Received: 26 February 2020 | Revised: 17 December 2020 | Accepted: 4 January 2021

DOI: 10.1111/mec.15791

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Timing strains of the marine insect *Clunio marinus* diverged and persist with gene flow

Tobias S. Kaiser[1,2,3,4] | **Arndt von Haeseler**[2,5] | **Kristin Tessmar-Raible**[3] | **David G. Heckel**[4]

[1]Max Planck Research Group Biological Clocks, Max Planck Institute for Evolutionary Biology, Plön, Germany

[2]Center for Integrative Bioinformatics Vienna, Max Perutz Laboratories, University of Vienna and Medical University of Vienna, Vienna, Austria

[3]Max Perutz Laboratories, University of Vienna, Vienna, Austria

[4]Department of Entomology, Max Planck Institute for Chemical Ecology, Jena, Germany

[5]Bioinformatics and Computational Biology, Faculty of Computer Science, University of Vienna, Vienna, Austria

Correspondence
Tobias S. Kaiser, Max Planck Research Group Biological Clocks, Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany.
Email: kaiser@evolbio.mpg.de

Funding information
H2020 European Research Council, Grant/Award Number: 802923; Max-Planck-Gesellschaft; FP7 Ideas: European Research Council, Grant/Award Number: 227799 and 337011

Abstract

Genetic divergence of populations in the presence of gene flow is a central theme in speciation research. Theory predicts that divergence can happen with full range overlap – in sympatry – driven by ecological factors, but there are few empirical examples of how ecologically divergent selection can overcome gene flow and lead to reproductive isolation. In the marine midge *Clunio marinus* (Diptera: Chironomidae) reproduction is ecologically restricted to the time of the lowest tides, which is ensured through accurate control of development and adult emergence by circalunar and circadian clocks. As tidal regimes differ along the coastline, locally adapted timing strains of *C. marinus* are found in different sites across Europe. At the same time, ecologically suitable low tides occur at both full and new moon and twice a day, providing *C. marinus* with four nonoverlapping temporal niches at every geographic location. Along the coast of Brittany, which is characterized by a steep gradient in timing of the tides, we found an unusually large number of differentially adapted timing strains, and the first known instances of sympatric *C. marinus* strains occupying divergent temporal niches. Analysis of mitochondrial genotypes suggests that these timing strains originated from a single recent colonization event. Nuclear genotypes show strong gene flow, sympatric timing strains being the least differentiated. Even when sympatric strains exist in nonoverlapping temporal niches, timing adaptations do not result in genome-wide genetic divergence, suggesting timing adaptations are maintained by permanent ecological selection. This constitutes a model case for incipient ecological divergence with gene flow.

KEYWORDS
allochrony, circadian clocks, circalunar clocks, ecological speciation, local adaptation, sympatric speciation

1 | INTRODUCTION

Evolutionary biologists have long debated whether speciation can happen only with geographic isolation (allopatric speciation) or also with full range overlap between diverging populations (sympatric speciation) (Bolnick & Fitzpatrick, 2007; Coyne & Orr, 2004; Fitzpatrick et al., 2008, 2009; Foote, 2018; Mallet et al., 2009; Maynard Smith, 1966; Mayr, 1947; Via, 2001). Today, the occurrence of sympatric speciation is accepted, backed by mathematical models (Barton, 2010; Bolnick & Fitzpatrick, 2007; Dieckmann & Doebeli, 1999; Kondrashov & Kondrashov, 1999; Maynard Smith, 1966), laboratory experiments (Rice & Hostert, 1993; Thoday & Gibson, 1962) and empirical examples of sister species in isolated habitats such as crater lake cichlids (Barluenga et al., 2006; Kautt et al., 2016; Malinsky et al., 2015; Schliewen et al., 1994) or island palms (Savolainen et al., 2006). The focus of speciation research has expanded from specific geographic settings to a more general notion of population divergence with gene flow (Rice & Hostert, 1993; Richards et al., 2019; Smadja & Butlin, 2011), which may, but need not lead to speciation and in which allopatric speciation and sympatric speciation can be considered the extreme end points on a continuum of levels of gene flow. Recent advances in speciation genomics (Campbell et al., 2018; Nosil & Feder, 2012; Seehausen et al., 2014), have highlighted well-studied examples of population divergence with gene flow (e.g. Butlin et al., 2014; Doellman et al., 2018; Martin et al., 2013; Riesch et al., 2017). However, there are many open questions, particularly regarding the incidence of different forms of population divergence with gene flow and the driving evolutionary forces (Foote, 2018; Richards et al., 2019).

The concept of population divergence with gene flow intersects with studies of local adaptation (Kawecki & Ebert, 2004; Savolainen et al., 2013), as divergence with gene flow is generally assumed to start with divergent natural selection (Feder et al., 2012; Seehausen et al., 2014; Smadja & Butlin, 2011). For this process to result in speciation, reproductive isolation must evolve. However, it is still largely unclear how adaptive ecological divergence is maintained against gene flow at early stages of the process, when no other isolating factors exist. In principle, occupying different ecological niches can by itself reduce gene flow between diverging populations. As all abiotic or biotic factors can be part of a species' ecological niche, countless factors can potentially drive divergent ecological adaptation and many have been considered to play a role in population divergence with gene flow, e.g., soil parameters (Savolainen et al., 2006) or biotic interactions with host plants (Dres & Mallet, 2002).

An important dimension of the ecological niche is an organism's timing of activity, life history and reproduction, its temporal niche (see e.g. Häfker & Tessmar-Raible, 2020; Hut et al., 2012). Allochrony, i.e., differences in timing between individuals, populations or species, can lead to isolation by time (IBT, Hendry & Day, 2005) and has been considered a major factor to facilitate sympatric speciation (Taylor & Friesen, 2017). Allochrony in reproductive timing automatically entails part of the reproductive isolation needed for speciation. Reproductive timing can thus be a magic trait, i.e., a trait for which ecological divergence and reproductive isolation are directly linked (Gavrilets, 2004). In the context of ecological divergence of populations with gene flow, most studies consider allochrony due to seasonal timing (Helm & Womack, 2018; Ragland et al., 2017; Tauber & Tauber, 1977; Taylor & Friesen, 2017) or daily timing (Devries et al., 2008; Fukami et al., 2003; Hänniger et al., 2017). Here, we demonstrate ecological divergence with gene flow based on daily and lunar-phase timing in the marine midge *Clunio marinus*. We add lunar phase as a time scale to the concept of the temporal niche and underline the importance of temporal niches in the process of population divergence with gene flow.

Clunio marinus (Diptera: Chironomidae) is primarily known as a model organism for studying circalunar clocks (Kaiser et al., 2016; Neumann, 2014), i.e., endogenous biological timekeeping mechanisms that limit reproduction and behaviour of many marine organisms to distinct lunar phases. While the existence of circalunar clocks has been repeatedly demonstrated (Neumann, 2014), their molecular basis is unknown. In *C. marinus* the circalunar clock has a clear ecological relevance (Kaiser, 2014): The marine larvae of *C. marinus* develop at the lowest levels of the intertidal zone, where they are almost permanently submerged by the sea. Adult females can only access these sites for oviposition when the tides recede maximally, which is predictably around full moon and new moon (Figure 1a, bottom panel), and on these days during specific hours (Figure 1a, left and central panel). Adult emergence of *C. marinus* is timed to these occasions by combining a circalunar clock controlling development (Figure 1a , x-dimension) with a circadian clock controlling adult emergence (Figure 1a, y-dimension). As adults live only for 2–3 h, precise timing of development and adult emergence automatically limits mating and oviposition to the suitable low tide events.

While at a given geographic location specific tidal conditions always recur at the same time during the lunar month and daily cycle, the timing of these conditions gradually shifts along the coastline. *C. marinus* populations from different geographic origins are adapted to the local tides in their exact circadian and circalunar emergence times (Kaiser et al., 2011; Neumann, 1967). These timing differences are particularly interesting for studying local adaptation and allochrony. They are precise within a population, but differ widely between populations. The correlation between the local time of spring low tide and *C. marinus'* emergence time is striking (Kaiser et al., 2011). Additionally, timing adaptations can be measured in the laboratory under common garden conditions, they are genetically determined and some of the underlying loci have been identified (Kaiser & Heckel, 2012; Kaiser et al., 2011, 2016; Neumann, 1967).

In widely separated sites, the evolution of local timing adaptations is facilitated by geographic isolation, which largely arises due to the fact that *C. marinus* is restricted to rocky coasts (Kaiser et al., 2010). While populations in previous studies can be considered allopatric because of their occurrence in different stretches of rocky coast, we assumed that there must be parapatric *Clunio* populations within continuous stretches of rocky coast. The current study was motivated by the question: Can genetic timing adaptations persist in the absence of geographic isolation if ecological gradients are

