

Article

Heterogeneous Genomic Divergence Landscape in Two Commercially Important European Scallop Species

David L. J. Vendrami ^{1,*} , Joseph I. Hoffman ^{1,2}  and Craig S. Wilding ³ ¹ Department of Animal Behaviour, University of Bielefeld, Postfach 100131, 33615 Bielefeld, Germany² British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK³ School of Biological and Environmental Sciences, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK

* Correspondence: david.vendrami@uni-bielefeld.de

Abstract: Two commercially important scallop species of the genus *Pecten* are found in Europe: the north Atlantic *Pecten maximus* and the Mediterranean *Pecten jacobaeus* whose distributions abut at the Almeria–Orán front. Whilst previous studies have quantified genetic divergence between these species, the pattern of differentiation along the *Pecten* genome is unknown. Here, we mapped RADseq data from 235 *P. maximus* and 27 *P. jacobaeus* to a chromosome-level reference genome, finding a heterogeneous landscape of genomic differentiation. Highly divergent genomic regions were identified across 14 chromosomes, while the remaining five showed little differentiation. Demographic and comparative genomics analyses suggest that this pattern resulted from an initial extended period of isolation, which promoted divergence, followed by differential gene flow across the genome during secondary contact. Single nucleotide polymorphisms present within highly divergent genomic regions were located in areas of low recombination and contrasting patterns of LD decay were found between the two species, hinting at the presence of chromosomal inversions in *P. jacobaeus*. Functional annotations revealed that highly differentiated regions were enriched for immune-related processes and mRNA modification. While future work is necessary to characterize structural differences, this study provides new insights into the speciation genomics of *P. maximus* and *P. jacobaeus*.

Keywords: *Pecten*; scallop; RAD sequencing; genome differentiation; speciation; chromosomal inversion; inversion



Citation: Vendrami, D.L.J.; Hoffman, J.I.; Wilding, C.S. Heterogeneous Genomic Divergence Landscape in Two Commercially Important European Scallop Species. *Genes* **2023**, *14*, 14. <https://doi.org/10.3390/genes14010014>

Academic Editors: Francesca Bertolini and Luca Fontanesi

Received: 28 October 2022
Revised: 15 December 2022
Accepted: 19 December 2022
Published: 21 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

How genomes diverge during speciation is a central question in evolutionary genetics. Speciation is generally considered as a continuum where separate genetic lineages accumulate differences over time after a vicariant event due to mutation, genetic drift, and selection [1,2]. During periods of prolonged geographical isolation, genomic divergence can build up simply as a consequence of genetic drift. Additionally, when two separating lineages are exposed to different selective pressures, differentiation may be promoted by divergent selection [3]. Under this scenario, new beneficial mutations may undergo strong positive selection, which will increase their frequency and decrease genetic diversity at nearby sites via selective sweeps [4,5], further promoting divergence. In contrast, gene flow counteracts genetic divergence by homogenizing allele frequencies between diverging lineages. However, gene flow may not influence the entire genome equally, as chromosomal rearrangements such as inversions may favour reproductive isolation [6]. Notably, these areas are often hotspots for the accumulation of adaptive divergence [7]. Thus, the opposing forces of selection, genetic drift, and gene flow can give rise to heterogeneous landscapes of genomic divergence between separating taxa, with elevated differentiation at loci underlying divergent selection and loci involved in reproductive isolation [8,9]. Disentangling the mechanisms that led to such heterogeneous patterns is not trivial especially because apparently equivalent landscapes of genomic differentiation may arise from

markedly different demographic histories, such as ecological divergence with continuous gene flow and allopatric divergence followed by secondary contact [8].

Relative measures of genetic divergence, such as F_{ST} [10], are sensitive to within-species levels of genetic diversity. Therefore, when using these kinds of measures, peaks of elevated genomic divergence along an heterogeneous landscape of differentiation will be identified both when these derive from differential gene flow across the genome, and also when these are the result of differential selection, even in the absence of gene flow [11]. This is because linked selection will reduce genetic diversity locally, resulting in unusually high estimates of divergence [11]. In contrast, absolute measures of genetic divergence, such as d_{xy} [12], are independent of levels of genetic diversity and will therefore identify peaks of divergence only in the presence of differential gene flow, as local reductions in genetic diversity caused by selection will not affect these kinds of estimates. Cruickshank and Hahn [11] proposed the comparison of relative and absolute measures of genetic divergence as an effective means of discriminating between the effects of differential selection and migration, while still identifying genomic regions potentially involved in adaptive divergence. This approach, together with demographic inference, should in principle allow the detailed reconstruction of the events and mechanisms responsible for generating heterogeneous landscapes of divergence during the speciation process.

The commercially exploited scallops *Pecten maximus* and *P. jacobaeus* are two Pectinids inhabiting different parts of the coast of Europe. *P. maximus* has a north Atlantic distribution and can be found from Norway to the western Mediterranean [13]. In contrast, *P. jacobaeus* is restricted entirely to the part of the Mediterranean that lies to the east of the Almeria–Orán front (AOF), a recognized biogeographic boundary [14] known to represent a major barrier for gene flow in both sessile species with planktonic larval phases (e.g., the urchin *Paracentrotus lividus* [15], the clam *Donax trunculus* [16], the mussel *Mytilus galloprovincialis* [17], the snail *Stramonita haemastoma* [18]), and species with actively swimming larvae and juveniles (e.g., *Dicentrarchus labrax* [19]). However, the extent to which the AOF is preventing and has prevented gene flow between *P. maximus* and *P. jacobaeus* remains unclear.

These two scallop taxa are currently considered as two separate species on the basis of morphological traits and nonoverlapping geographical distributions, but this view has been put into question by a suite of genetic studies published over the last two decades with different levels of genetic resolution offered by a variety of genotyping approaches including allozymes [13,20], mitochondrial DNA sequencing [20], microsatellites [21] and single nucleotide polymorphisms (SNPs) [22]. None of these studies identified any fixed differences between the two species and all of them reported genetic distances well below the magnitude expected under the original supposition that *P. maximus* and *P. jacobaeus* separated around five million years ago (MYA), shortly after the Messinian salinity crisis [23].

Vendrami et al. [22] examined genomic divergence within and between *P. maximus* samples collected from 12 sites along the Atlantic coast of Europe and *P. jacobaeus* samples collected from two sites along the Mediterranean coast. Clear genetic differences were found between *P. jacobaeus* and *P. maximus* yet the two species were estimated to have diverged much earlier than originally thought, sometime around the end of the Pleistocene (~95,000 generations or 0.5 MYA). Climatic conditions strongly fluctuated during that period and many bivalve species went extinct or shifted their distributions [24]. It is therefore reasonable to infer that newly established populations of the common ancestor of these two *Pecten* species might have begun to experience novel environmental pressures, leading them to diverge from populations located in their native range. Moreover, it is likely that divergence was also promoted by genetic drift and the consequent potential increase in reproductive isolation caused by the rise, for example, of Bateson–Dobzhansky–Muller incompatibilities [25–28]. Furthermore, Vendrami et al. [22], in line with the findings of Morvezen et al. [21], showed that *P. maximus* is composed of two major genetic lineages, one located along the coast of Norway and the other along the rest of the Atlantic coast of Europe, which diverged more recently during the last glacial maximum (~18,000 years

ago). This finding was recently corroborated by Hollenbeck et al. [29] who also showed that chromosomal inversions appear to play an important role in maintaining adaptive variation in *P. maximus* populations inhabiting different regions along a European latitudinal gradient.

Recently, a high-quality chromosome level genome assembly for *P. maximus* was published by Kenny et al. [30] and this was subsequently refined by Zeng et al. [31]. Here we have reanalysed the RADseq data produced by Vendrami et al. [22] by mapping them to the reference genome of Zeng et al. [31] to characterize the genomic landscape of the divergence between *P. maximus* and *P. jacobaeus* and investigate its likely determinants. Our aims were, (i) to compare a series of demographic models of speciation to understand which are the mostly likely circumstances under which the divergence between *P. maximus* and *P. jacobaeus* occurred; (ii) to characterize and quantify relative and absolute divergence along the genome while also identifying, if present, any fully diagnostic loci; and (iii) to characterize the regions of these species' genomes showing the highest levels of differentiation, as these may underlie important adaptive divergence.

2. Materials and Methods

2.1. RAD Sequencing Data

The RAD sequencing data originally described by Vendrami et al. [22] were downloaded from the Sequence Read Archive (SRA) using the accession number PRJNA526674. These included sequencing data from all 280 samples in Vendrami et al. [22], comprising both great scallops (*P. maximus*), collected from 12 locations from the Atlantic coast of Europe, and Mediterranean scallops (*P. jacobaeus*), collected from two locations in the Mediterranean Sea (Figure 1, Table S1). Briefly, the whole genomic DNA was extracted from each sample from the adductor muscle tissue using an adapted phenol–chloroform protocol [32] and shipped to the Beijing Genomics Institute for RAD sequencing [33]. Libraries were constructed using the restriction enzyme PstI and 50 bp SE sequenced on an Illumina HiSeq 4000 (Illumina Hayward, Hayward, CA, USA).

