Received: 4 January 2021 | Revised: 22 July 2021 | Accepted: 26 July 2021

DOI: 10.1111/mec.16099

SPECIAL ISSUE

MOLECULAR ECOLOGY WILEY

Transposable elements and introgression introduce genetic variation in the invasive ant *Cardiocondyla obscurior*

Mohammed Errbii[1] | Jens Keilwagen[2] | Katharina J. Hoff[3,4] | Raphael Steffen[1] | Janine Altmüller[5,6,7] | Jan Oettler[8] | Lukas Schrader[1]

[1]Institute for Evolution and Biodiversity, University of Münster, Münster, Germany

[2]Institute for Biosafety in Plant Biotechnology, Julius Kühn Institute (JKI) – Federal Research Centre for Cultivated Plants, Quedlinburg, Germany

[3]Institute of Mathematics and Computer Science, University of Greifswald, Greifswald, Germany

[4]Center for Functional Genomics of Microbes, University of Greifswald, Greifswald, Germany

[5]Cologne Center for Genomics, Institute of Human Genetics, University of Cologne, Cologne, Germany

[6]Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Core Facility Genomics, Berlin, Germany

[7]Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), Berlin, Germany

[8]Lehrstuhl für Zoologie/ Evolutionsbiologie, University Regensburg, Regensburg, Germany

Correspondence
Lukas Schrader, Institute for Evolution and Biodiversity, University of Münster, Münster 48149, Germany.
Email: lukas.schrader@uni-muenster.de

and
Jan Oettler, Lehrstuhl für Zoologie/ Evolutionsbiologie, University Regensburg, Regensburg 93053, Germany.
Email: joettler@gmail.com

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: 403813881, SCHR 1554/2-1 and OE 549/4-1

Abstract

Introduced populations of invasive organisms have to cope with novel environmental challenges, while having reduced genetic variation caused by founder effects. The mechanisms associated with this "genetic paradox of invasive species" has received considerable attention, yet few studies have examined the genomic architecture of invasive species. Populations of the heart node ant *Cardiocondyla obscurior* belong to two distinct lineages, a New World lineage so far only found in Latin America and a more globally distributed Old World lineage. In the present study, we use population genomic approaches to compare populations of the two lineages with apparent divergent invasive potential. We find that the strong genetic differentiation of the two lineages began at least 40,000 generations ago and that activity of transposable elements (TEs) has contributed significantly to the divergence of both lineages, possibly linked to the very unusual genomic distribution of TEs in this species. Furthermore, we show that introgression from the Old World lineage is a dominant source of genetic diversity in the New World lineage, despite the lineages' strong genetic differentiation. Our study uncovers mechanisms underlying novel genetic variation in introduced populations of *C. obscurior* that could contribute to the species' adaptive potential.

KEYWORDS
Cardiocondyla obscurior, introgression, invasive species, population genomics, rapid adaptation, transposable elements

1 | INTRODUCTION

Genetic variation underlies adaptation and species evolution (Frankham et al., 2002; Lande & Shannon, 1996). Genetically depleted populations, as a consequence of genetic drift, are more likely to suffer from a sudden change of their environment (Frankham, 2005a; Frankham & Ralls, 1998; Lacy, 1997). However, incipient populations of invasive species readily adapt to new habitats, despite their low genetic variation caused by the founder effect. The mechanisms underlying this unexpected adaptive capacity, known as "the genetic paradox of invasive species" (Frankham, 2005b), remain largely unclear. A vast number of species have been distributed around the world by human activities, allowing for the study of founder effects in such rapid adaptation processes.

Standing genetic variation and/or variation generated through multiple introduction events can facilitate adaptation (Barrett & Schluter, 2008; Kolbe et al., 2004; Lai et al., 2019). Particularly, hybridization between divergent non-native subspecies or lineages at an introduced range can lead to introgression of potentially adaptive variants (Hedrick, 2013; Nolte et al., 2009). Another potential source for adaptation is genetic variation generated by transposable elements (TEs) (Casacuberta & González, 2013; Marin et al., 2020). Since transposition can result in genomic changes ranging from simple point mutations to large structural variations (Horváth et al., 2017; Schrader & Schmitz, 2019; Stapley et al., 2015), TEs can in principle rapidly generate phenotypic variation, also including novel adaptive variants (Niu et al., 2019). Although extensive effort has been employed to identify and characterize TEs, our understanding of the extent to which they contribute to the process of rapid adaptation is still poor (Biémont, 2010; Feschotte & Pritham, 2007; Lisch, 2013).

The ant *Cardiocondyla obscurior* is an excellent model to disentangle sources of genetic variation (Heinze, 2017; Oettler, 2020). Originally from Southeast Asia, it has established populations in warm climates around the world. In this species, newly emerged queens mate inside the maternal nest with related nondispersing males and stay, or found a new colony in the vicinity, accompanied by some workers. This propensity for inbreeding together with the genetic bottlenecks imposed by founder events, is predicted to substantially reduce genetic variation in populations of this species (Heinze, 2017). Populations of *C. obscurior* belong to two (so far described) phenotypically and genetically distinct lineages (Schrader et al., 2014), which can interbreed asymmetrically due to *Wolbachia* infection (Ün et al., 2021). The Old World lineage is found globally (e.g., Taiwan, Japan, USA, Spain, and in greenhouses in Northern Europe), whereas a second lineage has been found only in Brazil so far (Heinze et al., 2006). Colonies of the Old World lineage on average contain more queens and show less intraspecific aggression (Schrader et al., 2014), which are both characteristic traits known from other invasive ants as well (Tsutsui & Suarez, 2003).

The genome of *C. obscurior* contains regions enriched in TEs ("TE islands"), in a background of low TE content ("low-density regions, LDRs") (Schrader et al., 2014). A similar genomic organization is so far only known from plant pathogenic fungi (Faino et al., 2016; Möller & Stukenbrock, 2017). In *C. obscurior*, TE islands are enriched for genes involved in the synthesis (fatty acid synthases) and perception (olfactory receptors) of cuticular hydrocarbons (Schrader et al., 2014), which play a particular prominent role in social Hymenoptera (Sprenger & Menzel, 2020). The distribution of TEs in the genome of *C. obscurior* has been suggested to play a prominent role in the evolution of this species (Schrader et al., 2014; Schrader & Schmitz, 2019).

Based on a new genome assembly and a population genomic and TE dynamics analysis, we explore the demographic history and different sources of genetic variation in populations of *C. obscurior*. We find that TEs are associated with genetic variation. Moreover, the data show recent TE activity in the Old World lineage, and introgression from the Old World into the New World lineage. Our study provides support that TEs and introgression contribute considerably to novel genetic variation in populations of *C. obscurior*.

2 | MATERIALS AND METHODS

2.1 | Sample collection, DNA extraction and sequencing

To study genetic variation, we performed short read pool-seq of two samples; a pool of 16 workers collected from 16 colonies from Tenerife, Spain, in 2019 (average distance between sites: 24 km), representing the Old World lineage and a pool of 30 workers collected from 30 colonies from Itabuna, Brazil, in 2018 (average distance between sites: 3.32 km), representing the New World lineage. To investigate the population structure and demographic history of *C. obscurior*, we analysed whole genome sequencing data of 12 single individual workers collected in 2018 from Itabuna ($n = 4$), Una ($n = 4$) and Guarujá ($n = 4$), Brazil, plus four workers collected from Leiden, Netherlands, and Taipei, Taiwan, two samples each (Table S1).

DNA from the pooled and single individual samples was extracted with the CTAB method and Illumina NovaSeq sequencing (paired-end 150-bp reads) was done at the Cologne Center for Genomics. Pooled samples were sequenced to an average coverage of >73×, while the individual samples were sequenced to a coverage of >12× (Table S1).

As backbone for the analyses, we performed two independent long read sequencing runs using Oxford Nanopore sequencing. To this end, high-molecular-weight DNA was extracted from a pool of 50 pupae, all coming from a single colony from Brazil, kept under strict inbreeding conditions in the laboratory for over 10 years. After quality control (agarose gel, Nanodrop) and quantification (Qubit), 1 µg of DNA template was used for 1D library construction with a Ligation Sequencing kit SQK-LSK109 (Oxford Nanopore). Next, 1 µg of library was loaded on a R9.4.1 FLO-MIN106 flowcell and sequenced on a MinION for 48 h. Base-calling with GUPPY 3.3.0 yielded 1,912,000 reads of 5.36 Gb (N50 = 6.51 kb) in the first run and 11,548,590 reads of 12.31 Gb (N50 = 1.67 kb) in the second run.

2.2 | Genome assembly, scaffolding, and polishing

After quality control with MINIONQC (Lanfear et al., 2019), Nanopore reads were filtered for adapter contaminations with PORECHOP (version 0.2.4). For genome assembly, we used 6 Gb of filtered long read data (~30× expected genome coverage). Here, we selected only the longest and highest quality with FILTLONG (version 0.2.0) (-split 500 --min_length 3000 --keep_percent 90 --target_bases 6000000000 --trim), using previously published Illumina short-read data (SRR1564444) for quality filtering of long reads.

Genome assembly was done using CANU (version 1.9) (Koren et al., 2017) (-genomeSize = 190 m, -correctedErrorRate = 0.085) resulting in 291 contigs (cN50 = 4.62 MB). Subsequently, we scaffolded the assembly with SSPACE-LONGREAD (version 1.1) (Boetzer & Pirovano, 2014) and SSPACE-STANDARD (version 3.0) (Boetzer et al., 2011) using long read data and previously published 20- and 8-kb 454 long-insert libraries (SRR1565732.1, SRR1565732.1) into 156 scaffolds (sN50 = 5.72 Mb). After gap-filling with LR_CLOSER, we performed a final round of scaffolding with SSPACE using a previously published short-read Illumina paired-end library (SRR1564444).

The assembly was polished with short-read data (SRR1564444) using NTEDIT (Warren et al., 2019) and four rounds of PILON (Walker et al., 2014). Quality control of each assembly step was done with BUSCO using the arthropoda_odb10 data set and AUGUSTUS starting parameter set fly (Simão et al., 2015) and QUAST (Gurevich et al., 2013).

We flagged and removed four scaffolds of the final assembly as prokaryotic, based on short-read coverage, GC content and homology to bacterial genomes as inferred with MEGABLAST (Morgulis et al., 2008). Three of the scaffolds (scf37, scf51, scf83) were identified as *Wolbachia* and one scaffold (scf36) was identified as *Candidatus* Westeberhardia cardiocondylae (Klein et al., 2016). Further, we flagged scaffold106 as mitochondrial based on BLASTN comparisons to the mitochondrial genome of *C. obscurior* (KX951753.1) (Liu et al., 2019).

2.3 | Primary annotation of transposable elements

We used a primary and several different alternative approaches to annotate TEs in the genome of *C. obscurior*. For the primary annotation, we compiled a TE library by first collecting 22,775 published arthropod-specific repeats from RepBase (release 20181026) and 11,972 Hymenoptera-specific repeats from ArTEdb (Wu & Lu, 2019). Next, we generated repeat libraries for 20 published ant genomes (Table S2) and the genome of *C. obscurior* using REPEATSCOUT (version 1.0.5) (Price et al., 2005). REPEATSCOUT produced between 2681 (*Pogonomyrmex barbatus*) and 16,100 (*Solenopsis invicta*) repeats for the published ant genomes and 2468 repeats for the genome of *C. obscurior*. Each species-specific de novo library was annotated by classifying contained TEs using PASTECLASSIFIER (Hoede et al., 2014) and TECLASS (Abrusán et al., 2009). Remaining unclassified elements were removed, if they showed significant similarity to annotated protein sequences of the respective species, according to BLASTX (with –task = megablast –evalue = 1e-10 -best_hit_score_edge = 0.05 -best_hit_overhang = 0.25 -outfmt = 7 -perc_identity = 50 -max_target_seqs = 1). After combining repeats from classified and filtered de novo libraries, RepBase and ArTEdb, we removed redundant sequences with CD-HIT (with -T 0 -c 0.8 -*n* 5).

