



# Extraordinarily rapid speciation in a marine fish

Paolo Momigliano<sup>a,1</sup>, Henri Jokinen<sup>b,2</sup>, Antoine Fraimout<sup>c,2</sup>, Ann-Britt Florin<sup>d</sup>, Alf Norkko<sup>b,e</sup>, and Juha Merilä<sup>a</sup>

<sup>a</sup>Ecological Genetics Research Unit, Department of Biosciences, University of Helsinki, FI-00014 Helsinki, Finland; <sup>b</sup>Tvärminne Zoological Station, University of Helsinki, FI-10900 Hanko, Finland; <sup>c</sup>Institut National de la Recherche Agronomique UMR, Centre de Biologie pour la Gestion des Populations (Institut National de la Recherche Agronomique/Institut de Recherche pour le Développement/Cirad/Montpellier SupAgro), FR-34988 Montpellier-sur-Lez, France; <sup>d</sup>Department of Aquatic Resources, Swedish University of Agricultural Sciences, SE-74242 Öregrund, Sweden; and <sup>e</sup>Baltic Sea Centre, Stockholm University, SE-10691 Stockholm, Sweden

Edited by Dolph Schluter, University of British Columbia, Vancouver, BC, Canada, and accepted by Editorial Board Member Douglas Futuyma April 25, 2017 (received for review September 20, 2016)

Divergent selection may initiate ecological speciation extremely rapidly. How often and at what pace ecological speciation proceeds to yield strong reproductive isolation is more uncertain. Here, we document a case of extraordinarily rapid speciation associated with ecological selection in the postglacial Baltic Sea. European flounders (*Platichthys flesus*) in the Baltic exhibit two contrasting reproductive behaviors: pelagic and demersal spawning. Demersal spawning enables flounders to thrive in the low salinity of the Northern Baltic, where eggs cannot achieve neutral buoyancy. We show that demersal and pelagic flounders are a species pair arising from a recent event of speciation. Despite having a parapatric distribution with extensive overlap, the two species are reciprocally monophyletic and show strongly bimodal genotypic clustering and no evidence of contemporary migration, suggesting strong reproductive isolation. Divergence across the genome is weak but shows strong signatures of selection, a pattern suggestive of a recent ecological speciation event. We propose that spawning behavior in Baltic flounders is the trait under ecologically based selection causing reproductive isolation, directly implicating a process of ecological speciation. We evaluated different possible evolutionary scenarios under the approximate Bayesian computation framework and estimate that the speciation process started in allopatry ~2,400 generations ago, following the colonization of the Baltic by the demersal lineage. This is faster than most known cases of ecological speciation and represents the most rapid event of speciation ever reported for any marine vertebrate.

ecological speciation | genomics | evolution | rapid speciation | Baltic Sea

Divergent selection may generate barriers to gene flow and ultimately lead to the evolution of distinct species—a process referred to as ecological speciation (1). Partial reproductive isolation can arise as a by-product of local adaptation within dozens to a few hundred generations, suggesting that divergent selection may initiate speciation over ecological timescales (2); more uncertain, however, is how often and at what pace divergence proceeds along the speciation continuum until strong reproductive isolation is established (2, 3). Partial reproductive isolation does not necessarily lead to the evolution of distinct species; even in cases when divergence continues, it may persist as a “quasiequilibrium state” for millions of years (4).

Ecological selection on traits that directly cause reproductive isolation favors the completion of the speciation process and influences the pace at which it proceeds (4). Such “magic traits” may play a pivotal role in facilitating speciation in the absence of impermeable geographic barriers (5). This can happen when ecological adaptations cause assortative mating, such as among sister species of *Heliconius* butterflies with contrasting color pattern mimicry (6), or when selection leads to habitat specialization and divergent populations mate within their preferred habitats leading to reproductive segregation (7). Selection on habitat preferences, specifically, is expected to lead to rapidly evolving barriers to gene flow (2). Early onset of reproductive isolation (through assortative mating, reproductive segregation, or an initial allopatric/parapatric phase) may thus favor rapid speciation, because it could theoretically initiate a positive-feedback loop where restricted migration

enhances divergence at increasingly weakly selected loci, which in turn reduces gene flow (4).

In the marine environment, evidence for ecological speciation is scarce (8). This is somehow surprising: barriers to gene flow are rarely absolute in the sea, hence models of speciation that can operate in the presence of gene flow, such as ecological speciation, are likely to be important in explaining marine biodiversity (8). The sympatric speciation of coral-dwelling gobies of the genus *Gobiodon* represents the most compelling case and is estimated to have happened ~200 kya (9). The fastest known case of speciation in a marine animal is the evolution of the sea star *Cryptasterina hystera*, which occurred 1–22 kya, although it is unclear whether ecological selection played any role in this speciation event (10).

Here, we present a case of extraordinarily rapid speciation in the postglacial Baltic Sea, a large body of brackish water that became connected to the North Sea at the end of the last glaciation, ~8.5 kya (11). The narrow connection between the Baltic Proper and the North Sea presents a steep gradient in temperature and salinity: a barrier that most marine fish cannot cross and that reduces gene flow for the few that can (12). Among the few marine fish that thrive in the Baltic Sea is the European flounder (*Platichthys flesus*), an important fishery target. European flounders show two distinct reproductive behaviors within the Baltic Sea: offshore spawning of pelagic eggs (hereafter, “pelagic flounders”), as seen in the rest of their distribution, and coastal spawning of demersal eggs (hereafter, “demersal flounders”),

## Significance

Divergent selection can lead to the evolution of distinct species, a process known as ecological speciation. Evidence for ecological speciation in the marine environment is scarce, and the few known examples have happened within a time frame of hundreds of thousands to millions of years. We present evidence that European flounders in the Baltic Sea exhibiting different breeding behaviors are a species pair arising from a recent event of ecological speciation. The two lineages diverged within less than 3,000 generations. This is the fastest event of speciation ever reported for any marine vertebrate. Extraordinarily rapid speciation driven by natural selection can therefore happen even in the marine environment.

Author contributions: P.M. and J.M. designed research; P.M. and H.J. performed research; H.J., A.-B.F., A.N., and J.M. contributed new reagents/analytic tools; P.M. and A.F. analyzed data; and P.M. wrote the paper.