2.2. Raw Read Processing, SNP Calling, and Genotyping

Raw read quality was assessed using FastQC v0.11.9 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, accessed on 17 December 2021), which revealed that trimming was unnecessary as read quality scores were uniform along the entire length of the reads. Sequencing reads were then mapped to the *P. maximus* reference genome [31] using BWA MEM v0.7.13 [34] with default parameters. The resulting alignment files were used as input within the ref_map.pl module of the Stacks pipeline [35] for SNP calling and genotyping. Next, raw genotypes were filtered to retain only SNP calls with genotype quality and depth of coverage greater than five using VCFTools v0.1.17 [36]. Subsequently, we retained only SNPs genotyped in at least 80% of the individuals and discarded samples with more than 80% of missing data using VCFTools. The resulting VCF file was converted into a folded joint site frequency spectrum (SFS) using easySFS (<https://github.com/isaacovercast/easySFS>, accessed on 28 November 2022) to be used for demographic modelling (see Section 2.3). We note that the resulting SFS contained information on rare variants since no minor allele frequency (MAF) filtering was implemented at this point. Following this, VCFTools was utilized to further filter our SNP dataset for downstream analyses to retain only variants with an MAF greater than 0.05. We did not implement a Hardy–Weinberg equilibrium filtering step, as this would have removed highly divergent and fully diagnostic loci. In addition to this dataset, we generated five alternative SNP datasets to explore how sensitive the inference of genomic divergence (see Section 2.4) was to different filtering strategies. Specifically, we produced a dataset where only SNPs with genotyping rates below 80% in both species (as opposed to a global 80% threshold) were excluded. Two additional SNP datasets were obtained by removing SNPs with MAF below 1% and by not implementing MAF filtering, respectively. Finally, the remaining two datasets were generated by excluding SNPs that showed significant

departures from Hardy–Weinberg equilibrium (HWE), using a p -value threshold of 0.05, either in all of the sampled populations or only in half of the sampled populations.

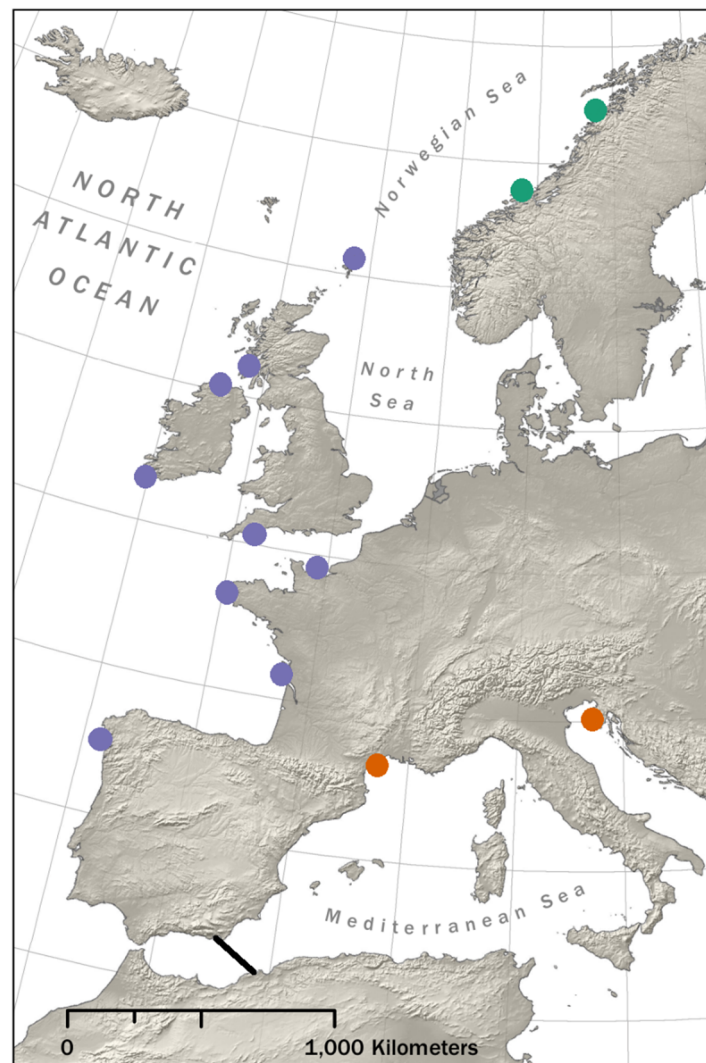


Figure 1. Map showing sampling sites (circles) and the location of the Almeria–Orán front (AOF, black line). Orange circles indicate sampling sites for *Pecten jacobaeus*, while the remaining circles indicate sampling sites for *P. maximus*. The *P. maximus* populations belonging to the previously identified Atlantic and Norwegian genetic clusters are indicated in purple and green [21,22], respectively.

2.3. Demographic Modelling

We implemented demographic analyses using the coalescent simulator *fastsimcoal2* [37] in combination with the joint folded SFS obtained from our data, prior to MAF filtering, to reconstruct the divergence history of *P. maximus* and *P. jacobaeus*. Specifically, we compared five alternative models of speciation that differ in the absence, presence and timing of gene flow. The first model (SI, Figure S1a) represents a model of divergence without gene flow and consists of two genetic lineages, one corresponding to *P. maximus* and one corresponding to *P. jacobaeus*, which diverged T_S generations ago. Constant population size was assumed throughout the model. The second model (IM, Figure S1b) represents a model of isolation with constant gene flow and differs from the SI model by allowing constant gene flow between the two lineages after divergence. The third model (AM, Figure S1c) represents a model of ancestral gene flow followed by isolation. In this case, gene flow occurs immediately after divergence, but stops in the past, T_1 generations ago. The fourth model (SC, Figure S1d) represents a model of secondary contact where after an

initial period of allopatric divergence, gene flow is established T_1 generations ago and continues until the present. Finally, the fifth model (SCS, Figure S1e) represents another model of secondary contact which differs from the SC model by interrupting the gene flow T_2 generations before the present. For all models, in addition to T_S , T_1 , and T_2 , we estimated the effective population size of the *P. maximus* lineage (N_{pma}), of the *P. jacobaeus* lineage (N_{pja}), and of the ancestral lineage that later split into the two *Pecten* species (N_{anc}). Moreover, in all models except SI, we also estimated the gene flow as the proportion of migrants per generation from the *P. maximus* lineage to the *P. jacobaeus* lineage (mmj) and vice versa (mjm). mmj and mjm were estimated independently from one another to allow for the possibility of asymmetrical gene flow to occur. All priors for these analyses are given in Table S2.

For each model, we implemented 50 independent *fastsimcoal2* runs with 100,000 simulations and 40 cycles of the likelihood maximization algorithm. The run associated with the highest maximum likelihood value for each model was then selected to compute Akaike's information criteria (AIC) as described in Varin and Vidoni [38] to allow for model comparison. We then extracted parameter estimates from the model receiving the highest support and computed their 95% confidence intervals (CI) via parametric bootstrapping, as described by Excoffier et al. [37]. Specifically, we simulated 100 folded SFS based on the parameter estimates from the best model and used them to recalculate the model parameters using the same workflow described above.

2.4. Comparative Genomics

In order to evaluate the degree of genetic differentiation between Mediterranean and great scallops along their genomes, we calculated F_{ST} values within sliding windows using VCFTools. Specifically, a 100 kb window was used, as recommended by Hohenlohe et al. [39] for RAD sequencing data, and only the windows containing at least five SNPs were considered. This analysis was replicated for five additional SNP datasets (see Section 2.2) to assess how sensitive the pattern of genomic divergence between the two species was to different filtering criteria. Windows whose F_{ST} was above the 95th percentile were considered as highly divergent genomic regions. For comparison, we repeated the same analysis where F_{ST} values were calculated between the Atlantic and Norwegian genetic lineages within *P. maximus*. Next, separately for the two scallop species, we used VCFTools to quantify nucleotide diversity within the same sliding windows used to calculate F_{ST} and tested for the association between these two measures by performing a Spearman correlation test. Subsequently, we employed the software pixy [40], in combination with a VCF file which also included nonvariant sites, to calculate d_{xy} between the two scallop species within the same genomic windows described above to verify whether genomic regions characterized by elevated F_{ST} values also showed high divergence regardless of local levels of genetic diversity. The correlation between these two measures was assessed using a Spearman correlation test. Moreover, we compared highly divergent genomic windows against the remaining genomic areas at their levels of divergence quantified using both F_{ST} and d_{xy} . Agreement between these two measures would imply differential gene flow across the *Pecten* genome, while disagreement would indicate that heterogeneous patterns of F_{ST} along the genome were primarily caused by differential selection [11]. Finally, we aimed at identifying fully diagnostic SNPs for the two scallop species. Specifically, we searched for SNPs that were fixed for the reference allele in great scallops and fixed for the alternative allele in Mediterranean scallops.

2.5. Linkage Disequilibrium

In order to evaluate how linkage disequilibrium (LD) decays with physical distance along the genome in *P. maximus* and *P. jacobaeus*, we calculated r^2 values among all pairs of loci located within the same chromosome, separately for the two species, using PLINK [41]. LD decay was then visualized by fitting a nonlinear regression curve, where the expected value of r^2 under drift recombination was expressed according to Hill and Weir [42]. Next,

we repeated the same analysis but focusing exclusively on SNPs located within highly divergent genomic windows to verify whether loci located on highly differentiated areas of the *Pecten* genome were characterized by a different LD decay pattern. Subsequently, we employed the R package 'LDheatmap' [43] to generate LD heatmaps showing LD among all pairs of SNPs separately for each chromosome and separately for the two species. This was performed in order to graphically assess whether any genomic regions were characterized by increased levels of LD, which would be indicative of a locally reduced recombination.

2.6. Gene Annotations and Enrichment Analysis

In order to functionally characterize highly divergent genomic regions of the scallop genome, we focused on the SNPs located within windows whose F_{ST} values were above the 95th percentile as well as on diagnostic SNPs (from now on collectively referred to as 'highly divergent SNPs'). First, we used the great scallop reference genome annotations provided by Zeng et al. [31] to annotate highly divergent SNPs based on the genes they were located in or on the closest gene. Second, we retrieved gene ontology (GO) terms for all highly divergent SNPs using the annotations provided by Zeng et al. [31]. For SNPs not located within a gene, we used the GO annotation of the closest gene. These were used to perform a GO enrichment analysis using the R package topGO [44] to investigate whether any specific biological processes were overrepresented among the genes containing, or located in close proximity to, highly divergent SNPs. This was implemented while using a background list of GO terms containing all of the GO-annotated genes of the great scallop. In order to account for hierarchical relationships and nonindependence among the GO terms, we used the "weight01" algorithm with a Fisher statistic. GO terms with p -value < 0.05 were considered significantly enriched. We then repeated the GO enrichment analysis considering only SNPs located within genes. Finally, we used the program SnpEff v5.1 [45], together with a custom annotation file built using the great scallop reference genome annotations, to categorize highly divergent SNPs into different mutation classes. Specifically, this procedure allowed us to verify whether nucleotide changes within genes were synonymous, nonsynonymous, or loss-of-function mutations.