The final library contained 91,492 sequences, including 1228 de novo repeats for *C. obscurior* and between 400 (*Atta cephalotes*) and 8204 (*S. invicta*) for other ants. Across the entire library, 3807 repeats remained unclassified, including 50 unclassified repeats from *C. obscurior* and between 25 (*A. cephalotes*) and 446 (*S. incivta*) for other ants. Using this library, we annotated TEs in the genome of *C. obscurior* with REPEATMASKER (version 4.0.8) in sensitive mode (with -s -gff -a -excln -cutoff 250 -nolow) and used One-code-to-find-them-all (Bailly-Bechet et al., 2014) to further assemble REPEATMASKER hits.

REPEATMASKER annotated 41,338 loci in the genome of *C. obscurior*, of which 20,756 loci were classified as DNA transposons, 9310 as LTRs, 4946 as LINE elements, 2696 as retroelements and 1636 as SINEs. In addition, 1994 annotated repeat loci remained unclassified.

We used these repeat annotations for calculating Kimura distance-based TE landscapes and for analysing the genomic distribution of TEs in the genome. TE landscapes were generated from REPEATMASKER's align and summary output files, with *calcDivergenceFromAlign.pl* from REPEATMASKER and *parseRM.pl* available from https://github.com/4ureliek/Parsing-RepeatMasker-Outputs assuming a mutation rate of 3.6×10^{-9}. The genomic distribution of TEs and exon content was assessed by processing gene and repeat annotations in gff format using BEDTOOLS (Quinlan & Hall, 2010) and BEDOPS (Neph et al., 2012), with subsequent visualization in R.

2.4 | Alternative approaches for the annotation of TEs confirming the primary annotation

To establish whether the results from the above described repeat annotations are robust, we repeated TE annotation using four alternative annotation strategies. First, we annotated TEs using the REPET pipeline as previously described (Dennis et al., 2020). Second, we generated de novo predictions for *C. obscurior* using two tools, EDTA and REPEATMODELER2. These de novo repeats were further classified with DEEPTE (Yan et al., 2020), TESORTER (https://github.com/zhangrengang/TEsorter), and PASTECLASSIFIER (Hoede et al., 2014), before removing unclassified repeats with significant (*e*-value < 1e-10) similarities to arthropod proteins in Uniprot (retrieved April 27, 2018) using BLASTX. The resulting library of de novo repeats was combined with arthropod-specific repeats from RepBase (release 20181026) and used to annotate repeats in the genome of *C. obscurior* with REPEATMASKER (with -s -gff -a -inv -small -excln -cutoff 250 -nolow) and REPEATCRAFT (https://github.com/niccw/repeatcraftp). Third, we used REPEATMASKER's default RepBase library (combined with the Dfam library) for annotating repeats in the genome, without including any de novo predicted repeats. In addition, we generated dedicated annotations for LTR transposons with GENOMETOOL's LTRharvest (http://genometools.org/tools/gt_ltrharvest.html).

Visual inspection of the different TE annotations in IGV (version 2.8.13) showed that all pipelines consistently recovered the uncommon genomic distribution of TEs in the genome of *C. obscurior* (Figure S1). Further, we found that our primary approach (REPEATSCOUT of 21 ants + REPEATMASKER) outperformed the other strategies as REPET-based annotations tended to combine independent loci to chimeric TEs and the REPEATMODELER/EDTA-based approach tended to miss loci consistently identified by the three other approaches (Figure S2). The large number of de novo predictions, however, precluded careful manual curation of the repeat library. To compensate, our stringent filtering of putative host proteins should, however, considerably reduce the number of falsely annotated TEs. Comparing Kimura distance-based TE landscape plots for different annotations produced with REPEATMASKER further showed identical patterns, with a high frequency of (nearly) identical TE copies in the genome of *C. obscurior*.

2.5 | Annotation of TE islands

To define TE islands, we calculated TE content for 500-kb sliding windows (sliding by 100 kb), based on our primary *C. obscurior* TE annotation. Windows containing over 50% TE content were annotated as TE-enriched. Consecutive windows of enriched TE content were merged into larger discrete TE islands. Finally, we manually refined boundaries of all TE islands by overlaying TE annotations in IGV (version 2.8.13).

2.6 | Gene annotation

Genes were annotated by combining homology and evidence-based approaches, integrating RNA-seq data. RNA-seq data of *C. obscurior* were downloaded from NCBI (accessions PRJNA237579, PRJNA284224, PRJNA293450 and PRJNA309926). Adapter sequences and short reads (<40 bp) were removed using TRIMMOMATIC (version 0.38) (Bolger et al., 2014) and rRNA reads were filtered out with SORTMERNA (Kopylova et al., 2012). The resulting reads were then mapped to the softmasked *C. obscurior* genome (Cobs2.1) using HISAT2 (version 2.1.0) (Kim et al., 2015) with default parameters.

The alignment files were then provided to GEMOMA (Keilwagen et al., 2016) and BRAKER1 (Hoff et al., 2016). Six species were used as reference in GEMOMA comprising *Wasmannia auropunctata*, *Linepithema humile*, *Camponotus floridanus*, *Harpegnathos saltator*, *Ooceraea biroi* and *Solenopsis invicta*. The corresponding annotations and genomes were downloaded from NCBI. MMSEQS2 (Steinegger & Söding, 2017) was used to search for homologues of coding exons in the target genomes, which were then combined into gene predictions by GEMOMA.

BRAKER1, which combines GENEMARK-ET (Lomsadze et al., 2014) with AUGUSTUS (Stanke et al., 2006; Stanke & Waack, 2003) for RNA-seq based annotation, was applied to the softmasked genome (Cobs2.1) in combination with the RNA-seq alignment file.

INTERPROSCAN (version 5.39-77.0) (Jones et al., 2014) was run to functionally annotate the resulting gene predictions files and an additional, binary gff attribute functionalAnnotation was added to each file using the module AddAttribute from GEMOMA.

The gene prediction files were then combined with the module GAF from GEMOMA using the filter "start=='M' and stop=='*' and (functionalAnnotation=true or score/AA>=0.75)" to obtain the final annotation. We removed genes coding for proteins with functional annotations related to TEs (Table S3), with significant (*e*-value < 1e-10) similarities to RepBase proteins (repbase20.05) according to BLASTP, or genes overlapping by more than 50% with predicted TEs.

2.7 | Mapping, SNP calling and filtering

FASTQC (version 0.11.7) was used to inspect the quality of the raw reads. TRIMMOMATIC (version 0.38) was used to remove short and low-quality reads (Bolger et al., 2014). The resulting reads were mapped to the genome using BWA-MEM (version 0.7.17) with default parameters (Li & Durbin, 2009). The quality of generated alignments was assessed using QUALIMAP (version 2.2.1) (Okonechnikov et al., 2016). For single nucleotide polymorphism (SNP) calling, duplicates were marked using PICARD's MARKDUPLICATES (version 2.20.0) and joint-variant calling was performed using GATK's HAPLOTYPECALLER (version 4.1.2.0) with default parameters across all samples (Van der Auwera et al., 2013). InDels were excluded from the raw variant set (1,124,968 variants) and the resulting SNPs (841,937) were hard-filtered using the GATK VariantFiltration tool based on the following parameters: the variant confidence (QUAL < 30.0), the variant confidence normalized by depth (QD < 2.0), estimates of the variant strand bias test (FS > 60.0 and SOR > 3.0), mapping quality (MQ < 40.0 and MQRankSum < −5.0) and the rank sum test for the positioning of variants within reads (ReadPosRankSum < −5.0). We based the choice of these parameters on a visual inspection of the distribution of annotation values of the raw SNPs as recommended by the GATK documentation (Depristo et al., 2011).

The 715,833 SNPs that passed the applied hard filters were further filtered using VCFTOOLS (version 0.1.16) (Danecek et al., 2011) excluding SNPs that did not satisfy one of the following criteria: mean coverage greater than or equal to 5× for the single individuals and 10× for the pooled samples. Furthermore, we only considered biallelic SNPs (for the pooled samples, only the two most likely alleles were used) with an MAF >0.05 that have been genotyped in at least 85% of our samples, resulting in 541,252 SNPs. These cut-offs were chosen based on the distribution of these parameters in the SNPs left after applying the hard filters. Finally, we excluded short (<1 Mb) scaffolds with low mapping quality (Figure S3) and considered 522,267 SNPs identified in the 30 largest scaffolds (representing 93% of the assembly ~179.63 Mb), in all downstream analyses.

2.8 | Population structure analyses

To investigate genetic structure in our data set, we conducted a principal component analysis (PCA) using PLINK (version 1.90p) (Chang et al., 2015; Purcell et al., 2007) and ran the model-based clustering algorithm in ADMIXTURE (version 1.3.0) (Alexander et al., 2009). Prior to perform analyses, PLINK was used to LD-prune the 522,267 SNPs at a threshold of $r^2 < .2$ (*--indep-pairwise* 1 kb 1 0.2). The resulting 115,334 LD-pruned SNPs were then used to perform a PCA using the *--pca* 4 command of PLINK. The LD-pruned SNP data set was used to run ADMIXTURE considering values of *K* from 2 to 4 and kept only the two *k* values showing the lowest cross-validation error.

2.9 | Demographic population history analysis

We inferred demographic population history using Multiple Sequential Markovian Coalescence (MSMC2; Schiffels & Durbin, 2014; Schiffels & Wang, 2020). As representatives of the New World lineage, we used sequencing data of four single individuals (i.e., eight haplotypes) collected from Itabuna and Una, Brazil, while for Old World lineage, we used two single individuals each from Leiden, Netherlands, and Taipei, Taiwan. Their membership to each of the lineages was confirmed using ADMIXTURE and PCA-based analyses.

MSMC2 requires phased haplotype data of single chromosomes as input, so we used SHAPEIT (version 2.r904) (Delaneau et al., 2012) to statistically phase the 30 largest scaffolds of the *C. obscurior* genome. Due to the absence of a reference panel for *C. obscurior*, we used the 16 single individual genomes produced in this study for phasing. Moreover, to produce a reliable set of phased genomes, we only considered 328,167 SNPs with no missing data and an MAF >0.1 for phasing. Prior to phasing, we separated the SNPs genotyped across all 16 individuals into scaffolds using BCFTOOLS (version 1.9). Phasing was then performed by running SHAPEIT (*-check*, *-phase* and *-convert* commands) on each scaffold separately.

We generated mapability masks for each scaffold using SNPABLE (http://lh3lh3.users.sourceforge.net/snpable.shtml).

Effective population size (N_e) and the relative cross-coalescence rate (rCCR) were then calculated using MSMC2 (Schiffels & Wang, 2020). Five hundred bootstrap replicates were run by randomly sampling 60 Mb for each of the analysed genomes.

Since there are no available mutation rates for ants, we used estimates from another social hymenopteran: 3.6×10^{-9} mutations per generation per site, estimated in bumblebees (Liu et al., 2017), which is similar to that of honeybees (Yang et al., 2015). Additionally, our demographic model assumes a similar mutation rate and generation time for the two lineages. Therefore, the results should be taken with some caution as the evolutionary rates might differ among species and even between populations.

2.10 | Population genomics metrics estimation

To describe the genome-wide patterns of variation and differentiation in populations of *C. obscurior*, we estimated three parameters commonly used in population genetics from the pool-seq data: pairwise nucleotide diversity (π) and Tajima's *D* using POPOOLATION (Kofler, Orozco-terWengel, et al., 2011) and genetic differentiation (F_{ST}) using POPOOLATION2 (version 1.201) (Kofler, Pandey, et al., 2011). Both tools are suitable for analysing pool-seq data.

For π and Tajima's *D*, we first generated an mpileup file from the alignment bam file of each pool using SAMTOOLS (version 1.7) (Li et al., 2009). The mpileup files were filtered from indels using two perl scripts (*identify-genomic-indel-regions.pl* and *filter-pileup-by-gtf.pl*) available under POPOOLATION. Both population genetic parameters were then estimated using the POPOOLATION *Variance-sliding.pl* script with a minimum allele count of two. When it calculates π and Tajima's *D*, POPOOLATION accounts for the bias introduced by pooling and sequencing errors (Kofler, Orozco-terWengel, et al., 2011).