The authors declare no conflict of interest.

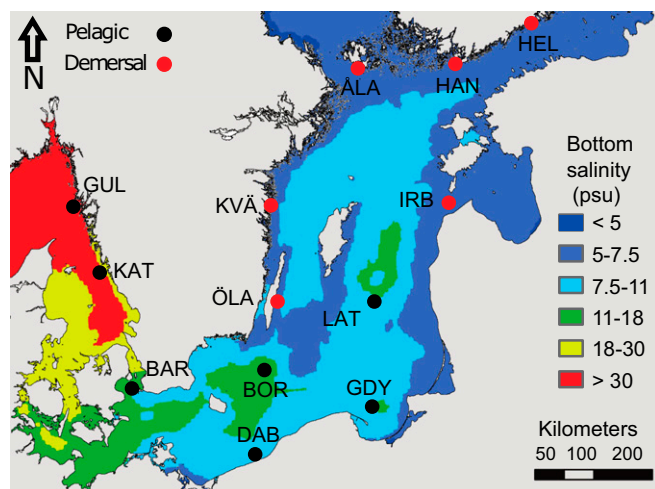
This article is a PNAS Direct Submission. D.S. is a guest editor invited by the Editorial Board.

Data deposition: The short sequence reads reported in this paper have been deposited in GenBank Short Read Archive (Bioproject PRJNA382467, Short Read Study SRP103564). The sequences from the outliers’ flanking regions have been deposited in the GenBank database (accession nos. KY933571–KY933582). In addition, the codes for the ABC scenarios and additional datasets have been deposited in the DRYAD Digital Repository database, [datadryad.org](http://datadryad.org) (doi: 10.5061/dryad.f6154).

<sup>1</sup>To whom correspondence should be addressed. Email: [paolo.momigliano@helsinki.fi](mailto:paolo.momigliano@helsinki.fi).

<sup>2</sup>H.J. and A.F. contributed equally to this work.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1615109114/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1615109114/-DCSupplemental).



**Fig. 1.** Map of the sampling locations showing modeled mean bottom salinity in the Baltic Sea from Bendtsen et al. (62). Areas theoretically suitable for pelagic spawning are colored in green, yellow, and red. Locations are color-coded based on expected spawning behaviors. ÅLA, Åland Archipelago; BAR, Barsebäck; BOR, Bornholm Basin; DAB, Dabki; GDY, Gdynia; GUL, Gullmaren; HAN, Hanko (Western Gulf of Finland); HEL, Helsinki (Gulf of Finland); IRB, Irbe Strait; KAT, Kattegat; KVA, Kvädöfjärden; LAT, Latvian Sea; and ÖLA, Öland Island. More than 90% (Table S1) of samples from BOR (pelagic habitat) and ÖLA (demersal habitat) were ready for spawning at the time of capture, hence these locations were used as reference pelagic and demersal populations, respectively.

a behavior observed only in the Baltic Sea (13). Pelagic spawning occurs below the 11-practical salinity unit (psu) isohaline in deep, offshore basins of the southern and central Baltic Sea where eggs can achieve neutral buoyancy (13). Demersal flounders spawn in shallow coastal waters and reproduce successfully at salinity as low as 6 psu (13), which enabled them to expand their distribution into the coastal waters of the northern Baltic Sea (Fig. 1). Demersal and pelagic spawners exhibit differences in morphological and physiological reproductive traits; demersal eggs are smaller, but with thicker chorions, a potential adaptation to the rougher conditions of shallow demersal habitats, whereas pelagic eggs are larger, a potential adaptation to increase buoyancy (13, 14). Sperm activation is salinity dependent: sperm from samples collected in pelagic spawning sites show no motility below 10–9 psu, whereas sperm activity in demersal spawners is still observed at 3 psu, hence fertilization can successfully occur at much lower salinity (13).

There is evidence that pelagic and demersal flounders represent genetically distinct ecotypes (15, 16), although genetic differentiation among putative demersal and pelagic populations is extremely weak ( $F_{ST} \sim 0.01$ – $0.03$ ). Weak differentiation may suggest ongoing gene flow or reflect recent divergence and large effective population sizes, depending on whether populations have reached migration-drift equilibrium (17). Even if demersal and pelagic flounders were genetically distinct taxa, gene flow could potentially still occur if males opportunistically fertilize eggs of the opposite ecotype, or if spawning behavior and/or reproductive traits show some degree of plasticity, as has been suggested for another flatfish species (18). An artificial fertilization experiment (19) suggests the two ecotypes can produce viable larvae, although it is unclear what their fitness would be under natural conditions. However, if reproductive traits such as egg buoyancy and sperm motility were under ecological selection they could act as magic traits, favoring rapid speciation via reproductive segregation (6, 20).

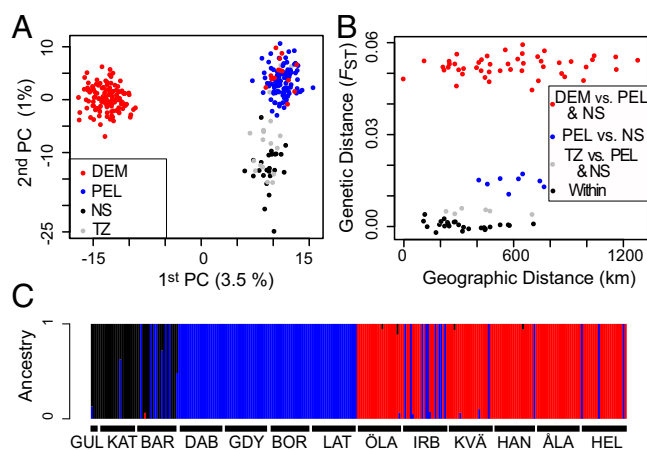
Here, we tested whether pelagic and demersal flounders represent two genetically divergent taxa, and investigated how far divergence has proceeded along the speciation continuum. We used some of the criteria listed by Nosil et al. (3) and Nosil and Feder (21) to determine how far speciation has progressed, and

we predict that, if speciation has already resulted in strong reproductive isolation, (i) genomic divergence should be heterogeneous, showing strong signatures of divergent selection in some regions of the genome (a predicted outcome of ecological speciation) and low levels of divergence elsewhere (the latter being a reflection of recent divergence); (ii) genotypic clustering should be strongly bimodal, regardless of geographical distance; and (iii) there should be limited introgression among the two taxa, even where they co-occur, and reciprocal monophyly of the pelagic and demersal ecotypes.