3. Results

We used the RADseq data of Vendrami et al. [22] to characterize patterns of genomic divergence between the two European scallop species, *P. maximus* and *P. jacobaeus*. RAD sequencing produced a total of 823,811,871 50 bp SE sequence reads. Alignment rates to the *P. maximus* reference genome were high (>90%) for both scallop species. The aligned reads were used to call a total of 1,009,368 raw SNPs. Application of the filtering criteria described in the materials and methods resulted in a final dataset consisting of 45,704 SNPs genotyped in 262 samples, including 235 great scallops and 27 Mediterranean scallops.

3.1. Demographic Inference

We used the coalescent simulator *fastsimcoal2* together with the empirical folded joint SFS to implement demographic reconstructions and to infer the divergence history of *P. maximus* and *P. jacobaeus*. All of the models that included gene flow between the two lineages received higher support than the SI model (Table S3). Among them, the best supported model was the secondary contact SCS model. The resulting parameters estimates together with their 95% CI are reported in Table S4. *P. maximus* and *P. jacobaeus* were estimated to have diverged 299,001 (95% CI: 235,482–329,076) generations ago. After a relatively long period of initial isolation, secondary contact occurred 14,799 (95% CI: 14,020–20,850) generations ago, with gene flow continuing until 58 (95% CI: 20–286) generations before present. Gene flow occurred in both directions but was more pronounced from the *P. jacobaeus* lineage to the *P. maximus* lineage.

3.2. Genomic Divergence

We calculated F_{ST} values between *P. maximus* and *P. jacobaeus* within 4460 100 kb sliding windows containing at least five SNPs. The resulting F_{ST} values revealed a highly variable divergence landscape between the two scallop species, with F_{ST} values ranging from -0.013 to 0.967 across the genome (Figure 2a) and an overall mean F_{ST} value of 0.118 (Table 1). This pattern was highly consistent across differently filtered SNP datasets (Figure S2), implying that our results are rather robust. A total of 223 windows showed F_{ST} values above the 95th percentile ($F_{ST} = 0.623$). These were located on chromosomes 1–3, 7, 9–15, and 17–19, with chromosomes 2 and 10 carrying the largest numbers of highly divergent genomic regions (18.38% and 28.46% of the total number of sliding windows, respectively) and having the highest chromosome-wise mean F_{ST} values (Table 1). In contrast, chromosomes 4–6, 8, and 16, as well as the unplaced scaffolds, did not carry any windows with F_{ST} values above the 95th percentile (Table 1). This genomic divergence landscape contrasts with that obtained when comparing the Atlantic and Norwegian genetic lineages of *P. maximus* (Figure 2b). Here, the mean, maximum, and 95th percentile values of F_{ST} were considerably lower (mean $F_{ST} = 0.03$, maximum $F_{ST} = 0.65$, 95th percentile = 0.11) and we could retrieve only a handful of peaks of high divergence, which were localized within comparably short genomic regions.

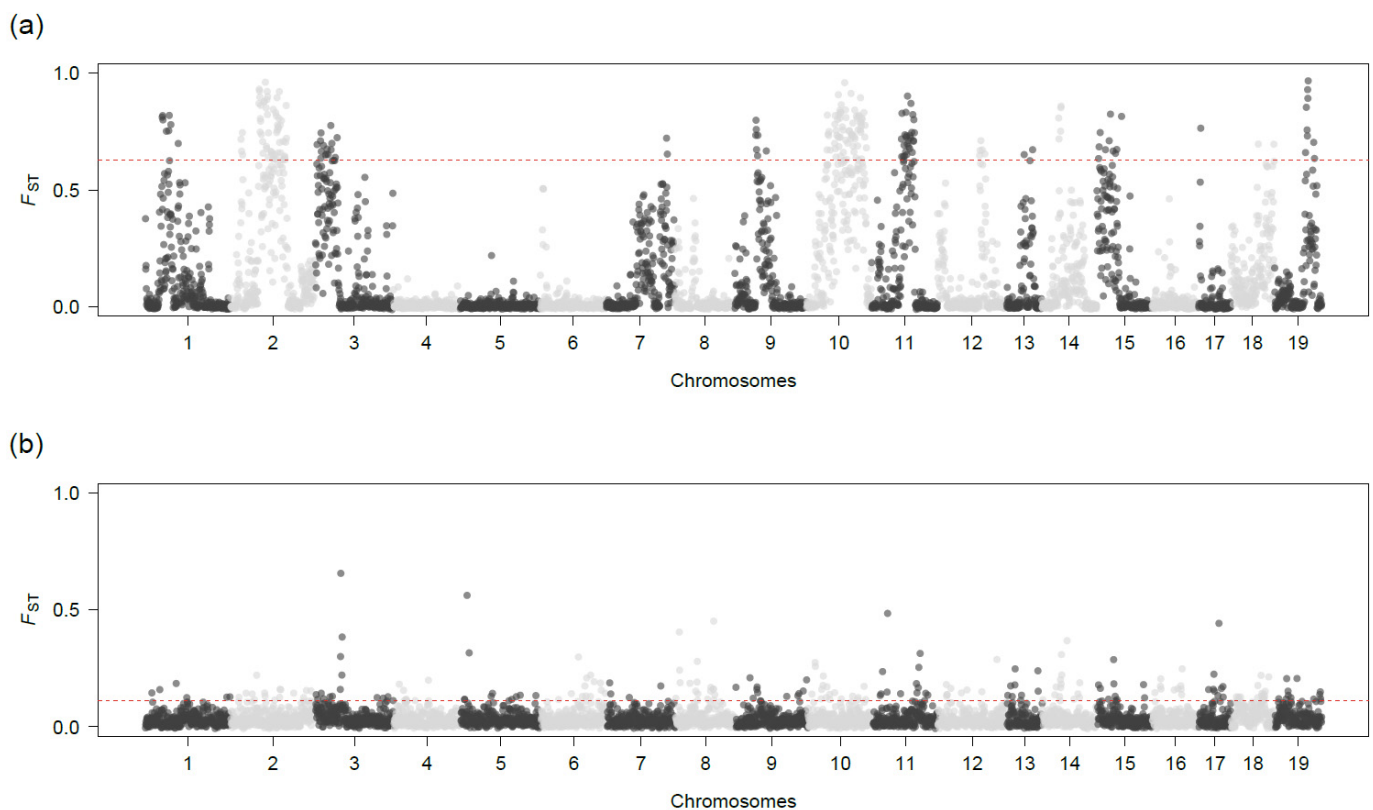


Figure 2. F_{ST} values calculated between samples within 4460 100 kb sliding windows. Panel (a) shows the comparison between *P. maximus* and *P. jacobaeus*, while panel (b) shows the comparison between Atlantic *P. maximus* and Norwegian *P. maximus*. Each point represents the F_{ST} value of a given window located on a given chromosome, as shown on the x axis. Light and dark grey blocks correspond to the 19 chromosomes. The red dashed lines represent thresholds corresponding to the 95th percentile F_{ST} values.

Table 1. Number of sliding windows with an F_{ST} value above the 95th percentile ('highly divergent') by chromosome. The total number of sliding windows drawn within each chromosome (N) and the percentage of highly divergent windows ('% highly divergent') are shown together with the mean F_{ST} value and mean d_{xy} value. The number of diagnostic single nucleotide polymorphisms (SNPs) within each chromosome is also reported.

Chromosome	N	Highly Divergent	% Highly Divergent	Mean F_{ST}	Mean d_{xy}	Diagnostic SNPs
1	323	8	2.48	0.107	0.0045	0
2	321	59	18.38	0.259	0.0058	4
3	296	15	5.07	0.164	0.0051	0
4	253	0	0	0.007	0.0043	0
5	303	0	0	0.006	0.0040	0
6	250	0	0	0.017	0.0046	0
7	259	2	0.77	0.121	0.0046	0
8	228	0	0	0.031	0.0045	0
9	272	7	2.57	0.1	0.0045	0
10	246	70	28.46	0.398	0.0070	2
11	250	28	11.2	0.164	0.0054	54
12	262	5	1.91	0.079	0.0046	0
13	133	2	1.5	0.087	0.0046	0
14	209	5	2.39	0.108	0.0048	1
15	209	10	4.78	0.168	0.0058	0
16	176	0	0	0.02	0.0049	0
17	122	1	0.82	0.034	0.0047	0
18	166	2	1.21	0.169	0.0057	1
19	179	9	5.03	0.127	0.0057	0
Unplaced scaffolds	3	0	0	0.016	0.0049	0
Total	4460	223	5.0	0.118	0.0050	62

3.3. F_{ST} , Nucleotide Diversity, and d_{xy} Values

We calculated nucleotide diversity separately for the two scallop species within the same sliding windows used to calculate F_{ST} values. Genome-wide nucleotide diversity (π) was 1.8×10^{-5} in *P. maximus* and 1.69×10^{-5} in *P. jacobaeus* (per-chromosome measures are available in Table S5). The two species showed fairly similar distributions of nucleotide diversity across loci, with *P. jacobaeus* showing a larger number of windows with lower π (Figure 3 and Figure S3). F_{ST} and π were significantly negatively correlated in both species (*P. maximus*: $\rho = -0.06$, $p < 0.01$; *P. jacobaeus*: $\rho = -0.33$, $p < 0.01$; Figure 3). This indicates that the most divergent genomic regions tended to be located within genomic regions characterized by lower genetic diversity. Nevertheless, high F_{ST} values do not appear to be a by-product of locally reduced genetic diversity, as d_{xy} values calculated for *P. maximus* and *P. jacobaeus* within the same genomic windows showed a comparable pattern. First, chromosomes showing the highest average values of F_{ST} were also characterized by the highest average d_{xy} values (Table 1). Second, F_{ST} values were overall significantly positively correlated with d_{xy} values ($\rho = 0.29$, $p < 0.01$). Third, highly divergent genomic windows, when compared to the remaining windows, were characterized by higher levels of divergence as quantified by both F_{ST} and d_{xy} (Figure S4). This suggests that the gene flow at highly divergent genomic windows was substantially reduced in comparison to the remaining genomic regions.