To estimate pairwise F_{ST}, we combined the alignment files of the two pools to produce a single mpileup file using SAMTOOLS. The resulting file was then converted into a synchronized file following POPOOLATION2's manual. Finally, F_{ST} was calculated using *fst-sliding.pl* available under POPOOLATION2 (Kofler, Pandey, et al., 2011).

All three estimates were computed in 100-kb nonoverlapping windows across the 30 largest scaffolds. To reduce the number of falsely called variants in duplicated and repetitive regions in the genome, we only considered positions with a coverage ranging from a minimum of 10× to the top 2% in each pool (i.e., Tenerife: 79× and Itabuna: 75×).

2.11 | Repeat quantification and TE insertions identification

We used DNAPIPETE (version 1.3.1) (Goubert et al., 2015) to assemble, annotate and quantify the repeatome in the two pool data sets using raw reads from each sample.

DNAPIPETE was run on low-coverage read samples (*-genome_coverage* 0.1×) from each population with four iterations (*-sample_number* 4). To avoid reference biases, we used a TE library combining the RepBase (version 25.04) (Jurka et al., 2005) and the Dfam library (release Dfam_3.2) (Hubley et al., 2016). We applied an increasing number of iterations (two to five) and found that four iterations performed best as this value maximized the N50 of the assembled repeat contigs.

To estimate TE abundance and frequencies of both reference and de novo TE insertions in the two studied populations we used POPOOLATIONTE2 (Kofler et al., 2016). First, the reference genome was masked using REPEATMASKER. Then, the masked genome was combined with the TE library to produce a TE-merged reference genome. POPOOLATIONTE2 also requires a TE hierarchy that was generated using for every entry in the TE library its id, order and family information.

Paired-end reads from the two populations were mapped to the TE-merged reference genome with BWA-BWASW (Li & Durbin, 2009) and paired-end information was recovered with POPOOLATIONTE2 *se2pe*. The resulting bam files were then used to generate a ppileup (physical pileup) file using POPOOLATIONTE2 *ppileup* with the *--map-qual 15* option. To correct for insert size differences between populations (Figure S4), we subsampled the physical coverage to equal levels (target coverage = 20) in the two populations, as recommended by Kofler et al. (2016).

Next, we identified TE insertions and their frequency in each sample separately by running POPOOLATIONTE2 with the following changes to the default parameters: (i) identifySignatures (*--min-count 2*), (ii) frequency, (iii) filterSignatures (*--max-otherte-count* 2 *--max-structvar-count 2*) and (iv) pairupSignatures (*--min-distance* -200 *--max-distance* 300).

All TE insertions were then categorized based on their frequency in each population, as low frequency (<25%), common (25%–95%) and fixed (>95%) TE insertions.

We performed Pearson's Chi-squared analyses to test for differences between genomic regions and populations.

All statistical analysis and figures were produced using R (version 4.0.2) (R Core Team, 2020).

3 | RESULTS

We generated a new assembly Cobs2.1 (193.05 Mb) that is more contiguous than Cobs1.4 (177.9 Mb), comprising only 127 scaffolds (sN50 = 6.29 Mb, IN50 = 11; Table S4) instead of 1854 scaffolds and containing 98.3% complete BUSCOs (0.5% duplications).

By combining homology- and RNA-seq-based predictions and after filtering for putative TE-encoded genes, we produced a final gene set containing 20,966 coding genes with 36,624 transcripts.

We identify 34 TE enriched regions between 41.7 kb and 2.05 Mb in size (24.57 Mb in total) (Figure 1a; Figure S1; Table

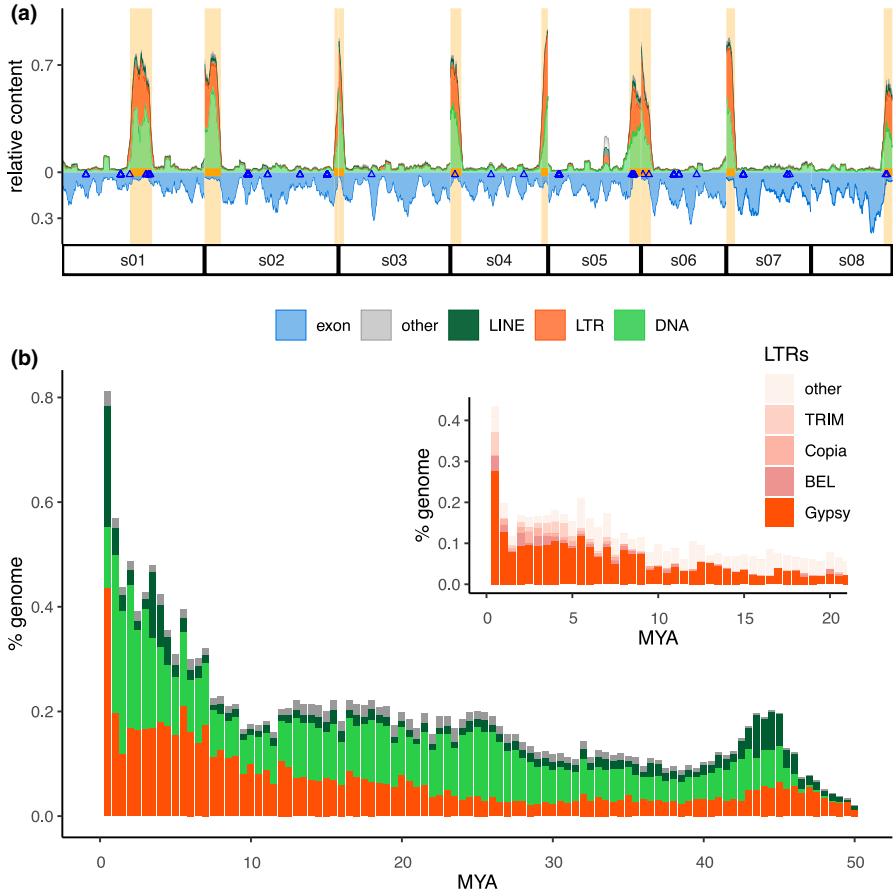

FIGURE 1 Genomic organization and architecture in *Cardiocondyla obscurior*. (a) Relative content of exonic and TE-derived sequences along the eight largest scaffolds of the *C. obscurior* genome. Shown are DNA transposons (DNA), long interspersed nuclear element (LINE) and long terminal repeat (LTR) retrotransposons, as well as other TEs (other). The genome is well structured into TE-poor regions ("low-density regions," LDRs) and TE-rich regions ("TE islands," orange highlights). Blue triangles indicate the genomic position of odorant receptor (OR) genes. (b) Kimura distance-based copy divergence analysis of TEs in *C. obscurior*. The smaller inset panel shows Kimura distance-based divergence of different LTR families. Low Kimura distance values (*x*-axis) correspond to TE copies (as proportion of the genome) that did not diverge from the consensus sequence (i.e., recent copies), while high *k*-values correspond to more divergent copies (i.e., old copies). The shape of the distribution with a peak at 0 divergence suggests recent activity of TEs, particularly the LTR/Gypsy family

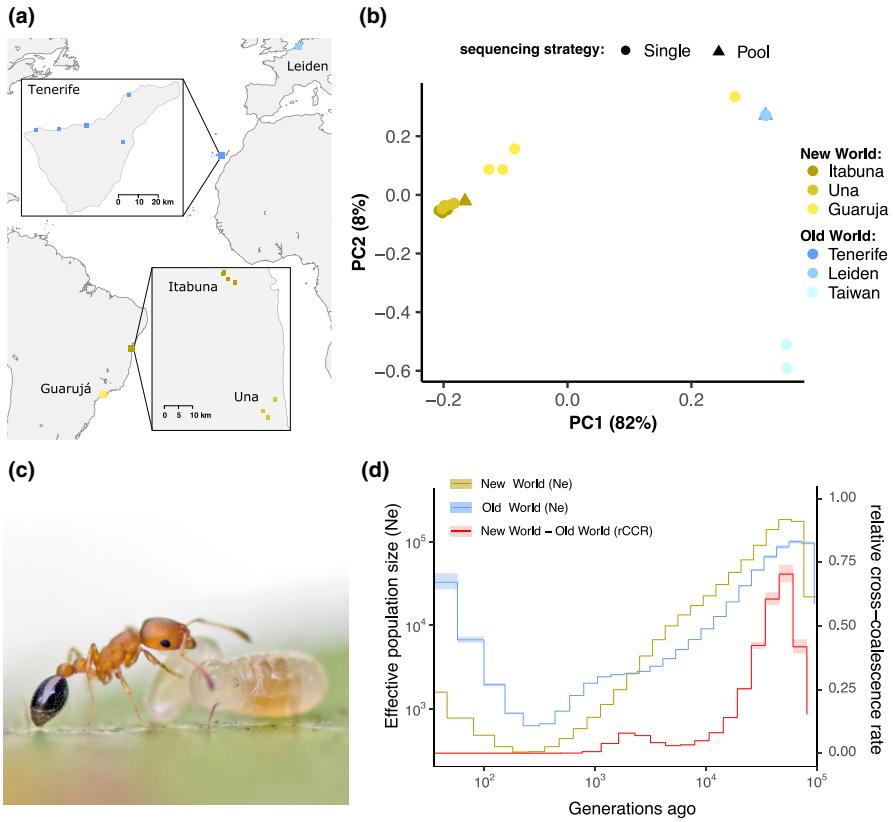

FIGURE 2 Genetic divergence and demographic population history in *Cardiocondyla obscurior*. (a) Sampling locations of the New World lineage populations collected from Brazil, and populations of the Old World lineage collected from Tenerife, Netherlands and Taiwan (not shown on the map). (b) Principal component analysis based on 115,334 SNPs of single and pooled samples belonging to the Old and New World lineages. PC 2, which only explains 8% of the variation, separates the Old World lineage into two discrete clusters. (c) Worker of *C. obscurior* with two larvae. (d) Population history in *C. obscurior* estimated using MSMC2 with eight phased haplotypes, representing the New and Old World lineages respectively. Lines and shaded areas are means and 95% confidence intervals, respectively

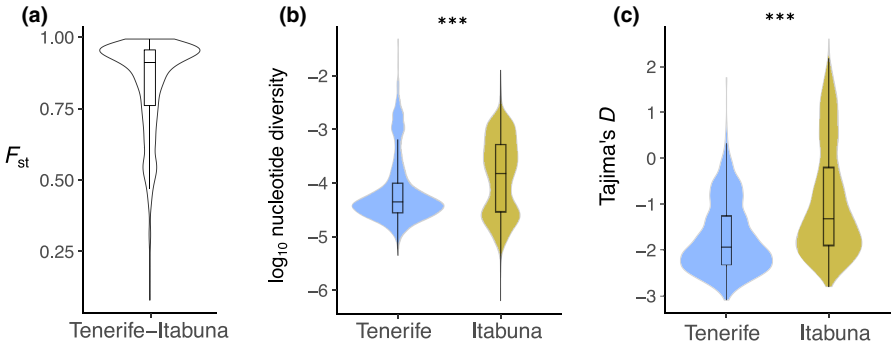

FIGURE 3 Genetic diversity and differentiation averaged per population in *Cardiocondyla obscurior*. (a) Genetic differentiation, (b) nucleotide diversity (π) and (c) Tajima's *D* in two populations of *C. obscurior*, from Tenerife and Itabuna. All estimates were calculated in 100-kb nonoverlapping windows across the genome and averaged per population (***$p < .0001$)

S5). With only 2111 protein coding genes, gene density is significantly reduced in TE islands relative to LDRs (Fisher's exact test, $p < 2.2e\text{-}16$), but they are enriched for odorant receptor (OR) genes (e.g., GO:0004984, olfactory receptor activity, $p = .00012$, parent-child tests) (Figure 1a) and fatty acid synthase (FAS) genes (e.g., GO:0050660, flavin adenine dinucleotide binding, $p = .00056$, parent–child tests) (Table S6), confirming previous findings (Schrader et al., 2014).

To explore the history of TE proliferation in the genome, we analysed divergence among TE copies using Kimura distances. The presence of conserved TE copies in the genome indicates TE activity in recent times, particularly of the LTR retrotransposons Gypsy family (LTR/Gypsy) (Figure 1b).

