



## Research

**Cite this article:** Mérot C, Berdan EL, Babin C, Normandeau E, Wellenreuther M, Bernatchez L. 2018 Intercontinental karyotype–environment parallelism supports a role for a chromosomal inversion in local adaptation in a seaweed fly. *Proc. R. Soc. B* **285**: 20180519. <http://dx.doi.org/10.1098/rspb.2018.0519>

Received: 7 March 2018

Accepted: 24 May 2018

**Subject Category:**

Evolution

**Subject Areas:**

evolution, ecology, genetics

**Keywords:**

chromosomal inversions, environmental gradient, local adaptation, balancing selection, parallelism, Diptera

**Authors for correspondence:**

Claire Mérot

e-mail: [claire.merot@gmail.com](mailto:claire.merot@gmail.com)

Louis Bernatchez

e-mail: [louis.bernatchez@bio.ulaval.ca](mailto:louis.bernatchez@bio.ulaval.ca)

<sup>†</sup>These authors share equal authorship.

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.4127378>.

# Intercontinental karyotype–environment parallelism supports a role for a chromosomal inversion in local adaptation in a seaweed fly

Claire Mérot<sup>1</sup>, Emma L. Berdan<sup>2</sup>, Charles Babin<sup>1</sup>, Eric Normandeau<sup>1</sup>, Maren Wellenreuther<sup>3,4,†</sup> and Louis Bernatchez<sup>1,†</sup>

<sup>1</sup>Département de biologie, Université Laval, Quebec, Canada

<sup>2</sup>Department of Marine Sciences, University of Gothenburg, Gothenburg, Sweden

<sup>3</sup>School of Biological Sciences, University of Auckland, New Zealand

<sup>4</sup>Seafood Research Unit, Port Nelson, Nelson, New Zealand

CM, 0000-0003-2607-7818

Large chromosomal rearrangements are thought to facilitate adaptation to heterogeneous environments by limiting genomic recombination. Indeed, inversions have been implicated in adaptation along environmental clines and in ecotype specialization. Here, we combine classical ecological studies and population genetics to investigate an inversion polymorphism previously documented in Europe among natural populations of the seaweed fly *Coelopa frigida* along a latitudinal cline in North America. We test if the inversion is present in North America and polymorphic, assess which environmental conditions modulate the inversion karyotype frequencies, and document the relationship between inversion karyotype and adult size. We sampled nearly 2000 flies from 20 populations along several environmental gradients to quantify associations of inversion frequencies to heterogeneous environmental variables. Genotyping and phenotyping showed a widespread and conserved inversion polymorphism between Europe and America. Variation in inversion frequency was significantly associated with environmental factors, with parallel patterns between continents, indicating that the inversion may play a role in local adaptation. The three karyotypes of the inversion are differently favoured across microhabitats and represent life-history strategies likely to be maintained by the collective action of several mechanisms of balancing selection. Our study adds to the mounting evidence that inversions are facilitators of adaptation and enhance within-species diversity.

## 1. Introduction

Adaptation to heterogeneous environments is a major driver of evolution and in the diversification of life [1,2]. When a species occurs over a large geographical range, it experiences spatially variable conditions. Under limited migration, each population may follow its own evolutionary trajectory driven by local environmental conditions [1,3]. This can result in genetic and phenotypic polymorphism, variation at adaptive traits between habitats or along environmental clines, and, ultimately, diversification into ecotypes or species [3–5].

The ability to undergo polygenic specialization and local adaptation is related to the amount of genetic exchange between populations, because gene flow mixes unfavourable immigrant alleles with resident alleles. Negative modifiers of recombination can play an important role by limiting allele shuffling in parts of the genome [6–9]. Chromosomal inversions modify recombination because the heterozygote gene order is reversed between the standard and the inverted arrangements, resulting in limited recombination within the inverted part [10,11]. Various models have argued that inversions can facilitate

local adaptation when they trap a set of co-adaptive alleles [12–15], conditions which are not unusual when inversions span hundreds of genes [16].

Empirical evidence supporting a link between inversions and local adaptation comes from clines of inversion frequencies along environmental gradients [13,17,18] or chromosomal rearrangements associated with ecotype divergence [19–21]. The pattern of inversion frequency distributions can be variable, ranging from fixation between distinct habitats [20,22], clinal modulation of intermediate frequencies [17] to widespread polymorphism [23–25]. The fate of an inversion depends on the selective mechanisms at play and on the genes trapped within the inversion. Strong selection and steep differences between habitats can drive inversions near fixation when locally adaptive alleles are involved [14] and lower heterokaryotype fitness may drive divergence when populations carry alternative rearrangements [20,22,26]. By contrast, polymorphisms may be maintained by gene flow or if balancing selection is involved [25–29]. To disentangle these mechanisms, it is thus useful to combine knowledge of inversion effects on phenotypes with ecological and molecular data on natural populations harbouring the inversion across environmental gradients.

The seaweed fly *Coelopa frigida* provides a suitable model to understand how inversions facilitate adaptation to heterogeneous environments because it carries a large inversion, with the two forms called  $\alpha/\beta$  [30], that is polymorphic in all populations sampled so far. Studies on European *C. frigida* have shown that the inversion frequency varies along a latitudinal cline in Scandinavia which follows a natural gradient of temperature, salinity and seaweed composition, that the inversion has large phenotypic effects on male size, development time and fertility, and that heterokaryotes generally have a higher egg-to-adult viability than homokaryotes [31–36]. Here, we investigate American populations of *C. frigida* along the North Atlantic coast. The sampled area follows a latitudinal cline, a gradient of salinity into the St Lawrence Estuary, and spans a heterogeneous seaweed distribution. This allowed investigating separately the effects of different environmental parameters on the inversion frequency and testing the extent of parallelism with Europe. Specifically, we set out to determine if the inversion is present in North America and polymorphic, which environmental conditions modulate the inversion karyotype frequencies, and the relationship between inversion karyotype and adult size.

## 2. Methods

### (a) Study species, field sampling

*Coelopa frigida* belongs to the acalyprate flies and occupies a wide geographical range from Cape Cod (USA) to Greenland on the west coast of Atlantic Ocean and from Brittany (France) to Svalbard (Norway) on the east coast. Both larvae and adult flies are restricted to decomposing seaweed (wrackbed) for both food and habitat.

We sampled about 90–120 *C. frigida* individuals per population within a three-week period during September/October 2016 at 20 locations, spanning over 10° of latitude (figure 1a). Adult flies were collected with nets and preserved individually in ethanol or RNAlater. Environment at each location was described by three categories of variables: local wrackbed seaweed composition, local wrackbed abiotic characteristics and large-scale climatic/abiotic conditions (electronic supplementary

material, table S3). Wrackbed composition was an estimation of the relative proportions of Laminariaceae, Fucaceae, Zoosteraeaceae, plant debris and other seaweed species. Wrackbed abiotic characteristics included an estimation of the surface and a measure of depth, internal temperature and salinity. The three latter variables were estimated by averaging five measurements made at randomly selected points of the wrackbed with a salinity multimeter Aquaterr EC350.

Large-scale climatic/abiotic conditions were extracted for each location from public databases. These included the annual mean in precipitations and air temperature obtained from the Worldclim database with the R package *Raster* [37,38], the annual mean in sea surface temperature and sea surface salinity obtained from Marspec [39] (except for sites within the St Lawrence River Estuary: mareograph data from the OGSL). For annual mean tidal amplitude, we extracted hourly water level data from the closest station recorded by NOAA (USA) or Fisheries and Ocean (Canada) and then calculated the difference between the highest and the lowest water level each day and averaged over the year.

### (b) Fly sex determination, size measurements and inversion genotyping

Adult flies were examined under a binocular magnifier (Zeiss Stemi 2000C) to confirm species identification and to determine sex. For 1967 flies, the size was estimated using wing length as a proxy because wings can be mounted, flattened for photography and measured in a standardized way (electronic supplementary material, figure S6).