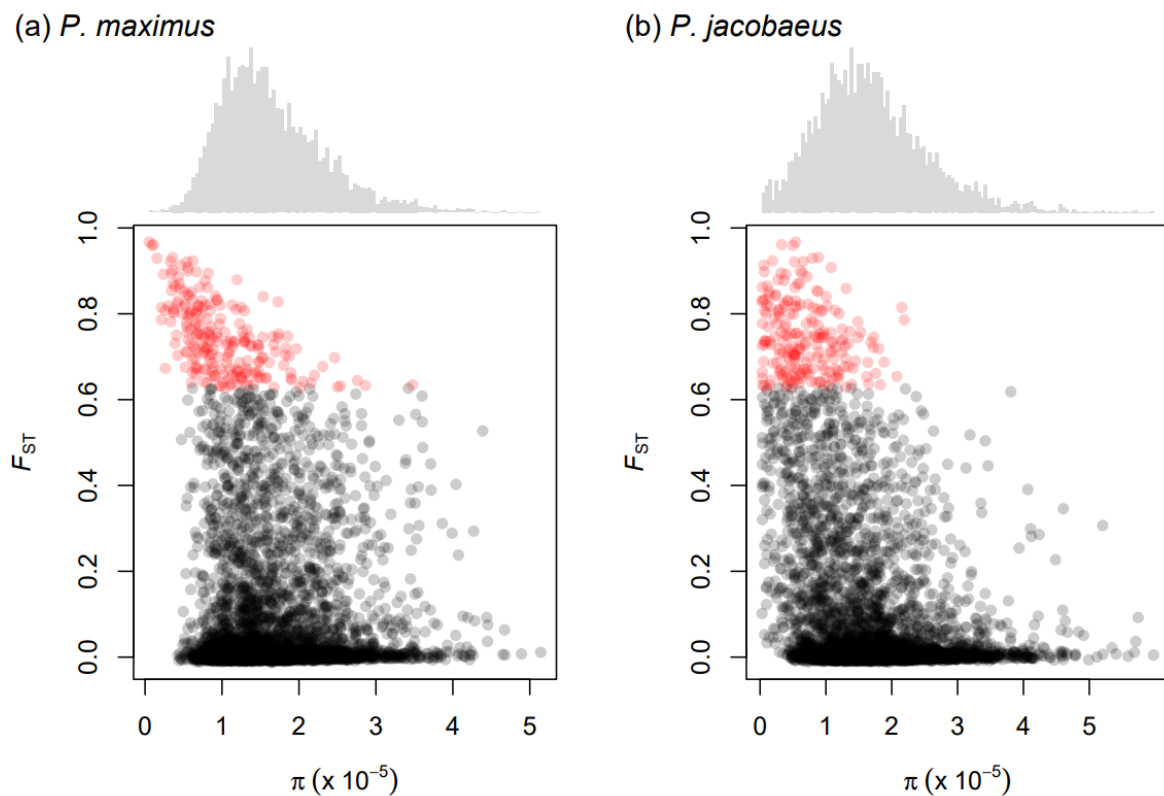


Figure 3. Relationship between nucleotide diversity (π) and F_{ST} values calculated in 100 kb sliding windows. Panels (a,b) shows results for *P. maximus* and *P. jacobaeus*, respectively. The histograms above the scatter plots show the distribution of nucleotide diversity calculated within sliding windows in the two scallop species. In the scatter plots, each point represents a sliding window. Red points indicate sliding windows with F_{ST} values equal to or greater than the 95th percentile.

3.4. Diagnostic Loci

Comparison of the genotypes from the two scallop species uncovered a total of 62 (0.13%) fully diagnostic SNPs (Table S6). Notably, 54 (87%) of them were located in a single genomic region spanning approximately 10 Mb of chromosome 11 (Table 1). The other diagnostic SNPs were located on chromosomes 2 (four SNPs), 10 (two SNPs), 14 (one SNP), and 18 (one SNP, Table 1). A fasta file containing 100 bp of flanking sequence on either side of each fully diagnostic SNP is available at: https://figshare.com/articles/online_resource/P_maximus_vs_P_jacobaeus_fully_diagnostic_loci_flanking_sequences/21311916 (last accessed on 20 December 2022), while the genotypes characteristic of each species can be found in Table S6.

3.5. Linkage Disequilibrium

The pattern of LD decay was quantified separately for the two scallop species using only those SNPs that mapped to the 19 *P. maximus* chromosomes. SNPs located within unplaced scaffolds were excluded from this analysis. In *P. maximus*, LD was found to decay rapidly, with r^2 reaching the background level ($r^2 = 0.01$) by around 2 kb (Figure S5a). This pattern did not change notably when restricting the analysis to SNPs located in highly divergent genomic regions, although the background level of LD was higher (Figure S5c). LD appeared to decay rather rapidly also in *P. jacobaeus*, although it showed a slightly higher background level ($r^2 = 0.015$) that was reached more slowly, at around 3 kb (Figure S5b). In contrast, when analysing only SNPs located within highly divergent genomic regions in *P. jacobaeus*, r^2 did not reach the background level until over ~400 kb (Figure S5d). This suggests that differences in chromosome structure may be present in this species in the proximity of highly differentiated regions.

Visualization of LD values calculated among all pairs of loci separately for all 19 chromosomes revealed that highly divergent genomic regions were located in areas of elevated LD, suggesting low local recombination (Figure 4 and Figure S6). Moreover, in *P. jacobaeus*, regions of low recombination were often accompanied by reduced genetic diversity (in the absence of polymorphism at a given SNP, PLINK returns a NA value for LD which is plotted in white in Figure 4 and Figure S6). Notably, chromosomes lacking highly divergent genomic windows were not characterized by any blocks of high LD (Figure S6). Finally, chromosomes 2 (terminal part, Figure 4), 17, and 19 (Figure S6) contained genomic regions that showed elevated LD only in *P. maximus*.

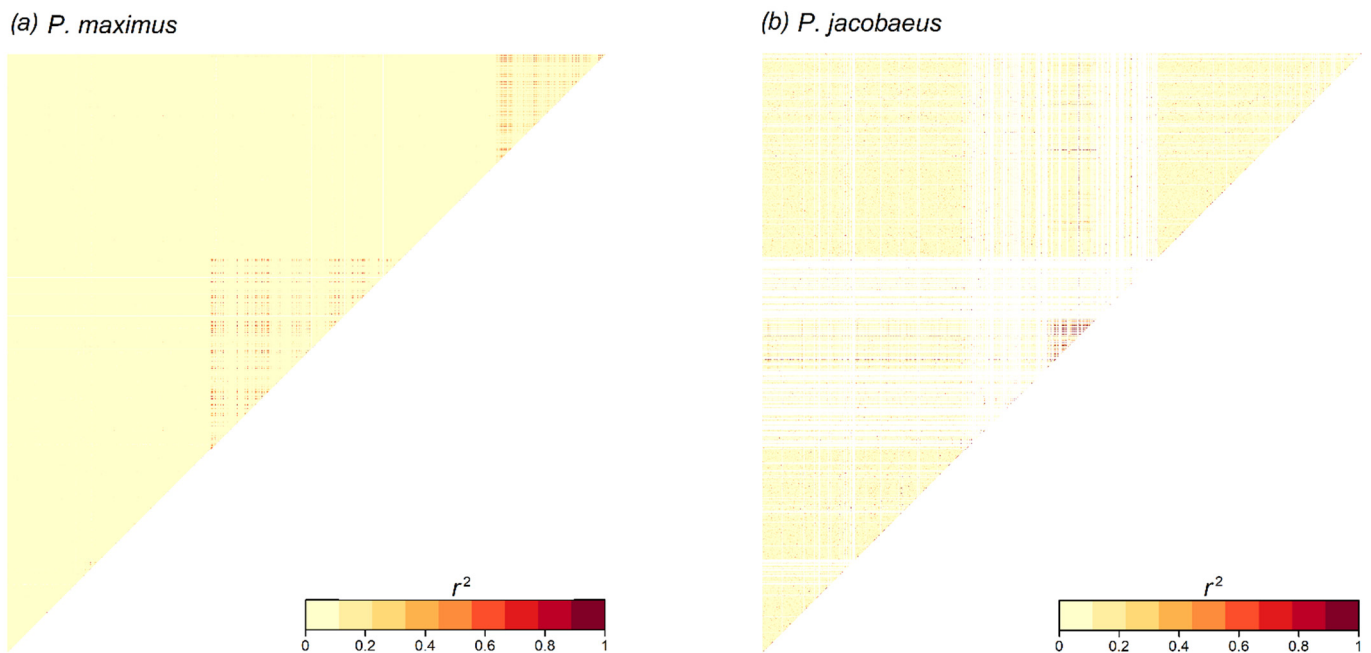


Figure 4. Heatmaps showing patterns of linkage disequilibrium (LD, r^2) along chromosome 2, with darker areas corresponding to genomic regions characterized by elevated values of LD, as shown in the colour key. White lines refer to single nucleotide polymorphisms (SNPs) that were not polymorphic in a given species which resulted in a NA value when used to calculate LD. Panels (a,b) refer to *P. maximus* and *P. jacobaeus*, respectively. Red triangles highlight genomic regions characterized by higher levels of LD.