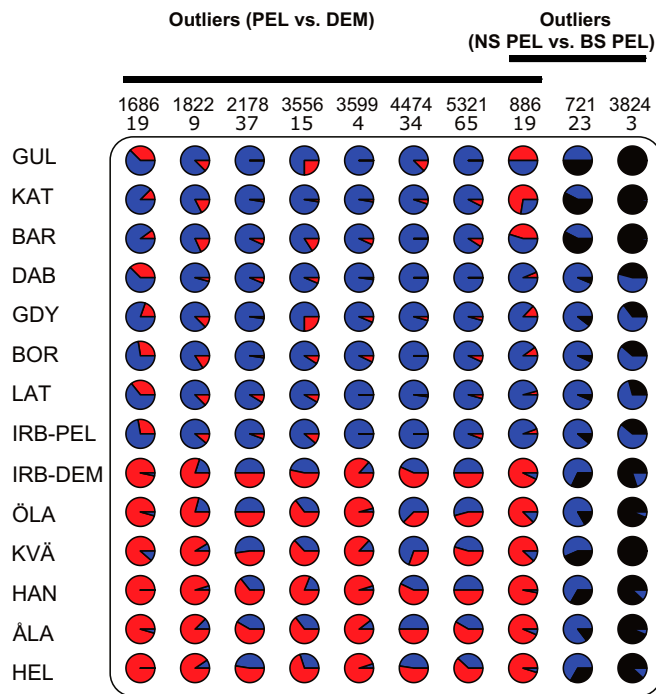
Using genomic data from flounders sampled in 13 locations from the Baltic Sea, the North Sea, and the transition zone between these two regions, we show that each one of these expectations is clearly met, providing convincing evidence of selection and strong reproductive isolation. By evaluating different evolutionary scenarios within an approximate Bayesian computation (ABC) (22) framework, we demonstrate that speciation has started following an early colonization of the post-glacial Baltic Sea. This represents the most rapid speciation event ever reported for any marine vertebrate.

## Results

The final dataset consisted of 2,051 biallelic SNPs genotyped for 282 individuals. A principal-component analysis (PCA) performed with all loci revealed two genetic clusters along the first axis (Fig. 2A), which largely coincide with the expected distribution of demersal and pelagic spawners (Table S1 and Fig. 1). Samples from the two reference populations (Bornholm for pelagic spawners, Öland for demersal spawners) clustered as expected with the “pelagic” and “demersal” clusters, respectively (Fig. 2A). A few individuals (nine from Irbe, one from Kvädöfjärden, one from Hanko, and three from Helsinki) sampled in putative demersal locations appear to have genotypes characteristic of pelagic Baltic flounders, suggesting the two taxa co-occur in some locations (Fig. 1). PCAs run only with loci from the core  $F_{ST}$  distribution (i.e., loci unlikely to be under selection) and excluding  $F_{ST}$  outliers revealed nearly identical patterns (Fig. S1A and C), with samples from the Baltic Sea showing a clear bimodal distribution on the first axis (Fig. S1B and D). There was no evidence of isolation by distance, regardless of whether the dataset was analyzed as a whole or pelagic and demersal flounders were analyzed separately (Fig. 2B), and regardless of whether loci potentially under selection were excluded (Fig. S1E



**Fig. 2.** (A) PCA of allele frequency data from all loci. Codes are as follows: DEM, putative demersal locations; NS, North Sea samples; PEL, putative pelagic locations within the Baltic; TZ, transition zone between the Baltic Sea and the North Sea (Barsebäck samples). (B) Relationship between geographic and genetic distance among sampling locations. “Within” refers to pairwise comparisons among locations belonging to the three major genetic clusters identified by PCA, fastSTRUCTURE, and  $K$ -means clustering. (C) Individual ancestries from fastSTRUCTURE analysis ( $K = 3$ ).



**Fig. 3.** Frequencies of outlier loci in each sampling location. SNPs are identified by the locus number (above) and the SNP position within the locus in base pairs (below). The first eight loci are outliers between pelagic and demersal flounders, and the last three are outliers between North Sea and Baltic pelagic locations. Individuals in Irbe were grouped separately in pelagic (above) and demersal (below) groups based on PCA, fastSTRUCTURE, and *K*-mean clustering results. Blue alleles are characteristic of pelagic Baltic populations; red alleles are characteristic of demersal Baltic populations. Black alleles are characteristic of North Sea populations.

and *F*). Genetic diversity indices were very similar for each sampling location and for the three major genetic clusters (Table S2), suggesting that no major demographic events affecting genetic diversity have occurred since the time of common ancestry.

Genotypic clustering using *K*-means and fastSTRUCTURE suggests that values of *K* of 2 and 3 are useful models of the data (Fig. S2A–D). Both fastSTRUCTURE and discriminant analysis of principal components (DAPC) identified the same genetic clusters as the PCA. At *K* = 2, both fastSTRUCTURE and DAPC separate the putative demersal and pelagic taxa (Fig. S2E). When *K* = 3, the third cluster represented the North Sea samples (Fig. 2C and Fig. S2E). Individuals from the demersal and pelagic clusters co-occur in similar proportions in one location (Irbe, Fig. 2C), and a small number of individuals with pelagic genotypes were sampled from the northern Baltic Sea coast in Finland and Sweden (Fig. 2C). We found no sign of hybridization between the two Baltic taxa as none of the 282 sampled individuals showed intermediate genotypes (Fig. 2A and C). There was some evidence of hybridization between pelagic flounders from the Baltic Sea and the North Sea (Fig. 2C), particularly in the transition zone (Barsebäck). The private allele method (23) suggests less than one migrant per generation between each of the major genetic clusters; five independent runs of BAYESASS reported migration rates of less than 0.5% between demersal and pelagic flounders, and the 95% CI of these estimates always overlapped with zero, providing no evidence for contemporary gene flow.