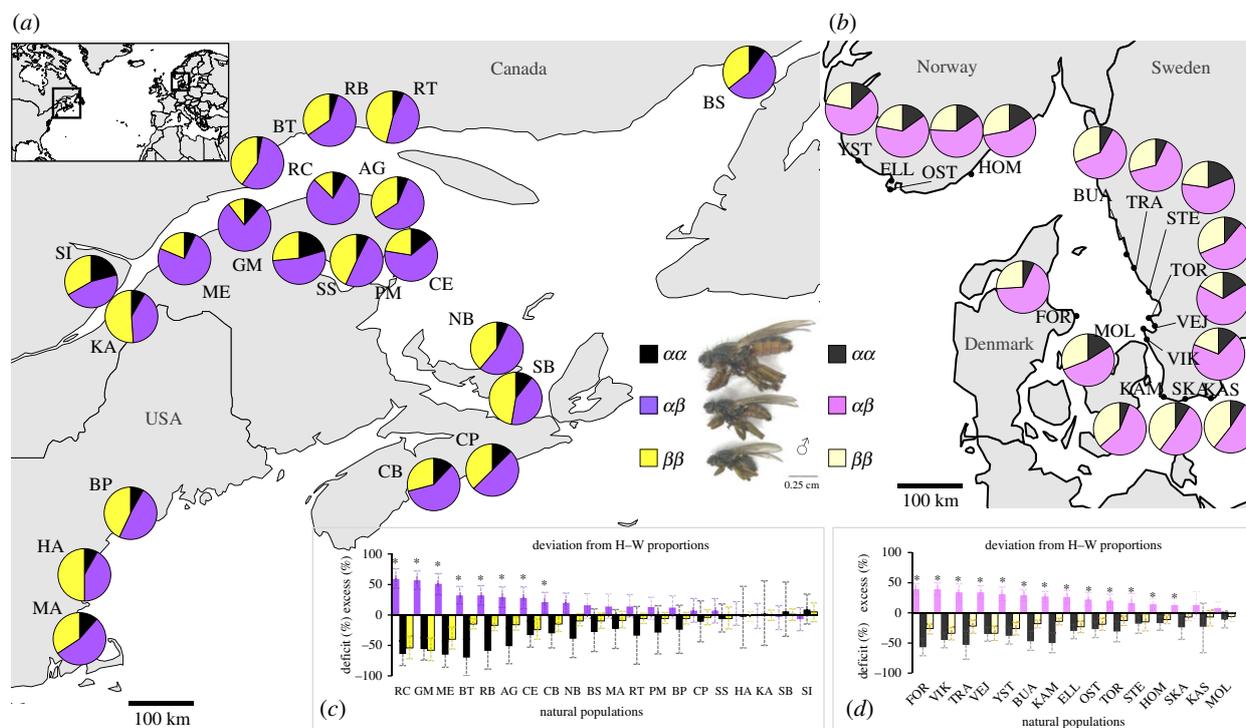
Previous work showed a strong correlation between the chromosome I inversion karyotype ( $\alpha/\beta$ ) and two alleles (B/D) of the alcohol dehydrogenase (*Adh*) allozyme [40]. We used this association to develop an inversion-specific DNA marker and targeted three coding regions within the inversion (*Adh* and two adjacent loci) on which we analysed linkage disequilibrium (LD) and haplotypic variation in American and European samples. LD was calculated as a squared allelic correlation  $R^2$  between unphased polymorphic sites (biallelic site with frequency higher than 5%), tested with  $\chi^2$ -test and visualized using the R package *LDheatmap* [41]. Haplotype phasing was inferred using coalescent-based Bayesian methods in DNAsp [42]. Haplotype networks were constructed with median joining in NETWORK 5.0.0 (<http://www.fluxus-engineering.com>).

The DNA marker consisted of two single-nucleotide polymorphisms that were associated with the different inversion rearrangements, and these were genotyped with two restriction enzymes (detailed in electronic supplementary material, table S1 and figure S1). It was validated with 44 samples previously karyotyped with the allozyme procedure as described by Edward *et al.* [34] (electronic supplementary material, table S2) and subsequently used to characterize the karyotype of 1988 wild American samples of *C. frigida* (89–117 individuals/population; electronic supplementary material, table S4). For each population, the frequency of  $\alpha$  rearrangement and the proportion of each karyotype were calculated in males and females separately, and then estimated for both sexes pooled at a sex-ratio of 1 : 1.

### (c) Statistical analyses

#### (i) Inversion frequencies and Hardy–Weinberg equilibrium

Heterogeneity in inversion frequencies and karyotypes frequencies was tested using an analysis of deviance on a generalized linear model (GLM) with binomial logistic transformation, followed by a comparison of contrasts, and a pairwise  $\chi^2$ -test (electronic supplementary material, figure S3) adjusted following [43]. Within each population, Hardy–Weinberg equilibrium (HWE) was tested using a  $\chi^2$ -test. Meta-analysis of HWE was



**Figure 1.** Inversion polymorphism along the North American Atlantic coast and Scandinavian coast. Map of sampled locations and proportions of the three karyotypes in (a) North America and (b) Scandinavia. The inset shows the location of the two areas in the North Atlantic area. Deviation from Hardy–Weinberg proportions for each karyotype by populations (calculated as the ratio of observed karyotype frequency over the expected karyotype frequency under HWE) in (c) America and (d) Scandinavia. Bars represent 95% confidence intervals. Asterisks denote significant deviation from HWE ( $\chi^2$ -test,  $p < 0.05$ ). All data from Scandinavia are extracted from Day *et al.* [36]. (Online version in colour.)

tested on this set of  $p$ -values using weighted Z-method. Deviation from HWE was calculated for each karyotype as the ratio of observed frequency over the expected frequency. Confidence intervals were drawn by bootstrapping (electronic supplementary material, table S4).

### (ii) Association inversion frequency/environment

Correlation between environmental variables was tested with a Pearson correlation test. For correlated environmental variables within the same category ( $p < 0.05$ , Pearson  $R^2 > 0.4$ ), a summary variable was drawn by retaining the first significant PC of a principal component analysis on original variables relying on the Kaiser–Guttman and broken stick criteria (electronic supplementary material, figure S4) [44]. Using the summary variables, variance inflation factor was lower than 2.72, indicating the absence of multicollinearity.

Associations between inversion frequencies and environmental variables were first tested for each variable alone, using a GLM with a logistic link function for binomial data, the response variable being the number of individuals carrying/not carrying the  $\alpha$  arrangement and the explanatory variable being an environmental variable, correcting for multiple testing following [43]. Then, the combined effect of the environmental variables was investigated by model selection. For each combination of explanatory variables, two kinds of model were implemented: a GLM, as described earlier, and a  $\beta$ -regression with response variables being the inversion frequency value. To identify the best model(s), several indicators were used following [45,46]. First, the models were ranked using AICc values (small-sample-size-corrected version of Akaike information criterion). Second, on the best 25 models, a jackknife (leave-one-out) procedure was used by repeatedly building the  $\beta$ -regression model on 19 populations and measuring its predictive fit over the 20 populations. Third, the adjusted  $R^2$  was compared between the most plausible models.