3.6. Gene Annotations and Enrichment Analysis

The 223 sliding windows whose F_{ST} values were equal to or greater than the 95th percentile contained a total of 1777 SNPs, which included all of the fully diagnostic SNPs. These were found within or in close proximity to 438 different genes, for which we retrieved GO annotation terms. A GO enrichment analysis carried out against a background list of GO terms containing all of the GO-annotated genes of the great scallop revealed a total of seven significantly enriched biological process GO terms (Table 2 and Table S7). Repeating the analysis considering only highly divergent SNPs located within genes yielded very similar significantly enriched GO categories (Table S8).

Table 2. List of significantly enriched gene ontology (GO) terms relative to genes containing or located in the proximity of highly divergent SNPs.

GO Term ID	Description	p-Value
GO:0002755	MyD88-dependent toll-like receptor signalling pathway	<0.01
GO:0009058	biosynthetic process	<0.01
GO:0034472	snRNA 3'-end processing	<0.05
GO:0032456	endocytic recycling	<0.05
GO:0036265	RNA (guanine-N7)-methylation	<0.05
GO:0006777	Mo-molybdopterin cofactor biosynthetic process	<0.05
GO:0006465	signal peptide processing	<0.05

The snpEff analysis revealed that the majority of the 1777 SNPs located within highly divergent genomic regions were located in non-coding regions, with 61% being situated outside of genic regions, 30% being located within introns, and 2.3% in UTRs (Table 3). Equivalent results were obtained when considering the diagnostic SNPs alone, with 63%, 29%, and 6% of the SNPs being located outside genic regions, in introns, and in UTRs, respectively (Table 3). Of the 107 highly divergent SNPs located within exons, 67, 40, and 2 represented synonymous, missense, and loss of function mutations, respectively (Table 3). Only one diagnostic SNP was located within a gene, and this was a synonymous mutation (Table 3). A full report of the snpEff annotations is given in Table S9.

Table 3. Details of the 1777 highly divergent SNPs, including the 62 fully diagnostic SNPs. For each type of SNP, we indicate whether the SNP is located within a gene and whether it is transcribed. We also indicate how many highly divergent SNPs and fully diagnostic SNPs are attributed to each mutation category. Proportions are reported in parenthesis.

SNP Type	Within Gene	Transcribed	Highly Divergent	Diagnostic
intergenic region	N	N	724 (0.41)	24 (0.39)
downstream gene variant	N	N	166 (0.09)	10 (0.16)
upstream gene variant	N	N	196 (0.11)	5 (0.08)
3' UTR variant	Y	N	34 (0.02)	4 (0.06)
5' UTR variant	Y	N	6 (0.003)	0 (0)
intron variant	Y	N	531 (0.3)	18 (0.29)
splice region variant and intron variant	Y	N	11 (0.01)	0 (0)
synonymous variant	Y	Y	67 (0.04)	1 (0.02)
missense variant	Y	Y	40 (0.02)	0 (0)
loss of function	Y	Y	2 (0.001)	0 (0)

4. Discussion

The extent of genetic divergence between *P. maximus* and *P. jacobaeus* has been previously assessed, but most studies have used a limited number of loci (e.g., [13,20,21]). Even for the single study based on RADseq data [22], it was not possible to locate highly divergent regions due to the absence at the time of a reference genome. Here, we reanalysed RADseq data from Vendrami et al. [22] and mapped them to a new chromosome-level reference genome [31]. We demonstrate that genome-wide divergence between the two scallop species is highly heterogeneous, being characterized by peaks of divergence scattered throughout the genome, a pattern which contrasts sharply with that obtained when comparing Atlantic and Norwegian lineages within *P. maximus*, which are known to have diverged much more recently [22]. However, five of the 19 chromosomes showed little to no divergence. By implementing demographic and comparative genomic analyses, we were able to demonstrate that this heterogeneous pattern probably resulted from an initial prolonged period of isolation after divergence followed by secondary contact where

genomic regions not involved in reproductive isolation have likely been homogenized by gene flow.

Our demographic analysis suggests that *P. maximus* and *P. jacobaeus* diverged around 300,000 generations ago and remained isolated until approximately 15,000 generations ago when secondary contact occurred. Gene flow then appears to have continued until around 60 generations before the present. Given a generation time of two to five years in *P. maximus* [42] (<https://www.marlin.ac.uk/species/detail/1398>, last accessed on 8 December 2022) this would imply that the two lineages split between around 0.6 and 1.5 million years ago (MYA) and that secondary contact started approximately between 30,000 and 75,000 years ago and lasted until a few centuries before the present. This is consistent with previous studies that estimated the split between the two *Pecten* species to have occurred in the Pleistocene during a period of fluctuating climatic conditions [13,22,46] and suggests that secondary contact might have occurred during the latter part of the last glacial period.

Similar models of allopatric divergence followed by secondary contact have been described for other Atlantic–Mediterranean sister lineages, including mussels [47] and sea bass [48]. Furthermore, both of these studies reported divergence to have occurred during the second half of the Pleistocene and secondary contact to have taken place around the end of the last glacial period. Additionally, a model of allopatric divergence followed by secondary contact yielding comparable estimates for the timings of both divergence and secondary contact, was also reported when studying the divergence between coastal and marine ecotypes of European anchovy populations [49]. Together, these studies and the present work suggest that shifts in range distributions during the Pleistocene had an important role in promoting evolutionary diversification in marine species with Atlantic and Mediterranean distributions. Moreover, the end of the last glacial period appears to have corresponded with geological and oceanographical changes within this geographic area, which provided the opportunity for marine lineages that diverged during an initial phase of allopatry to re-establish gene flow.

We recognize that our demographic parameter estimates are somewhat different from those obtained by Vendrami et al. [22], even though they are based on the same RADseq dataset. This is not necessarily surprising, as the current analyses differ from those of Vendrami et al. [22] in several ways. For example, our SFS was produced from a SNP dataset obtained from a reference-based genotyping approach, and not from a de novo based approach. Moreover, easySFS can incorporate SNPs with missing genotypes when computing the SFS via the down projection procedure described by [50] and is therefore likely to be more accurate. In addition, the demographic scenarios evaluated by Vendrami et al. [22] did not include gene flow, so the divergence time between the two *Pecten* species may have been underestimated. Consequently, we believe that our current demographic analyses represent an improvement over those of Vendrami et al. [22].

Based on our demographic inference, it appears that *P. maximus* and *P. jacobaeus* went through a period of isolation that lasted for around 285,000 generations, corresponding to around 0.5–1.5 million years. This indicates that divergence between the two species accumulated over a period of time ~20 times longer than the duration of secondary contact. Given that ecological parameters such as temperature, salinity, and water density clearly differ across the AOF [51], it is likely that strong divergent selection for different alleles between the two species, as well as positive selection for new beneficial mutations, may have been the primary drivers of divergence. The observation that genomic regions showing the highest F_{ST} values were also characterized by lower levels of genetic variability is consistent with this, and is suggestive of the occurrence of selective sweeps [4,5].

Genetic incompatibilities may also have increased during the prolonged period of isolation. The fact that we obtained consistent patterns of divergence using both relative (F_{ST}) and absolute (d_{xy}) measures of genomic divergence provides evidence that differential gene flow occurred after secondary contact and supports the presence of reproductive barriers within the *Pecten* genome. This implies that the observed highly divergent genomic regions are not simply a by-product of locally reduced genetic diversity [11], but may also

be the consequence of reduced gene flow, a feature that has also been uncovered in the sea bass genome when comparing Atlantic and Mediterranean lineages [48,52]. Taken together, our results suggest that the highly differentiated genomic regions between the two *Pecten* species likely accumulated important adaptive divergence and were reproductively isolated. As a consequence, gene flow during secondary contact could not homogenize allele frequencies at these regions, resulting in a mosaic of highly and lowly differentiated genomic regions.

Loci under divergent selection are known to accumulate in regions of low recombination [53] such as centromeres [54] and the sites of chromosomal rearrangements [7,55]. Among them, chromosomal inversions have been linked to adaptation in several marine organisms, such as the Atlantic cod (*Gadus morhua*; [55,56]), three spined stickleback (*Gasterosteus aculeatus*, [7]), rough periwinkle (*Littorina saxatilis* [57]), and, notably, the great scallop (*P. maximus* [29]). In the latter study, which used a different reference genome (that of Kenny et al. [30]) and sequencing approach (ddRADseq) than our study, the authors identified inversions on three chromosomes that were associated with adaptive divergence among *P. maximus* populations sampled from different latitudes within Europe. These results may help to explain why we observed three genomic regions characterized by low levels of recombination exclusively in *P. maximus* and not in *P. jacobaeus*. Specifically, it is possible that these genomic regions, located on chromosomes 2, 17, and 19, may be of adaptive importance only in *P. maximus*.

Here we focused on the comparison between *P. maximus* and *P. jacobaeus* and tested whether SNPs located in highly divergent regions exhibited higher levels of LD. We found that, whilst overall LD decays rapidly in both species, when only divergent SNPs are considered in *P. jacobaeus*, LD extends beyond 400 kbp, indicating the likely presence of structural variants, such as chromosomal inversions [58]. Consistent with this hypothesis, all of the highly divergent regions were located in genomic areas showing high levels of LD and therefore low recombination. Taken together, the results of our study and that of Hollenbeck et al. [29] suggest that chromosomal inversions, or more generally genomic areas characterized by low levels of recombination, are important in European *Pecten* to maintain adaptive variation at different scales, both within different *P. maximus* lineages and also between *Pecten* species, where they act as intrinsic barriers to gene flow and maintain locally adapted variation.