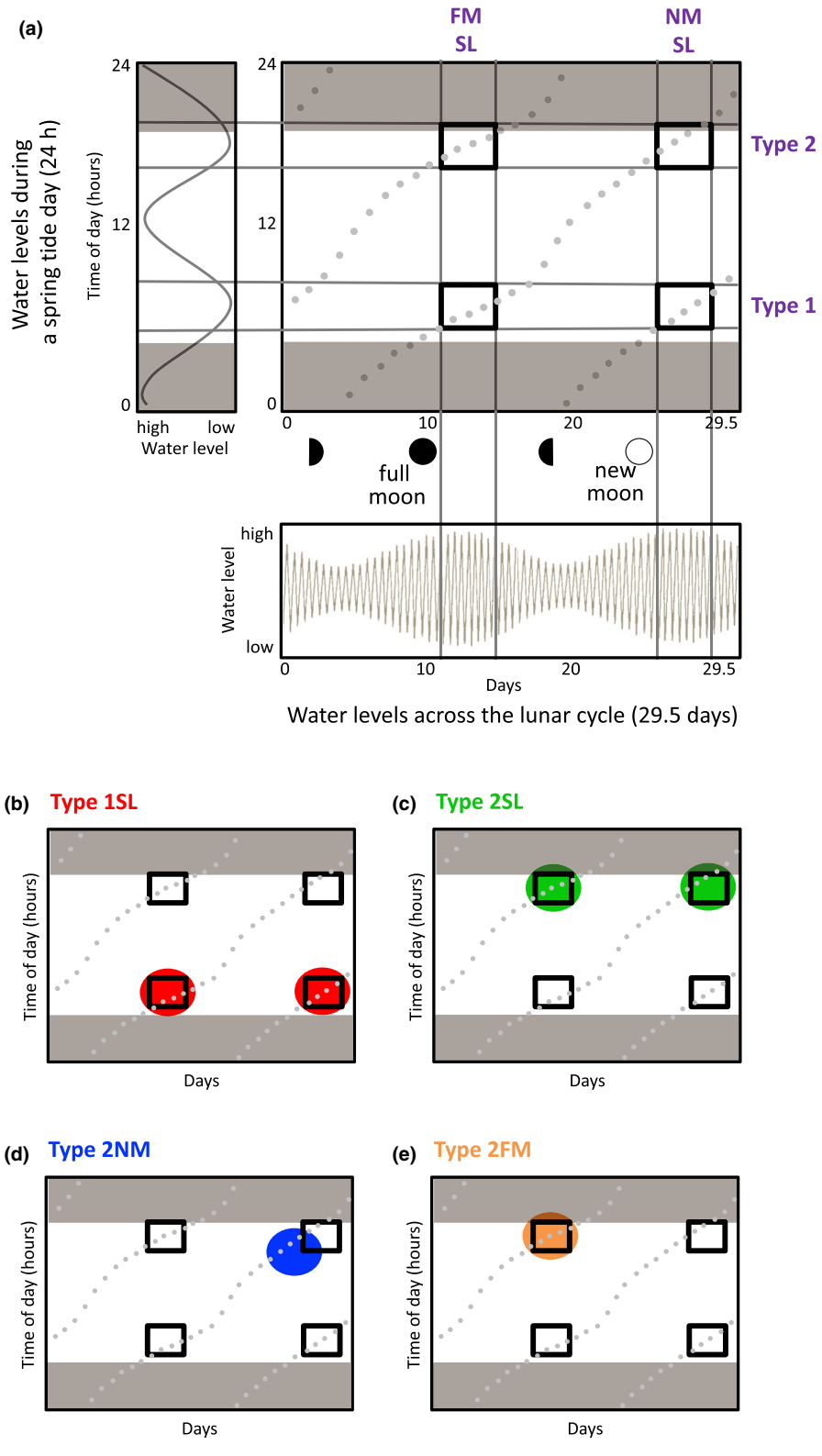

FIGURE 1 Temporal niches and timing types of *Clunio marinus*. (a) Tidal amplitudes are bimodal across the lunar cycle, with lowest low tides occurring around full and new moon (bottom panel in a). Tides are also bimodal during the day, with two low tides a day (left panel in a). The time of low tide shifts across the lunar cycle (grey dots in central panel of a). Grey shading indicates night time. Superposition of the lunar cycle (x-dimension) and the daily cycle (y-dimension) results in four time points with extreme low tides (thick-lined squares in central panel of a), which are distinct temporal niches for reproduction of *Clunio marinus*. In *C. marinus* a circadian clock restricts reproduction to one of the two daily temporal niches, leading to Type 1 and Type 2 strains. A circalunar clock restricts reproduction to the lunar temporal niches, resulting in full moon strains (FM), new moon strains (NM), or strains emerging at both occasions (SL for "semilunar"). (b–e) Timing strains of *C. marinus* have specific combinations of circadian and circalunar timing. Strains can therefore be classified by temporal niche into the timing Types 1SL (b), 2SL (c), 2NM (d) and 2FM (e). Type 2NM strains generally emerge slightly earlier than optimal (d). Type 1NM or 1FM strains have not been found

strong, i.e. along a continuously rocky coast with large differences in the timing of the tides? Consulting tide tables of the European Atlantic Coast, we identified the coastline of Brittany and Normandy in France as a suitable study area. The tidal amplitude in this region reaches up to 12 m, there is a steep gradient in timing of the tides, and the coastline is predominantly rocky. In a long-term project that started in 2009, we explored eight sites along the coastline (Figure 2). We not only found the expected parapatric *C. marinus* populations with gradual timing adaptations, but also several sympatric populations which occupy divergent temporal niches and which we therefore introduce here as different timing types of *C. marinus*. As the English Channel did not exist during the last ice age (Patton et al.,

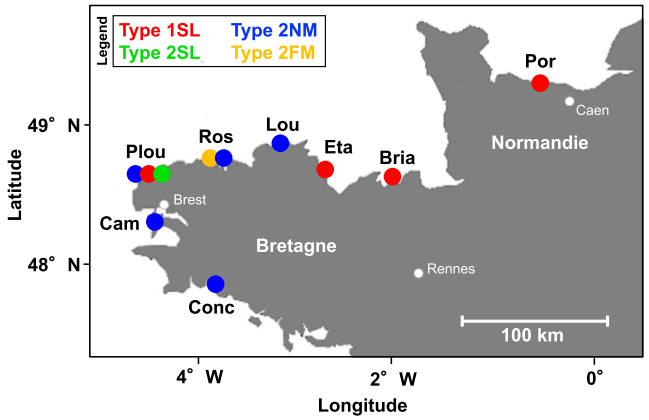

FIGURE 2 Sampling sites and *C. marinus* timing types found in these sites. A map of Brittany and Normandy shows the eight sampling sites at Concarneau (Conc), Camaret-sur-Mer (Cam), Plouguerneau (Plou), Roscoff (Ros), Louannec (Lou), Etables-sur-Mer (Eta), St. Briac-sur-Mer (Bria) and Port-en-Bessin (Por). Coloured dots represent the timing strains that were found in these sites. The colours correspond to the different timing types as described in Figure 1. In Plouguerneau and Roscoff several timing types co-exist in sympatry

2017), sympatric coexistence of timing types must have established recently. We hypothesize that either different allopatric timing types of *C. marinus* have colonized Brittany independently and are maintained by reproductive isolation through allochrony, or that Brittany was colonized only once and sympatric timing types have evolved on site with gene flow. To discriminate between these scenarios, first, we established laboratory strains from all populations we found and validated the existence of distinct sympatric timing strains under common garden conditions. Second, we genetically analysed 24 individuals from each population to elucidate the evolutionary history of these timing strains, i.e. whether they came from one or several sources. Third, we assessed gene flow between parapatric and sympatric populations to see whether it is plausible that local adaptation in timing is maintained and may have evolved with gene flow. Fourth, in order to see if ecological timing adaptations result in reproductive isolation, we asked whether timing differences could explain part of the genetic differentiation between timing strains.

2 | MATERIALS AND METHODS

2.1 | Definition of temporal niches and nomenclature of *C. marinus* timing strains

Reproduction of *C. marinus* is confined to the lowest tides. The semidiurnal tides of the European Atlantic coast produce four distinct occasions with very low tides (Figure 1a). Low tides are particularly low during the spring tide days, which occur just after full moon and new moon (bottom panel in Figure 1a). Additionally, there are two low tides per day (left panel in Figure 1a). The time of low tide shifts from day to day, but at full moon and new moon they always occur at about the same time of day (grey dots in central panel of Figure 1a). Superposition of the lunar and daily temporal dimensions results in four nonoverlapping temporal niches that are suitable for reproduction of *C. marinus* (thick-lined black boxes in the central panel of Figure 1a).

Timing strains of *C. marinus* are characterised by distinct, genetically determined combinations of circadian and circalunar timing, which synchronize reproduction to a suitable temporal niche. Timing strains were previously only reported for separate geographic sites and thus named by geographic origin. Here, we report timing strains that co-exist in the same site but reproduce in different temporal niches, which requires amending their nomenclature and distinguishing them by the temporal niche(s) they occupy.

When classified by temporal niche, the timing strains fall into four distinct groups, which we define here as ecological timing types. Firstly, as the circadian clock limits adult emergence of *C. marinus* to one of the two daily low tides (y-dimension in Figure 1a), we can distinguish timing strains that emerge during the first low tide after sunrise ("Type 1") from timing strains that emerge during the second low tide after sunrise ("Type 2"; see Figure 1a). Secondly, the circalunar clock limits adult emergence to a specific lunar phase (x-dimension in Figure 1a). Some *C. marinus* timing strains emerge only during full moon (lunar rhythm; "Type FM"), some only during new moon (lunar rhythm; "Type NM") and some during both (a so-called semilunar rhythm; therefore "Type SL"; see Figure 1a). Combining all possible circadian and circalunar phenotypes theoretically results in six possible timing types (1FM, 2FM, 1NM, 2NM, 1SL, 2SL) but so far only 1SL, 2SL, 2NM and 2FM strains have been observed (Figure 1b–e; Table S1). We name timing strains based on a shorthand for their geographic origin and their timing type (Table S1). As the timing of the tides shifts gradually along the coastline, timing strains within each timing type differ slightly in circadian and circalunar emergence times in adaptation to the local tides.

On a side note, the SL type is not a mixture of FM and NM types nor the product of a genetic polymorphism. Laboratory experiments show that SL type strains have an endogenous circalunar clock with a ~15-day period, whereas NM and FM strains have circalunar clocks with a ~30-day period (Neumann, 1966). As a result, irrespective of whether SL type parents emerge during full moon or new moon, their offspring emerge during full moon and new moon in a roughly 1:1 ratio, over the course of 1–2 months.

2.2 | Fieldwork and establishment of laboratory strains

Field samples and laboratory strains for this study originate from five field trips undertaken between 2009 and 2017 (Table S2). From a total of eight sites we identified 11 different timing strains (Table S2, Figure 2).

We applied two methods for sampling. First, we collected larvae by collecting small patches of algae and sand and transferred them

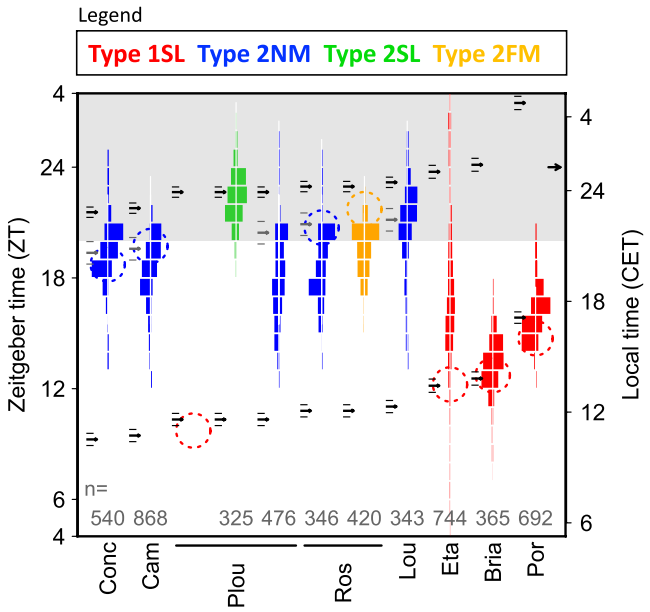

FIGURE 3 Circadian timing of the timing strains found in Brittany and Normandy. Daily emergence times were recorded from laboratory strains under common garden conditions (LD 16:8, 20°C) without tidal timing cues. Numbers of emerging males are plotted on the left side of a white line in 1 h intervals. Females are plotted on the right. Total numbers of individuals (n) are given at the bottom. Plots are coloured according to timing types. Males generally emerge 1–2 h before females (protandry). Grey shading indicates dark phase in the laboratory light-dark (LD) cycle (left *y*-axis). For comparison, we plotted the time of low tide in local time (CET, right *y*-axis). Laboratory LD cycle (left *y*-axis) and local time (right *y*-axis) are matched so that the middle of dark phase corresponds to the middle of the night (black arrow on the right y-axis, 1:16 CET). Small black arrows and lines in the plot are the mean and standard deviations of time of low tide on full and new moon days at the place of origin of the respective timing strain. For type 2NM strains, which generally emerge a few days before new moon, the mean time of low tide at three days before new moon is given as an additional reference point (small grey arrows and lines). Field observations of swarming adults are given in local time as dotted circles. For the Plou-1SL strain only field observations are available