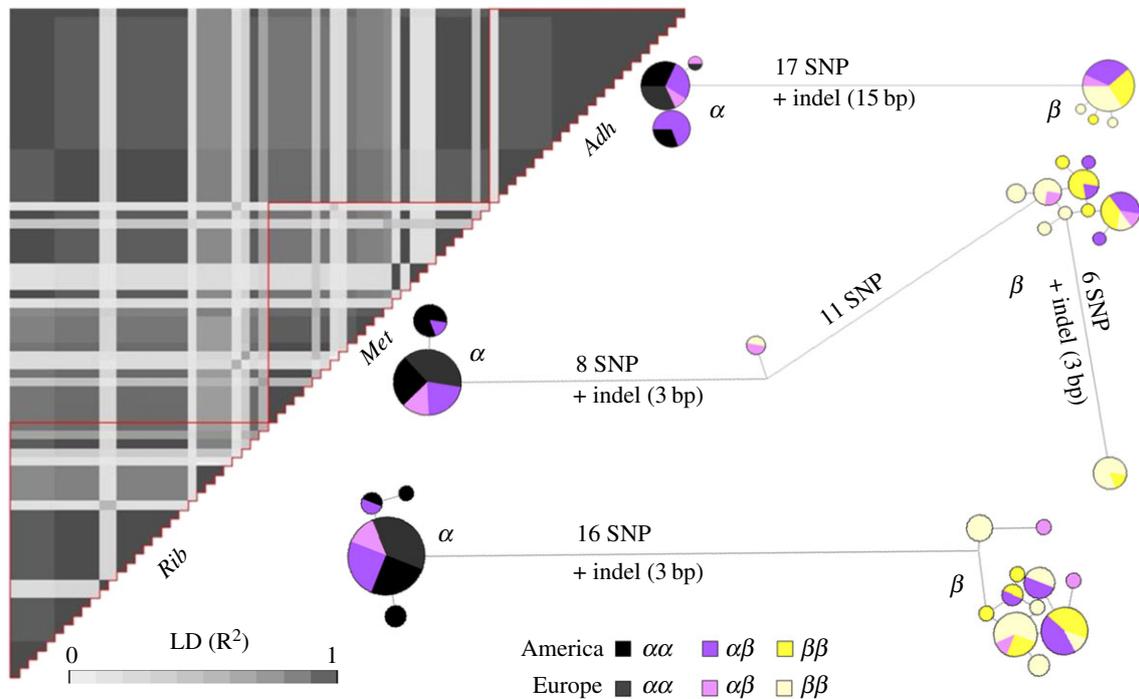
Association between the three karyotypes and environmental predictors was first modelled with a Dirichlet regression. The best Dirichlet model, however, could not predict the relative proportions of the three karyotypes with high accuracy ( $R^2$  being either small or negative). Comparing the predictive value of the different alternative Dirichlet models showed that each karyotype frequency was best predicted by a different combination of variables (electronic supplementary material, table S6). Therefore, the karyotype data were translated into binomial proportions (consisting of counts of each category divided by total counts) and analysed separately for each karyotype with binomial GLMs and  $\beta$ -regression models as described above. As these three models are not independent, they are interpreted accordingly.

Similar analyses were also performed considering as a response variable either the frequencies in males, females or with a sex-ratio as observed in the sampling and led to similar conclusions (electronic supplementary material, figure S5 and table S5). We also tested for spatial auto-correlation by building models of redundancy analysis which include environmental predictors and variables describing geographical proximity between populations (electronic supplementary material, table S7) [47].

For comparison, we re-analysed within the same framework data from Day *et al.* [36] which examines the cline in the frequency of the inversion in Scandinavia (figure 1*b,d*; electronic supplementary material, figure S9–S10 and tables S8–S10).

### (iii) Relationships between size, inversion and environment

Size variation in relation to karyotype and sex were analysed with a linear model and *post hoc* pairwise  $t$ -tests (adjusted following [43]). Residuals from this model are a measure of size variation between individuals, controlled for the combined effect of sex/karyotype and were kept for future analyses (called 'Residuals'). Size differences between populations were analysed using a linear model with sex/karyotype as covariates



**Figure 2.** Linkage disequilibrium and haplotype polymorphism. Heatmap representing LD ( $R^2$ ) within and between the three coding regions adjacent to the marker (29 unphased sequences). Haplotypes networks representing, for each coding region, similarity and differences between haplotypes (Adh: 42 samples, 84 haplotypes; Met/Rib: 31 samples, 62 haplotypes). Circle areas are proportional to the number of haplotypes with the same sequence. Links are proportional to the number of substitutions. For each locus, two main haplotype groups were found corresponding to the  $\alpha$  and the  $\beta$  rearrangements, as labelled. After phasing, heterokaryotypes typically have one haplotype in each group. (Online version in colour.)

and a *post hoc* *t*-test between each pair of populations on the *Residuals* (adjusted following [43]). The association between size and each environmental factor or inversion frequency was analysed using a linear mixed model with size as the response variable, environmental variables/inversion frequencies as explanatory variables, population as random factor and karyotype/sex as co-variables. As male mating success may be related to a male size advantage over females, we calculated, for each population, the mean size difference between each male karyotypic group and females, and tested with a linear model whether male–female size difference correlated with environmental variables/inversion frequencies.

All analyses were performed in R v. 3.4.2 [48] using the packages *lme4* [49], *AICcmodavg* [50], *corrplot* [51], *metap* [52], *HardyWeinberg* [53], *lsmeans* [54], *betareg* [55], *vegan* [56], *DirichletReg* [57] and *lmertest* [58].

### 3. Results

#### (a) A DNA marker of the inversion

The *Adh* gene and the two adjacent coding regions showed a characteristic pattern of low-recombination, consistent with an inversion. First, they were characterized by a very high linkage-disequilibrium within and between the three regions over approximately 8 kb with 89% of the SNPs being in significant linkage-disequilibrium (figure 2). Second, the three regions showed two distinct haplotype groups, which strongly differed by a total of 41 SNPs and three indels (figure 2). In the *Met* regions, 2 haplotypes (out of 62) that elsewhere belonged to the  $\beta$  group shared 11 SNPs characteristics of the  $\alpha$  haplotype, suggesting a possible (rare) event of recombination or gene conversion over at least 600 bp. Both  $\alpha$  and  $\beta$  haplotype groups included samples from Europe and America. Mean divergence was stronger between inversion

rearrangements (2.4%) than between populations from two continents (0.2%). The haplotype groups and the SNP targeted as a marker showed 100% of concordance with inversion rearrangement karyotypes as determined with the proven allozyme method (4/4( $\alpha\alpha$ ), 17/17( $\alpha\beta$ ), 23/23( $\beta\beta$ ); electronic supplementary material, table S2).

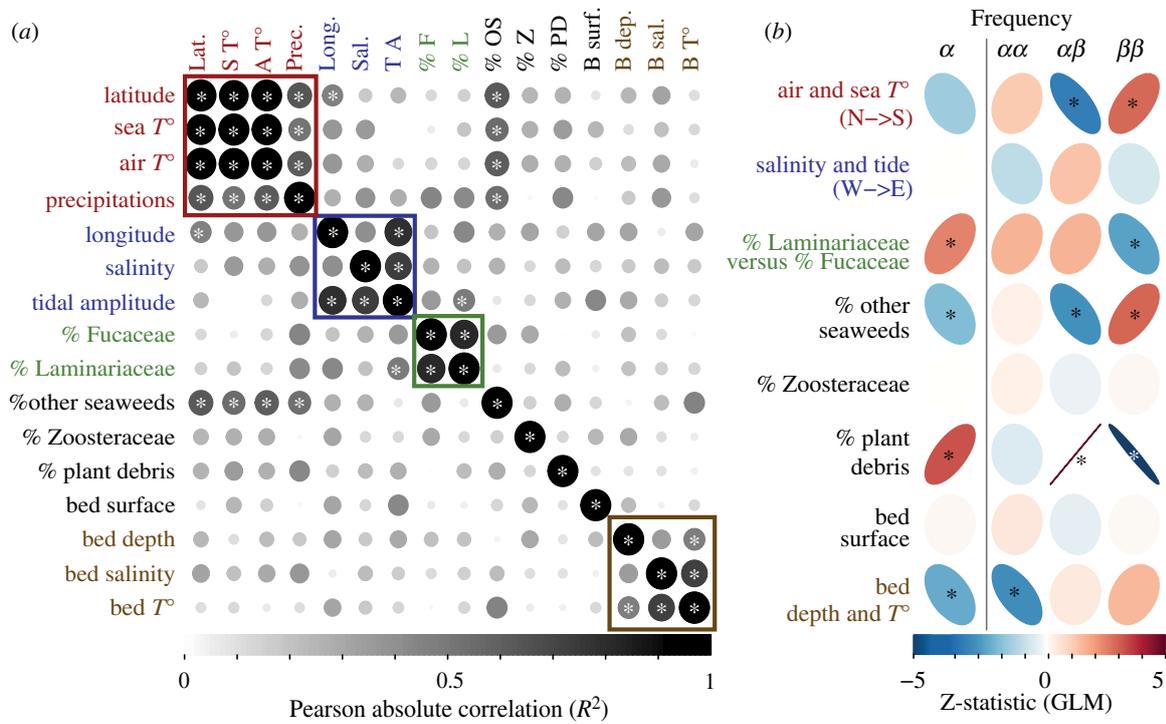
#### (b) Inversion and karyotype frequencies

All American populations were polymorphic for the inversion (figure 1a) and displayed the same global pattern, with  $\alpha$  being less frequent than  $\beta$  ( $\alpha$  mean frequency = 38% [28–51%]) and  $\alpha\alpha$  being the rarest karyotype ([5–21%]). Yet, inversion and karyotype frequencies were significantly heterogeneous between populations (deviance = 75( $\alpha$ ), 41( $\alpha\alpha$ ), 102( $\alpha\beta$ ), 123( $\beta\beta$ ); d.f. = 19,  $p < 0.001$ ; electronic supplementary material, figures S2–S3 and table S4).