In support of the argument for locally adapted variation, we found that highly divergent genomic regions between the two species appeared to be enriched for several biological processes. Furthermore, the genes enriched for GO-terms were distributed throughout the genome, discounting the possibility of this signal being produced by tandem repeats of the same gene. Among them, we found the “MyD88-dependent toll-like receptor signalling pathway” which is involved in the innate immune response of nonmammal species [59], including the scallop *Chlamys farreri* [60]. The evolution of the immune system is known to have influenced speciation in teleost fishes [61], while parasitism has also been proposed to play a role in ecological speciation [62]. It is therefore possible that differences in the immunological challenges encountered within the contrasting environments where *P. maximus* and *P. jacobaeus* evolved might have shaped patterns of genetic divergence between these two species. Furthermore, toll-like receptors also participate in development [63,64] and have been demonstrated to play a role in larval development in *Drosophila* [65]. If this is also the case for *Pecten*, divergence at genes associated with this pathway could potentially promote reproductive isolation, as larvae of one species dispersing within the distribution range of the other, may not be adapted to survive and develop under contrasting environmental conditions.

Genes involved in RNA modification also appeared to be over-represented in our enrichment analysis (GO terms: “snRNA 3'-end processing” and “RNA (guanine-N7)-methylation”). In recent years, mRNA modifications have been implicated as regulatory mechanisms that control gene expression [66], which is known to be important for adaptation to novel environments [67]. Additionally, enrichment for the term “Mo-molybdopterin

cofactor biosynthetic process”, which relates to the biosynthesis of the molybdenum cofactor, the active part of all molybdenum-containing enzymes [68], suggests that differences in the availability of molybdenum within the geographical ranges of the two scallop species, if present, might also have played a role in the speciation process.

Similar enriched GO terms were obtained when repeating the GO enrichment analysis considering only SNPs located within genes, corroborating our conclusions. Whilst highly divergent loci are located in low-recombination genomic regions, such that these SNPs could potentially be associated with any gene in the region where recombination is restricted, the fact that the majority of SNPs are outside coding regions suggests that most of the identified loci are likely to have regulatory roles rather than producing changes at the protein level. Similar results were obtained by Pujolar et al. [69] when studying patterns of divergence between the American eel (*Anguilla rostrata*) and the European eel (*A. anguilla*), and by Jones et al. [7], who investigated genomic divergence between marine and freshwater threespine sticklebacks. These authors found that 95.7% and 83%, respectively of the observed divergence was present in noncoding regions. Furthermore, divergence at regulatory elements plays a key role in speciation by promoting hybrid dysfunction via negative interactions between alleles [70]. Regulatory divergence between species is in fact widespread and has also been documented in mice, birds, flies, yeast, and plants [71–75].

Our demographic analyses further suggest that there currently is no gene flow occurring between *P. maximus* and *P. jacobaeus*. Consistent with this, we are not aware of any studies that have identified hybrid individuals based upon genetic data and this is also true of the current study [22], although none of our sampled populations were geographically close to the AOF. Whilst *P. jacobaeus* and *P. maximus* hybrids can be produced artificially [76] and measures of fertility (e.g., the gonadotrophic index) suggest that both species are reproductively mature at a similar time [77,78], opportunities for cross-species fertilization are likely to be limited. This is because spawning in scallops is triggered by temperature [79] and waters on the two sides of the AOF differ by 1–2 °C [14]. Moreover, among the genes within highly divergent genomic regions, we identified serine and threonine kinases, which are upregulated in scallop gonads [80,81]. This hints at the possible presence of differences in the reproductive cycles of these two species. Together, these arguments support the outcome of our demographic inference, as they suggest that gene flow is unlikely to occur at the present moment.

5. Conclusions

RADseq data for the commercially exploited European scallop species *P. maximus* and *P. jacobaeus* revealed a highly heterogeneous genomic divergence landscape between the two species. Our analyses indicate that this pattern likely originated as the consequence of a long period of isolation after divergence, during which the two lineages accumulated both adaptive divergence and reproductive barriers, followed by secondary contact, during which only those genomic regions permeable to gene flow were homogenized. Highly divergent genomic regions were associated with reduced levels of genetic diversity, which is indicative of the occurrence of selective sweeps and is consistent with the notion that positive selection imposed by contrasting environmental conditions played an important role in determining the pattern of genomic divergence observed today. Moreover, our analyses indicate that highly divergent genomic regions are located in areas of low recombination, hinting at the possible presence of chromosomal inversions, which are known to be important for maintaining adaptive genetic variation and promoting reproductive isolation. In line with this, highly divergent genomic regions also showed elevated d_{xy} values indicating that differential gene flow across the genome occurred during secondary contact. Finally, the genes located within highly differentiated regions are related to immunity, development, mRNA modification, and the biosynthesis of molybdenum cofactors, with the majority of SNPs likely affecting regulatory elements. Taken together, these results provide new insights into the divergence history of these two scallop species, while

also identifying specific genomic regions that potentially contribute to the maintenance of adaptively important variation.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes14010014/s1>, Figure S1: Schematic representation of the alternative demographic models; Figure S2: F_{ST} values calculated using alternative SNP datasets; Figure S3: Nucleotide diversity calculated within 100 kb windows separately for *P. maximus* and *P. jacobaeus*; Figure S4: Differences in F_{ST} and d_{xy} between highly divergent genomic windows and the remaining genomic windows; Figure S5: Linkage disequilibrium decay when considering all SNPs and only SNPs located within highly divergent genomic regions separately for *P. maximus* and *P. jacobaeus*; Figure S6: linkage disequilibrium heatmaps for all chromosomes separately for *P. maximus* and *P. jacobaeus*; Table S1: Sampling location information; Table S2: Priors for all estimated demographic parameters; Table S3: Summary statistics for the demographic models; Table S4: Estimated demographic parameter values together with their 95% confidence intervals estimated for the best supported model; Table S5: Per-chromosome measures of nucleotide diversity calculated separately for *P. maximus* and *P. jacobaeus*; Table S6: List of fully diagnostic loci; Table S7: Full results of the gene ontology enrichment analysis; Table S8: Gene ontology enrichment analysis based only on SNPs located within genes; Table S9: Full annotation report of all SNPs located within highly divergent genomic regions, including the snpEff results.

Author Contributions: Conceptualization, D.L.J.V. and C.S.W.; methodology, D.L.J.V. and J.I.H.; formal analysis, D.L.J.V.; investigation, D.L.J.V. and C.S.W.; data curation, D.L.J.V.; writing—original draft preparation, D.L.J.V. and C.S.W.; writing—review and editing, D.L.J.V., C.S.W. and J.I.H.; supervision, J.I.H.; funding acquisition, J.I.H. All authors have read and agreed to the published version of the manuscript.

Funding: The research leading to these results received funding from the European Union Marie Curie Seventh Framework Programme under Grant Agreement No. 605051 for the CACHE ITN. D.L.J.V. was also supported by supported by the German Research Foundation (DFG) as part of the SFB TRR 212 (NC³) (Project Numbers 316099922 and 396774617).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The sequencing data analysed here are available (SRA, accession number PRJNA526674) and were originally described by Vendrami et al. [22]. The code for this study is available at: https://github.com/DavidVendrami/Pmaxiums_jacobaeus_genomic_divergence (last accessed on 20 December 2022).

Acknowledgments: We would like to thank the two anonymous reviewers for their constructive feedback. We acknowledge support for the publication costs by the Open Access Publication Fund of Bielefeld University and the Deutsche Forschungsgemeinschaft (DFG).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Roux, C.; Fraïsse, C.; Romiguier, J.; Anciaux, Y.; Galtier, N.; Bierne, N. Shedding Light on the Grey Zone of Speciation along a Continuum of Genomic Divergence. *PLoS Biol.* **2016**, *14*, e2000234. [[CrossRef](#)]
2. Seehausen, O.; Butlin, R.K.; Keller, I.; Wagner, C.E.; Boughman, J.W.; Hohenlohe, P.A.; Peichel, C.L.; Saetre, G.P.; Bank, C.; Brännström, Å.; et al. Genomics and the Origin of Species. *Nat. Rev. Genet.* **2014**, *15*, 176–192. [[CrossRef](#)] [[PubMed](#)]
3. Rundle, H.D.; Nasil, P. Ecological Speciation. *Ecol. Lett.* **2005**, *8*, 336–352. [[CrossRef](#)]
4. Via, S.; West, J. The Genetic Mosaic Suggests a New Role for Hitchhiking in Ecological Speciation. *Mol. Ecol.* **2008**, *17*, 4334–4345. [[CrossRef](#)]
5. Nielsen, R.; Williamson, S.; Kim, Y.; Hubisz, M.J.; Clark, A.G.; Bustamante, C. Genomic Scans for Selective Sweeps Using SNP Data. *Genome Res.* **2005**, *15*, 1566–1575. [[CrossRef](#)] [[PubMed](#)]
6. Noor, M.A.F.; Gratos, K.L.; Bertucci, L.A.; Reiland, J. Chromosomal Inversions and the Reproductive Isolation of Species. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12084–12088. [[CrossRef](#)] [[PubMed](#)]
7. Jones, F.C.; Grabherr, M.G.; Chan, Y.F.; Russell, P.; Mauceli, E.; Johnson, J.; Swofford, R.; Pirun, M.; Zody, M.C.; White, S.; et al. The Genomic Basis of Adaptive Evolution in Threespine Sticklebacks. *Nature* **2012**, *484*, 55–61. [[CrossRef](#)]