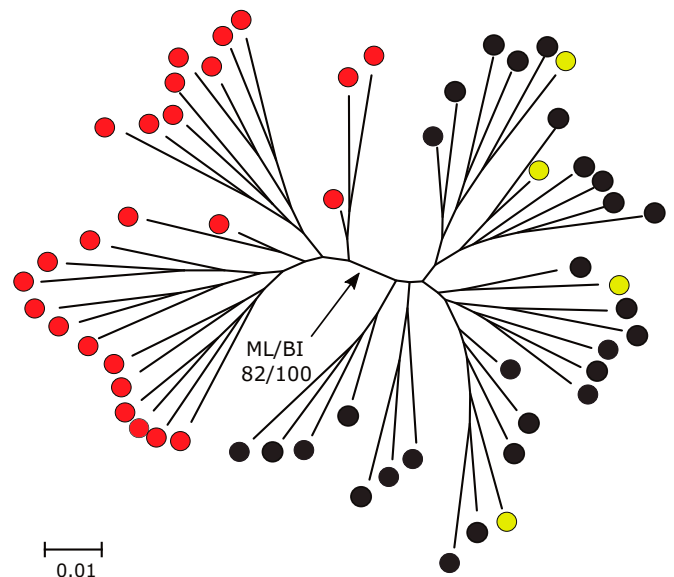
Multiple outlier tests (BAYESCAN, Arlequin, OutFLANK, and FLK) identified a set of 10 outlier loci that show higher than expected  $F_{ST}$  (Fig. S3 and Table S3). Eight of these loci show evidence of divergent selection between demersal and pelagic locations in the Baltic, whereas only three show evidence of divergent

selection among locations on either side of the transition zone (Fig. 3 and Fig. S3). There was no evidence of spatially diversifying selection among populations of the same taxon within the Baltic Sea (Fig. 3 and Fig. S3G and H). In Irbe (IRB), we present frequencies of outlier loci for putative demersal and putative pelagic individuals separately (Fig. 3), both of which show outlier frequencies similar to other demersal and pelagic populations, respectively, regardless of geographical proximity (Fig. 3). Of the six outliers for which we obtained flanking regions, four matched protein-coding genes (Table S4). One (a protocadherin FAT1-like gene) has been shown to be involved in response to salinity stress in oysters (24, 25) and teleosts (26). Locus 3599 has a nonsynonymous mutation within the dynein axonemal heavy chain 5 gene, which codes for a protein playing a central role in flagellar movement and sperm motility (27)—one of the traits that differs between the two taxa (13).

Maximum-likelihood and Bayesian phylogenetic analyses based on the sequences of 2,051 concatenated loci demonstrate that the two Baltic taxa are reciprocally monophyletic (Fig. 4). However, there is no support for a Baltic pelagic clade distinct from North Sea flounders (Fig. 4). Using an ABC-random forest (ABC-RF) approach (28, 29), we evaluated different potential evolutionary scenarios (Fig. 5 and Tables S5–S9). The demersal and pelagic flounders within the Baltic originated via two independent invasions (Fig. 5, scenario 1), which happened 2,400 (95% CI: 1,450–9,110) and 1,490 (95% CI: 596–4,780) generations ago, respectively. Evidence was also found for a demographic expansion in the ancestral population (Table S9), for a mild bottleneck in the demersal population (Table S8), and for secondary introgression—albeit weak—among the two species in the Baltic (Table S9).

## Discussion

**Ecological Speciation.** Genomic analyses revealed that pelagic and demersal spawning flounders in the Baltic Sea represent two reciprocally monophyletic taxa exhibiting limited-to-no introgression, strongly bimodal genotypic clustering, and signatures of divergent selection. Genetic differentiation among Baltic locations closely mirrored expected breeding behavior and was uncorrelated

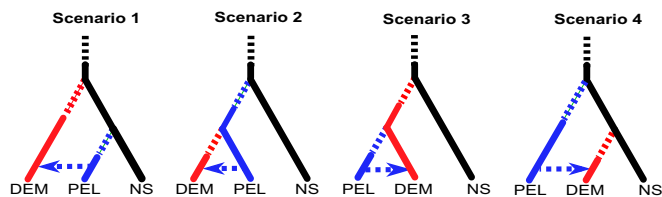


**Fig. 4.** Unrooted ML tree based on the concatenated sequences from 2,051 loci for a subset of 57 individuals. Black filled circles represent putative pelagic individuals. Red filled circles represent putative demersal individuals. Yellow circles represent individuals from a putative demersal location (Irbe), which clustered in all previous analyses with pelagic individuals. Branch support values represent ML bootstrap support and Bayesian clade credibility values, respectively. Only the single node that shows high support (>80) in both ML and Bayesian analyses is labeled.

with geographic distance. Four of the six outliers for which we obtained flanking sequences matched protein-coding genes, one of which has been implicated in salinity stress response in oysters (24, 25) and fish (26). A SNP in locus 3599 causes an amino acid change in a gene governing sperm motility (27), one of the key traits that have previously been shown to differ between demersal and pelagic spawners (13), reinforcing the interpretation that the outliers identified in this study are likely targets of ecological selection. Interpreted together, these findings are indicative of strong reproductive isolation and suggestive of a well-advanced process of ecological speciation (3). Along with previous studies showing that reproductive traits are population specific (13, 30, 31), the genetic differentiation and signatures of selection between pelagic and demersal spawners uncovered here strongly suggest that breeding behavior in *P. flesus* is not a plastic trait, as has been previously suggested for another flatfish species (18), but rather the result of ecological adaptations to the low salinities experienced in the coastal waters of the Baltic Sea (13, 15).

Pelagic and demersal flounders have a parapatric distribution and share the same feeding and wintering grounds in the central Baltic (30). Because the two taxa can co-occur around the time of spawning in the same shallow coastal waters (for example, as they do in Irbe), it should be possible for pelagic spawners to opportunistically fertilize demersal eggs. However, a study has shown that pelagic spawners' sperm motility is very low at salinities below 11 psu (13). Because demersal spawners lay their eggs in coastal waters where salinity rarely exceeds 7.5 psu—and never exceeds 11 psu (Fig. 1)—opportunistically fertilization by pelagic males is unlikely. Conversely, demersal spawners could remain in deeper areas after wintering and opportunistically fertilize pelagic eggs, although to what extent this may happen under natural conditions is unknown. Breeding behavior is expected to be under strong ecological selection in the northern Baltic region, and this expectation is supported by our findings of signatures of selection. Ecological selection on spawning behavior could lead to reproductive isolation via spawning segregation. As noted by Nosil (4), a trait under ecological selection causing reproductive isolation implicates a process of ecological speciation.

