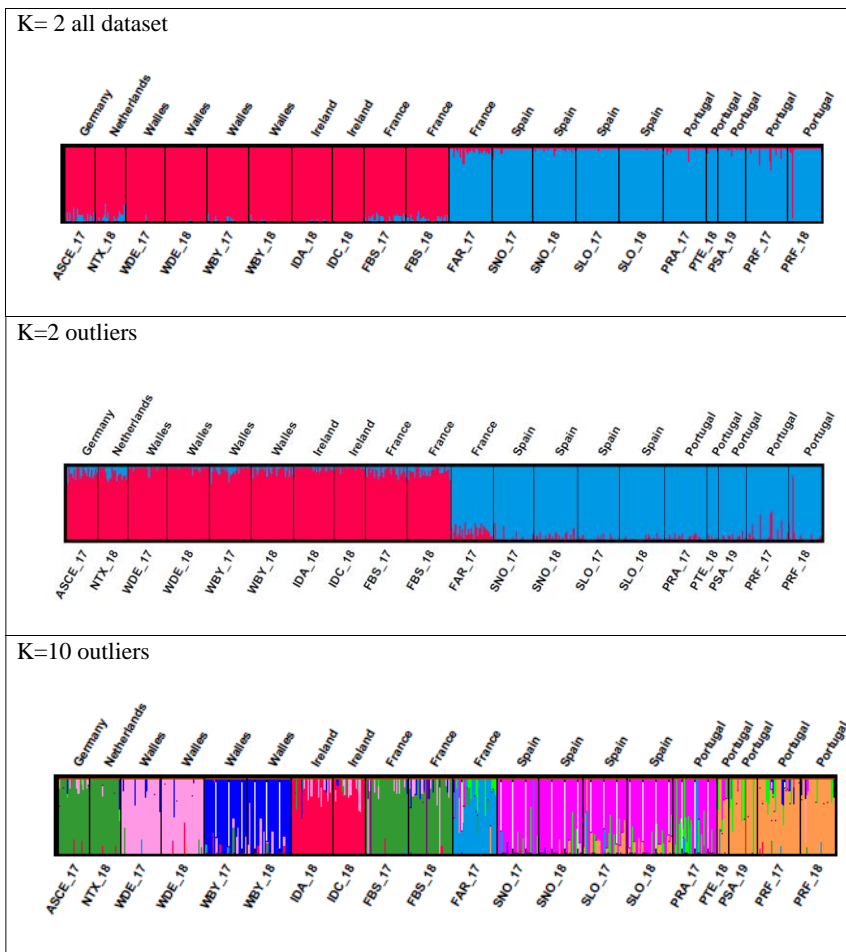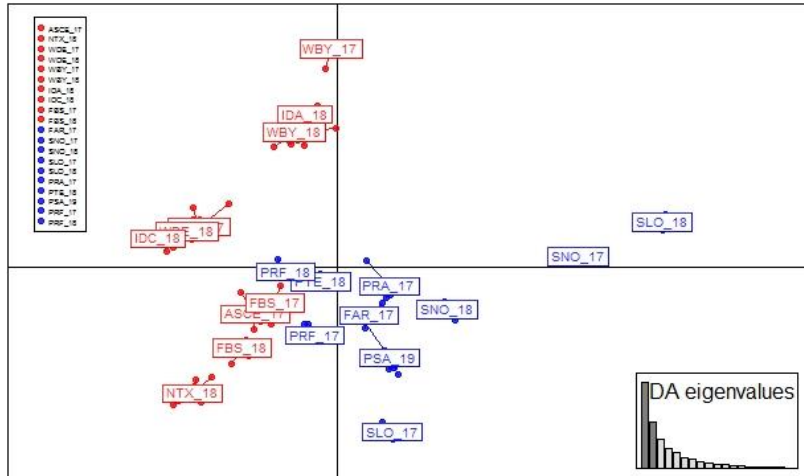

Two groups (Ushant front)