Significant deviation from HWE was observed among all American populations (combined probabilities,  $p < 0.001$ ) translating into a mean excess of heterokaryotypes of 20%, due to a deficit of both homokaryotypes ( $\alpha\alpha$ : –30%,  $\beta\beta$ : –15%, electronic supplementary material, table S4). When considering the 20 populations individually, eight populations showed significant deviation from HWE with heterozygotes in excess, eight showed a slight excess of heterozygotes (non-significant) whereas four populations were at HWE (figure 1b).

#### (c) Inversion distribution and environmental variability

American populations spanned heterogeneous environments, whose variations could be described by two large-scale gradients and heterogeneity in local wrackbed characteristics (figure 3a; electronic supplementary material, figure S4).



**Figure 3.** Association between environmental variables and inversion karyotype frequencies. (a) Matrix of Pearson's correlations between environmental variables in North America. Coloured squares delimit groups of variables that were clustered in 'summary variables' for subsequent analyses (electronic supplementary material, figure S4). (b) Statistical associations between each environmental predictor and the frequency of  $\alpha$  rearrangement or the frequency of each karyotype. Strength and direction of the statistical association (GLM) are indicated by the shape and orientation of the ellipse, and its colour (red, positive; blue, negative). Asterisks denote significance at 0.05 level, corrected following [43]. (Online version in colour.)

The two gradients included a climatic north–south cline, along which covaried air temperature, sea temperature and precipitations, and a west–east cline, with lower sea salinity and higher tidal amplitude in the western part of St Lawrence Estuary. Sampled wrackbeds varied in the seaweed composition, generally dominated by either Laminariaceae or Fucaceae and whose proportions correlated significantly negatively. Zoosteraceae and plant debris were present in 6 out of 20 locations and represented less than 50% (electronic supplementary material, table S3). The abundance of other seaweed species correlated positively with sea temperature (figure 3a). Abiotic characteristics of the wrackbed were split into two independent dimensions: wrackbed surface and a summary variable associating wrackbed depth, temperature and salinity.

Variation in inversion frequencies were associated with variation in environmental parameters, namely local biotic and abiotic characteristics of the wrackbed, and more marginally, north–south climatic variation. In fact, the best significant predictors of inversion frequency were the composition of the wrackbed and the depth/ $T^\circ$  of the wrackbed, which, respectively, explained 30% and 9% of the variance in the best models (table 1; electronic supplementary material, table S5). Overall, the  $\alpha$  rearrangement was more frequent in shallow and cold wrackbeds, with a high proportion of Laminariaceae or plant debris while the  $\beta$  rearrangement was more frequent in deep and warm wrackbeds dominated by Fucaceae (figure 3b). Inversion frequency was also marginally associated with north–south climatic variation, or the correlated presence of other seaweeds, which both explained an additional 3% of variance in alternative models (figure 3b and table 1). The  $\alpha$  rearrangement frequency decreased in the south, in warmer areas that

contained a high proportion of other seaweed species. This result mirrored the parallel decrease in  $\alpha$  frequencies along the Scandinavian north–south thermic cline, in association with warmer air temperature and higher proportion of other seaweeds [36] (electronic supplementary material, figures S9–S10 and tables S8–S10).

More detailed analyses to investigate variations in karyotype composition underlying variations in rearrangement frequency showed that each environmental predictor was differentially associated with karyotype proportions (figure 3b and table 1; electronic supplementary material, figure S5 and tables S5–S7). The decrease in  $\alpha$  frequency in deeper, warmer wrackbeds was linked to a decrease of  $\alpha\alpha$  proportions (mostly relative to  $\beta\beta$ ). The increase in  $\alpha$  frequency with higher proportions of Laminariaceae (versus Fucaceae) was underlined by higher proportions of  $\alpha\alpha$  (and to a lesser extent  $\alpha\beta$ ) relatively to  $\beta\beta$ . The increase of  $\alpha$  frequency with the abundance of plant debris was related to higher proportions of  $\alpha\beta$  relatively to  $\beta\beta$ . The association between  $\alpha$  frequency and the north–south cline (climate/other seaweeds) was mostly due to higher proportions of  $\beta\beta$  relatively to  $\alpha\beta$  karyotypes in the south. Local conditions of the wrackbed were again the best predictors of karyotype proportions with the depth/ $T^\circ$  of the wrackbed or its composition (Laminariaceae, plant debris) predicting 28–34% of variance while climate (or the correlated abundance of other seaweeds) improved alternative model fit by 3–6% (electronic supplementary material, tables S5–S7). The positive association between Laminariaceae and  $\alpha\alpha$  proportions, as well as the positive association between other seaweeds/warmer air temperature and  $\beta\beta$  proportions showed parallelism on the Scandinavian cline (electronic supplementary material, figures S9–S10 and tables S9–S10).

**Table 1.** Best models explaining the distribution of inversion frequency by a combination of environmental variables. Grey line indicates the most plausible model minimizing the AICc of both the  $\beta$ -regression and GLM models. %LF = % Laminariaceae versus Fucaeeae, %PD = % plant debris, %OS = % other seaweeds.

model	$\beta$ -regression			GLM			$R^2$ adjusted	jackknife difference
	AICc	$\Delta i$	wAICc	AICc	$\Delta i$	wAICc		
frequency $\alpha \sim$ %LF + %PD + bed depth and $T^\circ$	-48	1	0.07	164	0	0.10	39%	5%
frequency $\alpha \sim$ %LF + %PD	-49	0	0.10	170	5	0.01	30%	5%
frequency $\alpha \sim$ %LF + %PD + Bed depth and $T^\circ$ + %OS	-45	4	0.01	164	0	0.10	42%	5%
frequency $\alpha \sim$ %LF + %PD + bed depth and $T^\circ$ + climate	-45	4	0.01	164	0	0.08	42%	5%

#### (d) Inversion type and size variation

Size was significantly associated with the inversion karyotypes, more strikingly for males than for females (figure 4a). For both sexes,  $\alpha\alpha$  was the largest karyotype and  $\beta\beta$  the smallest, with heterokaryotypes being intermediates. Size also varied significantly among populations ( $F_{19,1849} = 29$ ,  $p < 0.001$ ), with significant differences in 67% (128/190) pairwise comparisons between populations (electronic supplementary material, figure S7C). The mean size of each karyotype and sex observed at a given sampling location were significantly correlated (electronic supplementary material, figure S7AB), suggesting that local conditions similarly affect size in the three karyotypes and both sexes.

Size variation was not significantly associated with environment (electronic supplementary material, figure S7D). Size variation, controlling for karyotype and sex, was marginally associated with variation in karyotype frequencies ( $F_{1,19} = 3.8$ ,  $p = 0.06$ ; electronic supplementary material, figure S7D), with larger flies observed in populations with higher  $\alpha\alpha$  frequency and smaller flies found in populations with higher  $\alpha\beta$  frequency.

