Supplementary Materials and Methods

Sampling sites in the Northern Hemisphere

The two closest species, *M. edulis* and *M. galloprovincialis*, meet in a mosaic hybrid zone along the French coast (Bierne *et al.* 2003). Samples come from populations enclosed within the mosaic zone (the Bay of Biscay for *M. edulis* and Brittany for *M. galloprovincialis*) and from peripheral populations (the North Sea for *M. edulis* and the Iberian Coast for *M. galloprovincialis*). Within *M. galloprovincialis*, previous studies have reported a genetic break between Atlantic and Mediterranean populations (Quesada *et al.* 1995), which was recently described as being a mosaic hybrid zone along the Algerian coast (El Ayari *et al.* 2017). Samples from either side of the Siculo-Tunisian strait in the Mediterranean Sea (in Sete for the western basin and Crete for the eastern basin) have been collected. The third species of the mussel complex, *M. trossulus*, met independently twice with *M. edulis* (Riginos & Cunningham 2005): once in Europe in a clinal hybrid zone at the entrance of the Baltic Sea and once in North America from Maine to Nova Scotia. In Europe, individuals of *M. trossulus* were sampled at the far end of the Baltic Sea, in the gulf of Finland. In North America, *Mytilus trossulus* mussels were obtained from the Saint Lawrence river in Canada, and *M. edulis* mussels were sampled from three localities in Rhode Island (USA). Sampling information is provided in Table S1.

Sampling sites in the Southern Hemisphere

Thirty-five sites were sampled between 1999 and 2003 (Gérard *et al.* 2015) in the Kerguelen Islands. Five samples came from the North coast (PCh, PMt, PCx, I3B, AJ), two from the North-East coast (PMo, RdA), twenty-two from the Gulf of Morbihan (PAF, PR1, PR2, IH, IM, IGn, PF, BOCRD, BOCentre, BOCRG, BO100am, BO200am, BO100av, BO200av, BOCAB, BOFF, Ar1, Ar2, HdS, PJDA, PB, IS), four from the South coast (BdS, BM, FPN, BM) and one from the West coast (PCu). Sampling information is provided in Table S3. We additionally included other samples from the Southern Hemisphere (Gérard *et al.* 2008): one sample from Western Australia in Nedlands (AUS, n=12 individuals); two samples from New Zealand, Dunedin (DUN, n=8) and Wellington Harbour (WHL, n=10); two samples from Tasmania, the Simpson's Bay (SIM, n=8) and Hobart (HOB, n=9); one sample from Chile in Maullin (MAU, n=15).

Genotyping-by-sequencing in the Southern Hemisphere

Eight individuals were collected from Baie de la Mouche (BM, Table S1) in the Kerguelen Islands. Genomic DNA was extracted from adults using the DNeasy Blood and Tissue Kit (Qiagen) following the manufacturer's protocol. Genomic DNA was then enriched for 3 Mb of target regions using a SureSelectXT Custom system (Agilent Technologies, Santa Clara, CA) comprising ~55,000 RNA probes of 120 bp (Mamanova *et al.* 2010) designed on BAC and cDNA sequences (Fraïsse *et al.* 2016). BAC sequences (totalling 2 Mb) were obtained from a BAC library constructed based on whole genomic DNA of three *M. edulis* individuals, and composed of 32 tagged clones and a pool of 192 untagged clones. The targeted coding regions were of two types: (i) a random panel of 338 cDNA contigs (1.6 kb on average, totalling 0.5 Mb) from *M. galloprovincialis* produced by RNA sequencing (Romiguier *et al.* 2014); and (ii) a set of 553 publicly released expressed sequence tags (1.0 kb on average, totalling 0.5 Mb) including 262 immunity genes, 133 genes involved in cytonuclear interactions, 30 reproduction-related genes, 20 habitat-related genes, six nucleoporins and 102 housekeeping genes. Illumina paired-end sequencing libraries with insert sizes of 300–600 bp were prepared for each individual. All individuals were pooled in equimolar proportions prior to being subjected to TruSeq Custom Enrichment (http://support.illumina.com) and sequenced on a HiSeq 2000 instrument. Paired-end reads of 101 bp were produced, and trimmed for index sequences and low-quality terminal bases; low-quality reads were discarded. Reads were mapped onto the reference BAC and cDNA sequences with bwa-mem v0.7.5a (Li & Durbin 2009) using the following non-default parameters: a clipping penalty of 3, a mismatch penalty of 2, a gap open penalty of 3 and 10-bp minimal seed length. SNP calling was performed with bcftools (Li 2011), and various quality filters were applied using vcftools v0.1.12a (Danecek *et al.* 2011). We excluded calls from positions with an average depth across individuals below 10 reads. We also applied an upper depth limit corresponding to the 98.5th percentile of the depth distribution across all positions to exclude unmasked repeated regions. To filter out paralogous regions, we excluded sites deviating from Hardy–Weinberg equilibrium (P-value <0.05) using an exact test implemented in vcftools. For each variant retained, the genotype with the highest likelihood (i.e. with a null phred-scaled likelihood score) was assigned to each individual assuming Hardy–Weinberg proportions in genotype prior probabilities. Only genotype calls with a quality score above 10 were retained (error probability of 10%), otherwise missing data was applied. Any position with more than 20% of missing data were