Seven main groups:
- North Sea up to English Channel
- WDE
- WBY
- Ireland
- Biscay Gulf,
- North-western Spain (Galicia) together with N Portugal
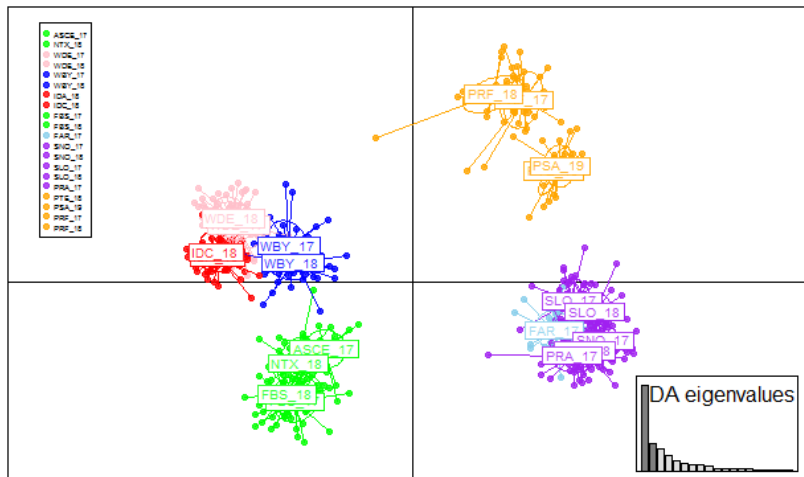- S Portugal

Effect of biogeographical barriers?

Population structure (DAPC):
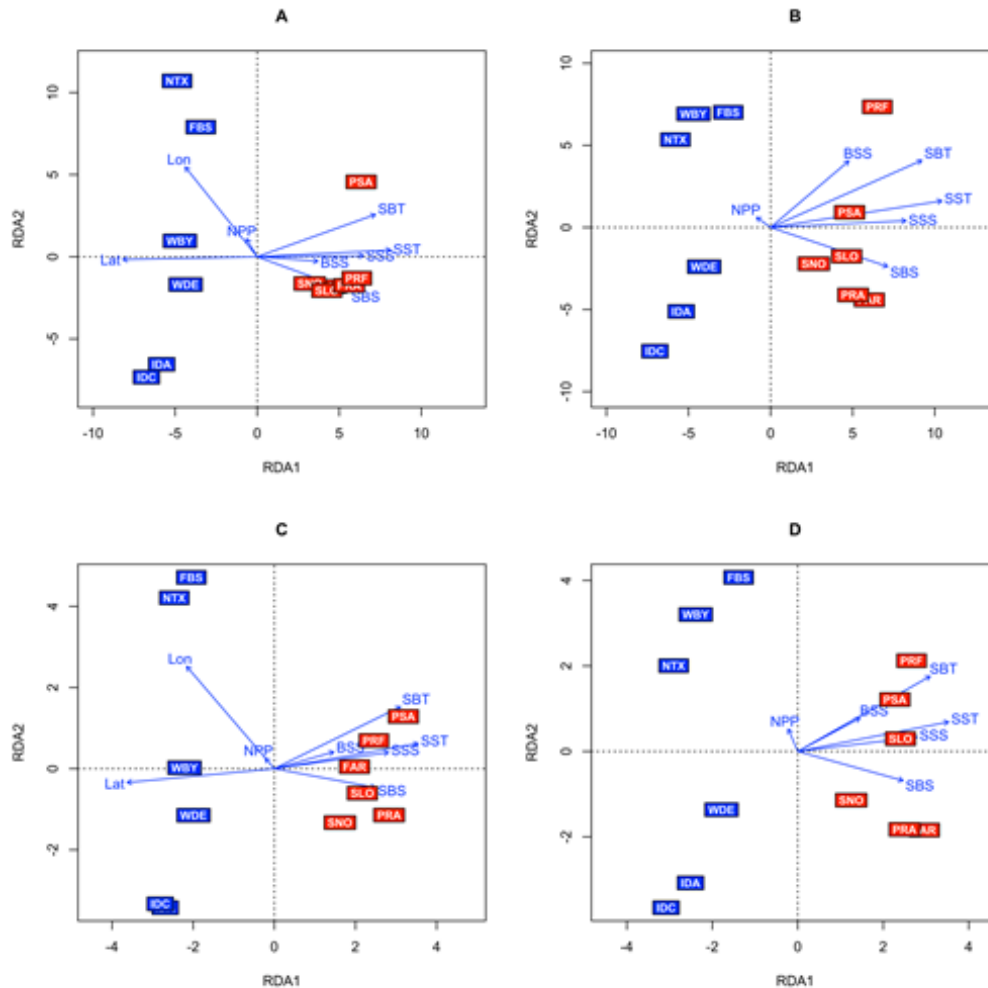

9309 SNPs

441 divergent SNPs

Redundancy Analyses (RDA):