The mean size difference between male and female, a potential indicator of mating success [59], was constant for the smallest  $\beta\beta$  karyotype but varied between populations for  $\alpha\alpha$  and  $\alpha\beta$  males, correlating with the  $\alpha$  frequency in males (figure 4b). Male–female size difference increased at high  $\beta\beta$  frequencies (+100% ( $\alpha\alpha$ ), +70% ( $\alpha\beta$ )), but decreased at high  $\alpha\alpha$  frequencies (-30% ( $\alpha\alpha$ ); electronic supplementary material, figure S8).

## 4. Discussion

Investigating North American natural populations of *C. frigida* revealed the presence of a conserved  $\alpha/\beta$  inversion polymorphism, previously known in European populations [30]. Our results highlighted the importance of local and larger-scale environmental variation in explaining inversion karyotype frequencies, consistent with the prediction that the  $\alpha/\beta$  inversion may contribute to local adaptation [60]. Parallelism in the association between inversion distribution and environment, as well as the shared strong haplotype divergence between continents, supports the hypothesis that the inversion polymorphism has been subjected to

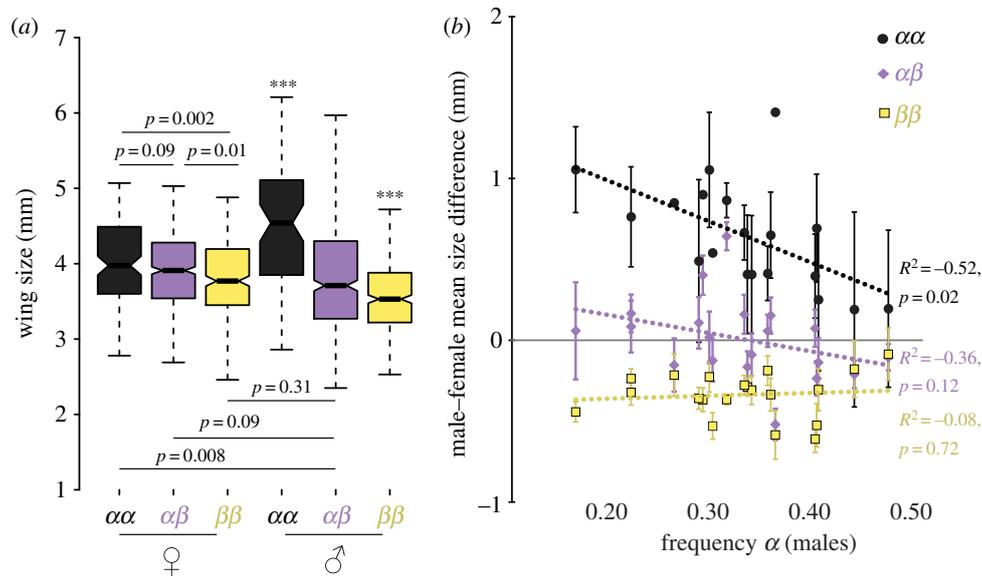
comparable evolutionary processes over an extended range of the species. We discuss hereafter how our results allow new insights into this intercontinental inversion polymorphism and how our data suggest a collective role for several mechanisms of balancing selection.

#### (a) A conserved intercontinental inversion polymorphism

As a result of the reduced recombination rate within the inversion, nucleotide sequences within the inversion are generally characterized by high LD and strong divergence between the different rearrangements [24,25,61]. In *C. frigida*, we identified these characteristics in three adjacent coding regions whose haplotypes were perfectly associated with the *Adh* allozyme marker of the  $\alpha/\beta$  inversion. This shows that recombination is strongly reduced between the two rearrangements, and provides the first reliable SNP marker for genotyping the  $\alpha/\beta$  inversion in *C. frigida*. Examining haplotype variations at the three regions further revealed that both haplotypes are found in Europe and America and that haplotype divergence is much stronger than intercontinental variations between populations 5000 km apart. Thus, although further genomic studies are needed to confirm whether this holds true along the inversion, our results point towards a conserved inversion haplotype block throughout the species range. European and American populations also display a similar relationship between inversion karyotype and adult wing size [31] as well as a parallel natural distribution of the inversion [36,40], thus indicating that the  $\alpha/\beta$  inversion in *C. frigida* represents a widespread polymorphism, with similar features conserved throughout the range of the species.

#### (b) A role for the inversion in local adaptation to heterogeneous environments

Large climatic gradients or heterogeneous habitats impose spatially variable selection which favours the evolution of differently adapted phenotypes, selected for local environmental conditions [1,3]. Inversions are particularly prone to be involved in such local adaptation because they may hold together sets of locally adapted alleles in the face of gene flow [14,62]. Consistent with these predictions, our results



**Figure 4.** Wing size in relation to karyotype and inversion frequency. (a) Wing size variation by sex and karyotype. Boxes indicate quartiles, notches are 95% confidence intervals of the median, whiskers extend to maximal values. \*\*\* denotes significant size differences with all other groups ( $p < 0.001$ ). For other comparisons, the  $p$ -value of the pairwise  $t$ -test is indicated by ' $p =$ '. (b) Male–female mean size difference within each population, as a function of  $\alpha$  frequency in males. Lines indicate Pearson's correlations. (Online version in colour.)

show that the *C. frigida*  $\alpha/\beta$  inversion frequencies covary in parallel with the climatic cline and wrackbed composition between both continents (electronic supplementary material, figure S9). Although not a hard proof, parallel patterns of genetic or phenotypic variation are considered to be strong indirect evidence for local adaptation shaped by natural selection [63].

Along large-scale latitudinal climatic gradients, both *C. frigida* European and American clines exhibit a slight increase of the  $\beta$  arrangement at southern locations. This frequency shifts may be a direct effect of increased air temperature, if  $\beta\beta$  are less cold-tolerant or if their smaller size/shorter development time is an advantage when warmer temperatures speed up wrackbed decomposition. The latter hypothesis has been used to explain the latitudinal size cline in *Drosophila*, with a smaller size and faster development time being favourable when larval food resources are ephemeral at warmer temperatures [64]. In *C. frigida*, this hypothesis is also supported by the significantly higher frequencies of  $\beta\beta$  in warm wrackbeds. The latitudinal cline of frequencies could also be an indirect effect due to the different kinds of seaweeds associated with the southern part of both clines.

Admittedly, clinal patterns of variation in allele frequencies could result from isolation by distance without the need to invoke selection. This, however, would unlikely result in similar directionality of the clinal variation across the two continents. Moreover, in Scandinavia, neither population differentiation nor any pattern of isolation-by-distance was observed with SSCP neutral markers over 300 km, suggesting that when the habitat is continuous, population structure is weak [65]. This may be surprising considering that the whole life cycle of *C. frigida* is subjugated to wrackbeds, however, occasional mass migratory flight have also been reported which could maintain regular migration between colonies, up to a few hundreds of kilometres [66,67]. Given the relatively continuous habitat observed in North America, little population structure is expected but this needs to be properly tested with neutral markers at a scale appropriate for the cline studied herein (1400 km).

Moreover, it is noteworthy that, rather than large-scale gradients, the best predictors explaining variation in *C. frigida* inversion frequencies were local wrackbed characteristics, such as the depth, temperature and composition of the wrackbed. The influence of the wrackbed composition is of particular interest considering its scale of heterogeneity. The global ratio of Laminariaceae/Fucaceae changes at an order of magnitude along a spatial scale of 100–200 km, a scale at which dispersal is expected [66], and at which local adaptation, related to wrackbed composition, has been observed in Scandinavian populations [60]. Further, the association between those two predictors and inversion frequencies remained when controlling for spatial auto-correlation (electronic supplementary material, table S7), suggesting that the environment–karyotype associations are not driven by environmental and genetic similarity between neighbour populations. Our data suggest that the relative proportions of seaweed species vary with habitat on which each karyotype is preferentially adapted. Both in Europe and America, an increased abundance of Laminariaceae is associated with an increased proportion of  $\alpha\alpha$  karyotypes (electronic supplementary material, figure S9), a result which is consistent with better survival of  $\alpha\alpha$  on Laminariaceae in the laboratory [34]. Mixed wrackbeds or plant debris favours heterokaryotypes while wrackbeds of Fucaceae or other seaweeds are associated with increased proportions of  $\beta\beta$ . The amount of resources available in each substrate may be one of the factors explaining the different karyotype proportions. In fact, in the laboratory, Laminariaceae sustain a greater viability and a larger size than Fucaceae [34], which suggests it is a richer substrate, facilitating the long larval development and the large size of  $\alpha\alpha$ . In the wild populations investigated here, higher  $\alpha\alpha$  proportions are also found in populations with large flies (all karyotypes/sexes), and insect size is generally a good indicator of larval growth conditions [68]. Adaptation to different substrates has also been found in the inversion rearrangements in cactophilic *Drosophila*, possibly linked to host-plant chemical compounds or microbiome fauna [69].