The extent of reproductive isolation seems to be stronger in the flounder species pair than in other cases of recent ecological speciation. Sympatric species of cichlid fish (*Amphilophus citrinellus* and *Amphilophus zaliostus*) in Lake Apoyo are not reciprocally monophyletic (the two species share many mtDNA haplotypes) and show stronger signs of introgression and less clearly bimodal genotypic clustering (32). Similarly, three-spined



**Fig. 5.** Scenarios tested via ABC-RF. Scenario 1: double-colonization scenario with early invasion by demersal flounders. Scenario 2: single colonization scenario with early invasion of pelagic flounders. Scenario 3: single-colonization scenario with early invasion of demersal flounders. Scenario 4: double-colonization scenario with early invasion by pelagic flounders. Each lineage is color coded: black lines represent the North Sea lineage (NS), blue lines represent the Baltic pelagic lineage (PEL), and red lines represent the Baltic demersal lineage (DEM). Broken lines represent possible demographic changes in the ancestral population (black broken line) as well as potential bottlenecks at the time of invasion (blue and red broken lines). The blue dotted line with arrowhead represents secondary introgression among the two Baltic species. All divergence scenarios were tested against each other under an orthogonal combination of all other demographic parameters (changes in  $N_e$  and introgression). Full details of all models tested via ABC-RF are given in [Supporting Information](#). The most likely model was scenario 1, including a demographic expansion in the ancestral population and a mild bottleneck in the demersal lineage.

stickleback limnetic/benthic species pairs do occasionally mate in the wild; for example, roughly 1% of individuals in Paxton Lake are hybrids (33)—although hybrids have reduced fitness in the natural environment (34). Despite sampling throughout the Baltic Sea, we found no evidence of first-generation hybrids or of intermediate genotypes. If hybridization ever occurs under natural conditions, it must be either rare or geographically restricted.

Our data suggest the existence of another barrier to gene flow and of spatially diversifying selection between pelagic flounders from the North Sea and the Baltic Sea, coincident with a steep gradient in temperature and salinity. Reproductive isolation between the North Sea and Baltic Sea pelagic flounders is partial, as documented by a number of individuals showing admixture in the transition zone. However, it is possible that a process of speciation may have initiated between these two populations, possibly driven by selection for lower temperatures and salinity in the Baltic Sea. This hypothesis is supported by findings from reciprocal transplant experiments showing that pelagic flounders on each side of the transition zone have different expression patterns in candidate genes for osmoregulation, stress resistance, and heme protein biosynthesis (35).

**Possible Evolutionary Scenarios.** The Baltic Sea became connected to the North Sea roughly 8.5 kya. Salinity remained well below the lowest salinity at which demersal flounders can reproduce until ~7.5 kya and remained below the minimum requirements for pelagic spawning until ~6.5 kya (11). During the period between 6.5 and 5 kya, salinity in the Baltic remained between 10 and 15 psu, creating an opportunity for colonization of the basin by pelagic spawners (11), and since has been steadily declining (11).

We propose that flounders adapted to a new ecological niche by developing a demersal spawning behavior in the early days of the Litorina Sea stage, possibly as early as 7.5 kya. Under this scenario, demersal flounders initially diverged from the pelagic flounders during a first allopatric phase lasting until ~6 kya (11). Pelagic flounders could have invaded the Baltic Sea when conditions were suitable for pelagic spawning in most of the area (6–5 kya) and not exclusively, as today, in a few deep offshore basins. Such allopatric phase lasting 1,500/2,500 y might have prevented any homogenizing effect of gene flow in the early days of speciation. Other cases of extremely rapid ecological speciation, such as the evolution of benthic and limnetic species pairs of three-spined sticklebacks in postglacial lakes, are also thought to have started via double colonization and early divergence in allopatry (36).

This hypothesis is compatible with the double-invasion scenario identified by ABC-RF. The estimated time of divergence between demersal flounders and the North Sea population has a mode of 2,400 generations (95% CI: 1,450–9,110). Within the Baltic, flounders reach sexual maturity around 2 y of age (37), although in the northernmost locations sexual maturity in females is delayed for another 1–2 y (38). Hence the coalescence estimates suggest an early colonization of the Baltic Sea (7.2 kya, assuming a generation time of 3 y). Divergence time estimates of pelagic flounders from the ancestral population are also consistent with the hypothesis of an invasion during peak salinity 6–5 kya.

The speciation of a marine vertebrate in such a short time frame has never been reported before. Documented examples of ecological speciation in the marine environment are extremely rare (8) and have usually required much longer time frames (9). The speciation of the Baltic flounder species pair must have initiated no earlier than 8.5 kya, when the Baltic Sea first became connected to the North Sea. Other ecological speciation events have been recorded over similar time frames (<10,000 y), but such events involve species that reach sexual maturity within 1 or 2 y and none are from marine environments (32, 33). Hence, in terms of number of generations, speciation of the Baltic flounder species has likely happened faster than in those examples. To the best of our knowledge, this would place the speciation of the Baltic flounder species pair as one of the most rapid events of ecological speciation, and the fastest speciation event ever recorded for a marine vertebrate.

**Conservation Implications.** The discovery that European flounders in the Baltic Sea represent a pair of closely related species calls for a reassessment of the species pair's conservation status by the International Union for the Conservation of Nature. The fact that both species co-occur during feeding in a part of their range suggests that harvesting should be managed within the framework of a multispecies fishery.