3.1 | Populations of *C. obscurior* belong to two discrete lineages

We performed short-read pool sequencing of two samples (pool sizes: New World lineage, 30 workers; Old World lineage, 16 workers) (Figure 2a; Table S1) as well as single individual sequencing of 12 workers sampled from Itabuna ($n = 4$), Una ($n = 4$) and Guarujá ($n = 4$), Brazil, in addition to four workers collected from Leiden ($n = 2$), Netherlands, and Taipei ($n = 2$), Taiwan (Table S1, see Methods for details on sampling and sequencing).

We conducted a PCA on 115,334 LD-pruned SNP markers, showing that PC 1, which explains 82% of the variation, separates the data set into two distinct clusters: the New World lineage

(Itabuna and Una) and the Old World lineage (Leiden, Taiwan and Tenerife, Figure 2b). Surprisingly, while individuals from Guarujá predominantly cluster with the New World lineage, they are shifted on PC 1 towards the Old World lineage (with one sample clustering with the Old World lineage), suggesting ongoing genetic admixture between the Old and the New World lineage in this population. ADMIXTURE-based analyses further support this finding (Figure S5).

3.2 | Old and New World lineages diverged 40,000 generations ago

Modelling divergence history and demographic dynamics of *C. obscurior* (Figure 2c) revealed similar demographic trajectories. Both lineages show a continuous decline in effective population size (N_e) from approximately 50,000 to 300 generations ago, consistent with a single *C. obscurior* ancestral population that underwent an extended bottleneck during this period (Figure 2d). At 200 generations ago, N_e increases in both lineages, but particularly in the Old World lineage.

Estimates of rCCR (Schiffels & Durbin, 2014) indicate that the two lineages started to diverge at least 40,000 generations ago, coincidental with the onset of the continuous decline in N_e (Figure 2d). Around 5000 generations ago, both lineages were almost completely isolated (rCCR ≈ 0). A recent increase in rCCR around 4000 generations ago points to secondary contacts between the lineages. Repeating the analysis using four individuals all coming from Itabuna resulted in a similar population demographic history (Figure S6).

3.3 | Populations from Itabuna and Tenerife experienced a bottleneck recently and are expanding

While there is little genetic structure between populations of the New World lineage from Itabuna and Una (Figure S7), suggesting a single source population, the pooled samples from Tenerife and Itabuna show strong genetic differentiation ($F_{ST\ median}$ = 0.91) (Figure 3a).

F_{ST} compares the levels of genetic variance within populations relative to that between populations. Therefore, it is sensitive to demographic processes (e.g., bottlenecks or population expansions) affecting genetic variation within populations (Chakraborty & Nei, 1977; Charlesworth, 1998; Hedrick, 1999; Nei, 1973). We thus also estimated nucleotide diversity (π) (Nei & Li, 1979) in both populations using POPOOLATION, correcting for pool size and depth of coverage (Kofler, Orozco-terWengel, et al., 2011). Nucleotide diversity is significantly lower in Tenerife (π_{median} = 4.43e-5) compared to Itabuna (π_{median} = 1.5e-4) (Figure 3b; Wilcoxon test: W = 1,250,531, $p < .0001$), possibly reflecting that the Tenerife population has been introduced more recently.

To better understand the factors that might have caused this difference between Tenerife and Itabuna, we examined deviation from neutrality by estimating Tajima's D in the two populations. Tajima's D is a test based on the difference between π and its expected value (Watterson's θ [θ_W]; Watterson, 1975), commonly employed in population genetics to detect signatures of selection, but it also informs about demographic processes (Tajima, 1989a, 1989b).

On average, Tajima's D is more negative in Tenerife (median = −1.93) compared to Itabuna (median = −1.31) (Figure 3c), with more genomic regions evolving neutrally (Tajima's $D \approx 0$) or under balancing selection (Tajima's $D > 0$) in Itabuna (Figure S8; Fisher's exact test, $p < 2.2e\text{-}16$).

The reduced levels of genetic diversity together with an excess of low-frequency variants (as indicated by a negative Tajima's D) confirm that both populations are recovering from a recent genetic bottleneck, particularly the Tenerife population.

3.4 | Evolutionary rate in TE islands differs between lineages

We previously found that TE islands diverge faster than the rest of the genome of *C. obscurior*, with TE islands harbouring a large fraction of sequence variation identified between the Old and New World lineages (Schrader et al., 2014). To investigate this further, we analysed intragenomic variation of diversity, neutrality and differentiation.

F_{ST} is significantly reduced in TE islands (median = 0.67) compared with LDRs (median = 0.92) (Figure 4a; Wilcoxon test: W = 306,473, $p < .0001$), concordant with genetic variation within populations being higher in TE islands.

In Tenerife, genetic diversity within TE islands is 32-fold higher than in LDRs (Figure 4b, Kruskal–Wallis rank sum test, χ^2 = 883.63, df = 3, $p < .0001$; results of pairwise Wilcoxon rank sum post hoc tests are displayed on the figure). In the Itabuna population, the increase is only 2.4-fold.

When comparing both populations, we find opposing patterns of local genomic variation. In Tenerife, genetic variation in TE islands is more than three times higher than in Itabuna. However, we find the inverse in LDRs, with Itabuna having almost four times more genetic variation than Tenerife (Figure 4b; results of pairwise Wilcoxon rank sum post hoc tests are displayed on the figure).

Consistent with this observation, Tajima's D is significantly higher in TE islands (median = −0.75) compared to LDRs (median = −2.04) only in Tenerife (Figure 4c; Kruskal–Wallis rank sum test, χ^2 = 790.01, df = 3, $p < .0001$; results of pairwise Wilcoxon rank sum post hoc tests are displayed on the figure), but not in Itabuna ($LDRs_{median}$ = −1.32; TE islands $_{median}$ = −1.27) (Figure 4c).

3.5 | Introgression introduces genetic variation into New World populations

Irrespective of TE content, our genome scans further reveal the existence of large genomic spans (e.g., on scaffolds 1, 4 and 8) with a

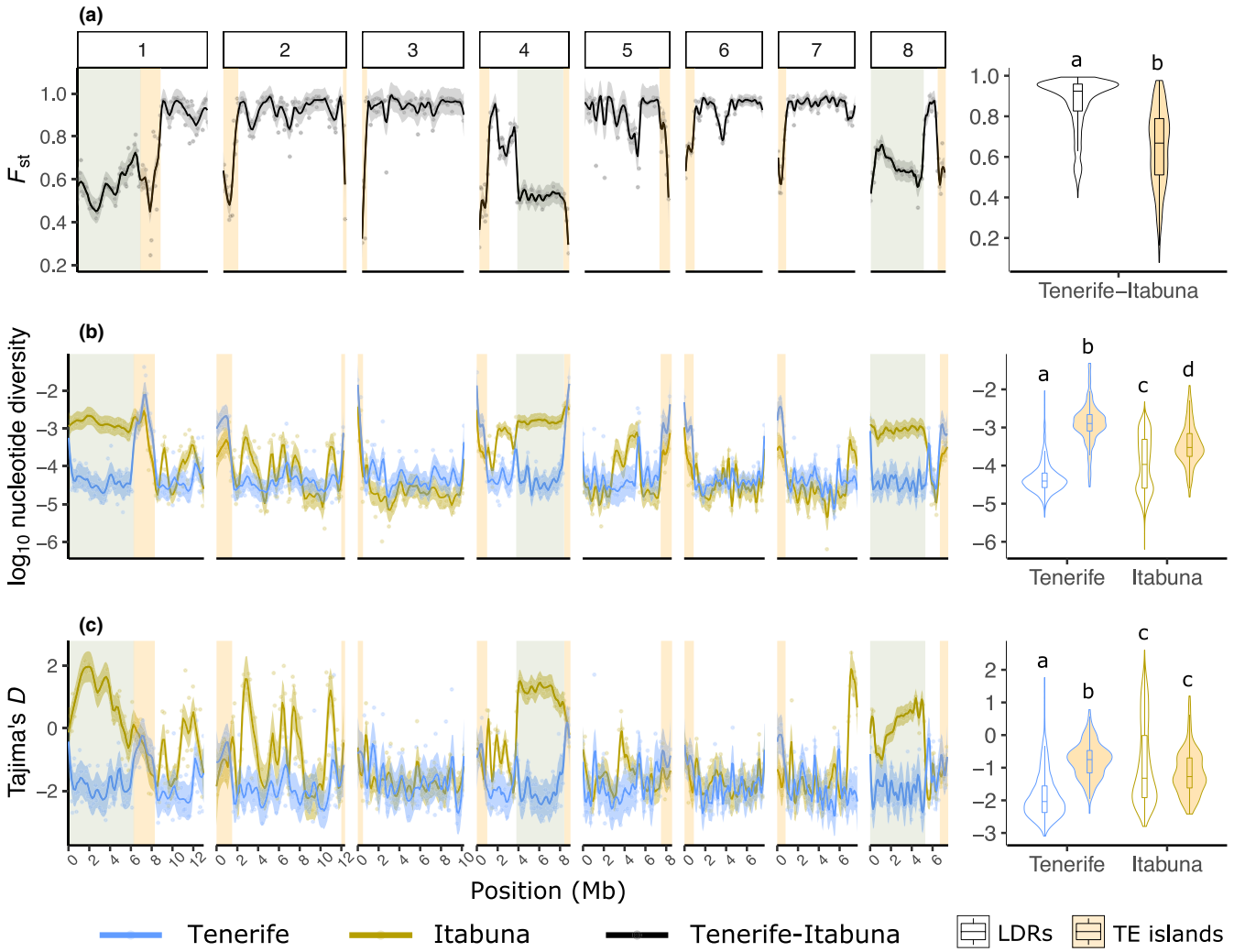

FIGURE 4 Genome-wide distribution of genetic diversity and differentiation of the eight largest *Cardiocondyla obscurior* genome scaffolds. (a) Genetic differentiation (F_{ST}), (b) nucleotide diversity (π) and (c) Tajima's *D* and in two populations of *C. obscurior*, from Tenerife and Itabuna. All estimates were calculated in 100-kb nonoverlapping windows. Regions highlighted in orange are TE islands and in green are potentially introgressed regions. Lines and shaded areas are means and 95% confidence intervals, respectively. Violin plots show the distribution of each estimate within LDRs and TE islands in each population. Different letters represent significant differences according to pairwise Wilcoxon rank sum post hoc tests. Data for other scaffolds (1–30) are shown in Figures S9–S11

distinct pattern of genetic variation and differentiation within LDRs in Itabuna, but not in Tenerife (Figure 4a–c; Figures S9–S11).

These blocks range between 4.6 and 6.3 Mb in size, show 1.6-fold less differentiation (Wilcoxon test: $W = 281{,}531$, $p < .0001$), harbour 6-fold more nucleotide diversity (Wilcoxon test: $W = 4529$, $p < .0001$) and have a positive Tajima's *D* (median = 0.72; Wilcoxon test: $W = 33{,}948$, $p < .0001$), compared to the LDRs average (Figure S12).

Such heterogeneous patterns of genetic variation and differentiation could hint towards intraspecific introgression from the Old World lineage into the New World lineage. In fact, new alleles brought in after a secondary contact between formerly isolated populations can create an excess of common variants. Under such a scenario, genetic diversity is expected to increase within the recipient population, causing Tajima's *D* to become positive and genetic differentiation between donor–recipient populations to be lower in introgressed regions compared to the remainder of the genome (Alcala et al., 2013; Burgarella et al., 2019).

The detected patterns of genetic diversity and differentiation within and among populations described above were not affected by excluding all potentially introgressed genomic blocks (Figure S13).

3.6 | TEs are more active in the Old World lineage

Given the prominent signature of TEs in the genome of the species, we explored whether divergent activity and proliferation of TEs could contribute to the structured patterns of genetic diversification of both lineages.