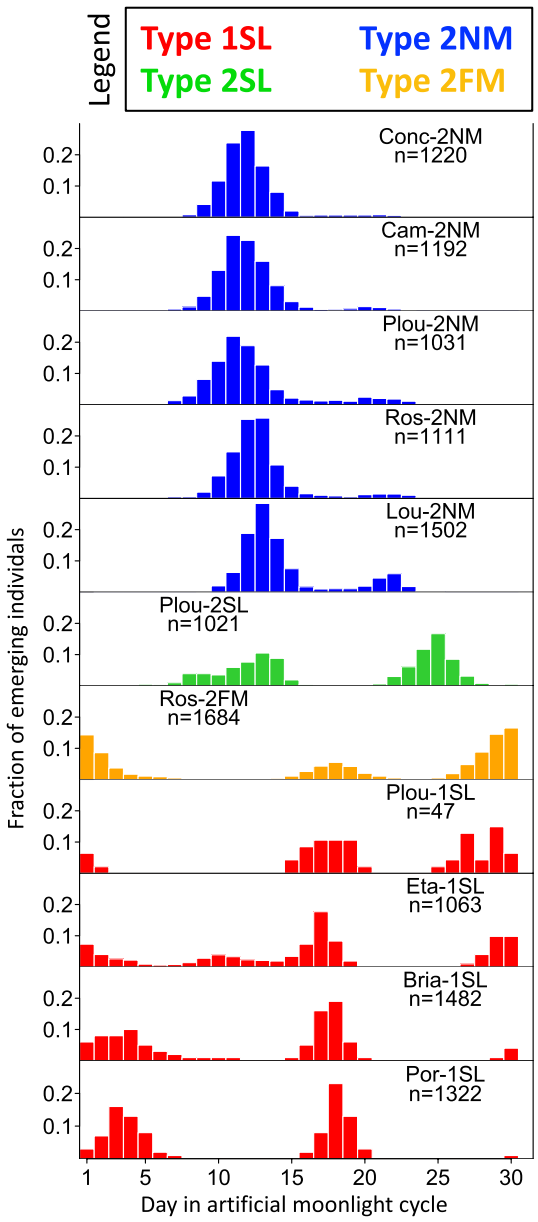

FIGURE 4 Circalunar timing of the timing strains found in Brittany and Normandy. Lunar emergence times were recorded from laboratory strains under common garden conditions (LD 16:8, 20°C). Artificial moonlight was presented during four nights (days 1–5) in a 30-day artificial moonlight cycle (*x*-axis). As biological rhythms are cyclic, day 30 is equivalent to the day before day 1. The bars show the fraction of emerging individuals on each day of the artificial moonlight cycle. The plots are coloured according to timing type. Within timing types the strains are sorted by geography, highlighting the gradual shift of circalunar emergence times along the coast, corresponding to the shift in the local tides. The small peaks around day 22 in the 2NM strains and day 18 in the 2FM strain are artefacts of the artificial moonlight treatment and have not been observed in the field (see Figure S1 for details)

to the laboratory. Adults emerging from these samples reproduce readily and served to establish laboratory strains. Second, we sampled swarming adults in the field for genetic analysis. If these samples included copulae, we also established laboratory strains from their egg clutches. The reasoning behind this two-fold strategy and necessary deviations is described below.

We sampled larvae for all sites except for Concarneau and Camaret-sur-Mer (Table S2). As developmental time of *C. marinus* varies between 6–12 weeks, even for siblings (Neumann, 1966), generations of different timing types and timing strains overlap. Consequently, at sites with sympatric timing types, larvae are present throughout the lunar cycle in ratios largely proportional to the population sizes of the different timing types. Collecting larval substrates allows us to identify all common timing types at a site, irrespective of the time point of collection.

In the case of sympatric timing types, establishment of laboratory strains from larvae regularly leads to mixed laboratory strains. We applied two different strategies to obtain pure laboratory strains for each timing type. In cases where at least one of the temporal

niches was exclusively occupied by one of the strains (Plou-1SL, Ros-2FM, Ros-2NM), we went to the field during that specific time, caught copulating pairs and established pure laboratory strains from their egg clutches. However, the Plou-2NM strain would always co-emerge with Plou-2SL (see Figures 3 and 4). In this case we sampled larval substrates again, but first established a laboratory strain only from the egg clutches obtained during full moon, resulting in a pure Plou-2SL strain. Egg clutches obtained during new moon were reared individually and resulted in hybrid strains, pure Plou-2SL strains and the required pure Plou-2NM strains.

Hybrid offspring of different timing types have intermediate timing phenotypes (Kaiser et al., 2011; Neumann, 1967). During extensive fieldwork, over the years we have only observed a handful of swarming adults at times during which such hybrids would swarm. In contrast, each day we have observed hundreds to many thousands of swarming adults during the pure timing types' temporal niches. We conclude that timing types are distinct units in nature and that establishing pure timing-type laboratory strains is not an artefact of laboratory selection.

Field samples for genetic analysis consisted of swarming males caught on site. Females could not be found in sufficient numbers, as they are immobile and thus virtually invisible in the field. For Bria-1SL, Lou-2NM, Plou-2NM and Plou-2SL it was logistically impossible to collect adults in the field. Instead, field-caught larvae were transferred to the laboratory and the emerging adults were collected. For Bria-1SL and Lou-2NM this is unproblematic, as there is only one timing strain present in each site. For Plou-2NM and Plou-2SL this procedure was necessary because of the above-mentioned overlap in their emergence time. After rearing these field-caught larvae in the laboratory, for genetic analyses we used only those which had pure Plou-2SL type offspring or pure Plou-2NM type offspring respectively. To avoid sampling genetically related adults, in all cases several hundred larvae were collected from >10 different places in an area >5,000 m^2 and a random subset of 24 individuals was subject to genetic analysis.

2.3 | Characterization of laboratory strains

Timing adaptations, timing types and timing strains were assessed based on laboratory strains under common garden conditions. Except for the Plou-1SL strain, all laboratory strains were maintained for a minimum of six, but usually >20 generations. They were kept in standard culture conditions (Neumann, 1966) at 20°C and subject to a light-dark cycle of 16 h light and 8 h darkness (LD 16:8). Artificial moonlight (from a compact fluorescent lamp) was presented for four subsequent nights every 30 days, resulting in a 30-day artificial moonlight cycle, which entrained the circalunar rhythm in all strains. Lunar emergence times were recorded from the second laboratory generation onward by counting the number of adults that emerged daily. Counts were summed up over several artificial moonlight cycles. Daily emergence times were recorded from the third laboratory generation onward, after the strains had been identified based on their circalunar rhythms. Daily emergence times were recorded in 1-hr intervals with the help of a fraction collector (Honegger, 1977). Emergence was recorded over several generations until for each strain >300 individuals had been assayed for circadian emergence times and >1,000 individuals had been assayed for lunar emergence times. The Ros-2FM and Por-1SL laboratory strains have been described (Kaiser et al., 2011; Neumann, 1966, 1989). The re-established laboratory strains correspond to the original reports. Timing phenotypes were stable throughout the lifetime of a laboratory strain.

2.4 | Acquisition of data on the local tides

For assessment of the match between circadian emergence times and the local tides, we calculated the average time of low tide on full moon and new moon days. We read the times of low tides from tide tables of the year 2009 (Gezeitentafeln, 2008: Europäische Gewässer 2008) and calculated the average time of day and standard deviation across that year (n = 25). As 2NM type strains emerge about three days before new moon, we additionally obtained the average time of the evening low tide three days before new moon in the same way (n = 12). The offset between our laboratory light cycle and local time (CET) in the field was corrected by matching the middle of the dark phase in both reference frameworks. The middle of the dark phase in the laboratory is defined by the experimenter as zeitgeber time 0 (ZT0; Figure 3, left axis). The middle of the dark phase in the field was calculated by obtaining the times of sunset and sunrise in Roscoff for 25 dates evenly distributed over 2009, then calculating the midpoint between sunset and sunrise and finally averaging over the 25 dates. The middle of the dark phase in Roscoff is 1:16 CET (*SD* 9 min; Figure 3, right axis).

2.5 | Acquisition of genetic data

In order to test the evolutionary history and genetic structure of the timing strains and sympatric timing types, we obtained mitochondrial haplotypes and nuclear SNP data for 24 males from all timing strains (23 for Plou-1SL). DNA was extracted with a salting out method (Reineke et al., 1998), and subjected to whole genome amplification (REPLI-g Mini kit, Qiagen) according to the manufacturer's protocol.

For analysis of mitochondrial haplotypes we picked a fragment of cytochrome oxidase subunit I (COI), which discriminated *C. marinus* populations in previous studies (Fuhrmann & Kaiser, 2020; Kaiser et al., 2010). COI sequences were amplified with the primers C1-J-2183 and TL2-N-3014 (Simon et al., 1994), cleaned up with Exonuclease I and Shrimp Alkaline Phosphatase and sequenced from both ends on an ABI PRISM 3730xl DNA Analyzer (Applied Biosystems) with standard sequencing chemistry. For one Lou-2 NM individual, amplification failed repeatedly (resulting in n = 262 individuals). COI sequences were assembled in the CLC Main Workbench

(CLC bio, Aarhus). From a previous study which included the Por-1SL strain (Kaiser et al., 2010) a number of COI haplotypes were already known (GenBank Accessions GU943253.1, GU943254.1, GU943258.1, GU943261.1, GU943262.1). All new haplotypes were submitted to ENA-EBI (Accessions LN851805–LN851838).

For analysis of the nuclear genome, Restriction-site Associated DNA sequencing (RAD-seq) was performed following Baird et al., (2008), using *BamHI* as restriction enzyme and size selection for fragments of 300–800 bp. A combinatorial barcoding approach was used, i.e., both P1 and P2 adapters were barcoded. The 263 individuals were distinguished by the different combinations of barcodes on the two adaptors. A set of 48 custom P1-adapters with 6 bp-barcodes were ligated to the sticky ends and individuals were pooled into groups of 48 or 24 individuals, so that each P1 barcode was unique in each of the pools. These pools were then labelled with different 4-bp barcodes on the P2 adapter. Four pools with 48 samples each were run two by two in two lanes of an Illumina HiSeq2000 sequencer (100-bp paired-end reads). Pools with 24 individuals were run (with other samples) in three lanes of an Illumina HiSeq2000 (100-bp single-end reads). Only forward reads were available for all individuals, so that analysis was restricted to forward reads. Raw data were submitted to ENA-EBI under the project accession PRJEB9361.

The reads were quality trimmed with cutadapt (Martin, 2011; parameters -e 0.1 -n 1 -m 15 -O 8 -q 13). Trimmed reads were aligned to the CLUMA_1.0 reference assembly (Kaiser et al., 2016) with the Burrows-Wheeler-Aligner (*BWA*) (Li & Durbin, 2009), using the aln and samse functions with default parameters. Alignments were filtered for phred-scaled mapping quality of 20 and merged into one sorted and indexed alignment with SAMTOOLS (Li et al., 2009), using the functions view (-q 20 -bS), sort, merge (-r -h) and index. Coverage was determined with SAMTOOLS depth (default parameters). Median read coverage over RAD tags and individuals was 91-fold (Table S6). Single nucleotide polymorphisms (SNPs) and genotypes were called using the *UnifiedGenotyper* implemented in the Genome Analysis Toolkit (GATK) (McKenna et al., 2010), using the parameter -glm SNP to restrict the analysis to SNPs.