Interestingly, ecological factors identified as good predictors of *C. frigida* karyotype frequencies at a regional scale also vary with time and at a finer scale (i.e. within a heterogeneous wrackbed or between neighbouring beaches). Wrackbed depth and temperature can be patchy; the deposition of plant debris is linked to nearby river or storm events. Further work is needed to test whether the association between ecological predictors and karyotypes proportions observed at a regional scale hold true at a finer scale. If each of the three karyotypes is differentially favoured in each micro-habitat, they may represent a form of specialization maintained by micro-spatially varying selection balanced by gene flow. Balancing selection between micro-niches has been proposed in *Timema cristinae* stick-insects, in which an inversion underlies a green morph and a dark morph, respectively, cryptic on leaves or stems of the same host-plant [24]. With a scale of micro-habitat heterogeneity below dispersal distances, inversion structure can be even more important by linking together adaptive alleles in the face of gene flow and facilitating the coexistence of different ecotypes.

### (c) Additional mechanisms of balancing selection contributing to inversion polymorphism

Our results associating inversion karyotype frequencies with latitudinal gradients and seaweed habitats mirrors the well-described clines of inversion frequencies established along eco-climatic gradients in *Anopheles* mosquitoes [18], the mosaic of inversion karyotypes associated with soil moisture in *Mimulus guttatus* [20] and along altitudinal gradients in Kenyan *Apis mellifera* bees [21]. Yet, in those examples, inversion rearrangements are almost fixed at the end of the cline or between habitats, which suggests that selection for each rearrangement is locally very strong. In those cases, spatially heterogeneous selection, balanced by migration, appears to be the main mechanism determining overall polymorphism [22]. By contrast, clinal variation in *C. frigida* is more modest as all populations are polymorphic and frequencies remain at an intermediate range. Such pattern is not unusual in the literature, with clinal variation of frequencies in *D. melanogaster* spanning between 20 and 40% [17]; yet it also suggests that additional factors contribute in maintaining inversion polymorphism.

Heterozygote advantage is one the earliest explanation for the persistence of genetic polymorphisms in natural populations [70]. In *C. frigida*, our data and the literature suggest a fitness advantage for heterokaryotypes, which is likely to be the main mechanism underlying the persistence of this widespread polymorphism across space and time [71]. Indeed,  $\alpha\beta$  is in excess on both continents and, in the laboratory higher  $\alpha\beta$  survival is higher than homokaryotypes survival [32]. In the case of an inversion, higher viability of heterokaryotypes can be due to the inversion structure itself. Inversion breakpoints can disrupt important genes, which can be lethal for one homokaryotype, as shown in the case of the Ruff *Philomachus pugnax* [25]. Reduced recombination over such a large segment of the genome may also prevent the purge of deleterious effects within each rearrangement [26]. In fire ants *Solenopsis invicta*, one homokaryotype is lethal because of the accumulation of repetitive elements and deleterious mutations [72]. In the case of *C. frigida*, none of the homokaryotypes is lethal, but the parallel, repeatedly striking deficit of the homokaryotypes

suggests the presence of moderately deleterious effects. This is supported by evidence for genic selection in a series of inter- and intra-population crosses [73], and remains to be investigated at the genome level.

Variability in the excess of heterokaryotypes (0–60%) indicates that heterozygote advantage may be modulated by local biotic and abiotic conditions. For instance, higher proportions of  $\alpha\beta$  correlated with smaller sizes for all karyotypes/sexes, the latter relationship probably being mediated by density. In the laboratory, smaller size is linked to higher density of larva and density increases overdominance, with a 2.6-fold viability difference between heterokaryotypes and both homokaryotypes at high density but only 1.2-fold at low density [32]. Why heterokaryotypes are better competitors in some environmental contexts is unknown and remains to be investigated.

Heterokaryotype advantage can also result from reproduction [74]. Notably, analyses of wild *C. frigida* progeny suggest an excess of disassortative mating relatively to inversion karyotype, which may contribute to heterokaryote excess [75]. Polymorphism may also be maintained by opposing viability and sexual advantage. In Soay sheep (*Ovis aries*), one allele confers higher reproductive success while the alternative allele increases survival, resulting in increased overall fitness in the heterozygote [76]. This mechanism has also been suggested in *C. frigida* [59]. Because larger adults have a higher fertility,  $\alpha\alpha$  (and  $\alpha\beta$ ) have a sexual advantage over  $\beta\beta$ , particularly in males [35]. Yet, larger size and longer development time may not be easily achieved under low-resource conditions, competition or in ephemeral wrackbeds, giving  $\beta\beta$  (and  $\alpha\beta$ ) an egg-to-adult viability advantage over  $\alpha\alpha$  [32,59]. Such opposing selective pressures may explain the overall heterokaryotype advantage, but also the fluctuations of frequencies between low-resource and high-resource substrates. Moreover, in populations with high  $\beta\beta$  proportions, our results show that  $\alpha\alpha$  (and  $\alpha\beta$ ) males are not only rarer, facing less adult competition from large-size males, but they are also larger, possibly because of lower larval competition from similar karyotypes. Thus,  $\alpha$  rearrangement could benefit from the ‘advantage of the rare’, a form of frequency-dependent selection, which is frequently involved in protecting polymorphisms [32]. By contrast, in populations with high  $\alpha\alpha$  proportions, the observed lower male–female size difference and the numerous same-size competitors are expected to reduce  $\alpha\alpha$  males sexual advantage. As such, previous studies on European populations along with our results raise the hypothesis that the three karyotypes found in *C. frigida* may represent alternative life-history strategies with different relative investment in the trade-off between growth and reproduction, and for which balanced polymorphism could be maintained by a form of negative frequency-dependence selection [77], something that could be experimentally tested.

## 5. Conclusion

Our study shows that the  $\alpha/\beta$  inversion polymorphism is conserved between Europe and North America. Significant associations between karyotype frequencies and environmental variables on both continents provide strong indirect evidence for the role of the  $\alpha/\beta$  inversion in local adaptation. As such, *C. frigida* represents an excellent system to elucidate

the multifarious evolutionary mechanisms involved in the maintenance of structural variants. Our data indicate that the three inversion karyotypes are differently favoured by ecological conditions, and raises the hypothesis that they may represent three alternative life-history strategies, particularly in males. This joins theoretical predictions and accumulating evidence that within-species diversification and specialization can be made possible by the genomic architecture of the inversion itself [11]. Future work in *C. frigida* will focus on population genomics to investigate the contribution of drift and demographic factors, to assess the age of the inversion and to identify which loci are the targets of selection and how linkage increases (or decreases) fitness, depending on the selective process involved. Our analysis combined with the abundant life-history literature on *C. frigida* suggests that several balancing selection mechanisms (e.g. heterosis, genic selection, antagonistic sexual/natural selection, spatially varying and negative frequency-dependent selection) interact to maintain this polymorphism. Further modelling could help to disentangle the relative contributions of these processes in shaping geographical patterns of inversion frequencies and spatial associations. Interestingly, several recent studies also highlighted the combined effects of several balancing selection mechanisms on inversion polymorphism [24,28]. This indicates that the specific architecture of inversions may make them more likely, compared with single-locus polymorphism, to be subjected to multiple

and opposing selective factors, and asks under which conditions this results in transient polymorphisms, long-term polymorphisms or speciation.

**Data accessibility.** Data supporting this study are available at Dryad Digital Repository (<http://dx.doi.org/10.5061/dryad.pn2mq57>) [78].