Redundancy analyses (RDA):
9309 SNPs (A and B)
441 divergent SNPs (C and D)

Model 1: all landscape variables (A and C)
Model 2: only abiotic factors (B and D)

SST: Sea Surface Temperature (driver)
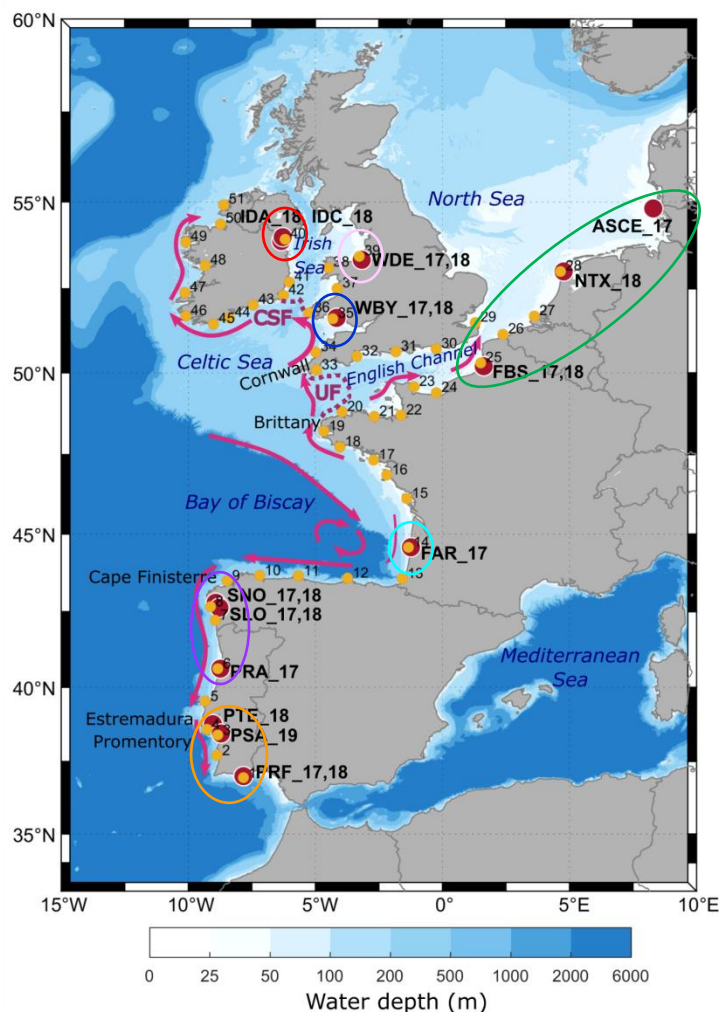SBT: Sea Bottom Temperature
SSS: Sea Surface Salinity
SBS: Sea Bottom Salinity
BSS: Bottom Shear Stress
NPP: Net Primary Productivity

Biogeographical barriers and fronts (UoB):

Study area for *C. edule* genetic analysis and larval dispersal modelling:

Ocean bathymetry is shaded in blue.

Summer surface currents are schematically represented by magenta-coloured arrows.

Locations of fronts are depicted by purple dotted lines (CSF = Celtic Sea Front; UF = Ushant Front). Beds are shown in dark red.

Particle release locations for larval dispersal modeling are shown in yellow and numbered from 1 to 51.

Microgeographical studies along barriers and fronts described

Microgeographical scale: Galicia

|	SBA_17	SMI_17	SAN_17	SNO_17	SNO_18	SLO_17	SLO_18	SSA_17	SVI_17	SCA_17	SMO_17	SBA_19
SBA_17	0.0000											
SMI_17	0.0015	0.0000										
SAN_17	0.0014	0.0018	0.0000									
SNO_17	0.0004	0.0009	0.0003	0.0000								
SNO_18	0.0007	0.0015	0.0004	-0.0005	0.0000							
SLO_17	0.0012	0.0010	0.0005	-0.0004	0.0018	0.0000						
SLO_18	0.0014	0.0008	0.0006	0.0001	0.0002	0.0013	0.0000					
SSA_17	0.0016	0.0010	0.0009	0.0002	0.0000	-0.0002	0.0001	0.0000				
SVI_17	0.0004	0.0007	0.0002	-0.0024	-0.0005	-0.0010	-0.0007	-0.0011	0.0000			
SCA_17	0.0011	0.0012	-0.0003	-0.0007	0.0003	0.0015	-0.0005	0.0008	-0.0014	0.0000		
SMO_17	0.0020	0.0018	0.0002	-0.0019	0.0005	0.0018	-0.0002	-0.0006	-0.0005	-0.0015	0.0000	
SBA_19	0.0024	0.0023	0.0016	0.0004	0.0016	0.0011	-0.0009	0.0017	-0.0010	-0.0006	0.0022	0.0000