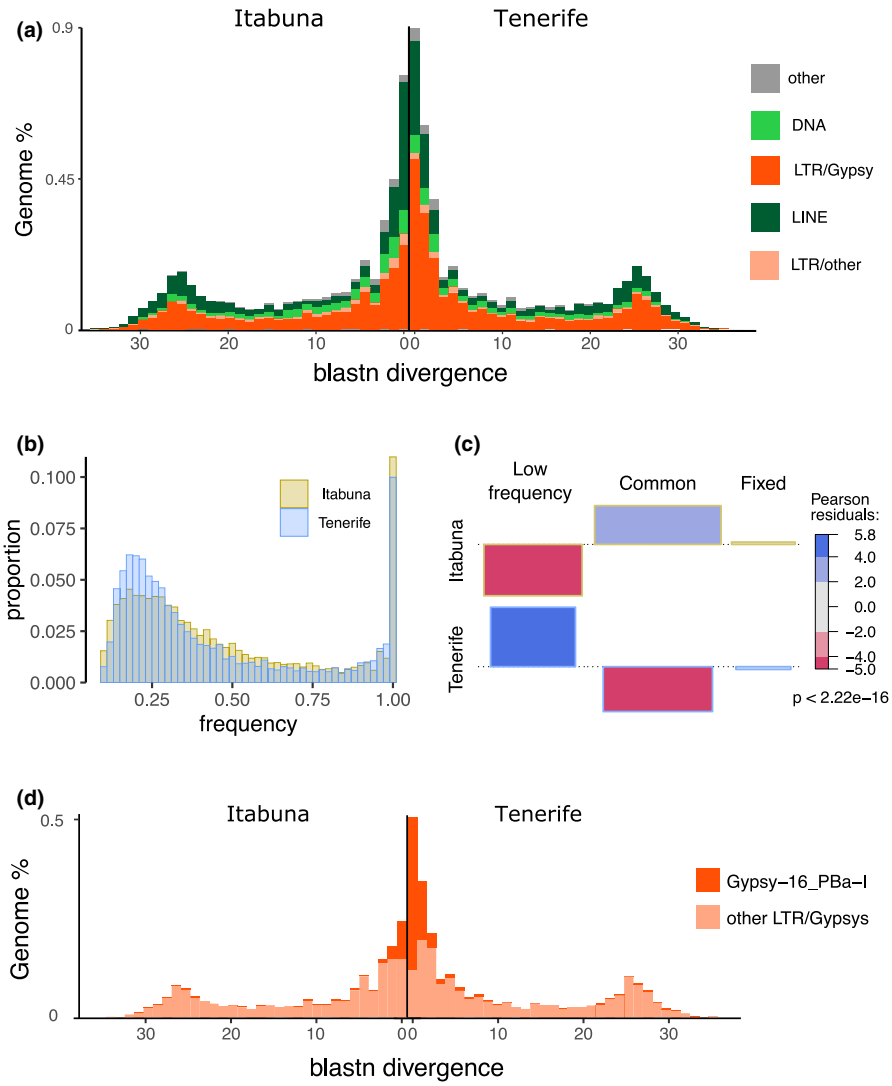

FIGURE 5 Repeat content and dynamics in populations of *Cardiocondyla obscurior*. (a) Divergence between TE copies of the main families found in populations of *C. obscurior* from Itabuna (left) and Tenerife (right), using DNAPIPETE with 0.1× coverage per samples. The shape of the distribution with a peak at 0 divergence suggests recent activity of TEs in both populations. The relatively higher peak in Tenerife suggests TE activity to recently have been higher in this population. (b) Frequency distribution of TE insertions identified by POPOOLATIONTE2 in populations of *C. obscurior* from Tenerife and Itabuna. The significant (see panel c) shift towards more rare variants in Tenerife suggests current transposon activity in the population. (c) Association plot showing the signed contribution to Pearson's χ^2 (left panels) and the Pearson's χ^2 standardized residuals (right panels) calculated for the identified TE insertions. For each population, the identified TE insertions were categorized based on their frequency as low frequency (<25%), common (25%–95%) and fixed (>95%) TE insertions. Blue colour of the rectangles indicates a positive correlation, while dark pink indicates a negative correlation between variables on the *y*- and *x*-axis. The area of the rectangles is proportional to the difference between observed and expected frequencies of each cell in Table S8. (d) Divergence between TE copies of the LTR/Gypsy found in populations of *C. obscurior* from Tenerife and Itabuna showing the *Gypsy-16_PBa-I* element, showing a higher abundance of nearly identical copies in the Tenerife population

To identify signatures of recent and/or past TE activity in the two lineages, we explored the mobilomes of Tenerife and Itabuna using DNAPIPETE (Figure 5a; Table S7). The total genomic coverage of repeats is similar in both lineages (~21%), with approximately 50% of the assembled repeats remaining unclassified (Table S7). The presence of unclassified repeats indicates that some repeat elements might be specific to *C. obscurior* (i.e., not represented in the combined TE libraries used in this analysis).

LTR/Gypsy elements cover 40% more bases in the Tenerife population compared to Itabuna, largely explained by an overrepresentation of nearly identical copies (i.e., recent TE insertions) (Figure 5a; Table S7). In contrast to LTR/Gypsy elements, patterns for other elements are similar in both populations (Figure 5a; Table S7).

Recent activity of TEs should increase the number of low-frequency TEs within a population. To test for this, we identified de novo as well as reference TE insertions in each population using POPOOLATIONTE2 (Kofler et al., 2016). We classified all detected TE insertions based on their frequency as low frequency (<25%), common (25%–95%) and fixed (>95%) TE insertions (Figure 5b; Table S8).

Overall, we identified more TE insertions in Itabuna (9380) compared to the Tenerife population (6988). However, the TE frequency distribution in Tenerife is biased towards low-frequency TEs (Figure 5b). While common TE insertions are more frequent than expected in Itabuna, the number of low-frequency TEs is higher than expected in the population from Tenerife (Figure 5c; Table S8; Pearson's Chi-squared test: $\chi^2(2) = 89.837$, $p < .0001$). This high prevalence of low-frequency TE insertions provides evidence for a recent activity of TEs in the Old World lineage population in Tenerife.

In both populations, low-frequency TE insertions are highly enriched in LDRs, while fixed insertions are predominantly found in TE islands (Table S8; Figure S14; Pearson's Chi-squared tests: $\chi^2(2)_{\text{Itabuna}} = 3{,}739.1$, $p < .0001$; $\chi^2(2)_{\text{Tenerife}} = 2620.5$, $p < 0.0001$). These results suggest that mechanisms controlling TEs (e.g., recombination and/or selection) act differently between LDRs and TE islands, consistent with previous results showing that unlike in TE islands, selection against TE insertions in LDRs is much stronger (Schrader et al., 2014).

3.7 | The LTR/Gypsy element *CobsR.176* is highly active

In-depth investigation of LTR/Gypsy elements revealed that the differences between lineages are driven to a large extent by a single element, *Gypsy-16_PBa-I*, that covers 747 kb (0.39%) of the genome in Tenerife compared to only 182.6 kb (0.1%) in Itabuna (Figure 5d). After including de novo predicted TEs for *C. obscurior*, we reannotated the *Gypsy-16_PBa-I* from RepBase as the full-length LTR/Gypsy *CobsR.176*. Phylogenetic analyses of the LTR/Gypsy *CobsR.176* element revealed the presence of 79 nearly identical copies across the genome of *C. obscurior*, indicating recent activity of this element (Figure S15).

Taken together, these results indicate a recent activity of TEs, particularly of LTR/Gypsy elements in the population from Tenerife compared to the population from Itabuna. Most notably, a single element, *CobsR.176*, appears to be more active in the Old World lineage compared to the New World lineage.

4 | DISCUSSION

As a consequence of genetic bottlenecks and inbreeding, introduced populations have reduced levels of genetic variation. In many cases, however, species can establish populations despite being genetically depleted. This process, known as the genetic paradox of invasions, has raised the questions of how founder populations evolve. Here, we show that in the ant *C. obscurior*, TE activity and intraspecific hybridization contribute to genetic diversity of introduced populations.

Combining short and long reads we produced a high-quality genome assembly for *C. obscurior* that is considerably more contiguous than the first published genome Cobs1.4 (Schrader et al., 2014) and most other sequenced ant genomes to date (http://antgenomics.dk/progress/; Bonasio et al., 2010; Boomsma et al., 2017; Gao et al., 2020; Hartke et al., 2019; Konorov et al., 2017; Nygaard et al., 2016; Oxley et al., 2014). The new assembly confirms the genome organization of *C. obscurior* with prominent TE-rich regions ("TE islands") in a TE-poor genomic background ("LDRs") (Schrader et al., 2014).

The whole genome sequencing data of pooled as well as single individuals deepen our understanding of genome organization, lineage divergence and population dynamics in *C. obscurior*. Principal component and ADMIXTURE analyses confirm two genetically distinct lineages (Oettler et al., 2010; Schrader et al., 2014). The population history inference indicates that both lineages diverged ~40,000 generations ago, and were almost completely separated 5000 generations ago. These estimates are scaled in generations assuming a mutation rate of 3.6×10^{-9} mutations per generation per site, as estimated for bumblebees (Liu et al., 2017). Assuming a generation time of 3 months (Oettler & Schrempf, 2016), the two lineages thus had diverged already by ~10,000 years ago and were almost completely separated ~1250 years ago. Using mutation rates inferred for other insects, that is 2.1×10^{-9} in *Chironomus riparius* (Oppold & Pfenninger, 2017), 2.8×10^{-9} in *Drosophila melanogaster* (Keightley et al., 2014) and 2.9×10^{-9} in *Heliconius Melpomene* (Keightley et al., 2015), our models consistently predicted divergence to begin at least 12,500 years ago. Hence, the split into two discrete lineages preceded anthropogenic factors (e.g., global trade) and it remains unclear what initiated the divergence of both lineages in the first place. One possible explanation is the emergence of a reproductive barrier following acquisition of a novel endosymbiont (Shropshire et al., 2020), as hybridization between Old and New World populations is drastically affected due to incompatible *Wolbachia* strains in *C. obscurior* (Ün et al., 2021).

We detected a secondary contact between the two lineages around 1000 years ago, before becoming fully separated again, around 200 years ago. Over this period, effective population size drastically decreased in the two lineages and started increasing only after the populations were completely separated. This recent increase in estimates of N_e might be correlated with increased trading towards the end of the 18th century (Dedinger & Girard, 2017). Thus, secondary contacts might have been driven by human commerce.

4.1 | Introduced populations of *C. obscurior* exhibit low levels of genetic variation

Overall, genetic variation is low in both studied introduced populations, particularly in the Tenerife population. Such reduction of genetic diversity in introduced ranges is a property that is shared among a number of invasive species (Comeault et al., 2020; Dlugosch & Parker, 2008; Puzey & Vallejo-Marín, 2014; Tsutsui et al., 2000). Given that F_{ST} estimates are sensitive to the level of within-population diversity (Chakraborty & Nei, 1977; Charlesworth, 1998; Hedrick, 1999; Nei, 1973), genetic differentiation between the Tenerife and Itabuna population is very pronounced. Such inflation of pairwise F_{ST} across populations can be a consequence of processes typical

for invasive species such as genetic bottlenecks and inbreeding that reduce a population's N_e and diversity. In addition, the populations in Tenerife and Itabuna belong to two lineages that diverged for most of the last 10,000 years, as shown by the population history analyses. The discovery of recent secondary contacts and introgression, however, indicates that hybridization between both lineages does occur. We so far only find unidirectional introgression from the Old World into the New World lineage, consistent with unidirectional cytoplasmic incompatibility (Ün et al., 2021). Such introgression events could lead to the displacement of the New World lineage by successive hybridization (Bobrov, 1982). Studying ecological interactions of Old World and New World colonies in natural populations, the fitness of F_1 Old World – New World hybrids, and the adaptive significance of the introgressed regions are necessary to understand these phenomena.

In both lineages, genetic diversity and Tajima's *D* are substantially higher in TE islands compared to the rest of the genome. In general, TE-rich regions are expected to have less diversity due to background selection (Charlesworth et al., 1993; Hudson & Kaplan, 1995) or hitchhiking (Kaplan et al., 1989) operating in low recombining regions. The pattern of increased diversity suggests that TE islands evolve more dynamically than the remainder of the genome. This would resemble patterns so far only described in some plant pathogenic fungi, where fast evolving parts of the genome are enriched in TEs in a similar manner (Möller & Stukenbrock, 2017). The fact that the TE islands of *C. obscurior* are enriched in ORs and FAS genes provides an interesting avenue to evaluate the significance of these high levels of diversity in *C. obscurior*.