**Authors' contribution.** C.M. designed the study, did fieldwork, laboratory work and photography, analysed sequences, performed the statistical analyses and drafted the manuscript. C.B. contributed to molecular laboratory work. E.N. and E.L.B. contributed to the development of the marker. L.B. and M.W. designed and coordinated the study, and helped draft the manuscript. All authors contributed in revising the manuscript and gave final approval for publication.

**Competing interests.** We declare we have no competing interests.

**Funding.** This research was supported by a discovery research grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to L.B., by the Canadian Research Chair in genomics and conservation of aquatic resources held by L.B. and by the Swedish Research Council grant 2012-3996 to M.W. C.M. was supported by a post-doctoral fellowship from the FRQNT and FRQS. E.L.B. was supported by a Marie-Curie Fellowship (H2020-MSCA-IF-2015, 704920).

**Acknowledgements.** We are very grateful to M. Lionard who sampled in Blanc-Sablon and to L. Johnson, E. Tamigneaux, D. Malloch for their advice during fieldwork. We thank G. Baigle, from the Statistic service (U. Laval), M. Laporte and J. Létourneau for their help with the statistical analysis. We thank the Microscopy service (IBIS, U. Laval), S. Bernatchez and B. Labbé for their help with wing photography. We thank three anonymous reviewers for detailed comments that improved the manuscript.

## References

- Savolainen O, Lascoux M, Merilä J. 2013 Ecological genomics of local adaptation. *Nat. Rev. Genet.* **14**, 807–820. (doi:10.1038/nrg3522)
- Stein A, Gerstner K, Kreft H. 2014 Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecol. Lett.* **17**, 866–880. (doi:10.1111/ele.12277)
- Hereford J. 2009 A quantitative survey of local adaptation and fitness trade-offs. *Am. Nat.* **173**, 579–588. (doi:10.1086/597611)
- Hedrick PW. 2006 Genetic polymorphism in heterogeneous environments: the age of genomics. *Annu. Rev. Ecol. Evol. Syst.* **37**, 67–93. (doi:10.1146/annurev.ecolsys.37.091305.110132)
- Ender JA. 1977 *Geographic variation, speciation, and clines*. Princeton, NJ: Princeton University Press.
- Nei M. 1967 Modification of linkage intensity by natural selection. *Genetics* **57**, 625–641.
- Nei M. 2003 Genome evolution: let's stick together. *Heredity* **90**, 411–412. (doi:10.1038/sj.hdy.6800287)
- Lenormand T, Otto SP. 2000 The evolution of recombination in a heterogeneous environment. *Genetics* **156**, 423–438.
- Charlesworth B. 1976 Recombination modification in a fluctuating environment. *Genetics* **83**, 181–195.
- Sturtevant A. 1921 A case of rearrangement of genes in *Drosophila*. *Proc. Natl Acad. Sci. USA* **7**, 235–237. (doi:10.1073/pnas.7.8.235)
- Kirkpatrick M. 2017 The evolution of genome structure by natural and sexual selection. *J. Hered.* **108**, 3–11. (doi:10.1093/jhered/esw041)
- Dobzhansky T. 1947 Genetics of natural populations. XIV. A response of certain gene arrangements in the third chromosome of *Drosophila pseudoobscura* to natural selection. *Genetics* **32**, 142–160.
- Dobzhansky T, Dobzhansky TG. 1970 *Genetics of the evolutionary process*. New York, NY: Columbia University Press.
- Kirkpatrick M, Barton N. 2006 Chromosome inversions, local adaptation and speciation. *Genetics* **173**, 419–434. (doi:10.1534/genetics.105.047985)
- Charlesworth B, Barton NH. 2018 The spread of an inversion with migration and selection. *Genetics* **208**, 377–382. (doi:10.1534/genetics.117.300426)
- Krimbas CB, Powell JR. 1992 *Drosophila inversion polymorphism*. Boca Raton, FL: CRC press.
- Kapun M, Fabian DK, Goudet J, Flatt T. 2016 Genomic evidence for adaptive inversion clines in *Drosophila melanogaster*. *Mol. Biol. Evol.* **33**, 1317–1336. (doi:10.1093/molbev/msw016)
- Ayala D, Acevedo P, Pombi M, Dia I, Boccolini D, Costantini C, Simard F, Fontenille D. 2017 Chromosome inversions and ecological plasticity in the main African malaria mosquitoes. *Evolution* **71**, 686–701. (doi:10.1111/evo.13176)
- Kirubakaran TG *et al.* 2016 Two adjacent inversions maintain genomic differentiation between migratory and stationary ecotypes of Atlantic cod. *Mol. Ecol.* **25**, 2130–2143. (doi:10.1111/mec.13592)
- Lowry DB, Willis JH. 2010 A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* **8**, e1000500. (doi:10.1371/journal.pbio.1000500)
- Wallberg A, Schöning C, Webster MT, Hasselmann M. 2017 Two extended haplotype blocks are associated with adaptation to high altitude habitats in East African honey bees. *PLoS Genet.* **13**, e1006792. (doi:10.1371/journal.pgen.1006792)
- Ayala D, Guerrero RF, Kirkpatrick M. 2013 Reproductive isolation and local adaptation quantified for a chromosome inversion in a malaria mosquito. *Evolution* **67**, 946–958. (doi:10.1111/j.1558-5646.2012.01836.x)
- Chouteau M, Arias M, Joron M. 2016 Warning signals are under positive frequency-dependent selection in nature. *Proc. Natl Acad. Sci. USA* **113**, 2164–2169. (doi:10.1073/pnas.1519216113)
- Lindtke D, Lucek K, Soria-Carrasco V, Villoutreix R, Farkas TE, Riesch R, Dennis SR, Gompert Z, Nosil P. 2017 Long-term balancing selection on chromosomal variants associated with crypsis in a stick insect. *Mol. Ecol.* **26**, 6189–6205. (doi:10.1111/mec.14280)
- Küpper C *et al.* 2016 A supergene determines highly divergent male reproductive morphs in the ruff. *Nat. Genet.* **48**, 79–83. (doi:10.1038/ng.3443)
- Kirkpatrick M. 2010 How and why chromosome inversions evolve. *PLoS Biol.* **8**, e1000501. (doi:10.1371/journal.pbio.1000501)
- Tuttle EM *et al.* 2016 Divergence and functional degradation of a sex chromosome-like supergene. *Curr. Biol.* **26**, 344–350. (doi:10.1016/j.cub.2015.11.069)