Climate change is predicted to increase freshwater runoff in the Baltic Sea, causing a reduction in salinity and hence contracting the distribution of many marine species by hundreds of kilometers (39). At the same time, eutrophication and climate change are causing a rapid increase in hypoxia and anoxia in bottom waters where salinity is suitable for pelagic spawning (40). Consequently, the spawning habitat of pelagic flounders, which is already geographically limited, will likely contract in the future, raising concerns over a possible local extinction of this species. The endemic demersal spawners might prove more resilient, despite local declines of northern populations (41), but could experience a possible southward range shift. Young species pairs arising from recent events of ecological speciation may be prone to species collapse if anthropogenic activities alter the very factors that led to speciation, a process known as reverse speciation (42). It is possible that a decrease in salinity in the Baltic Sea linked to climate change may provide a strong selective pressure on pelagic spawners, leading to a second behavioral shift to demersal spawning and the collapse of this species pair.

## Methods

**Sampling.** A total of 282 samples was collected from 13 locations spanning nearly the entire distribution of European flounders in the Baltic Sea, including samples from the North Sea and the transition zone between the two regions. Sample collection took place in 2002–2004, 2009, and 2012. Details of sampling locations are given in Table S1. Most samples were collected at the time of spawning in either coastal locations (targeting demersal spawners) or offshore spawning grounds (targeting pelagic spawners). In Bornholm and Öland, almost all individuals (100% and 92%, respectively) were spawning during sampling in a pelagic (Bornholm) and coastal benthic (Öland) environment. Hence, samples from these two locations can be used as references to test for genetic differentiation among flounders exhibiting contrasting spawning behaviors. In some locations (Gdynia, Dabki, and Hanko; Fig. 1), sampling took place shortly after spawning (Table S1). Samples from Swedish waters were collected under a permit issued by the Swedish Board of Fisheries and samples from Hanko under a permit granted by Tvärminne Zoological Station (University of Helsinki). Samples from other locations were obtained from commercial fisheries. The described scientific sampling did not require ethical permission according to the Finnish Animal Conservation Law (75/28.6.2013/498).

**SNP Genotyping.** DNA was extracted from fin clips or muscle tissues using either a standard salting out protocol or a DNeasy Blood and Tissue Kit (Qiagen). Library preparation and sequencing was carried out by Diversity Arrays Technology Pty Ltd using the standard DArTSeq. DArTSeq is a SNP genotyping-by-sequencing approach that combines genome complexity reduction via double-enzymatic digestion characteristic of DArT markers and sequencing on Illumina platforms (43). We initially tested four pairs of enzymes for complexity reduction and selected the PstI and SphI combination. Library preparation and sequencing were carried out as per Bookmythe et al. (44). Libraries were sequenced on three lanes of an Illumina HiSeq2000 platform. De novo assembly and SNP calling was performed using the pyRAD pipeline (45), using a minimum read depth of 10, a maximum read depth equal to the mean read depth plus two times the SD, a clustering threshold of 0.94, the strict quality filtering option, a minimum coverage of 95% of sampled individuals, and a maximum of five SNPs per locus. Only one random SNP per locus was retained for further analyses to avoid creating a set of tightly linked markers. We carried out the entire procedure, from library preparation to SNP calling, a second time for 60 technical replicates and retained only biallelic loci with 100% reproducibility (i.e., no genotyping errors). Loci that deviated from Hardy–Weinberg equilibrium in at least three locations (one representative of each genetic cluster: KAT from the North Sea and at least one demersal and one pelagic Baltic locations) were removed.

**Analyses.** A PCA on allele frequency data was performed in the R package *adegenet* (46). The analysis was repeated after removing loci from the right and left 5% tails of the  $F_{ST}$  distributions obtained when comparing the main genetic clusters (North Sea, Baltic pelagic, and Baltic demersal). This “neu-

tral” dataset should not include loci under strong divergent or stabilizing selection and should therefore be reflective of demographic processes such as genetic drift and migration. Genetic structure was further investigated using fastSTRUCTURE (47) and a DAPC (48). The most likely number of genetic clusters was determined by (i) running fastSTRUCTURE at multiple numbers of  $K$  using fivefold cross-validation and comparing the log-marginal likelihood lower bound (LLBO) and prediction error across increasing values of  $K$ ; and (ii) running  $K$ -means clustering at multiple values of  $K$  and comparing Akaike information criterion (AIC) and Bayesian information criterion (BIC), as per Jombart et al. (48). Weir and Cockerham  $F_{ST}$  among locations and their 95% CI were estimated with 100 bootstraps in the R package *diversity* (49), and geographic distance among locations was estimated as the least cost path distance over seawater in the R package *marmap* (50). An estimate of number of migrants per generation was obtained using the private allele method (23). Migration rates between pelagic and demersal flounders within the Baltic were estimated using BAYESASS (51). For the latter analyses, pelagic and demersal flounders were grouped according to the results from PCA, fastSTRUCTURE, and DAPC; we ran five independent runs, each 10 million generations long, using a burn-in of 2 million generations.

We used four outlier tests to look for signatures of selection: BAYESCAN (52); OutFLANK (53); the coalescent method from Beaumont and Nichols (54) implemented in the software Arlequin, version 3.5 (55) (referred to as the Fdist method); and FLK (56). The tests were carried out first on a dataset where the samples were grouped as two populations representing the two major genetic clusters to identify outlier loci under selection among pelagic and demersal flounders. The tests were then carried out on pelagic and demersal flounders separately, grouping individuals according to geographical location to look for signatures of spatially diversifying selection. The FLK tests were only carried out for the latter analyses, using 24 individuals from the opposite taxon as outgroup population. For BAYESCAN analyses, we set prior odds to 100 to minimize chances of false positives and ran 20 pilot runs, followed by 100,000 iterations (5,000 samples, a thinning interval of 10, and a burn-in of 50,000). For the Fdist analyses, we used the island model implemented in Arlequin in cases where hierarchical genetic structure were not discovered by previous analyses, running 100 simulated demes and 20,000 coalescent simulations. For the detection of outliers between pelagic flounders in the Baltic Sea and the North Sea, we used the hierarchical, coalescent-based approach outlined by Excoffier et al. (57). In this analysis, locations were split in three groups according to PCA results: North Sea, Baltic Pelagic, and samples from the transition zone (Barsebäck). We used 10 simulated groups, 100 simulated demes for each group, and ran 20,000 coalescent simulations. OutFLANK analyses were carried out as outlined by Whitlock and Lotterhos (53). FLK analyses were carried out using the R and Python codes provided by the authors (available at <https://qgsp.jouy.inra.fr/>). Loci that were identified as outliers by all tests were considered as putatively under selection.