Genotypes were filtered in several rounds. We first filtered with VCFTOOLS 0.1.14 (Danecek et al., 2011) for minimum phred-scaled genotype quality of 20 (--minGQ 20), maximum amount of missing data of 20% (--max-missing 0.8), biallelic sites (--max-alleles 2) and minor allele frequency >1% (--maf 0.01; which is minimally six alternative alleles in the 526 chromosomes sampled). This resulted in 5,076 SNPs. We then identified a set of 10 individuals with low sequence coverage and less than 2,500 called genotypes. These were removed from the analysis (5× Conc-2NM, 2× Lou-2NM, 1× Cam-2NM, 1× Bria-1SL and 1× Por-1SL), leaving 19 individuals for Conc-2NM and 22 to 24 individuals for all other populations. Re-filtering the remaining individuals with the above parameters resulted in a set of 5,275 SNPs. In order to account for linkage disequilibrium (LD) between SNPs, we applied a thinning procedure in vcftools 0.1.14, allowing only one SNP in 1,000 bp (--thin 1000). In *C. marinus* LD drops to $r^2 < 0.1$ within 1 kb, similar to *D. melanogaster* (N. Fuhrmann & T. S. Kaiser, unpublished data). Thinning resulted in a reduced set of 2,159 SNPs. The SNP data were transformed to the required file formats with VCFTOOLS 0.1.14 (Danecek et al., 2011) or PGDSpider 2.1.1.5 (Lischer & Excoffier, 2012).

With these genetic data we investigated the evolutionary history of the different timing types and strains in Brittany and Normandy, particularly the number of postglacial colonization sources, the extent of gene flow between them and the question of whether timing differences entail reproductive isolation.

2.6 | Population genetic analyses: Mitochondrial COI sequences

We inferred the number of maternal lineages that have colonized the study area by calculating a haplotype network. COI sequences were aligned in MEGA version 3.0 (Kumar et al., 2004) based on the CLUSTAL W algorithm (Thompson et al., 1994) and a haplotype network was calculated with NETWORK 4.6.1.1 software (Fluxus Technology Ltd), using the Median Joining method (Bandelt et al., 1999). Two multistate positions (i.e., positions with more than two alleles; positions 289 and 436) were excluded from the analysis.

We also assessed isolation by distance (IBD). Pairwise F_{ST} values for all population pairs were calculated in ARLEQUIN version 3.5 (Excoffier & Lischer, 2010) with default parameters. Linearized pairwise F_{ST} values ($F_{ST}/(1\text{-}F_{ST})$) were plotted against geographic distance. Geographic distances along the coastline were measured on a map of Northern France (scale 1:650,000) by tracing the coast in a series of straight lines of 1 cm length (equals 6.5 km), thereby smoothening the actual coastline (Table S3). We assessed correlation between the matrices of genetic differentiation and geographic distance by performing a Mantel test in ARLEQUIN version 3.5 with 1,000 permutations.

2.7 | Population genetic analyses: Nuclear SNP data

To assess the number of colonization sources and the genetic structure of populations we used principal component analysis (PCA) and the clustering algorithms implemented in STRUCTURE and ADMIXTURE. PCA was performed on the SNP data set using the R packages SNPRelate and gdsfmt (Zheng et al., 2012). Analysis was limited to 20 principal components, otherwise default options were used. The results were visualized in R (Crawley, 2007). A script with the detailed commands for analysis and plotting is given as a supplement. Genetic admixture of populations was tested with STRUCTURE 2.3.4 (Hubisz et al., 2009; Pritchard et al., 2000). Since we sampled eleven populations from eight localities, the test was carried out for *K* from 1 to 12. Admixed origin of populations was allowed (NOADMIX = 0). Due to the recent common origin of populations (see Section 3) allele frequencies were assumed to be correlated across populations (FREQSCORR = 1).

Burnin was set to 20,000 replications followed by 500,000 replications of data collection. All runs were iterated 10 times and convergence was assessed and visualized through the Cluster Markov Packager Across K (CLUMPAK) web server (Kopelman et al., 2015), with default options. Additionally, best K was obtained from STRUCTURE HARVESTER (Earl, 2012) by the method of Evanno et al. (2005). Given the weak population structure (see Section 3), we also performed STRUCTURE runs with location priors (LOCPRIOR = 1), but neither the model fit (LnP(K)) nor the cluster assignments changed notably (data not shown). For comparison with STRUCTURE, we ran ADMIXTURE (Alexander & Lange, 2011) with increased convergence stringency (–C 10^{-7}), 10-fold cross-validation (--cv=10) and otherwise default parameters. Results were again visualised through the CLUMPAK web server.

In order to specifically assess gene flow between timing strains we finally performed STRUCTURE runs for $K = 11$ with fixed population information (USEPOPINFO = 1), so that only the mixing between pre-defined population-specific genetic clusters was assessed. All other parameters were unchanged. We assessed convergence of the 10 replicates by calculating the standard deviation in the cluster assignments to each population (Table S4).

We further tested for gene flow between the populations by assessing if there was isolation by distance (IBD). Pairwise F_{ST} values and Mantel tests were calculated in ARLEQUIN version 3.5 (Excoffier & Lischer, 2010) as described above. For these calculations the maximum amount of missing genotypes per SNP was set to 5%.

To further investigate gene flow, we attempted to detect migration events with TreeMix 1.13 (Pickrell & Pritchard, 2012). We converted the SNP data to TreeMix format with the vcf2treemix.sh script (Ravinet). For 0 to 10 migration events, we ran five iterations each of TreeMix, with no blocks (as SNPs were already thinned for LD) and the southernmost population from Concarneau as root to the tree. As a second approach for testing migration, we tried to detect migrants in our samples from sympatric populations. Distinguishing genetic markers were obtained from the largely undifferentiated sympatric timing strains by filtering our SNP set to those with $F_{ST} > 0.2$ between sympatric timing strains ($n = 12$ for Roscoff; $n = 88$ for Plouguerneau). We then performed PCA on these subsets of SNPs as described above. Individuals clustering in a different cluster than expected could represent migrants. Genetic differentiation between timing strains was too low to obtain diagnostic markers for identifying hybrids via hybrid index and interstrain heterozygosity (data not shown).

In order to assess the contribution of geographic distance and/or timing differences to reproductive isolation, we tested whether genetic differentiation between the populations was influenced by these factors. The 11 timing strains were grouped according to geography or emergence timing (Table 1) and analysis of molecular variance (AMOVA) was performed for the SNP data set in ARLEQUIN version 3.5 (Excoffier & Lischer, 2010) with default settings and 1,000 permutations. From these ARLEQUIN runs we also obtained nucleotide diversity π per population for the SNP data set (Table S5).

TABLE 1 Groups of populations formed for testing genetic differentiation due to timing differences or geographic site in AMOVA

Test for	Groups	Timing strains
Circadian timing	Type 1	Plou-1SL, Eta-1SL, Bria-1SL, Por-1SL
	Type 2	Conc-2NM, Cam-2NM, Plou-2SL, Plou-2NM, Ros-2FM, Ros-2NM, Lou-2NM
Lunar timing	Type NM	Conc-2NM, Cam-2NM, Plou-2NM, Ros-2NM, Lou-2NM
	Type FM	Ros-2FM
	Type SL	Plou-2SL, Plou-1SL, Eta-1SL, Bria-1SL, Por-1SL
Geographic site	Conc	Conc-2NM
	Cam	Cam-2NM
	Plou	Plou-1SL, Plou-2SL, Plou-2NM
	Ros	Ros-2FM, Ros-2NM
	Lou	Lou-2NM
	Eta	Eta-1SL
	Bria	Bria-1SL
	Por	Por-1SL

Assuming that some of the populations in our sample may be very small or young, we assessed relatedness of individuals in our sample by calculating the inbreeding coefficient F with the --het command in VCFTOOLS 0.1.14 (Danecek et al., 2011). The results were visualized as population-specific box plots in R (Crawley, 2007).

3 | RESULTS

3.1 | Defining temporal niches and timing types of *C. marinus*

Finding sympatric *C. marinus* strains in distinct temporal niches required clear definitions of temporal niches, timing types and timing strains of *C. marinus*. Detailed definitions are given in the methods section, but are recapitulated here, as they are a result of this study and required for understanding all further results.

Reproduction of *C. marinus* is limited to the lowest tides. The semi-diurnal tides of the European Atlantic coast produce four distinct occasions per lunar cycle with very low tides, which represent four nonoverlapping temporal niches that are suitable for reproduction of *C. marinus* (thick-lined black boxes in the central panel of Figure 1a). These temporal niches occur at a specific lunar phase (bottom panel Figure 1a) and a specific time of day (left panel in Figure 1a). Timing strains of *C. marinus* are characterised

by distinct, genetically determined combinations of circadian and circalunar timing, which time reproduction to a suitable temporal niche. Depending on which temporal niche(s) they occupy, timing strains fall into four distinct groups, which we define here as ecological timing types. Along the circadian time axis, we distinguish timing strains that emerge during the first low tide after sunrise ("Type 1") from timing strains that emerge during the second low tide after sunrise ("Type 2"; see Figure 1a). Along the circalunar time axis, timing strains emerge either only during full moon (lunar rhythm; "Type FM"), or only during new moon (lunar rhythm; "Type NM") or during both (a so-called semilunar rhythm; therefore "Type SL"; see Figure 1a). Combining the two time axes results in six possible timing types, but only 1SL, 2SL, 2 NM and 2FM strains have been found (Figure 1 b–e; Table S1). In order to distinguish sympatric timing strains, we amend the previous convention of naming timing strains based on a shorthand for their geographic origin by adding their timing type (Table S1). As an example, "Ros-2FM" is a type 2FM strain from Roscoff. As the timing of the tides shifts gradually along the coastline, timing strains within each timing type differ slightly in circadian and circalunar emergence times in adaptation to the local tides.

3.2 | In Brittany and Normandy *C. marinus* timing types occur in sympatry

In this study, we explored the capabilities of *C. marinus* to genetically adapt to divergent timing of the tides on small geographic scales by studying eight locations along the coasts of Brittany and Normandy (Figure 2). In these eight sites we identified 11 distinct *C. marinus* timing strains (Figure 2), which represented all four known timing types of *C. marinus*, occupying all four available temporal niches (compare Figure 1 b–e with Figures 3 and 4). We characterized their daily and lunar emergence times under common garden conditions in the laboratory (Figures 3 and 4). The Plou-1SL laboratory strain persisted only briefly, so that for this timing strain circalunar data is limited and only field observations are available for the circadian rhythm (Figure 3).

In two locations, Roscoff and Plouguerneau, different timing types co-occurred (Figures 2 to 4), which is the first description of sympatric timing types for *C. marinus*. In Roscoff, the Ros-2FM and Ros-2NM strains are fully separated by lunar timing, one emerging only during full moon, the other only during new moon (Figure 4). The smaller second peaks around day 18 in Ros-2FM and around day 21 in Ros-2NM (Figure 4) are an artefact produced by the artificial laboratory moonlight (for details see Figure S1). During field work we have not observed individuals swarming at these times of the lunar month. In Plouguerneau three timing strains co-occur: Plou-1SL is fully separated from the other two by circadian emergence time (red dotted circle, Figure 3). Of the two type 2 populations, Plou-2NM emerges only during new moon, while Plou-2SL emerges during both full moon and new moon, which leads to a considerable overlap in emergence times (Figure 4).