28. Llaurens V, Whibley A, Joron M. 2017 Genetic architecture and balancing selection: the life and death of differentiated variants. *Mol. Ecol.* **26**, 2430–2448. (doi:10.1111/mec.14051)
29. Hoffmann AA, Sgrò CM, Weeks AR. 2004 Chromosomal inversion polymorphisms and adaptation. *Trends Ecol. Evol.* **19**, 482–488. (doi:10.1016/j.tree.2004.06.013)
30. Aziz JB. 1975 Investigations into chromosomes 1, 2 and 3 of *Coelopa frigida* (Fab.). Phd thesis, University of Newcastle upon Tyne, UK.
31. Butlin R, Read I, Day T. 1982 The effects of a chromosomal inversion on adult size and male mating success in the seaweed fly, *Coelopa frigida*. *Heredity* **49**, 51–62. (doi:10.1038/hdy.1982.64)
32. Butlin R, Collins P, Day T. 1984 The effect of larval density on an inversion polymorphism in the seaweed fly, *Coelopa frigida*. *Heredity* **52**, 415–423. (doi:10.1038/hdy.1984.49)
33. Day T, Dobson T, Hillier P, Parkin D, Clarke B. 1980 Different rates of development associated with the alcohol dehydrogenase locus in the seaweed fly, *Coelopa frigida*. *Heredity* **44**, 321–326. (doi:10.1038/hdy.1980.29)
34. Edward DA, Gilburn AS. 2013 Male-specific genotype by environment interactions influence viability selection acting on a sexually selected inversion system in the seaweed fly, *Coelopa frigida*. *Evolution* **67**, 295–302. (doi:10.1111/j.1558-5646.2012.01754.x)
35. Butlin R, Day T. 1985 Adult size, longevity and fecundity in the seaweed fly, *Coelopa frigida*. *Heredity* **54**, 107–110. (doi:10.1038/hdy.1985.14)
36. Day T, Dawe C, Dobson T, Hillier P. 1983 A chromosomal inversion polymorphism in Scandinavian populations of the seaweed fly, *Coelopa frigida*. *Heredity* **99**, 135–145. (doi:10.1111/j.1601-5223.1983.tb00738.x)
37. Hijmans RJ, van Etten J. 2014 Raster: Geographic data analysis and modeling. R package.
38. Fick SE, Hijmans RJ. 2017 WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315. (doi:10.1002/joc.5086)
39. Sbrocco EJ, Barber PH. 2013 MARSPEC: ocean climate layers for marine spatial ecology. *Ecology* **94**, 979. (doi:10.1890/12-1358.1)
40. Butlin R, Collins P, Skevington S, Day T. 1982 Genetic variation at the alcohol dehydrogenase locus in natural populations of the seaweed fly, *Coelopa frigida*. *Heredity* **48**, 45–55. (doi:10.1038/hdy.1982.5)
41. Shin J-H, Blay S, Lewin-Koh N, McNeney B, McNeney MB. 2016 Package 'LDheatmap'. R package.
42. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. 2017 DnaSP6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **34**, 3299–3302. (doi:10.1093/molbev/msx248)
43. Benjamini Y, Hochberg Y. 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300.
44. Borcard D, Gillet F, Legendre P. 2011 *Numerical ecology with R*. Heidelberg, Germany: Springer Science & Business Media.
45. Burnham KP, Anderson DR. 2003 *Model selection and multimodel inference: a practical information-theoretic approach*. Heidelberg, Germany: Springer Science & Business Media.
46. Létourneau J, Ferchaud A, Le Luyer J, Laporte M, Garant D, Bernatchez L. 2018 Predicting the genetic impact of stocking in brook charr (*Salvelinus fontinalis*) by combining RAD sequencing and modeling of explanatory variables. *Evol. Appl.* **11**, 577–592. (doi.org/10.1111/eva.12566)
47. Borcard D, Legendre P. 2002 All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol. Model.* **153**, 51–68. (doi:10.1016/S0304-3800(01)00501-4)
48. R Core Team. 2017 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See <https://www.R-project.org/>.
49. Bates D, Maechler M, Bolker B, Walker S. 2014 lme4: Linear mixed-effects models using Eigen and S4. R package.
50. Mazerolle MJ, Mazerolle MMJ. 2017 Package 'AICcmodavg'. R package.
51. Wei T, Simko V, Levy M, Xie Y, Jin Y, Zemla J. 2017 Package 'corrplot'. *Statistician* **56**, 316–324.
52. Dewey M. 2017 metap: meta-analysis of significance values. R package.
53. Graffelman J. 2015 Exploring diallelic genetic markers: the hardy weinberg package. *J. Stat. Softw.* **64**, 1–23. (doi:10.18637/jss.v064.i03)
54. Lenth RV. 2016 Least-squares means: the R package lsmeans. *J. Stat. Softw.* **69**, 1–33. (doi:10.18637/jss.v069.i01)
55. Zeileis A, Cribari-Neto F, Gruen B, Kosmidis I, Simas AB, Rocha AV, Zeileis MA. 2016 Package 'betareg'. R package.
56. Oksanen J *et al.* 2013 Package 'vegan'. R package.
57. Maier MJ. 2014 DirichletReg: Dirichlet regression for compositional data in R. R package.
58. Kuznetsova A, Brockhoff PB, Christensen RHB. 2015 Package 'lmerTest'. R package.
59. Gilburn AS, Day TH. 1994 Sexual dimorphism, sexual selection and the  $\alpha\beta$  chromosomal inversion polymorphism in the seaweed fly, *Coelopa frigida*. *Proc. R. Soc. Lond. B* **257**, 303–309. (doi:10.1098/rspb.1994.0130)
60. Wellenreuther M, Rosenquist H, Jaksons P, Larson KW. 2017 Local adaptation along an environmental cline in a species with an inversion polymorphism. *J. Evol. Biol.* **30**, 1068–1077. (doi:10.1111/jeb.13064)
61. Joron M *et al.* 2011 Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* **477**, 203–206. (doi:10.1038/nature10341)
62. Hoffmann AA, Rieseberg LH. 2008 Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annu. Rev. Ecol. Syst.* **39**, 21–42. (doi:10.1146/annurev.ecolsys.39.110707.173532)
63. Schluter D. 2000 The ecology of adaptive radiation. OUP.
64. James AC, Azevedo R, Partridge L. 1995 Cellular basis and developmental timing in a size cline of *Drosophila melanogaster*. *Genetics* **140**, 659–666.
65. MacDonald C, Brookfield JFY. 2002 Intraspecific molecular variation in the seaweed fly *Coelopa frigida* consistent with behavioural distinctness of British and Swedish populations. *Mol. Ecol.* **11**, 1637–1646. (doi:10.1046/j.1365-294X.2002.01559.x)
66. Egglshaw H. 1961 Mass Migrational Flights of *Coelopa frigida* (Fabricius) and *C. pilipes* Haliday (Diptera, Coelopidae) and *Thoracochoaeta zostera* Hal. (Diptera, Sphaeroceridae). *Entomologist* **94**, 11–19.
67. Dobson T. 1974 Mechanisms controlling species composition in natural populations of the seaweed fly *Coelopa*. *J. Nat. Hist.* **8**, 653–673. (doi:10.1080/00222937400770561)
68. Chown SL, Gaston KJ. 2010 Body size variation in insects: a macroecological perspective. *Biol. Rev.* **85**, 139–169. (doi:10.1111/j.1469-185X.2009.00097.x)
69. Fernández Iriarte P, Hasson E. 2000 The role of the use of different host plants in the maintenance of the inversion polymorphism in the cactophilic *Drosophila buzzatii*. *Evolution* **54**, 1295–1302.
70. Fisher RA. 1923 XXI.—on the dominance ratio. *Proc. R. Soc. Edinb.* **42**, 321–341. (doi:10.1017/S0370164600023993)
71. Butlin R, Day T. 1989 Environmental correlates of inversion frequencies in natural populations of seaweed flies (*Coelopa frigida*). *Heredity* **62**, 223–232. (doi:10.1038/hdy.1989.32)
72. Wang J, Wurm Y, Nipitwattanaphon M, Riba-Grognuz O, Huang Y-C, Shoemaker D, Keller L. 2013 A Y-like social chromosome causes alternative colony organization in fire ants. *Nature* **493**, 664–668. (doi:10.1038/nature11832)
73. Butlin R, Day T. 1985 Genie and karyotypic selection on an inversion polymorphism in the seaweed fly, *Coelopa frigida*. *Heredity* **54**, 267–274. (doi:10.1038/hdy.1985.36)
74. Wellenreuther M, Svensson EI, Hansson B. 2014 Sexual selection and genetic colour polymorphisms in animals. *Mol. Ecol.* **23**, 5398–5414. (doi:10.1111/mec.12935)
75. Day TH, Butlin RK. 1987 Non-random mating in natural populations of the seaweed fly, *Coelopa frigida*. *Heredity* **58**, 213–220. (doi:10.1038/hdy.1987.35)
76. Johnston SE, Gratten J, Berenos C, Pilkington JG, Clutton-Brock TH, Pemberton JM, Slate J. 2013 Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature* **502**, 93–95. (doi:10.1038/nature12489)
77. Roff D. 2000 Trade-offs between growth and reproduction: an analysis of the quantitative genetic evidence. *J. Evol. Biol.* **13**, 434–445. (doi:10.1046/j.1420-9101.2000.00186.x)
78. Mérot C, Berdan EL, Babin C, Normandeau E, Wellenreuther M, Bernatchez L. 2018 Data from: Intercontinental karyotype–environment parallelism supports a role for a chromosomal inversion in local adaptation in a seaweed fly. Dryad Digital Repository. (doi:10.5061/dryad.pn2mq57)