Maximum-likelihood (ML) and Bayesian phylogenies were constructed using a subsample of 57 individuals (32 putative pelagic and 25 putative demersal individuals). The alignment included the concatenated sequences of all 2,051 loci (for a total length of 135,164 bp). The ML Smart Model Selection (SMS) approach implemented in PHYML (58) was used to identify the evolutionary model (GTR+G), and ML and Bayesian phylogenetic analyses were carried out using the software PHYML (58) and MrBayes (59). ML analysis was performed using 100 bootstraps, 10 random starting trees and tree improvement by using the best of the nearest-neighbor interchange and subtree pruning and regrafting. Bayesian analyses were run for 3 million generations using two independent runs and 32 chains for each run. Convergence among runs was tested by checking SD of split frequencies fell below 0.05, and stability of each parameter was visually checked using the software TRACER (60). Convergence was reached around 2 million generations, and therefore any tree sampled before then was discarded.

We evaluated possible divergence scenarios and past demographic events under an ABC framework (see [Supporting Information](#) for a detailed description). We formulated competing scenarios describing the demographic history of *P. flesus* in the North Sea and the Baltic Sea (Fig. 5 and Table S5) and simulated datasets under these scenarios using the software DIYABC, version 2.1.0 (61). Model choice was performed using the recently developed ABC-RF approach (28) to evaluate the most probable scenario corresponding to our data, as well as its associated posterior probability. After ensuring the stability of model choice results for several demographic parameters (Fig. 5 and Tables S5–S9), we estimated divergence time for both demersal and pelagic flounders using the parameter estimation analysis implemented in DIYABC on the 1% closest simulated data. The adequacy of our final model was tested by performing the model-posterior checking analysis implemented in DIYABC (Fig. S4).

**ACKNOWLEDGMENTS.** We thank A. Stow, F. Calboli, A. Estoup, Z. Li, T. Shikano, B. Guo, and J. DeFaveri for discussion and comments. Thanks are due to the collectors of the 2003 samples in the study by Florin and Höglund (15). ABC-RF analyses were