Several nonexclusive mechanisms can underlie variation in diversity within genomes, such as local differences in mutation rate, the density of selection targets, gene flow or the rate of recombination (Ellegren & Galtier, 2016). Each of these mechanisms might indeed play a role in increasing genetic diversity in the gene-sparse but repeat-rich genomic islands in *C. obscurior*. The fact that differences between LDRs and TE islands are particularly strong in the Tenerife population further indicates that these mechanisms are shaped also by lineage- or population-specific factors. For recombination (Dapper & Payseur, 2017) and mutation rates (Baer et al., 2007; Krašovec et al., 2017; Sniegowski et al., 2000), environmental and population-level effects have been described in other species as well.

4.2 | Different mechanisms contribute to genetic variation in invasive populations of *C. obscurior*

A key question is how populations of *C. obscurior* can cope with the challenges imposed by limited genetic variation, resulting from genetic bottlenecks and inbreeding. Our genome-wide surveys in the Tenerife and Itabuna population revealed distinct mechanisms that contribute to genetic variation and differentiation: the activity of TEs and intraspecific introgression.

4.2.1 | Increased TE activity in the Old World lineage population

Our survey of TE dynamics revealed signatures of recent activity of TEs in both lineages. Notably, the recent and prominent proliferation of Gypsy and other LTR elements separates the genome of *C. obscurior* from all other ant genomes so far analysed (Petersen et al., 2019). Further, our study provides evidence for increased genetic variation in TE islands. TE proliferation is a mutagenic process (Bourque et al., 2018) that can induce structural variants with potentially large genomic and phenotypic effects (Kidwell & Lisch, 1997; Schrader & Schmitz, 2019). Most TE-induced mutations are expected to be neutral or deleterious, requiring the host to establish mechanisms to suppress or remove TEs from the genome. In ants and other social Hymenoptera, very high recombination rates (Wilfert et al., 2007) could function as an efficient defence mechanism against deleterious TE insertions, such as by removing TEs by nonhomologous recombination and by reducing the likelihood of genetic hitchhiking of TEs (Kent et al., 2017).

While most TE-induced mutations are deleterious or neutral, they can produce adaptive variants as shown in, for example, *Drosophila* (González et al., 2010; Lenkov et al., 2008; Mateo et al., 2014; Schmidt et al., 2010), but also plants (Li et al., 2018; Niu et al., 2019) and other organisms (Casacuberta & González, 2013; Santos et al., 2014; Schrader & Schmitz, 2019; Woronik et al., 2019). The peculiar genomic architecture with discrete TE islands and the enrichment of certain genes such as ORs in these regions suggest that TEs could indeed be particularly relevant for the evolution of this species. Intriguingly, genetic variation particularly in OR genes is associated with increased plastic olfactory perception, potentially facilitating the exploration of new habitats and exploitation of novel resources (Ache & Young, 2005; Anton & Rössler, 2021).

We find higher TE activity in the Tenerife population. Whether this increased activity of TEs is a lineage-specific trait or possibly unique for the population is unclear. Intriguingly, studies in *Drosophila suzukii* found similar patterns with a recent surge in TE activity coinciding with the range expansion of this species (Mérel et al., 2021). Further, in plant pathogenic fungi with a genomic compartmentalization similar to that *C. obscurior*, the role of TEs in the rapid emergence of adaptive changes is well documented (Möller & Stukenbrock, 2017).

4.2.2 | Introgression from the Old World lineage into the New World populations

The existence of genomic regions harbouring higher than average nucleotide diversity, showing a positive Tajima's *D* and a lower than average genetic differentiation, only in the Itabuna population indicates a genomic footprint of introgression between the lineages (Alcala et al., 2013; Burgarella et al., 2019). This is also consistent with the observation that the two lineages co-occur in Guarujá. Hence, intraspecific hybridization between the two lineages could

occur in this location. However, detailed statistical analyses using *D* and *f* statistics (Malinsky et al., 2021; Patterson et al., 2012) using individual-level data and screens of populations worldwide are needed to resolve the history of introgression events and the routes of introduction of this ant.

Genetic admixture through multiple introductions can provide gains of genetic diversity within introduced populations, thus increasing their fitness as well as their adaptive potential (Barker et al., 2017; Dlugosch & Parker, 2008; Facon et al., 2008; Keller & Taylor, 2010; Kolbe et al., 2004, 2007). Additionally, admixture can create new genotypes/phenotypes, increasing the adaptive potential of the colonizing populations (Nolte et al., 2009; Rius & Darling, 2014; van Boheemen et al., 2017; Verhoeven et al., 2011). The outcome of such hybridization not only would contribute to the successful establishment in the novel habitat, but also would allow these introduced populations to persist and even expand their range. Whether introgressed haplotypes from Old World populations offer adaptive advantages to the New World populations is unclear. However, introgression between closely related species can be adaptive (Hedrick, 2013; Martin & Jiggins, 2017) and selection can favour admixed individuals due to increased fitness (Nolte et al., 2009). This is particularly relevant for genetically depauperate incipient populations of invasive species. Several studies have provided insights on the role of introgression in promoting range expansion and adaptation to disturbed habitats. Examples include the mosquito *Anopheles gambiae* and the moth *Helicoverpa armigera*, where introgression promoted the expansion beyond their native ranges (Ayala et al., 2019; Sharakhov et al., 2006; Valencia-Montoya et al., 2020). In the latter example, it has also been shown that a region responsible for insecticide resistance was introgressed from the invasive into a local moth species, highlighting the potential role of introgression in adding to the native gene repertoire (Valencia-Montoya et al., 2020).

Pool-seq is a powerful technique for estimating allele frequencies (Schlötterer et al., 2014). However, its accuracy depends on several factors, including pool size, sequencing coverage and the true contribution of each individual to the final pool (Schlötterer et al., 2014; Zhu et al., 2012). Although we sequenced our pools to a sufficiently high coverage (>73×), unequal pool sizes and/or the possibility of unequal contributions of each individual to the pools could limit the resolution of this initial study on population-level TE dynamics in this invasive species. Further studies using single individual resequencing data will be necessary to gain a deeper understanding of these dynamics in *C. obscurior*.

5 | CONCLUSION

Our study provides evidence that TE activity and intraspecific hybridization between differentiated populations are important sources of genetic variation in genetically depleted invasive populations of *C. obscurior*. The uncommon genomic architecture paired with the recent TE activity could add to the invasive potential of this species. Unravelling whether stress-responsive mechanisms play a role in the activation of TEs or whether TEs are continuously active in the genome will help us to understand how TEs shape the genetic variation of invasive populations. While the role of TEs in invasion biology is acknowledged by a growing body of research (Lee & Wang, 2018; Stapley et al., 2015), we are far from differentiating general principles from special cases. There is also mounting evidence for the adaptive potential of introgression in a number of plant and animal species (Burgarella et al., 2019; Hedrick, 2013), encouraging further studies to unravel the impact of introgression and TEs in invasive species.

ACKNOWLEDGEMENTS

We would like to thank Eva Schultner for help collecting ants in the field. We thank Kathrin Brüggemann and Una Hadziomerovic for their help with Oxford Nanopore sequencing. We thank Xim Cerda for his support regarding collection of ants in Spain. We also thank three anonymous reviewers who provided constructive comments for improving the manuscript. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 403813881 with grants to L.S. (SCHR 1554/2-1) and J.O. (OE 549/4-1) under the priority programme "Rapid evolutionary adaptation: Potential and constraints" (SPP 1819). We further thank the DFG Research Training Group 2220 "Evolutionary Processes in Adaptation and Disease" for support to M.E. Research on ants collected in Spain was conducted under permission of Real Decreto 124/2017. We are grateful to the Brazilian Ministério do Meio Ambiente for permission to study *C. obscurior* in Bahia (permit 63371-1). Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

AUTHOR CONTRIBUTIONS

L.S. and J.O. conceived of the study. L.S., J.O. and M.E. designed the experiments and M.E., L.S. and J.O. wrote the paper; L.S. sequenced and assembled the genome; R.S., L.S. and M.E. were responsible for repeat annotations and J.K. and K.J.H. were responsible for gene annotations; J.A. was responsible for Illumina sequencing; population genomic analyses and analyses of TE dynamics were performed by M.E. All authors read and commented on the final version of the manuscript.

DATA AVAILABILITY STATEMENT

The genome assembly and the raw sequencing data prior to trimming and mapping are available at NCBI (BioProject: PRJNA680013). SNPs in plink format and other final data sets including DNAPIPETE and POPOOLATIONTE2 outputs are publicly available at https://doi.org/10.6084/m9.figshare.13318754 (Errbii, 2021). A detailed description of our bioinformatic analysis pipelines can be found at https://github.com/schraderL/CobscuriorGenome and https://github.com/merrbii/CobsPopGenomics.

ORCID

Mohammed Errbii https://orcid.org/0000-0001-7719-0998
Jens Keilwagen https://orcid.org/0000-0002-6792-7076
Katharina J. Hoff https://orcid.org/0000-0002-7333-8390
Raphael Steffen https://orcid.org/0000-0003-4947-3899
Jan Oettler https://orcid.org/0000-0002-8539-6029
Lukas Schrader https://orcid.org/0000-0001-9863-0526

REFERENCES

Abrusán, G., Grundmann, N., Demester, L., & Makalowski, W. (2009). TEclass – A tool for automated classification of unknown eukaryotic transposable elements. *Bioinformatics*, *25*(10), 1329–1330. https://doi.org/10.1093/bioinformatics/btp084

Ache, B. W., & Young, J. M. (2005). Olfaction: Diverse species, conserved principles. *Neuron*, *48*(3), 417–430. https://doi.org/10.1016/j.neuron.2005.10.022

Alcala, N., Streit, D., Goudet, J., & Vuilleumier, S. (2013). Peak and persistent excess of genetic diversity following an abrupt migration increase. *Genetics*, *193*(3), 953–971. https://doi.org/10.1534/genetics.112.147785

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Anton, S., & Rössler, W. (2021). Plasticity and modulation of olfactory circuits in insects. *Cell and Tissue Research*, *383*(1), 149–164. https://doi.org/10.1007/s00441-020-03329-z

Ayala, D., Zhang, S., Chateau, M., Fouet, C., Morlais, I., Costantini, C., Hahn, M. W., & Besansky, N. J. (2019). Association mapping desiccation resistance within chromosomal inversions in the African malaria vector *Anopheles gambiae*. *Molecular Ecology*, *28*(6), 1333–1342. https://doi.org/10.1111/mec.14880

Baer, C. F., Miyamoto, M. M., & Denver, D. R. (2007). Mutation rate variation in multicellular eukaryotes: Causes and consequences. *Nature Reviews Genetics*, *8*(8), 619–631. https://doi.org/10.1038/nrg2158

Bailly-Bechet, M., Haudry, A., & Lerat, E. (2014). "One code to find them all": A perl tool to conveniently parse RepeatMasker output files. *Mobile DNA*, *5*(1), 13. https://doi.org/10.1186/1759-8753-5-13

Barker, B. S., Andonian, K., Swope, S. M., Luster, D. G., & Dlugosch, K. M. (2017). Population genomic analyses reveal a history of range expansion and trait evolution across the native and invaded range of yellow starthistle (*Centaurea solstitialis*). *Molecular Ecology*, *26*(4), 1131–1147. https://doi.org/10.1111/mec.13998

Barrett, R. D. H. H., & Schluter, D. (2008). Adaptation from standing genetic variation. *Trends in Ecology and Evolution*, *23*(1), 38–44. https://doi.org/10.1016/j.tree.2007.09.008

Biémont, C. (2010). A brief history of the status of transposable elements: From junk DNA to major players in evolution. *Genetics*, *186*(4), 1085–1093. https://doi.org/10.1534/genetics.110.124180

Bobrov, E. G. (1982). On introgressive hybridization and its significance in the evolution of plants. *Folia Geobotanica et Phytotaxonomica*, *17*(1), 89–96. https://doi.org/10.1007/BF02852434

Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D., & Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*, *27*(4), 578–579. https://doi.org/10.1093/bioinformatics/btq683

Boetzer, M., & Pirovano, W. (2014). SSPACE-LongRead: Scaffolding bacterial draft genomes using long read sequence information. *BMC Bioinformatics*, *15*(1), 211. https://doi.org/10.1186/1471-2105-15-211

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bonasio, R., Zhang, G., Ye, C., Mutti, N. S., Fang, X., Qin, N., Donahue, G., Yang, P., Li, Q., Li, C., Zhang, P., Huang, Z., Berger, S. L., Reinberg, D., Wang, J., & Liebig, J. (2010). Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Science*, *329*(5995), 1068–1071. https://doi.org/10.1126/science.1192428

Boomsma, J. J., Brady, S. G., Dunn, R. R., Gadau, J., Heinze, J., Keller, L., Moreau, C. S., Sanders, N. J., Schrader, L., Schultz, T. R., Sundström, L., Ward, P. S., Wcislo, W. T., & Zhang, G., … The GAGA Consortium. (2017). The Global Ant Genomics Alliance (GAGA). *Myrmecological News*, *25*, 61–66.