3.3 | Assessing local adaptation to the timing of the tides

As the timing of the tides changes gradually along the coastline, we expect that within each timing type the different parapatric timing strains should be locally adapted. Indeed, within timing types, the lunar emergence peaks shift gradually when moving along the coastline (Figure 4), corresponding to the shift in the local timing of the tides. Daily emergence times also shift with the local time of the low tides (Figure 3). However, for some timing strains, daily emergence times appear suboptimal relative to the local tides (Figure 3). 2NM type strains generally emerge about three days before new moon, when the low tide is 2–3 h earlier (see Figure 1d). When correcting for this offset in circalunar timing by plotting the daytime of low tide at three days before new moon (grey arrows in Figure 3), the match of 2NM type strains' emergence with the local tides is much better. In some timing strains laboratory emergence times and field emergence times diverge, most notably in the Eta-1SL strain (compare histograms with dotted circles; Figure 3). In the field, circadian emergence time may depend not only on the circadian clock, but also additional external stimuli (see Section 4).

3.4 | All timing types emerged from a single recent colonization event

The existence of sympatric timing types and locally adapted timing strains raises the questions of how they have evolved and how they are maintained, particularly whether sympatric ecological divergence in the presence of gene flow is plausible. In order to address these questions, we genetically analysed 24 individuals for each of the 11 populations based on mitochondrial COI haplotypes and a nuclear SNP data set obtained by RAD sequencing (summary statistics in Table S6).

In a network of the mitochondrial COI haplotypes all 11 populations and 73% of all individuals share the same major haplotype (Figure 5a). The minor haplotypes are all directly derived from the major haplotype by one or rarely two mutations. This indicates recent colonization of the coast of Brittany and Normandy by a single maternal lineage and suggests that all populations originated from the same source. Additionally, *C. marinus* populations from the surrounding European coasts of Southern France, Ireland or Germany have completely different COI haplotypes (Figure 5b; adapted from Kaiser et al., 2010), rendering independent colonization of Brittany by the known 2NM and 1SL timing types from these coasts very unlikely.

Next, we assessed population structure based on the SNP data with principal component analysis (PCA, Figure 6a) and the STRUCTURE clustering algorithm (Figure 6b). Populations across Brittany and Normandy show limited genetic structure. In particular, the sympatric timing strains from Roscoff and Plouguerneau, which comprise all four timing types, appear highly similar in genetic composition in both analyses (Figure 6a and b). There is no apparent genetic distinction

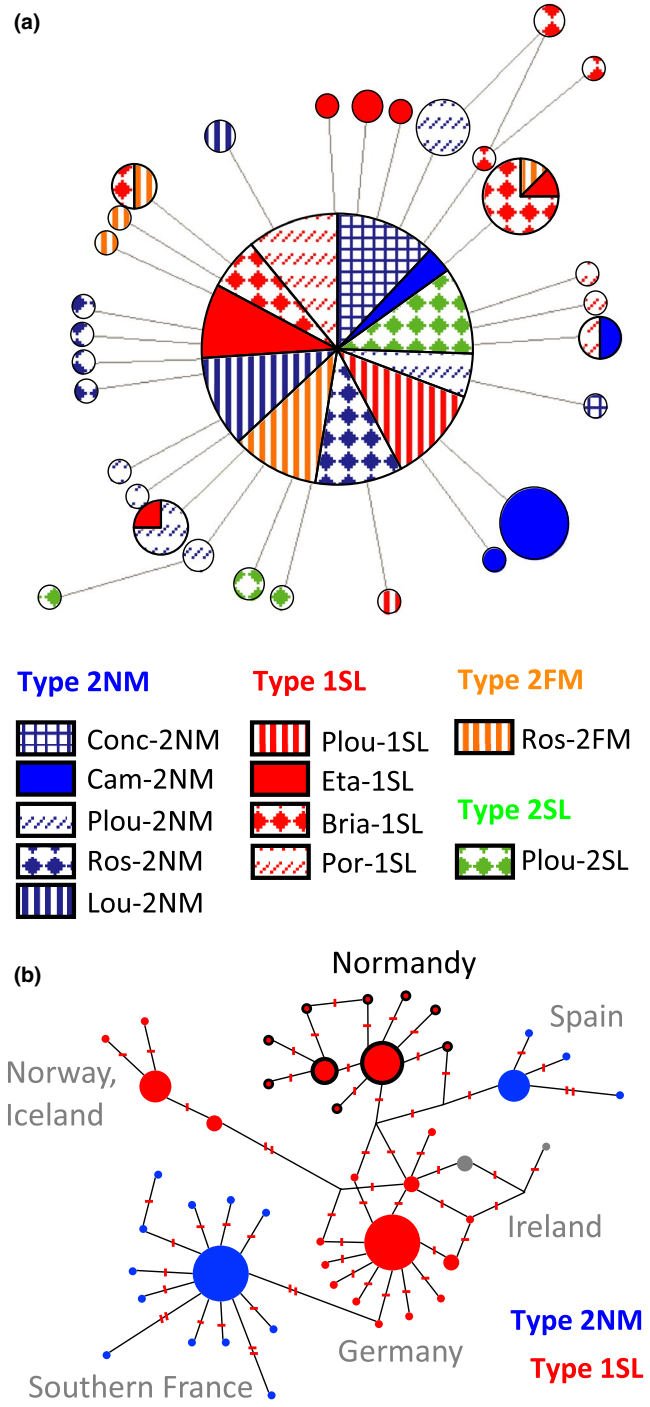

FIGURE 5 Mitochondrial Cytochrome oxidase I (COI) haplotype network. (a) Haplotype network of an 853 bp fragment of COI for 262 individuals from eleven populations in Brittany and Normandy. Lines indicate single nucleotide polymorphisms (SNPs) separating the haplotypes. Size of the circles corresponds to the fraction of individuals carrying the haplotype. Origin of individuals is coded by geography and timing type (see legend). (b) Haplotype network of the same COI fragment for *C. marinus* strains from neighbouring European coasts, adapted from (Kaiser et al., 2010). Small red lines indicate SNPs. Colour of the haplotypes represents the timing type (not known for Ireland). The haplotypes from Normandy (highlighted by black lines around the circles) include the Por-1SL population, which was sampled again for this study. Across the two studies, the major haplotype and many minor haplotypes are identical. Populations from other European coasts have entirely different haplotypes

between the four timing types, again dismissing the hypothesis that genetically distinct and reproductively isolated timing types may have colonized Brittany independently. In conclusion, the diversification into the four timing types and the establishment of local timing adaptations must have occurred within Brittany and Normandy.

3.5 | Geographic structure and gene flow

The PCA and STRUCTURE results based on the SNP data reveal weak geographic separation of the populations. In particular, the northernmost timing strains are separated along principal component 1 and the southernmost are separated along principal component 2 (Figure 6a). These principal components only explain 4.21% and 3.42% of the variation. Analysis of molecular variance (AMOVA) with groups formed according to geographic location attributes 5.96% of the variation to geography (Table 2). Similarly, genetic clusters obtained with STRUCTURE weakly follow geography, especially at low numbers of clusters (Figure 6b). The most likely number of groups in the sample, as determined with the ΔK method, are $K = 2$ and $K = 8$ (Figure S2). Results from an additional ADMIXTURE analysis are very similar to those obtained with STRUCTURE and have the lowest cross-validation error at $K = 8$ (Figure S3). Taken together, there is weak geographic structure imposed by the eight sampling sites.

Next, we fixed the population information in the STRUCTURE analysis, so that we could assess the mixing of pre-defined, population-specific clusters. In this analysis, each timing strain corresponds to a specific genetic cluster (Table S4). Additional genetic components come from sympatric timing strains or from adjacent geographic sites, which suggests that the weak geographic structure is a product of pervasive local gene flow along the coastline. We substantiated this finding by calculating pairwise genetic differentiation (F_{ST}; Table S7) based on the SNP data and comparing it to geographic distance (Figure 6c). The populations show very clear isolation by distance (IBD), suggesting that there is strong gene flow between the populations, which gradually decreases over the sampled geographic scale. Sympatric timing types are the least differentiated (Figure 6c).

In contrast to the SNP data, there is no IBD in the mitochondrial COI sequences (Table S7, Figure S4). Mitochondrial gene pools show highly variable differentiation, irrespective of geographic scale, suggesting they are geographically isolated and diverge at random by genetic drift. This finding is in line with the fact that *C. marinus* females are wingless and basically immobile. Asymmetry in nuclear vs. mitochondrial differentiation indicates that gene flow along the continuous rocky coast of Brittany and Normandy is primarily mediated by swarming adult males.

In order to test for migration events between the populations, we used the TreeMix algorithm, which assesses whether the observed correlation of allele frequencies between populations is better

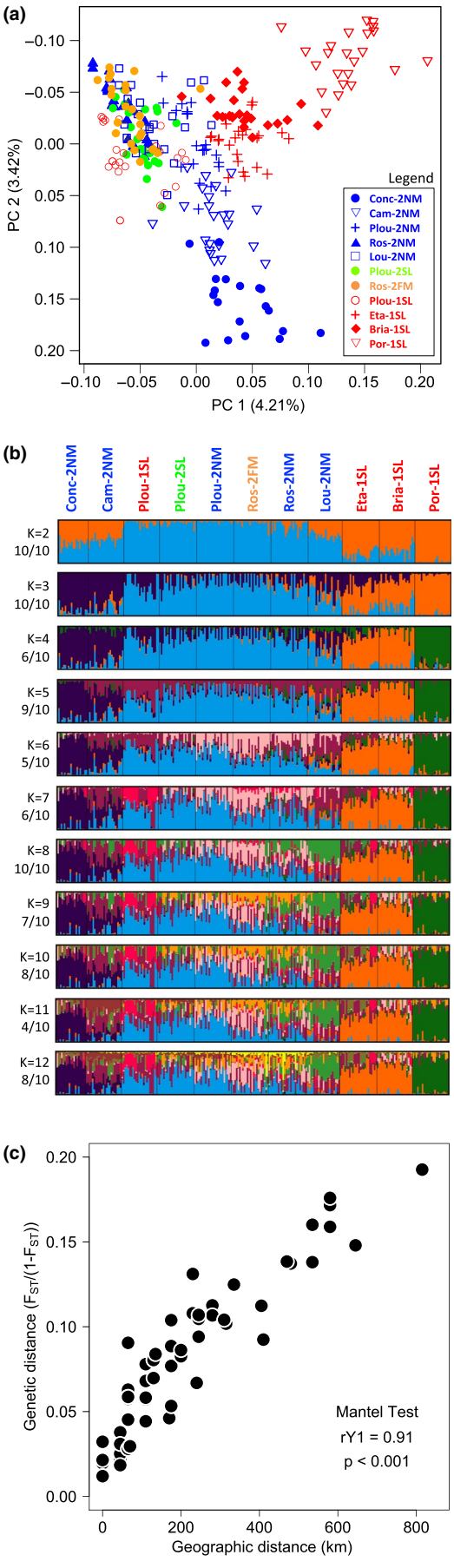

FIGURE 6 6Genetic structure and isolation by distance between timing strains of *C. marinus* in Brittany. Analyses are based on SNP data from restriction-site associated DNA sequencing (RAD-seq; 2,159 unlinked SNPs). (a) Principal component analysis (PCA) for 263 individuals from 11 populations. Different symbols represent timing strains, colours correspond to timing types as defined in Figure 1. A lack of clustering according to timing types suggests that timing types are not genetically distinct, reproductively isolated units. There is limited clustering according to geography. The small amount of variation explained by the principal components indicates that population structure is overall weak. (b) STRUCTURE analysis for 2–12 genetic clusters (*K*). Results from 10 independent replicates are summarized in CLUMPAK. The numbers of replicates that converged are given to the left of each panel ("x/10"). Replicates that did not fully converge only differ in small details (see Figure S7). There is no separation by timing type and limited geographic structure. (c) Linearized pairwise genetic differentiation (F_{ST}) between timing strains plotted against geographic distance indicates isolation by distance (IBD). Sympatric timing types are the least differentiated

captured if specific migration events are assumed. While the model likelihood was always better with migration, different iterations of the algorithm did not converge and suggested various sources, destinations and directions of migration (Figure S5). Population structure is so weak that migration events cannot be discerned.