- Schluter D, Rambaut A (1996) Ecological speciation in postglacial fishes. *Philos Trans R Soc Lond B Biol Sci* 351:807–814.
- Hendry AP, Nosil P, Rieseberg LH (2007) The speed of ecological speciation. *Funct Ecol* 21:455–464.
- Nosil P, Harmon LJ, Seehausen O (2009) Ecological explanations for (incomplete) speciation. *Trends Ecol Evol* 24:145–156.
- Nosil P (2012) *Ecological Speciation* (Oxford Univ Press, Oxford).
- Servedio MR, Van Doorn GS, Kopp M, Frame AM, Nosil P (2011) Magic traits in speciation: “Magic” but not rare? *Trends Ecol Evol* 26:389–397.
- Jiggins CD, Naisbit RE, Coe RL, Mallet J (2001) Reproductive isolation caused by colour pattern mimicry. *Nature* 411:302–305.
- Ogden R, Thorpe RS (2002) Molecular evidence for ecological speciation in tropical habitats. *Proc Natl Acad Sci USA* 99:13612–13615.
- Puebla O (2009) Ecological speciation in marine v. freshwater fishes. *J Fish Biol* 75: 960–996.
- Munday PL, van Herwerden L, Dudgeon CL (2004) Evidence for sympatric speciation by host shift in the sea. *Curr Biol* 14:1498–1504.
- Puritz JB, et al. (2012) Extraordinarily rapid life-history divergence between *Cryptasterina* sea star species. *Proc R Soc Lond Ser B Biol Sci* 279:3914–3922.
- Gustafsson BG, Westman P (2002) On the causes for salinity variations in the Baltic Sea during the last 8500 years. *Paleoceanography* 17:12–11–12–14.
- Johannesson K, André C (2006) Life on the margin: Genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. *Mol Ecol* 15:2013–2029.
- Nissling A, Westin L, Hjerne O (2002) Reproductive success in relation to salinity for three flatfish species, dab (*Limanda limanda*), plaice (*Pleuronectes platessa*), and flounder (*Pleuronectes flesus*), in the brackish water Baltic Sea. *ICES J Mar Sci* 59:93–108.
- Lønning S, Solemdal P (1972) The relation between thickness of chorion and specific gravity of eggs from Norwegian and Baltic flatfish populations. *Fiskeridirektoratets skrifter. Serie Havundersøkelse* 16:77–88.
- Florin AB, Höglund J (2008) Population structure of flounder (*Platichthys flesus*) in the Baltic Sea: Differences among demersal and pelagic spawners. *Heredity (Edinb)* 101:27–38.
- Hemmer-Hansen J, Nielsen EE, Gronkjaer P, Loeschcke V (2007) Evolutionary mechanisms shaping the genetic population structure of marine fishes; lessons from the European flounder (*Platichthys flesus* L.). *Mol Ecol* 16:3104–3118.
- Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: F<sub>ST</sub> not equal to 1/(4Nm + 1). *Heredity (Edinb)* 82:117–125.
- Florin AB, Höglund J (2007) Absence of population structure of turbot (*Psetta maxima*) in the Baltic Sea. *Mol Ecol* 16:115–126.
- Wallin I (2016) Opportunities for hybridization between two sympatric flounder (*Platichthys flesus*) ecotypes in the Baltic Sea. Master's thesis (Uppsala University, Uppsala, Sweden).
- Via S (1999) Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution* 53:1446–1457.
- Nosil P, Feder JL (2012) Genomic divergence during speciation: Causes and consequences. *Philos Trans R Soc Lond B Biol Sci* 367:332–342.
- Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics* 162:2025–2035.
- Slatkin M, Barton NH (1989) A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* 43:1349–1368.
- Zhang G, et al. (2016) Molecular basis for adaptation of oysters to stressful marine intertidal environments. *Annu Rev Anim Biosci* 4:357–381.
- Zhao X, Yu H, Kong L, Li Q (2012) Transcriptomic responses to salinity stress in the Pacific oyster *Crassostrea gigas*. *PLoS One* 7:e46244.
- Boutet I, Long Ky CL, Bonhomme F (2006) A transcriptomic approach of salinity response in the euryhaline teleost, *Dicentrarchus labrax*. *Gene* 379:40–50.
- Cosson J, et al. (2008) Studying sperm motility in marine fish: An overview on the state of the art. *J Appl Ichthyology* 24:460–486.
- Pudlo P, et al. (2016) Reliable ABC model choice via random forests. *Bioinformatics* 32: 859–866.
- Cornuet J-M, et al. (2008) Inferring population history with DIY ABC: A user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24:2713–2719.
- Nissling A, Dahlman G (2010) Fecundity of flounder, *Pleuronectes flesus*, in the Baltic Sea—reproductive strategies in two sympatric populations. *J Sea Res* 64:190–198.
- Nissling A, Thorsen A, da Silva FF (2015) Fecundity regulation in relation to habitat utilisation of two sympatric flounder (*Platichthys flesus*) populations in the brackish water Baltic Sea. *J Sea Res* 95:188–195.
- Barluenga M, Stöltting KN, Salzburger W, Muschick M, Meyer A (2006) Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* 439:719–723.
- McPhail JD (1992) Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): Evidence for a species-pair in Paxton Lake, Texada Island, British Columbia. *Can J Zool* 70:361–369.
- Hatfield T, Schluter D (1999) Ecological speciation in sticklebacks: Environment-dependent hybrid fitness. *Evolution* 53:866–873.
- Larsen PF, Nielsen EE, Williams TD, Loeschcke V (2008) Intraspecific variation in expression of candidate genes for osmoregulation, heme biosynthesis and stress resistance suggests local adaptation in European flounder (*Platichthys flesus*). *Heredity (Edinb)* 101:247–259.
- Taylor EB, McPhail JD (2000) Historical contingency and ecological determinism interact to prime speciation in sticklebacks, *Gasterosteus*. *Proc R Soc Lond Ser B Biol Sci* 267:2375–2384.
- performed on the CBGP HPC and UMS 2700 OMSI computational platforms. This project was funded by the the Maj & Tor Nessling Foundation, the Academy of Finland (project 292737), and the Walter and Andrée de Nottbeck Foundation.
- ICES (2016) *ICES WGBFAS REPORT 2016: Report of the Baltic Fisheries Assessment Working Group (WGBFAS)* (ICES, Copenhagen).
- Erlundsson J, Östman Ö, Florin A-B, Pekcan-Hekim Z (2017) Spatial structure of body size of European flounder (*Platichthys flesus* L.) in the Baltic Sea. *Fish Res* 189:1–9.
- Vuorinen I, et al. (2015) Scenario simulations of future salinity and ecological consequences in the Baltic Sea and adjacent North Sea areas-implications for environmental monitoring. *Ecol Indic* 50:196–205.
- Carstensen J, Andersen JH, Gustafsson BG, Conley DJ (2014) Deoxygenation of the Baltic Sea during the last century. *Proc Natl Acad Sci USA* 111:5628–5633.
- Jokinen H, et al. (2015) Decline of flounder (*Platichthys flesus* L.) at the margin of the species' distribution range. *J Sea Res* 105:1–9.
- Seehausen O (2006) Conservation: Losing biodiversity by reverse speciation. *Curr Biol* 16:R334–R337.
- Sansaloni C, et al. (2011) Diversity Arrays Technology (DArT) and next-generation sequencing combined: Genome-wide, high throughput, highly informative genotyping for molecular breeding of *Eucalyptus*. *BMC Proc* 5:P54.
- Bookmythe I, Head ML, Keogh JS, Jennions MD (2016) Fitness consequences of artificial selection on relative male genital size. *Nat Commun* 7:11597.
- Eaton DA (2014) PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30:1844–1849.
- Jombart T (2008) adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405.
- Raj A, Stephens M, Pritchard JK (2014) fastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics* 197:573–589.
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genet* 11: 94.
- Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA (2013) diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol Evol* 4:782–788.
- Pante E, Simon-Bouhet B (2013) marmap: A package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS One* 8:e73051.
- Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163:1177–1191.
- Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* 180:977–993.
- Whitlock MC, Lotterhos KE (2015) Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of F<sub>ST</sub>\*. *Am Nat* 186:S24–S36.
- Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proc R Soc Lond Ser B Biol Sci* 263:1619.
- Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10: 564–567.
- Bonhomme M, et al. (2010) Detecting selection in population trees: The Lewontin and Krakauer test extended. *Genetics* 186:241–262.
- Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity (Edinb)* 103:285–298.
- Guindon S, et al. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol* 59:307–321.
- Ronquist F, et al. (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61:539–542.
- Rambaut A, Suchard MA, Xie D, Drummon AJ (2014) Tracer, version 1.6. Available at [beast.bio.ed.ac.uk/Tracer](http://beast.bio.ed.ac.uk/Tracer). Accessed August 12, 2016.
- Cornuet J-M, et al. (2014) DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30:1187–1189.
- Bendtsen J, Söderkvist J, Dahl K, Hansen JL, Reker J (2007) Model simulations of blue corridors in the Baltic Sea. *BALANCE Interim Report* 9 (BALANCE, Copenhagen).
- Verdu P, et al. (2009) Origins and genetic diversity of pygmy hunter-gatherers from Western Central Africa. *Curr Biol* 19:312–318.
- Fraimout A, et al. (2017) Deciphering the routes of invasion of *Drosophila sukukii* by means of ABC random forest. *Mol Biol Evol* 34:980–996.
- Lombaert E, et al. (2014) Complementarity of statistical treatments to reconstruct worldwide routes of invasion: The case of the Asian ladybird *Harmonia axyridis*. *Mol Ecol* 23:5979–5997.
- Cornuet J-M, Ravigné V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics* 11:401.
- Siebert PD, Chenchik A, Kellogg DE, Lukyanov KA, Lukyanov SA (1995) An improved PCR method for walking in unclosed genomic DNA. *Nucleic Acids Res* 23:1087–1088.
- Tanoue T, Takeichi M (2005) New insights into Fat cadherins. *J Cell Sci* 118:2347–2353.
- Meyer zu Heringdorf D, Jakobs KH (2007) Lysophospholipid receptors: Signalling, pharmacology and regulation by lysophospholipid metabolism. *Biochim Biophys Acta* 1768:923–940.