Bourque, G., Burns, K. H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izsvák, Z., Levin, H. L., Macfarlan, T. S., Mager, D. L., & Feschotte, C. (2018). Ten things you should know about transposable elements. *Genome Biology*, *19*(1), 199. https://doi.org/10.1186/s13059-018-1577-z

Burgarella, C., Barnaud, A., Kane, N. A., Jankowski, F., Scarcelli, N., Billot, C., Vigouroux, Y., & Berthouly-Salazar, C. (2019). Adaptive introgression: An untapped evolutionary mechanism for crop adaptation. *Frontiers in Plant Science*, *10*, 4. https://doi.org/10.3389/fpls.2019.00004

Casacuberta, E., & González, J. (2013). The impact of transposable elements in environmental adaptation. *Molecular Ecology*, *22*, 1503–1517. https://doi.org/10.1111/mec.12170

Chakraborty, R., & Nei, M. (1977). Bottleneck effects on average heterozygosity and genetic distance with the stepwise mutation model. *Evolution*, *31*(2), 347. https://doi.org/10.2307/2407757

Chang, C., Chow, C., Tellier, L. C., Vattikuti, S., Purcell, S., & Lee, J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), 7. https://doi.org/10.1186/s13742-015-0047-8

Charlesworth, B. (1998). Measures of divergence between populations and the effect of forces that reduce variability. *Molecular Biology and Evolution*, *15*(5), 538–543. https://doi.org/10.1093/oxfordjournals.molbev.a025953

Charlesworth, B., Morgan, M. T., & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics*, *134*(4), 1289–1303. https://doi.org/10.1093/genetics/134.4.1289

Comeault, A. A., Wang, J., Tittes, S., Isbell, K., Ingley, S., Hurlbert, A. H., & Matute, D. R. (2020). Genetic diversity and thermal performance in invasive and native populations of African fig flies. *Molecular Biology and Evolution*, *37*, 1893–1906. https://doi.org/10.1101/800938

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. https://doi.org/10.1093/bioinformatics/btr330

Dapper, A. L., & Payseur, B. A. (2017). Connecting theory and data to understand recombination rate evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160469. https://doi.org/10.1098/rstb.2016.0469

Dedinger, B., & Girard, P. (2017). Exploring trade globalization in the long run: The RICardo project. *Historical Methods*, *50*(1), 30–48. https://doi.org/10.1080/01615440.2016.1220269

Delaneau, O., Marchini, J., & Zagury, J. F. (2012). A linear complexity phasing method for thousands of genomes. *Nature Methods*, *9*(2), 179–181. https://doi.org/10.1038/nmeth.1785

Dennis, A. B., Ballesteros, G. I., Robin, S., Schrader, L., Bast, J., Berghöfer, J., Beukeboom, L. W., Belghazi, M., Bretaudeau, A., Buellesbach, J., Cash, E., Colinet, D., Dumas, Z., Errbii, M., Falabella, P., Gatti, J.-L., Geuverink, E., Gibson, J. D., Hertaeg, C., … Gadau, J. (2020). Functional insights from the GC-poor genomes of two aphid parasitoids, *Aphidius ervi* and *Lysiphlebus fabarum*. *BMC Genomics*, *21*(1), 376. https://doi.org/10.1186/s12864-020-6764-0

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernytsky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., & Daly, M. J. (2011). A framework for variation discovery and genotyping using

next-generation DNA sequencing data. *Nature Genetics*, *43*(5), 491–501. https://doi.org/10.1038/ng.806

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, *17*(1), 431–449. https://doi.org/10.1111/j.1365-294X.2007.03538.x

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*(7), 422–433. https://doi.org/10.1038/nrg.2016.58

Errbii, M. (2021). Transposable elements and introgression introduce genetic variation in the invasive ant *Cardiocondyla obscurior*. https://doi.org/10.6084/m9.figshare.13318754

Facon, B., Pointier, J. P., Jarne, P., Sarda, V., & David, P. (2008). High genetic variance in life-history strategies within invasive populations by way of multiple introductions. *Current Biology*, *18*(5), 363–367. https://doi.org/10.1016/j.cub.2008.01.063

Faino, L., Seidl, M. F., Shi-Kunne, X., Pauper, M., Van Den Berg, G. C. M., Wittenberg, A. H. J., & Thomma, B. P. H. J. (2016). Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. *Genome Research*, *26*(8), 1091–1100. https://doi.org/10.1101/gr.204974.116

Feschotte, C., & Pritham, E. (2007). DNA transposons and the evolution of eukaryotic genomes. *Annual Review of Genetics*, *41*, 331–368. https://doi.org/10.1146/annurev.genet.40.110405.090448

Frankham, R. (2005a). Genetics and extinction. *Biological Conservation*, *126*(2), 131–140. https://doi.org/10.1016/j.biocon.2005.05.002

Frankham, R. (2005b). Resolving the genetic paradox in invasive species. *Heredity*, *94*(4), 385. https://doi.org/10.1038/sj.hdy.6800634

Frankham, R., Ballou, J. D., Briscoe, D. A., & McInnes, K. H. (2002). *Introduction to Conservation Genetics*. Cambridge University Press. https://doi.org/10.1017/CBO9780511808999

Frankham, R., & Ralls, K. (1998). Inbreeding leads to extinction. *Nature*, *392*, 441–442. https://doi.org/10.1038/33022

Gao, Q., Xiong, Z., Larsen, R. S., Zhou, L., Zhao, J., Ding, G., Zhao, R., Liu, C., Ran, H., & Zhang, G. (2020). High-quality chromosome-level genome assembly and full-length transcriptome analysis of the pharaoh ant *Monomorium pharaonis*. *GigaScience*, *9*, 1–14. https://doi.org/10.1093/gigascience/giaa143

González, J., Karasov, T. L., Messer, P. W., & Petrov, D. A. (2010). Genome-wide patterns of adaptation to temperate environments associated with transposable elements in Drosophila. *PLoS Genetics*, *6*(4), e1000905. https://doi.org/10.1371/journal.pgen.1000905

Goubert, C., Modolo, L., Vieira, C., Moro, C. V., Mavingui, P., & Boulesteix, M. (2015). De novo assembly and annotation of the Asian tiger mosquito (*Aedes albopictus*) repeatome with dnaPipeTE from raw genomic reads and comparative analysis with the yellow fever mosquito (*Aedes aegypti*). *Genome Biology and Evolution*, *7*(4), 1192–1205. https://doi.org/10.1093/gbe/evv050

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*, *29*(8), 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Hartke, J., Schell, T., Jongepier, E., Schmidt, H., Sprenger, P. P., Paule, J., Bornberg-Bauer, E., Schmitt, T., Menzel, F., Pfenninger, M., & Feldmeyer, B. (2019). Hybrid genome assembly of a neotropical mutualistic ant. *Genome Biology and Evolution*, *11*(8), 2306–2311. https://doi.org/10.1093/gbe/evz159

Hedrick, P. W. (1999). Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution*, *53*(2), 313. https://doi.org/10.2307/2640768

Hedrick, P. W. (2013). Adaptive introgression in animals: Examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology*, *22*(18), 4606–4618. https://doi.org/10.1111/mec.12415

Heinze, J. (2017). Life-history evolution in ants: The case of Cardiocondyla. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1850), 20161406. https://doi.org/10.1098/rspb.2016.1406

Heinze, J., Cremer, S., Eckl, N., & Schrempf, A. (2006). Stealthy invaders: The biology of Cardiocondyla tramp ants. *Insectes Sociaux*, *53*, 1–7. https://doi.org/10.1007/s00040-005-0847-4

Hoede, C., Arnoux, S., Moisset, M., Chaumier, T., Inizan, O., Jamilloux, V., & Quesneville, H. (2014). PASTEC: An automatic transposable element classification tool. *PLoS One*, *9*(5), e91929. https://doi.org/10.1371/journal.pone.0091929

Hoff, K. J., Lange, S., Lomsadze, A., Borodovsky, M., & Stanke, M. (2016). BRAKER1: Unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*, *32*(5), 767–769. https://doi.org/10.1093/bioinformatics/btv661

Horváth, V., Merenciano, M., & González, J. (2017). Revisiting the relationship between transposable elements and the eukaryotic stress response. *Trends in Genetics*, *33*, 832–841. https://doi.org/10.1016/j.tig.2017.08.007

Hubley, R., Finn, R. D., Clements, J., Eddy, S. R., Jones, T. A., Bao, W., Smit, A. F. A., & Wheeler, T. J. (2016). The Dfam database of repetitive DNA families. *Nucleic Acids Research*, *44*(D1), D81–D89. https://doi.org/10.1093/nar/gkv1272

Hudson, R. R., & Kaplan, N. L. (1995). Deleterious background selection with recombination. *Genetics*, *141*(4), 1605–1617. https://doi.org/10.1093/genetics/141.4.1605

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., & Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, *30*(9), 1236–1240. https://doi.org/10.1093/bioinformatics/btu031

Jurka, J., Kapitonov, V. V., Pavlicek, A., Klonowski, P., Kohany, O., & Walichiewicz, J. (2005). Repbase update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research*, *110*(1–4), 462–467. https://doi.org/10.1159/000084979

Kaplan, N. L., Hudson, R. R., & Langley, C. H. (1989). The "hitchhiking effect" revisited. *Genetics*, *123*(4), 887–899. https://doi.org/10.1093/genetics/123.4.887

Keightley, P. D., Ness, R. W., Halligan, D. L., & Haddrill, P. R. (2014). Estimation of the spontaneous mutation rate per nucleotide site in a *Drosophila melanogaster* full-sib family. *Genetics*, *196*(1), 313–320. https://doi.org/10.1534/genetics.113.158758

Keightley, P. D., Pinharanda, A., Ness, R. W., Simpson, F., Dasmahapatra, K. K., Mallet, J., Davey, J. W., & Jiggins, C. D. (2015). Estimation of the spontaneous mutation rate in *Heliconius melpomene*. *Molecular Biology and Evolution*, *32*(1), 239–243. https://doi.org/10.1093/molbev/msu302

Keilwagen, J., Wenk, M., Erickson, J. L., Schattat, M. H., Grau, J., & Hartung, F. (2016). Using intron position conservation for homology-based gene prediction. *Nucleic Acids Research*, *44*(9), 89. https://doi.org/10.1093/nar/gkw092

Keller, S. R., & Taylor, D. R. (2010). Genomic admixture increases fitness during a biological invasion. *Journal of Evolutionary Biology*, *23*(8), 1720–1731. https://doi.org/10.1111/j.1420-9101.2010.02037.x

Kent, T. V., Uzunović, J., & Wright, S. I. (2017). Coevolution between transposable elements and recombination. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*, 20160458. https://doi.org/10.1098/rstb.2016.0458

Kidwell, M. G., & Lisch, D. (1997). Transposable elements as sources of variation in animals and plants. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(15), 7704–7711. https://doi.org/10.1073/pnas.94.15.7704

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, *12*(4), 357–360. https://doi.org/10.1038/nmeth.3317

Klein, A., Schrader, L., Gil, R., Manzano-Marín, A., Flórez, L., Wheeler, D., Werren, J. H., Latorre, A., Heinze, J., Kaltenpoth, M., Moya, A., & Oettler, J. (2016). A novel intracellular mutualistic bacterium in the

invasive ant *Cardiocondyla obscurior*. *ISME Journal*, *10*(2), 376–388. https://doi.org/10.1038/ismej.2015.119