Given strong gene flow, we attempted to directly detect "migrants in time" from the sympatric populations. In order to obtain informative markers, we filtered our SNP data for F_{ST} > 0.2 in sympatric comparisons. PCA on these subsets of SNPs distinguishes the sympatric timing strains reasonably well (Figure 7). In Roscoff, two Ros-2FM individuals clearly cluster with the Ros-2NM individuals and could possibly be migrants (Figure 7a). For a few Ros-2NM individuals clustering is ambiguous (Figure 7a). In Plouguerneau, the Plou-1SL strain is separated along PC1 (Figure 7b), whereas Plou-2NM and Plou-2SL overlap along PC2. Conclusions on putative migrants are not possible. Given the low genetic differentiation between sympatric populations, there were no diagnostic markers available for identifying hybrids via ternary plots of hybrid index and inter-strain heterozygosity.

Taken together, the four timing types and the locally adapted timing strains in Brittany and Normandy persist and probably also diverged in the presence of very strong gene flow.

3.6 | Reproductive timing is not an effective reproductive isolation factor in *C. marinus*

For speciation to occur, ecological divergence between populations must be accompanied by the build-up of reproductive isolation. Given the large differences in reproductive timing of *C. marinus* timing types, we would assume that allochrony must play a role in this species. In order to assess to what extent temporal isolation results in genetic divergence between the timing types, we applied an analysis of molecular variance (AMOVA). Populations were

TABLE 2 AMOVA results for groups according to circadian timing, circalunar timing or geographic site

	Circadian timing				Circalunar timing				Geographic site			
Source of variation	**d.f.**	**SS**	**% var**	***p***	**d.f.**	**SS**	**% var**	***p***	**d.f.**	**SS**	**% var**	***p***
Among groups	1	829	1.59	.04	2	912	−0.14	.44	7	4061	5.96	.00
Among populations within groups	9	3830	6.74	.00	8	3747	7.71	.00	3	598	2.08	.00
Among individuals within populations	242	22686	−4.16	1.00	242	22686	−4.19	1.00	242	22686	−.17	1.00
Within individuals	253	25970	95.83	.01	253	25970	96.62	.01	253	25970	96.13	.01
Total	505	53314	100		505	53314	100		505	53314	100	

Abbreviations: d.f., degrees of freedom; *p*, *p*-vsalue obtained from 1,000 permutations; SS, sum of squares; %var, percentage of variance explained.

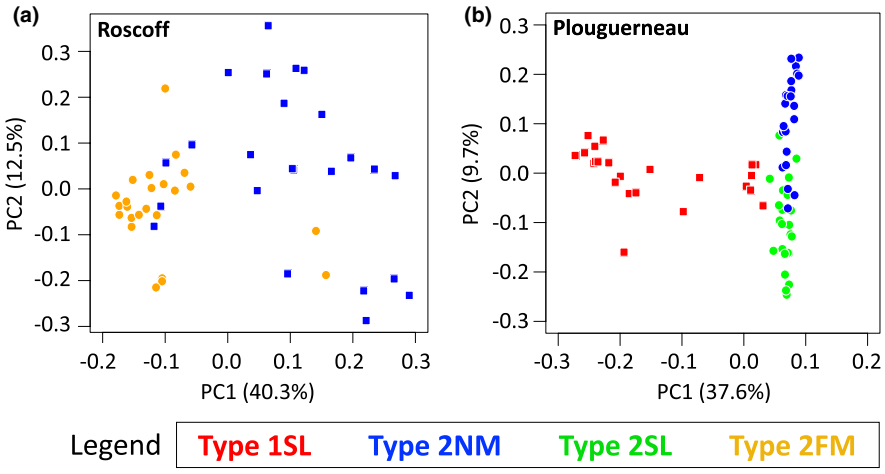

FIGURE 7 Putative "migrants in time" between sympatric timing types. Principal component analysis (PCA) on a subset of divergent SNPs (F_{ST} > 0.2) approximately discriminates sympatric timing strains. (a) In Roscoff, the Ros-2FM strain is genetically more uniform. Separated along the dominant first principal component (40.3%), two Ros-2FM individuals cluster with the Ros-2NM strain and could possibly be migrants. Reversely, some Ros-2NM individuals rather cluster with the Ros-2FM strain. (b) In Plouguerneau, the Plou-2NM and Plou-2SL strains are not well separated, but patterns do not allow speculation on migrants or hybrids

grouped according to circadian timing phenotype, circalunar timing phenotype or geography (Table 1), and we assessed which fraction of genetic variation is associated with these groups. In all three comparisons, the major component of genetic variation is within individuals (95.83%–96.62%) and there is some genetic variation between populations within each group (2.08%–7.71%). In addition to these general observations, 1.59% of genetic variation is associated with groups based on circadian timing (Table 2, left part), but no genetic variation is associated with groups based on circalunar timing (Table 2, middle part). Notably, 5.96% of genetic variation is associated with geographic groups (Table 2, right part) and AMOVA for timing is partially confounded by geography, as three out of four type 1SL strains are at one end of the sampled geographic range. Still, the results obtained in PCA on high F_{ST} SNPs for the sympatric Plouguerneau timing strains, i.e., without any possibility for geographic confounding, also suggests that circadian timing entails more genetic structure than circalunar timing (Figure 7b). The sympatric strains from Plouguerneau are well-separated by circadian timing (PC1; 37.6% of variation), but much less separated by circalunar timing (PC2; 9.7% of variation). In summary, a small but detectable fraction of genetic variation is associated with and possibly caused by differences in circadian timing. But overall, there is no strong genome wide genetic divergence that would be attributable to circadian allochrony or circalunar allochrony.

4 | DISCUSSION

4.1 | Local adaptation

In Brittany we found different sympatric and parapatric *C. marinus* timing types, i.e. strains occupying different temporal niches. It is unclear if these timing types represent local adaptations. Full moon and new moon low tides, as well as morning and evening low tides, do not differ consistently in water levels across the geographic range we assessed. Consequently, all four temporal niches should be equally suitable for reproduction of *C. marinus* and ecologically equivalent. In contrast, the site-specific circalunar and circadian

emergence times of the different timing strains, which exist within each timing type, must be considered local adaptations. They were shown to match the local timing of the tides extremely well, with correlation coefficients of up to 0.98 (Kaiser et al., 2011). They are genetically determined and stable under laboratory conditions without any tidal timing cues (this study; Kaiser et al., 2011; Neumann, 1967).

The circalunar timing phenotypes presented here (Figure 4) fulfil the expectations for local adaptation. Within each timing type they show a gradual shift along the coastline corresponding to the strong gradient in the timing of the tides. For circadian timing the match between local tides and laboratory emergence time is still good, but shows exceptions (Figure 3). There are two major reasons for this. First, the 2NM type strains generally emerge a few days before new moon when the low tide occurs 2–3 h earlier (Figure 1d). When correcting for this effect, the match between local low tide and laboratory emergence is much better (Figure 3). Variation in the daytime of the low tide is larger three days before new moon compared to new moon (compare standard deviations of black and grey arrows in Figure 3). This might explain the comparatively broad spread of circadian emergence times in 2NM type strains (Figure 3). Second, this study is the first for which emergence times have not only been recorded in the laboratory, but have also been observed in the field for many of the timing strains. In some cases, emergence times in the field and laboratory differ notably (Figure 3). The genetic timing adaptations, which are recorded in the laboratory without any tidal timing cues, may define a site-specific circadian emergence window within which environmental cues can trigger emergence and thereby further fine-tune its timing. Such phenotypic flexibility is advantageous, as the emergence of *C. marinus* takes place over several days during which the daytime of low tide shifts by several hours. Allowing for emergence to be triggered by tidal cues within the circadian emergence window would enable *C. marinus* to track this shift. Unlike other intertidal organisms, *C. marinus* does not possess a 12.4 h circatidal clock to achieve such tracking (Neumann, 2014).

Our field and laboratory observations indicate that environmental cues for directly triggering emergence might include light levels and exposure of the larval habitat by the tide. In the Plou-2NM, Ros-2NM and Ros-2FM laboratory strains most females emerge immediately when the light is turned off (Figure 3). In Eta-1SL field emergence was highly concentrated at the time when the larval habitat became exposed by the tide, whereas daily emergence in the laboratory is extremely broad and nonoverlapping with observed emergence in the field. Interestingly, the Eta-1SL strain has the lowest genetic diversity (Table S5) and very high inbreeding coefficients (Figure S6). Possibly, the Eta-1SL strain has established very recently and has not yet genetically adapted in circadian emergence time to the local tides, but can only persist at the site because of phenotypic flexibility in circadian emergence time. Phenotypic flexibility in timing may facilitate the colonization of new sites and genetic adaptation may follow secondarily.

The interplay of genetic adaptation and phenotypic flexibility, the fact that timing adaptations are very precise and accessible to laboratory experiments, and the one-dimensional distribution of *C. marinus* along the coastline (resembling Kimura's simple stepping stone model; Kimura & Weiss, 1964), make *C. marinus* an interesting model case for disentangling the evolutionary factors driving local adaptation.

4.2 | Evolution of sympatric timing types

The sympatric *C. marinus* timing types, as well as the locally adapted parapatric timing strains, form a single maternal lineage. Our genetic data rule out that several allopatric and reproductively isolated timing types have colonized Brittany and Normandy independently. Instead, we observe very strong gene flow along the coastline, against which the timing types and timing strains must be maintained. There are at least three evolutionary scenarios with various degrees of allopatry and secondary gene flow by which the sympatric timing types could have evolved. All of these scenarios are consistent with the currently available genetic data and the biology of the organism. Firstly, in Plouguerneau and Roscoff we may see cases of genuine sympatric divergence with gene flow, starting from no initial ecological divergence and a single panmictic population. Secondly, it is possible that timing types have evolved in separate sites in Brittany with some reduction in gene flow due to IBD, and then secondarily spread along the coast to become fully sympatric. Given the current strong gene flow along the entire coast, this would still represent a case of divergence with gene flow, but not starting from initial panmixia. Finally, it is possible that at least some of the alleles responsible for sympatric divergence in timing predate the postglacial colonization event and upon arrival in Brittany facilitated local divergence with gene flow. The source of pre-existing timing alleles may either be standing genetic variation in the colonizing population or yet undetected introgression events. The scenario of standing genetic variation is consistent with a previous study on five allopatric timing strains, which indicated that standing genetic variation at adaptive timing loci is common in *C. marinus* (Kaiser et al., 2016). The scenario of undetected introgression seems conceivable in the light of *C. marinus'* modes of dispersal (Kaiser et al., 2010). Adult dispersal is common, but geographically limited due to the short adult life span. As females are wingless, adult dispersal is largely mediated by males, resulting in strong geographic isolation of mitochondrial gene pools as also detected here (Figure S4). Dispersal of the benthic larvae via drifting algae appears to be rare, but can be long distance and is not expected to be sex-biased. If we assume that Brittany was initially colonized by only a single timing type, i.e., also a single maternal lineage, allopatric timing types from distant coasts may later have introgressed via larval dispersal. Upon local introgression the nuclear encoded timing alleles could spread along the coastline via adult male dispersal, while the introgressed mitochondrial lineages would be locally confined. Such introgressed and

likely divergent mitochondrial haplotypes might reside in sites that we have not sampled or they may have gone extinct again.