8. Ravinet, M.; Faria, R.; Butlin, R.K.; Galindo, J.; Bierne, N.; Rafajlović, M.; Noor, M.A.F.; Mehlig, B.; Westram, A.M. Interpreting the Genomic Landscape of Speciation: A Road Map for Finding Barriers to Gene Flow. *J. Evol. Biol.* **2017**, *30*, 1450–1477. [[CrossRef](#)]
9. Nosil, P.; Vines, T.H.; Funk, D.J. Perspective: Reproductive Isolation Caused by Natural Selection against Immigrants from Divergent Habitats. *Evolution* **2005**, *59*, 705–719. [[CrossRef](#)]
10. Wright, S. Evolution in Mendelian Populations. *Genetics* **1931**, *16*, 97. [[CrossRef](#)] [[PubMed](#)]
11. Cruickshank, T.E.; Hahn, M.W. Reanalysis Suggests That Genomic Islands of Speciation Are Due to Reduced Diversity, Not Reduced Gene Flow. *Mol. Ecol.* **2014**, *23*, 3133–3157. [[CrossRef](#)] [[PubMed](#)]
12. Nei, M.; Li, W.H. Mathematical Model for Studying Genetic Variation in Terms of Restriction Endonucleases. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 5269–5273. [[CrossRef](#)]
13. Ríos, C.; Sanz, S.; Saavedra, C.; Peña, J.B. Allozyme Variation in Populations of Scallops, *Pecten jacobaeus* (L.) and *P. maximus* (L.) (Bivalvia: Pectinidae), across the Almeria–Oran Front. *J. Exp. Mar. Biol. Ecol.* **2002**, *267*, 223–244. [[CrossRef](#)]
14. Patarnello, T.; Volckaert, F.A.M.J.; Castilho, R. Pillars of Hercules: Is the Atlantic-Mediterranean Transition a Phylogeographical Break? *Mol. Ecol.* **2007**, *16*, 4426–4444. [[CrossRef](#)]
15. Duran, S.; Palacín, C.; Becerro, M.A.; Turon, X.; Giribet, G. Genetic Diversity and Population Structure of the Commercially Harvested Sea Urchin Paracentrotus Lividus (Echinodermata, Echinoidea). *Mol. Ecol.* **2004**, *13*, 3317–3328. [[CrossRef](#)]
16. Marie, A.D.; Lejeune, C.; Karapatsiou, E.; Cuesta, J.A.; Drake, P.; Macpherson, E.; Bernatchez, L.; Rico, C. Implications for Management and Conservation of the Population Genetic Structure of the Wedge Clam *Donax Trunculus* across Two Biogeographic Boundaries. *Sci. Rep.* **2016**, *6*, 39152. [[CrossRef](#)]
17. El Ayari, T.; Trigui El Menif, N.; Hamer, B.; Cahill, A.E.; Bierne, N. The Hidden Side of a Major Marine Biogeographic Boundary: A Wide Mosaic Hybrid Zone at the Atlantic–Mediterranean Divide Reveals the Complex Interaction between Natural and Genetic Barriers in Mussels. *Heredity* **2019**, *122*, 770–784. [[CrossRef](#)]
18. El Ayari, T.; El Menif, N.T.; Saavedra, C.; Cordero, D.; Viard, F.; Bierne, N. Unexpected Mosaic Distribution of Two Hybridizing Sibling Lineages in the Teleplanically Dispersing Snail *Stramonita Haemastoma* Suggests Unusual Postglacial Redistribution or Cryptic Invasion. *Ecol. Evol.* **2017**, *7*, 9016–9026. [[CrossRef](#)]
19. Naciri, M.; Lemaire, C.; Borsa, P.; Bonhomme, F. Genetic Study of the Atlantic/Mediterranean Transition in Sea Bass (*Dicentrarchus labrax*). *J. Hered.* **1999**, *90*, 591–596. [[CrossRef](#)]
20. Wilding, C.S.; Beaumont, A.R.; Latchford, J.W. Are *Pecten Maximus* and *Pecten jacobaeus* Different Species? *J. Mar. Biol. Assoc. U. K.* **1999**, *79*, 949–952. [[CrossRef](#)]
21. Morvezen, R.; Charrier, G.; Boudry, P.; Chauvaud, L.; Breton, F.; Strand, Ø.; Laroche, J. Genetic Structure of a Commercially Exploited Bivalve, the Great Scallop *Pecten Maximus*, along the European Coasts. *Conserv. Genet.* **2016**, *17*, 57–67. [[CrossRef](#)]
22. Vendrami, D.L.J.; De Noia, M.; Telesca, L.; Handal, W.; Charrier, G.; Boudry, P.; Eberhart-Phillips, L.; Hoffman, J.I. RAD Sequencing Sheds New Light on the Genetic Structure and Local Adaptation of European Scallops and Resolves Their Demographic Histories. *Sci. Rep.* **2019**, *9*, 7455. [[CrossRef](#)] [[PubMed](#)]
23. Pérès, J.M. History of the Mediterranean Biota and the Colonization of the Depths. In *Key Environments: Western Mediterranean*; Pergamon Press: Oxford, UK, 1985; pp. 198–232.
24. Raffi, S.; Stanley, S.M.; Marasti, R. Biogeographic Patterns and Plio-Pleistocene Extinction of Bivalvia in the Mediterranean and Southern North Sea. *Paleobiology* **1985**, *11*, 368–388. [[CrossRef](#)]
25. Bateson, W. Heredity and Variation in Modern Lights. In *Darwin and Modern Science*; BoD—Books on Demand: Norderstedt, Germany, 1909.
26. Dobzhansky, T.H. Studies on Hybrid Sterility. II. Localization of Sterility Factors in *Drosophila Pseudoobscura* Hybrids. *Genetics* **1936**, *21*, 113. [[CrossRef](#)]
27. Muller, H. Isolating Mechanisms, Evolution and Temperature. *Biol. Symp.* **1942**, *6*, 71–125.
28. Cutter, A.D. The Polymorphic Prelude to Bateson-Dobzhansky-Muller Incompatibilities. *Trends Ecol. Evol.* **2012**, *27*, 209–218. [[CrossRef](#)]
29. Hollenbeck, C.M.; Portnoy, D.S.; Garcia, D.; Magnesen, T.; Matejusova, I.; Johnston, I.A.; Hollenbeck, C.M. Temperature-Associated Selection Linked to Putative Chromosomal Inversions in King Scallop (*Pecten maximus*). *Proc. R. Soc. B* **2022**, *289*, 20221573. [[CrossRef](#)]
30. Kenny, N.J.; McCarthy, S.A.; Dudchenko, O.; James, K.; Betteridge, E.; Corton, C.; Dolucan, J.; Mead, D.; Oliver, K.; Omer, A.D.; et al. The Gene-Rich Genome of the Scallop *Pecten Maximus*. *Gigascience* **2020**, *9*, 1–13. [[CrossRef](#)]
31. Zeng, Q.; Liu, J.; Wang, C.; Wang, H.; Zhang, L.; Hu, J.; Bao, L.; Wang, S. High-Quality Reannotation of the King Scallop Genome Reveals No ‘Gene-Rich’ Feature and Evolution of Toxin Resistance. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 4954–4960. [[CrossRef](#)]
32. Sambrook, J.; Fritsch, E.F.; Maniatis, T. *Molecular Cloning: A Laboratory Manual (No. Ed. 2)*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 1989.
33. Baird, N.A.; Etter, P.D.; Atwood, T.S.; Currey, M.C.; Shiver, A.L.; Lewis, Z.A.; Selker, E.U.; Cresko, W.A.; Johnson, E.A. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE* **2008**, *3*, e3376. [[CrossRef](#)]
34. Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
35. Catchen, J.; Hohenlohe, P.A.; Bassham, S.; Amores, A.; Cresko, W.A. Stacks: An Analysis Tool Set for Population Genomics. *Mol. Ecol.* **2013**, *22*, 3124–3140. [[CrossRef](#)]

36. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The Variant Call Format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [[CrossRef](#)]
37. Excoffier, L.; Marchi, N.; Marques, D.A.; Matthey-Doret, R.; Gouy, A.; Sousa, V.C. Fastsimcoal2: Demographic Inference under Complex Evolutionary Scenarios. *Bioinformatics* **2021**, *37*, 4882–4885. [[CrossRef](#)]
38. Varin, C.; Vidoni, P. A Note on Composite Likelihood Inference and Model Selection. *Biometrika* **2005**, *92*, 519–528. [[CrossRef](#)]
39. Hohenlohe, P.A.; Bassham, S.; Etter, P.D.; Stiffler, N.; Johnson, E.A.; Cresko, W.A. Population Genomics of Parallel Adaptation in Threespine Stickleback Using Sequenced RAD Tags. *PLoS Genet.* **2010**, *6*. [[CrossRef](#)]
40. Korunes, K.L.; Samuk, K. Pixy: Unbiased Estimation of Nucleotide Diversity and Divergence in the Presence of Missing Data. *Mol. Ecol. Resour.* **2021**, *21*, 1359–1368. [[CrossRef](#)]
41. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; De Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)]
42. Hill, W.G.; Weir, B.S. Variances and Covariances of Squared Linkage Disequilibria in Finite Populations. *Theor. Popul. Biol.* **1988**, *33*, 54–78. [[CrossRef](#)]
43. Shin, J.H.; Blay, S.; McNeney, B.; Graham, J. LDheatmap: An R Function for Graphical Display of Pairwise Linkage Disequilibria between Single Nucleotide Polymorphisms. *J. Stat. Softw.* **2006**, *16*, 1–9. [[CrossRef](#)]
44. Alexa, A.; Rahnenfuhrer, J. TopGO: Enrichment Analysis for Gene Ontology. R package version 2.48.0. 2022.
45. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A Program for Annotating and Predicting the Effects of Single Nucleotide Polymorphisms, SnpEff: SNPs in the Genome of *Drosophila Melanogaster* Strain W1118; Iso-2; Iso-3. *Fly* **2012**, *6*, 80–92. [[CrossRef](#)]
46. Saavedra, C.; Peña, J.B. Nucleotide Diversity and Pleistocene Population Expansion in Atlantic and Mediterranean Scallops (*Pecten maximus* and *P. jacobaeus*) as Revealed by the Mitochondrial 16S Ribosomal RNA Gene. *J. Exp. Mar. Biol. Ecol.* **2005**, *323*, 138–150. [[CrossRef](#)]
47. Boon, E.; Faure, M.F.; Bierne, N. The Flow of Antimicrobial Peptide Genes through a Genetic Barrier between *Mytilus edulis* and *M. galloprovincialis*. *J. Mol. Evol.* **2009**, *68*, 461–474. [[CrossRef](#)]
48. Tine, M.; Kuhl, H.; Gagnaire, P.A.; Louro, B.; Desmarais, E.; Martins, R.S.T.; Hecht, J.; Knaust, F.; Belkhir, K.; Klages, S.; et al. European Sea Bass Genome and Its Variation Provide Insights into Adaptation to Euryhalinity and Speciation. *Nat. Commun.* **2014**, *5*, 5770. [[CrossRef](#)]
49. Le Moan, A.; Gagnaire, P.A.; Bonhomme, F. Parallel Genetic Divergence among Coastal-Marine Ecotype Pairs of European Anchovy Explained by Differential Introgression after Secondary Contact. *Mol. Ecol.* **2016**, *25*, 3187–3202. [[CrossRef](#)]
50. Marth, G.T.; Czabarka, E.; Murvai, J.; Sherry, S.T. The Allele Frequency Spectrum in Genome-Wide Human Variation Three Large World Populations. *Genetics* **2004**, *372*, 351–372. [[CrossRef](#)] [[PubMed](#)]
51. Tintore, J.; La Violette, P.E.; Blade, I.; Cruzado, A. A Study of an Intense Density Front in the Eastern Alboran Sea: The Almeria–Oran Front. *J. Phys. Oceanogr.* **1988**, *18*, 1384–1397. [[CrossRef](#)]
52. Duranton, M.; Allal, F.; Fraïsse, C.; Bierne, N.; Bonhomme, F.; Gagnaire, P.A. The Origin and Remolding of Genomic Islands of Differentiation in the European Sea Bass. *Nat. Commun.* **2018**, *9*, 2518. [[CrossRef](#)]
53. Renaut, S.; Owens, G.L.; Rieseberg, L.H. Shared Selective Pressure and Local Genomic Landscape Lead to Repeatable Patterns of Genomic Divergence in Sunflowers. *Mol. Ecol.* **2014**, *23*, 311–324. [[CrossRef](#)]
54. Ellegren, H.; Smeds, L.; Burri, R.; Olason, P.I.; Backström, N.; Kawakami, T.; Künstner, A.; Mäkinen, H.; Nadachowska-Brzyska, K.; Qvarnström, A.; et al. The Genomic Landscape of Species Divergence in *Ficedula* Flycatchers. *Nature* **2012**, *491*, 756–760. [[CrossRef](#)] [[PubMed](#)]
55. Berg, P.R.; Star, B.; Pampoulie, C.; Sodeland, M.; Barth, J.M.I.; Knutsen, H.; Jakobsen, K.S.; Jentoft, S. Three Chromosomal Rearrangements Promote Genomic Divergence between Migratory and Stationary Ecotypes of Atlantic Cod. *Sci. Rep.* **2016**, *6*, 23246. [[CrossRef](#)] [[PubMed](#)]
56. Sodeland, M.; Jorde, P.E.; Lien, S.; Jentoft, S.; Berg, P.R.; Grove, H.; Kent, M.P.; Arnyasi, M.; Olsen, E.M.; Knutsen, H. “Islands of Divergence” in the Atlantic Cod Genome Represent Polymorphic Chromosomal Rearrangements. *Genome Biol. Evol.* **2016**, *8*, 1012–1022. [[CrossRef](#)]
57. Faria, R.; Chaube, P.; Morales, H.E.; Larsson, T.; Lemmon, A.R.; Lemmon, E.M.; Rafajlović, M.; Panova, M.; Ravinet, M.; Johannesson, K.; et al. Multiple Chromosomal Rearrangements in a Hybrid Zone between *Littorina Saxatilis* Ecotypes. *Mol. Ecol.* **2019**, *28*, 1375–1393. [[CrossRef](#)] [[PubMed](#)]
58. Wellenreuther, M.; Bernatchez, L. Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends Ecol. Evol.* **2018**, *33*, 427–440. [[CrossRef](#)] [[PubMed](#)]
59. Nie, L.; Cai, S.Y.; Shao, J.Z.; Chen, J. Toll-Like Receptors, Associated Biological Roles, and Signaling Networks in Non-Mammals. *Front. Immunol.* **2018**, *9*. [[CrossRef](#)]
60. Wang, M.; Yang, J.; Zhou, Z.; Ziu, L.; Wang, L.; Zhang, H.; Gao, Y.; Wang, X.; Zhang, L.; Zhao, J.; et al. A Primitive Toll-like Receptor Signaling Pathway in Mollusk Zhikong Scallop *Chlamys farreri*. *Dev. Comp. Immunol.* **2011**, *35*, 511–520. [[CrossRef](#)] [[PubMed](#)]

61. Malmstrøm, M.; Matschiner, M.; Tørresen, O.K.; Star, B.; Snipen, L.G.; Hansen, T.F.; Baalsrud, H.T.; Nederbragt, A.J.; Hanel, R.; Salzburger, W.; et al. Evolution of the Immune System Influences Speciation Rates in Teleost Fishes. *Nat. Genet.* **2016**, *48*, 1204–1210. [[CrossRef](#)]
62. Karvonen, A.; Seehausen, O. The Role of Parasitism in Adaptive Radiations—When Might Parasites Promote and When Might They Constrain Ecological Speciation? *Int. J. Ecol.* **2012**, *2012*. [[CrossRef](#)]
63. Brennan, J.J.; Gilmore, T.D. Evolutionary Origins of Toll-like Receptor Signaling. *Mol. Biol. Evol.* **2018**, *35*, 1576–1587. [[CrossRef](#)]
64. Zhang, L.; Li, L.; Guo, X.; Litman, G.W.; Dishaw, L.J.; Zhang, G. Massive Expansion and Functional Divergence of Innate Immune Genes in a Protostome. *Sci. Rep.* **2015**, *5*, 8693. [[CrossRef](#)]
65. Leulier, F.; Lemaitre, B. Toll-like Receptors—Taking an Evolutionary Approach. *Nat. Rev. Genet.* **2008**, *9*, 165–178. [[CrossRef](#)] [[PubMed](#)]
66. Zaccara, S.; Ries, R.J.; Jaffrey, S.R. Reading, Writing and Erasing mRNA Methylation. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 608–624. [[CrossRef](#)] [[PubMed](#)]
67. Fraser, H.B. Gene Expression Drives Local Adaptation in Humans. *Genome Res.* **2013**, *23*, 1089–1096. [[CrossRef](#)] [[PubMed](#)]
68. Mendel, R.R. The Molybdenum Cofactor. *J. Biol. Chem.* **2013**, *288*, 13165–13172. [[CrossRef](#)]
69. Pujolar, J.M.; Jacobsen, M.W.; Bertolini, F. Comparative Genomics and Signatures of Selection in North Atlantic Eels. *Mar. Genom.* **2022**, *62*, 100933. [[CrossRef](#)]
70. Mack, K.L.; Nachman, M.W. Gene Regulation and Speciation. *Trends Genet.* **2017**, *33*, 68–80. [[CrossRef](#)]
71. Goncalves, A.; Leigh-Brown, S.; Thybert, D.; Stefflova, K.; Turro, E.; Flicek, P.; Brazma, A.; Odom, D.T.; Marioni, J.C. Extensive Compensatory Cis-Trans Regulation in the Evolution of Mouse Gene Expression. *Genome Res.* **2012**, *22*, 2376–2384. [[CrossRef](#)]
72. Davidson, J.H.; Balakrishnan, C.N. Gene Regulatory Evolution during Speciation in a Songbird. *G3 Genes Genomes Genet.* **2016**, *6*, 1357–1364. [[CrossRef](#)]
73. McManus, C.J.; Coolon, J.D.; Duff, M.O.; Eipper-Mains, J.; Graveley, B.R.; Wittkopp, P.J. Regulatory Divergence in *Drosophila* Revealed by mRNA-Seq. *Genome Res.* **2010**, *20*, 816–825. [[CrossRef](#)]
74. Tirosh, I.; Reikhav, S.; Levy, A.A.; Barkai, N. A Yeast Hybrid Provides Insight into the Evolution of Gene Expression Regulation. *Science* **2009**, *324*, 659–662. [[CrossRef](#)]
75. Shi, X.; Ng, D.W.K.; Zhang, C.; Comai, L.; Ye, W.; Jeffrey Chen, Z. Cis- and Trans-Regulatory Divergence between Progenitor Species Determines Gene-Expression Novelty in *Arabidopsis allopolyploids*. *Nat. Commun.* **2012**, *3*. [[CrossRef](#)] [[PubMed](#)]
76. Huelvan, S. *Variabilité Génétique de Populations de Pecten Maximus L. En Bretagne*; Université de Bretagne Occidentale: Brest, France, 1985.
77. Popović, N.T.; Krbavčić, M.; Barišić, J.; Ljubić, B.B.; Strunjak-Perović, I.; Babić, S.; Lorencin, V.; Matulić, D.; Tomljanović, T.; Čož-Rakovac, R. Comparative Tissue Responses of Marine Mollusks on Seasonal Changes in the Northern Adriatic Sea. *Appl. Sci.* **2021**, *11*, 2874. [[CrossRef](#)]
78. Mestre, S.; Peña, J.B.; Farí, A.; Uriarte, I. The Spawning Period and Gametogenic Cycle of *Pecten jacobaeus* (L.). *Iberus* **1990**, *9*, 161–167.
79. Le Pennec, M.; Paugam, A.; Le Pennec, G. The Pelagic Life of the Pectinid *Pecten Maximus*—A Review. *ICES J. Mar. Sci.* **2003**, *60*, 211–223. [[CrossRef](#)]
80. Boutet, I.; Moraga, D.; Marinovic, L.; Obreque, J.; Chavez-Crooker, P. Characterization of Reproduction-Specific Genes in a Marine Bivalve Mollusc: Influence of Maturation Stage and Sex on mRNA Expression. *Gene* **2008**, *407*, 130–138. [[CrossRef](#)]
81. Sigang, F.; Yihui, G.; Youhou, X. Comparative Transcriptome Analysis of the Ovary and Testis in Noble Scallop (*Chlamys nobilis*). *Pak. J. Zool.* **2021**, *53*, 251. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.