All pairwise F_{ST} values non-significant (P > 0.05)

Global Galician F_{ST} value = 0.0019 (P > 0.05)

Panmictic unit (no effect of Cape Finisterre)

Microgeographical analyses (UoB, UCC, UCaen, CIMA, CETMAR):
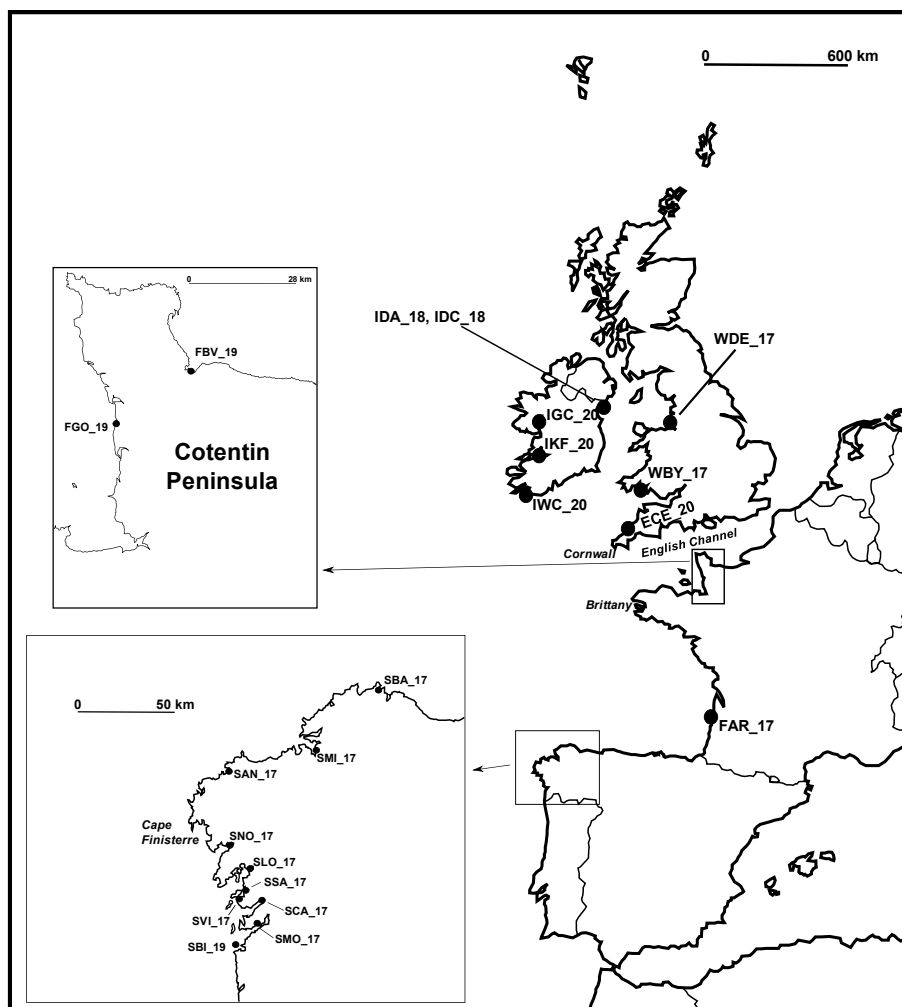
Areas under study (in process):
Celtic/Irish Sea
Normandy (English Channel)

Conclusions (take-home messages):

- Two main genetic groups were identified, northwards and southwards of French Brittany, in accordance with our modelling, which demonstrated a barrier for larval dispersal linked to the Ushant front.

- Further genetic subdivision was observed using outlier loci (under divergent selection) and considering larval behaviour (up to seven groups geographically distributed).

- Sea Surface Temperature (SST) could be an environmental driver explaining genetic differences (following a latitudinal axis).