Distinguishing between these scenarios will be very difficult given the strong gene flow and low degree of genetic differentiation. It will certainly require additional sampling and comprehensive genomic data. But taken together, all of these scenarios imply a mixing of timing types and divergence with gene flow, which poses questions to which role temporal isolation and ecological selection by the tides play in forming and maintaining sympatric timing types of *C. marinus*.

4.3 | Reproductive isolation and natural selection

Many prominent examples of divergence with gene flow are characterized by differential habitat or host use resulting in heterogeneous distributions within the range (Butlin et al., 2014; Filchak et al., 2000; Nosil et al., 2006; Schliewen et al., 2001), sometimes referred to as *mosaic sympatry* (Mallet et al., 2009) or *heteropatric speciation* (Maynard Smith, 1966). This does not seem to be the case for *C. marinus*, where we have isolated different timing types from the very same patch of algae.

In the light of *C. marinus*' biology and life history, reproductive timing must be expected to result in prezygotic reproductive isolation through allochrony and should eventually constitute a magic trait for speciation (Gavrilets, 2004). Given the short reproductive period of *C. marinus* of only a few hours, the large timing differences between sympatric timing types should minimize or even abolish any overlap in reproductive time. Our genetic analyses of sympatric timing types suggest that mating is certainly not random between timing types (Figures 6 and 7). In Roscoff we detected two or more putative migrants in a sample of 48 individuals, suggesting there is strong migration, but not panmixia. The correlation between timing differences and genetic divergence of populations is marginal for circadian timing and nonexistent for circalunar timing (Table 2). Clearly, prezygotic isolation by timing does not result in genome-wide genetic divergence.

The lack of isolation by circalunar timing may partially result from the existence of SL type strains, which emerge at full moon and new moon and may mediate gene flow between FM and NM strains. However, the strong gene flow in Roscoff is still surprising, as in the absence of an SL type strain, emergence of the sympatric timing strains at full moon vs. new moon results in no overlap in reproductive time. Field observations from Helgoland provide a potential explanation for this puzzle. In Helgoland, spring emergence of overwintering individuals was found to be triggered by the warming sea water, which overrode the effect of the circalunar clock (Krüger & Neumann, 1983). If this is also true for our study sites in Brittany, sympatric timing types could meet and interbreed every spring, i.e., every 3–4 generations, circumventing the temporal isolation imposed by their lunar timing phenotypes.

Another source of reproductive isolation may lie in selection against timing type hybrids, which are usually intermediate in timing between their parents (Heimbach, 1978; Kaiser et al., 2011; Neumann, 1967). We expect that Ros-2NM × Ros-2FM hybrids would emerge at neap tide high tides, a situation most unsuitable for reproduction. The situation in Plouguerneau would be the same, though slightly more complex given that there are three timing types. Thus, in the absence of notable temporal or geographic isolation, the existence and maintenance of distinct sympatric timing types probably relies on permanent strong selection against hybrids, an ecologically imposed form of extrinsic postzygotic isolation. As such an isolation mechanism drastically reduces the fitness of parents producing hybrids, we may expect strong selection for prezygotic isolating factors that would reinforce ecological divergence. Assortative mating based on other factors but timing is a possibility to consider in future research. A recent study on the olfactory system of *C. marinus* suggests that pheromone communication exists in this species (Missbach et al., 2020).

4.4 | Perspectives

Ultimately, studies on local adaptation and sympatric population divergence will be most powerful when the loci underlying the ecological adaptations of diverging populations are identified. For *C. marinus*, the populations described here may be key to identifying these loci, as the observed strong gene flow is expected to homogenize their genomes except for the few ecologically relevant loci. The current RAD sequencing does not have the resolution to identify such loci, as RAD tags in this study are on average 24 kb apart, while even in more divergent *C. marinus* populations genomic divergence peaks usually extend over 2–5 kb only (Kaiser et al., 2016). But genome scans and QTL mapping have proven a powerful tool in *C. marinus*. They may not only help to find genes involved in local adaptation of biological clocks and sympatric population divergence with gene flow, but may also give first insights into the yet enigmatic circalunar clockwork.

ACKNOWLEDGEMENTS

This work was supported by the Max Planck Society in the framework of a Max Planck Research Group and by a scholarship of the International Max Planck Research School "The Exploration of Ecological Interactions with Molecular and Chemical Techniques", both awarded to TSK. TSK was also supported by a Vienna International Postdoctoral Program (VIPS) scholarship. This work was further supported by the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7-2007-2013) with an ASSEMBLE grant (Grant agreement 227799) and under the Horizon 2020 research and innovation programme with an ERC Starting Grant (Grant agreement 802923), both awarded to TSK. This work was also supported by the European Research Council under the European Union's Seventh Framework Programme (FP7-2007–2013) with an ERC Starting Grant (Grant agreement 337011) to KTR and the research platform 'Rhythms of Life' of the University of Vienna to KTR and AvH. Illumina

Sequencing was performed at the Vienna Biocenter NGS Core Facility (https://www.viennabiocenter.org/facilities/next-generation-sequencing/). Andrej Belokurov of the Marine Facility at MFPL helped with animal husbandry. The *C. marinus* fraction collectors were built at the workshop of the Cologne Biocenter at the University of Cologne. We thank Hanna Kokko, Diethard Tautz, Jeffrey Feder and three anonymous reviewers for comments on the manuscript.

AUTHOR CONTRIBUTIONS

T.S.K. conceived and designed the study, performed field and laboratory work, analysed and interpreted the data and wrote the manuscript. A.V.H., K.T.R., and D.G.H. provided resources and edited the manuscript. All authors approved of the final manuscript.

DATA AVAILABILITY STATEMENT

Mitochondrial COI sequences are available under GenBank accession numbers GU943253.1, GU943254.1, GU943258.1, GU943261.1, GU943262.1, as well as ENA-EBI accession numbers LN851805–LN851838. Raw reads from RAD sequencing are deposited under ENA-EBI project accession number PRJEB9361. The script for conducting PCA is available in the Supporting Information.

ORCID

Tobias S. Kaiser https://orcid.org/0000-0002-4126-0533

REFERENCES

Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, *12*(1), 246. https://doi.org/10.1186/1471-2105-12-246

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. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, *3*(10), e3376.

Bandelt, H. J., Forster, P., & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, *16*(1), 37–48.

Barluenga, M., Stolting, K. N., Salzburger, W., Muschick, M., & Meyer, A. (2006). Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature*, *439*(7077), 719–723. https://doi.org/10.1038/nature04325

Barton, N. (2010). What role does natural selection play in speciation? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1547), 1825–1840.

Bolnick, D. I., & Fitzpatrick, B. M. (2007). Sympatric speciation: Models and empirical evidence. *Annual Review of Ecology Evolution and Systematics*, *38*, 459–487.

Butlin, R. K., Saura, M., Charrier, G., Jackson, B., André, C., Caballero, A., Coyne, J. A., Galindo, J., Grahame, J. W., Hollander, J., Kemppainen, P., Martínez-Fernández, M., Panova, M., Quesada, H., Johannesson, K., & Rolán-Alvarez, E. (2014). Parallel evolution of local adaptation and reproductive isolation in the face of gene flow. *Evolution*, *68*(4), 935–949. https://doi.org/10.1111/Evo.12329

Campbell, C. R., Poelstra, J., & Yoder, A. D. (2018). What is Speciation Genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. *Biological Journal of the Linnean Society*, *124*(4), 561–583.

Coyne, J. A., & Orr, H. A. (2004). *Speciation* (Vol. *37*): Sinauer Associates.

Crawley, M. J. (2007). *The R book*. John Wiley & Sons Ltd. https://doi.org/10.1002/9780470515075

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., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158.

Devries, P. J., Austin, G. T., & Martin, N. H. (2008). Diel activity and reproductive isolation in a diverse assemblage of Neotropical skippers (Lepidoptera: Hesperiidae). *Biological Journal of the Linnean Society*, *94*(4), 723–736.

Dieckmann, U., & Doebeli, M. (1999). On the origin of species by sympatric speciation. *Nature*, *400*(6742), 354–357. https://doi.org/10.1038/22521

Doellman, M., Ragland, G., Hood, G., Meyers, P., Egan, S., Powell, T., Lazorchak, P., Glover, M., Tait, C., Schuler, H., Hahn, D., Berlocher, S., Smith, J., Nosil, P., & Feder, J. (2018). Genomic differentiation during speciation-with-gene-flow: comparing geographic and host-related variation in divergent life history adaptation in *Rhagoletis pomonella*. *Genes*, *9*(5), 262.

Dres, M., & Mallet, J. (2002). Host races in plant-feeding insects and their importance in sympatric speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *357*(1420), 471–492. https://doi.org/10.1098/rstb.2002.1059

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, *14*(8), 2611–2620.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics*, *28*(7), 342–350.

Filchak, K. E., Roethele, J. B., & Feder, J. L. (2000). Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature*, *407*(6805), 739–742. https://doi.org/10.1038/35037578

Fitzpatrick, B. M., Fordyce, J., & Gavrilets, S. (2008). What, if anything, is sympatric speciation? *Journal of Evolutionary Biology*, *21*(6), 1452–1459.

Fitzpatrick, B., Fordyce, J., & Gavrilets, S. (2009). Pattern, process and geographic modes of speciation. *Journal of Evolutionary Biology*, *22*(11), 2342–2347.

Foote, A. D. (2018). Sympatric speciation in the genomic era. *Trends in Ecology & Evolution*, *33*(2), 85–95.

Fuhrmann, N., & Kaiser, T. S. (2020). The importance of DNA barcode choice in biogeographic analyses—a case study on marine midges of the genus Clunio. *Genome*, *999*, 1–11.

Fukami, H., Omori, M., Shimoike, K., Hayashibara, T., & Hatta, M. (2003). Ecological and genetic aspects of reproductive isolation by different spawning times in Acropora corals. *Marine Biology*, *142*(4), 679–684.

Gavrilets, S. (2004). *Fitness landscapes and the origin of species*, Vol. *41*. Princeton University Press. https://doi.org/10.2307/j.ctv39x541

Gezeitentafeln 2009 - Europäische Gewässer. (2008). Bundesamt für Seeschifffahrt und Hydrographie.

Häfker, N. S., & Tessmar-Raible, K. (2020). Rhythms of behavior: Are the times changin'? *Current Opinion in Neurobiology*, *60*, 55–66.