discarded. In total, 129,346 SNPs were called. For the genetic network analysis, we used a high-quality dataset (51,878 SNPs retained) to improve the reconstruction of genetic relationships while having a limited loss of information for such genome-wide analysis: genotype quality threshold was set at 30 (error probability of 0.1%) and missing data were not allowed . For the *TreeMix* and ADMIXTURE analyses, sites with a minor allele frequency below 5% over all populations were excluded (32,162 SNPs retained), as they are uninformative to infer admixture history.

Supplementary References

- Bierne N, Borsa P, Daguin C, Jollivet D, Viard F, Bonhomme F, David P (2003) Introgression patterns in the mosaic hybrid zone between *Mytilus edulis* and *M. galloprovincialis*. *Mol. Ecol.*, 12, 447–461.
- Danecek P, Auton A, Abecasis G *et al.* (2011) The variant call format and VCFtools. *Bioinformatics*, 27, 2156–2158.
- El Ayari T, Trigui El Menif N, Hamer B, Bierne N (submitted) The hidden side of a major marine biogeographic boundary: a wide endogenously maintained mosaic hybrid zone at the Atlantic-Mediterranean divide provides support for the coupling hypothesis.
- Fraïsse C, Belkhir K, Welch JJ, Bierne N (2016) Local interspecies introgression is the main cause of extreme levels of intraspecific differentiation in mussels. *Mol. Ecol.*, 25, 269–286.
- Gérard K, Bierne N, Borsa P, Chenuil A, Féral J-P (2008) Pleistocene separation of mitochondrial lineages of *Mytilus* spp. mussels from Northern and Southern Hemispheres and strong genetic differentiation among southern populations. *Mol. Phylogenet. Evol.*, 49, 84–91.
- Gérard K, Roby C, Bierne N, Borsa P, Féral JP, Chenuil A (2015) Does natural selection explain the fine scale genetic structure at the nuclear exon Glu-5′ in blue mussels from Kerguelen? *Ecol. Evol.*, 5, 1456–1473.
- Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data.

Bioinformatics , 27, 2987–2993.

- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* , 25, 1754–1760.
- Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J, Turner DJ (2010) Target-enrichment strategies for next-generation sequencing. *Nat. Methods*, 7, 111–118.
- Quesada H, Zapata C, Alvarez G (1995) A multilocus allozyme discontinuity in the mussel *Mytilus galloprovincialis*: the interaction of ecological and life-history factors. *Mar. Ecol. Prog. Ser.*, 116, 99–115.
- Riginos C, Cunningham CW (2005) Local adaptation and species segregation in two mussel (*Mytilus edulis* x *Mytilus trossulus*) hybrid zones. *Mol. Ecol.*, 14, 381–400.
- Romiguier J, Gayral P, Ballenghien M, Bernard A, Cahais V, Chenuil A, Chiari Y, Dernat R, Duret L, Faivre N, Loire E, Lourenco JM, Nabholz B, Roux C, Tsagkogeorga G, Weber AA, Weinert LA, Belkhir K, Bierne N, Glémin S, Galtier N (2014) Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, 515(7526), 261.