- Refined microgeographic analyses along Galician coast suggested no influence of Cape Finisterre as biogeographical barrier (single panmictic unit detected). Further studies along other barriers described previously in the AA will help to get a more comprehensive picture

Natural outbreak 2018 vs 2012

Population	N individuals	Affected by *Martelia* (until 2018)
SLO-18 (t_0)	14	No
SLO-18 (t_1-t_2)	24	Yes
SLO-12 (Lombos)	20	No
SSA-12 (Sarrido)	18	No

Natural outbreak 2018 vs 2012

Genetic diversity

Population	Ar	Ho	He	F_{IS}
SLO-18 (t_0)	1.328	0.072	0.072	0.0042
SLO-18 (t_1-t_2)	1.371	0.076	0.081	0.0702
SLO-12 (Lombos)	1.379	0.080	0.084	0.0482
SSA-12 (Sarrido)	1.380	0.079	0.085	0.0726

Identification of signatures of selection 2018 vs 2012: outliers

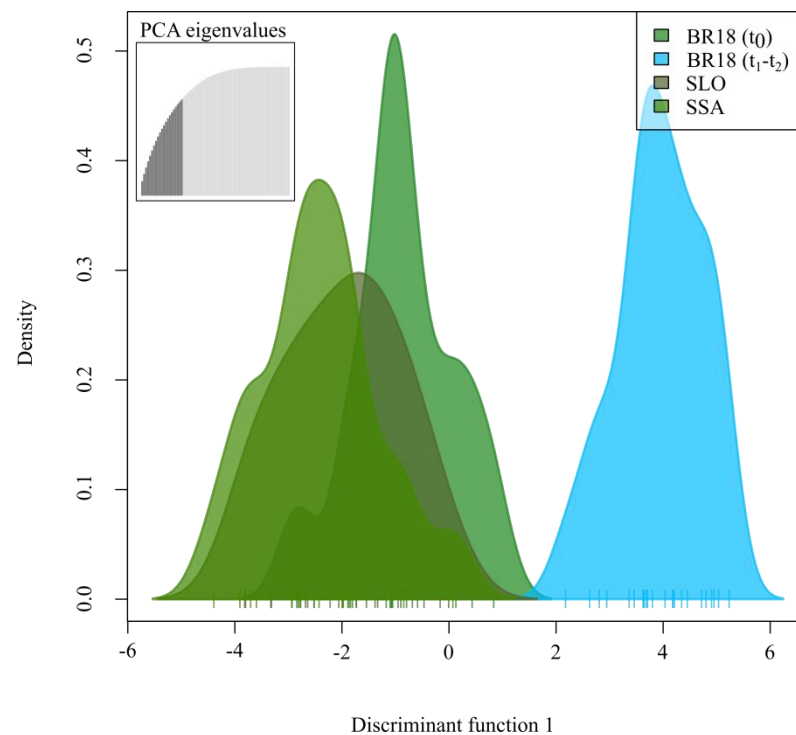| Approach	BayeScan (q-value <0.05)	Arlequin (p-value<0.01)
4 pop Arlequin		177
2 pop Arlequin		121

POPULATION GENOMICS MARTEILIOSIS

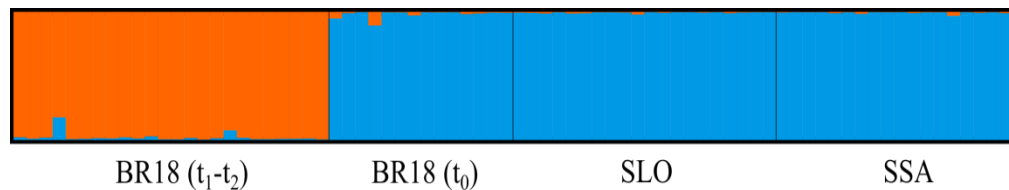### Natural outbreak 2018 vs 2012

21 PCs
74.5% explained variance

STRUCTURE (121 outliers)

K=2

Natural outbreak 2018 vs 2012

Conclusions:

- Genetic diversity slightly lower in 2018 likely as the impact of marteiliosis

- 121 outlier loci fully discriminate infected and non-infected cockles representing a useful tool for MAS to obtain a resistant strain

- 40 of these markers are linked or within differentially expressed genes (see WP7 tomorrow!) and will be validated in an ongoing experiment in Ria de Vilagarcía