Hänniger, S., Dumas, P., Schöfl, G., Gebauer-Jung, S., Vogel, H., Unbehend, M., Heckel, D. G., & Groot, A. T. (2017). Genetic basis of allochronic differentiation in the fall armyworm. *BMC Evolutionary Biology*, *17*(1), 68. https://doi.org/10.1186/s12862-017-0911-5

Heimbach, F. (1978). Sympatric species, *Clunio marinus* Hal. and *Cl. balticus* n. sp. (Dipt., Chironomidae), isolated by differences in diel emergence time. *Oecologia*, *32*(2), 195–202.

Helm, B., & Womack, R. (2018). Timing matters: Allochronic contributions to population divergence. In: D. Tietze (Ed.) *Bird species. Fascinating life sciences*, (pp. 95–107). Springer.

Hendry, A. P., & Day, T. (2005). Population structure attributable to reproductive time: isolation by time and adaptation by time. *Molecular Ecology*, *14*(4), 901–916.

Honegger, H. W. (1977). An automatic device for the investigation of the rhythmic emergence pattern of *Clunio marinus*. *International Journal of Chronobiology*, *4*, 217–221.

Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, *9*(5), 1322–1332. https://doi.org/10.1111/j.1755-0998.2009.02591.x

Hut, R. A., Kronfeld-Schor, N., van der Vinne, V., & De la Iglesia, H. (2012). In search of a temporal niche: environmental factors. In *Progress in brain research*, vol. *199* (pp. 281–304). Elsevier.

Kaiser, T. S. (2014). Local adaptations of Circalunar and circadian clocks: The case of *Clunio marinus*. In H. Numata, & B. Helm (Eds.), *Annual, lunar, and tidal clocks* (pp. 121–141). Springer.

Kaiser, T. S., & Heckel, D. G. (2012). Genetic architecture of local adaptation in lunar and diurnal emergence times of the marine midge *Clunio marinus* (Chironomidae, Diptera). *PLoS One*, *7*(2), e32092. https://doi.org/10.1371/journal.pone.0032092

Kaiser, T. S., Neumann, D., & Heckel, D. G. (2011). Timing the tides: Genetic control of diurnal and lunar emergence times is correlated in the marine midge *Clunio marinus*. *BMC Genetics*, *12*, 49. https://doi.org/10.1186/1471-2156-12-49

Kaiser, T. S., Neumann, D., Heckel, D. G., & Berendonk, T. U. (2010). Strong genetic differentiation and postglacial origin of populations in the marine midge *Clunio marinus* (Chironomidae, Diptera). *Molecular Ecology*, *19*(14), 2845–2857. https://doi.org/10.1111/j.1365-294X.2010.04706.x

Kaiser, T. S., Poehn, B., Szkiba, D., Preussner, M., Sedlazeck, F. J., Zrim, A., Neumann, T., Nguyen, L.-T., Betancourt, A. J., Hummel, T., Vogel, H., Dorner, S., Heyd, F., von Haeseler, A., & Tessmar-Raible, K. (2016). The genomic basis of circadian and circalunar timing adaptations in a midge. *Nature*, *540*(7631), 69–73. https://doi.org/10.1038/nature20151

Kautt, A. F., Machado-Schiaffino, G., & Meyer, A. (2016). Multispecies outcomes of sympatric speciation after admixture with the source population in two radiations of Nicaraguan crater lake cichlids. *PLoS Genetics*, *12*(6), e1006157. https://doi.org/10.1371/journal.pgen.1006157

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, *7*(12), 1225–1241.

Kimura, M., & Weiss, G. H. (1964). The Stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, *49*(4), 561–576.

Kondrashov, A. S., & Kondrashov, F. A. (1999). Interactions among quantitative traits in the course of sympatric speciation. *Nature*, *400*(6742), 351–354.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179–1191.

Krüger, M., & Neumann, D. (1983). Die Temperaturabhängigkeit semilunarer und diurnaler Schlüpfrhythmen bei der intertidalen Mücke *Clunio marinus* (Diptera, Chironomidae). *Helgolander Meeresuntersuchungen*, *36*(4), 427–464.

Kumar, S., Tamura, K., & Nei, M. (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, *5*(2), 150–163.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*(2), 298–299. https://doi.org/10.1093/bioinformatics/btr642

Malinsky, M., Challis, R. J., Tyers, A. M., Schiffels, S., Terai, Y., Ngatunga, B. P., Miska, E. A., Durbin, R., Genner, M. J., & Turner, G. F. (2015). Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science*, *350*(6267), 1493–1498.

Mallet, J., Meyer, A., Nosil, P., & Feder, J. L. (2009). Space, sympatry and speciation. *Journal of Evolutionary Biology*, *22*(11), 2332–2341. https://doi.org/10.1111/J.1420-9101.2009.01816.X

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet Journal*, *17*(1), 10–12.

Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Salazar, C., Walters, J. R., Simpson, F., Blaxter, M., Manica, A., Mallet, J., & Jiggins, C. D. (2013). Genome-wide evidence for speciation with gene flow in Heliconius butterflies. *Genome Research*, *23*(11), 1817–1828.

Maynard Smith, J. (1966). Sympatric speciation. *American Naturalist*, *100*(916), 637–650.

Mayr, E. (1947). Ecological factors in speciation. *Evolution*, 263–288.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Missbach, C., Vogel, H., Hansson, B. S., Große-Wilde, E., Vilcinskas, A., & Kaiser, T. S. (2020). Developmental and sexual divergence in the olfactory system of the marine insect *Clunio marinus*. *Scientific Reports*, *10*(1), 2125. https://doi.org/10.1038/s41598-020-59063-7.

Neumann, D. (1966). Die lunare und tägliche Schlüpfperiodik der Mücke *Clunio* - Steuerung und Abstimmung auf die Gezeitenperiodik. *Zeitschrift Fur Vergleichende Physiologie*, *53*(1), 1–61.

Neumann, D. (1967). Genetic adaptation in emergence time of *Clunio* populations to different tidal conditions. *Helgolander Wissenschaftliche Meeresuntersuchungen*, *15*(1–4), 163–171.

Neumann, D. (1989). Circadian components of semilunar and lunar timing mechanisms. *Journal of Biological Rhythms*, *4*(2), 285–294.

Neumann, D. (2014). Timing in Tidal, Semilunar, and Lunar Rhythms. In H. Numata, & B. Helm (Eds.), *Annual, lunar, and tidal clocks* (pp. 3–24). Springer.

Nosil, P., & Feder, J. L. (2012). Genomic divergence during speciation: causes and consequences Introduction. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *367*(1587), 332–342. https://doi.org/10.1098/rstb.2011.0263

Nosil, P., Sandoval, C. P., & Crespi, B. J. (2006). The evolution of host preference in allopatric vs. parapatric populations of *Timema cristinae* walking-sticks. *Journal of Evolutionary Biology*, *19*(3), 929–942. https://doi.org/10.1111/j.1420-9101.2005.01035.x

Patton, H., Hubbard, A., Andreassen, K., Auriac, A., Whitehouse, P. L., Stroeven, A. P., Shackleton, C., Winsborrow, M., Heyman, J., & Hall, A. M. (2017). Deglaciation of the Eurasian ice sheet complex. *Quaternary Science Reviews*, *169*, 148–172.

Pickrell, J., & Pritchard, J. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, *8*:e1002967.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Ragland, G. J., Doellman, M. M., Meyers, P. J., Hood, G. R., Egan, S. P., Powell, T. H. Q., Hahn, D. A., Nosil, P., & Feder, J. L. (2017). A test of genomic modularity among life-history adaptations promoting speciation with gene flow. *Molecular Ecology*, *26*(15), 3926–3942. https://doi.org/10.1111/mec.14178

Ravinet, M. Retrieved from https://github.com/speciationgenomics/scripts/blob/master/vcf2treemix.sh

Reineke, A., Karlovsky, P., & Zebitz, C. P. W. (1998). Preparation and purification of DNA from insects for AFLP analysis. *Insect Molecular Biology*, *7*(1), 95–99.

Rice, W. R., & Hostert, E. E. (1993). Laboratory experiments on speciation - What have we learned in 40 years. *Evolution*, *47*(6), 1637–1653. https://doi.org/10.2307/2410209

Richards, E. J., Servedio, M. R., & Martin, C. H. (2019). Searching for sympatric speciation in the genomic era. *BioEssays*, *41*(7), 1900047.

Riesch, R., Muschick, M., Lindtke, D., Villoutreix, R., Comeault, A. A., Farkas, T. E., Lucek, K., Hellen, E., Soria-Carrasco, V., Dennis, S. R., de Carvalho, C. F., Safran, R. J., Sandoval, C. P., Feder, J., Gries, R., Crespi, B. J., Gries, G., Gompert, Z., & Nosil, P. (2017). Transitions between phases of genomic differentiation during stick-insect speciation. *Nature Ecology &Amp. Evolution*, *1*, 0082. https://doi.org/10.1038/s41559-017-0082

Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, *14*(11), 807–820.

Savolainen, V., Anstett, M.-C., Lexer, C., Hutton, I., Clarkson, J. J., Norup, M. V., Powell, M. P., Springate, D., Salamin, N., & Baker, W. J. (2006). Sympatric speciation in palms on an oceanic island. *Nature*, *441*(7090), 210–213. https://doi.org/10.1038/nature04566

Schliewen, U., Rassmann, K., Markmann, M., Markert, J., Kocher, T., & Tautz, D. (2001). Genetic and ecological divergence of a monophyletic cichlid species pair under fully sympatric conditions in Lake Ejagham, Cameroon. *Molecular Ecology*, *10*(6), 1471–1488. https://doi.org/10.1046/j.1365-294X.2001.01276.x

Schliewen, U. K., Tautz, D., & Paabo, S. (1994). Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature*, *368*(6472), 629–632. https://doi.org/10.1038/368629a0

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., Brelsford, A., Clarkson, C. S., Eroukhmanoff, F., Feder, J. L., Fischer, M. C., Foote, A. D., Franchini, P., Jiggins, C. D., Jones, F. C., … Widmer, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, *15*(3), 176–192. https://doi.org/10.1038/Nrg3644

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., & Flook, P. (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene-sequences and a compilation of conserved polymerase chain-reaction primers. *Annals of the Entomological Society of America*, *87*(6), 651–701.

Smadja, C. M., & Butlin, R. K. (2011). A framework for comparing processes of speciation in the presence of gene flow. *Molecular Ecology*, *20*(24), 5123–5140. https://doi.org/10.1111/J.1365-294x.2011.05350.X

Tauber, C. A., & Tauber, M. J. (1977). Genetic model for sympatric speciation through habitat diversification and seasonal isolation. *Nature*, *268*(5622), 702–705.

Taylor, R. S., & Friesen, V. L. (2017). The role of allochrony in speciation. *Molecular Ecology*, *26*(13), 3330–3342. https://doi.org/10.1111/mec.14126

Thoday, J., & Gibson, J. (1962). Isolation by disruptive selection. *Nature*, *193*(4821), 1164–1166.

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). Clustal-W - Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*(22), 4673–4680.

Via, S. (2001). Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology & Evolution*, *16*(7), 381–390.

Zheng, X. W., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, *28*(24), 3326–3328. https://doi.org/10.1093/bioinformatics/bts606

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Kaiser TS, von Haeseler A, Tessmar-Raible K, Heckel DG. Timing strains of the marine insect *Clunio marinus* diverged and persist with gene flow. *Mol Ecol*. 2021;30:1264–1280. https://doi.org/10.1111/mec.15791