Kofler, R., Gómez-Sánchez, D., & Schlötterer, C. (2016). PoPoolationTE2: Comparative population genomics of transposable elements using pool-seq. *Molecular Biology and Evolution*, *33*(10), 2759–2764. https://doi.org/10.1093/molbev/msw137

Kofler, R., Orozco-terWengel, P., De Maio, N., Pandey, R. V., Nolte, V., Futschik, A., Kosiol, C., & Schlötterer, C. (2011). Popoolation: A toolbox for population genetic analysis of next generation sequencing data from pooled individuals. *PLoS One*, *6*(1), e15925. https://doi.org/10.1371/journal.pone.0015925

Kofler, R., Pandey, R. V., & Schlötterer, C. (2011). PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*, *27*(24), 3435–3436. https://doi.org/10.1093/bioinformatics/btr589

Kolbe, J. J., Glor, R. E., Schettino, L. R., Lara, A. C., Larson, A., & Losos, J. B. (2004). Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, *431*(7005), 177–181. https://doi.org/10.1038/nature02807

Kolbe, J. J., Larson, A., & Losos, J. B. (2007). Differential admixture shapes morphological variation among invasive populations of the lizard *Anolis sagrei*. *Molecular Ecology*, *16*(8), 1579–1591. https://doi.org/10.1111/j.1365-294X.2006.03135.x

Konorov, E. A., Nikitin, M. A., Mikhailov, K. V., Lysenkov, S. N., Belenky, M., Chang, P. L., Nuzhdin, S. V., & Scobeyeva, V. A. (2017). Genomic exaptation enables Lasius Niger adaptation to urban environments. *BMC Evolutionary Biology*, *17*(S1), 1–12. https://doi.org/10.1186/s12862-016-0867-x

Kopylova, E., Noé, L., & Touzet, H. (2012). SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*, *28*(24), 3211–3217. https://doi.org/10.1093/bioinformatics/bts611

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly via adaptive κ-mer weighting and repeat separation. *Genome Research*, *27*(5), 722–736. https://doi.org/10.1101/gr.215087.116

Krašovec, R., Richards, H., Gifford, D. R., Hatcher, C., Faulkner, K. J., Belavkin, R. V., Channon, A., Aston, E., McBain, A. J., & Knight, C. G. (2017). Spontaneous mutation rate is a plastic trait associated with population density across domains of life. *PLOS Biology*, *15*(8), e2002731. https://doi.org/10.1371/journal.pbio.2002731

Lacy, R. C. (1997). Importance of genetic variation to the viability of mammalian populations. *Journal of Mammalogy*, *78*, 320–335. https://doi.org/10.2307/1382885

Lai, Y.-T., Yeung, C. K. L., Omland, K. E., Pang, E.-L., Hao, Y. U., Liao, B.-Y., Cao, H.-F., Zhang, B.-W., Yeh, C.-F., Hung, C.-M., Hung, H.-Y., Yang, M.-Y., Liang, W., Hsu, Y.-C., Yao, C.-T., Dong, L. U., Lin, K., & Li, S.-H. (2019). Standing genetic variation as the predominant source for adaptation of a songbird. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(6), 2152–2157. https://doi.org/10.1073/pnas.1813597116

Lande, R., & Shannon, S. (1996). The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, *50*(1), 434–437. https://doi.org/10.1111/j.1558-5646.1996.tb04504.x

Lanfear, R., Schalamun, M., Kainer, D., Wang, W., & Schwessinger, B. (2019). MinIONQC: Fast and simple quality control for MinION sequencing data. *Bioinformatics*, *35*(3), 523–525. https://doi.org/10.1093/bioinformatics/bty654

Lee, C. C., & Wang, J. (2018). Rapid expansion of a highly germline-expressed mariner element acquired by horizontal transfer in the fire ant genome. *Genome Biology and Evolution*, *10*(12), 3262–3278. https://doi.org/10.1093/gbe/evy220

Lenkov, K., Petrov, D. A., González, J., Lipatov, M., & Macpherson, J. M. (2008). High rate of recent transposable element-induced adaptation in *Drosophila melanogaster*. *PLoS Biology*, *6*(10), e251. https://doi.org/10.1371/journal.pbio.0060251

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

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

Li, Z. W., Hou, X. H., Chen, J. F., Xu, Y. C., Wu, Q., Gonzalez, J., & Guo, Y. L. (2018). Transposable elements contribute to the adaptation of *Arabidopsis thaliana*. *Genome Biology and Evolution*, *10*(8), 2140–2150. https://doi.org/10.1093/gbe/evy171

Lisch, D. (2013). How important are transposons for plant evolution?. *Nature Reviews Genetics*, *14*(1), 49–61. https://doi.org/10.1038/nrg3374

Liu, H., Jia, Y., Sun, X., Tian, D., Hurst, L. D., & Yang, S. (2017). Direct determination of the mutation rate in the bumblebee reveals evidence for weak recombination-associated mutation and an approximate rate constancy in insects. *Molecular Biology and Evolution*, *34*(1), 119–130. https://doi.org/10.1093/molbev/msw226

Liu, L., Wu, Y., Chen, F., Wang, Q. X., Zhang, X. Y., Tang, Y., Li, F., & Qian, Z. Q. (2019). Characterization of the complete mitochondrial genome of the invasive tramp ant *Cardiocondyla obscurior* (Hymenoptera: Formicidae: Myrmicinae). *Mitochondrial DNA Part B: Resources*, *4*(1), 1496–1498. https://doi.org/10.1080/23802359.2019.1601522

Lomsadze, A., Burns, P. D., & Borodovsky, M. (2014). Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Research*, *42*(15), e119. https://doi.org/10.1093/nar/gku557

Malinsky, M., Matschiner, M., & Svardal, H. (2021). Dsuite – Fast D-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*, *21*(2), 584–595. https://doi.org/10.1111/1755-0998.13265

Marin, P., Genitoni, J., Barloy, D., Maury, S., Gibert, P., Ghalambor, C. K., & Vieira, C. (2020). Biological invasion: The influence of the hidden side of the (epi)genome. *Functional Ecology*, *34*(2), 385–400. https://doi.org/10.1111/1365-2435.13317

Martin, S. H., & Jiggins, C. D. (2017). Interpreting the genomic landscape of introgression. *Current Opinion in Genetics and Development*, *47*, 69–74. https://doi.org/10.1016/j.gde.2017.08.007

Mateo, L., Ullastres, A., & González, J. (2014). A transposable element insertion confers xenobiotic resistance in *Drosophila*. *PLoS Genetics*, *10*(8), e1004560. https://doi.org/10.1371/journal.pgen.1004560

Mérel, V., Gibert, P., Buch, I., Rada, V. R., Estoup, A., Gautier, M., Fablet, M., Boulesteix, M., & Vieira, C. (2021). The worldwide invasion of *Drosophila suzukii* is accompanied by a large increase of transposable element load and a small number of putatively adaptive insertions. *Molecular biology and evolution*, https://doi.org/10.1093/molbev/msab155

Möller, M., & Stukenbrock, E. H. (2017). Evolution and genome architecture in fungal plant pathogens. *Nature Reviews Microbiology*, *15*, 756–771. https://doi.org/10.1038/nrmicro.2017.76

Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R., & Schäffer, A. A. (2008). Database indexing for production MegaBLAST searches. *Bioinformatics*, *24*, 1757–1764. https://doi.org/10.1093/bioinformatics/btn322

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, *70*(12), 3321–3323. https://doi.org/10.1073/pnas.70.12.3321

Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(10), 5269–5273. https://doi.org/10.1073/pnas.76.10.5269

Neph, S., Kuehn, M. S., Reynolds, A. P., Haugen, E., Thurman, R. E., Johnson, A. K., Rynes, E., Maurano, M. T., Vierstra, J., Thomas, S., Sandstrom, R., Humbert, R., & Stamatoyannopoulos, J. A. (2012). BEDOPS: High-performance genomic feature operations. *Bioinformatics*, *28*(14), 1919–1920. https://doi.org/10.1093/bioinformatics/bts277

Niu, X. M., Xu, Y. C., Li, Z. W., Bian, Y. T., Hou, X. H., Chen, J. F., Zou, Y. P., Jiang, J., Wu, Q., Ge, S., Balasubramanian, S., & Guo, Y. L. (2019). Transposable elements drive rapid phenotypic variation in *Capsella rubella*. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(14), 6908–6913. https://doi.org/10.1073/pnas.1811498116

Nolte, A. W., Gompert, Z., & Buerkle, C. A. (2009). Variable patterns of introgression in two sculpin hybrid zones suggest that genomic isolation differs among populations. *Molecular Ecology*, *18*(12), 2615–2627. https://doi.org/10.1111/j.1365-294X.2009.04208.x

Nygaard, S., Hu, H., Li, C., Schiøtt, M., Chen, Z., Yang, Z., Xie, Q., Ma, C., Deng, Y., Dikow, R. B., Rabeling, C., Nash, D. R., Wcislo, W. T., Brady, S. G., Schultz, T. R., Zhang, G., & Boomsma, J. J. (2016). Reciprocal genomic evolution in the ant-fungus agricultural symbiosis. *Nature Communications*, *7*(1), 1–9. https://doi.org/10.1038/ncomms12233

Oettler, J. (2020). Cardiocondyla: Heart node ants. In C. K. Starr (Ed.), *Encyclopedia of social insects* (pp. 1–3). Springer International Publishing. https://doi.org/10.1007/978-3-319-90306-4_19-1

Oettler, J., & Schrempf, A. (2016). Fitness and aging in *Cardiocondyla obscurior* ant queens. *Current Opinion in Insect Science*, *16*, 58–63. https://doi.org/10.1016/j.cois.2016.05.010

Oettler, J., Suefuji, M., & Heinze, J. (2010). The evolution of alternative reproductive tactics in male cardiocondyla ants. *Evolution*, *64*(11), 3310–3317. https://doi.org/10.1111/j.1558-5646.2010.01090.x

Okonechnikov, K., Conesa, A., & García-Alcalde, F. (2016). Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*, *32*(2), 292–294. https://doi.org/10.1093/bioinformatics/btv566

Oppold, A. M., & Pfenninger, M. (2017). Direct estimation of the spontaneous mutation rate by short-term mutation accumulation lines in *Chironomus riparius*. *Evolution Letters*, *1*(2), 86–92. https://doi.org/10.1002/evl3.8

Oxley, P. R., Ji, L. U., Fetter-Pruneda, I., McKenzie, S. K., Li, C., Hu, H., Zhang, G., & Kronauer, D. J. C. (2014). The genome of the clonal raider ant *Cerapachys biroi*. *Current Biology*, *24*(4), 451–458. https://doi.org/10.1016/j.cub.2014.01.018

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., & Reich, D. (2012). Ancient admixture in human history. *Genetics*, *192*(3), 1065–1093. https://doi.org/10.1534/GENETICS.112.145037

Petersen, M., Armisén, D., Gibbs, R. A., Hering, L., Khila, A., Mayer, G., Richards, S., Niehuis, O., & Misof, B. (2019). Diversity and evolution of the transposable element repertoire in arthropods with particular reference to insects. *BMC Evolutionary Biology*, *19*(1), 1–15. https://doi.org/10.1186/s12862-018-1324-9

Price, A. L., Jones, N. C., & Pevzner, P. A. (2005). De novo identification of repeat families in large genomes. *Bioinformatics*, *21*(Suppl 1), 351–358. https://doi.org/10.1093/bioinformatics/bti1018

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Puzey, J., & Vallejo-Marín, M. (2014). Genomics of invasion: Diversity and selection in introduced populations of monkeyflowers (*Mimulus guttatus*). *Molecular Ecology*, *23*(18), 4472–4485. https://doi.org/10.1111/mec.12875

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

R Core Team. (2020). *A language and environment for statistical computing*. R Foundation for Statistical Computing (4.0.2, Vol. 2). http://www.r-project.org/

Rius, M., & Darling, J. A. (2014). How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology and Evolution*, *29*, 233–242. https://doi.org/10.1016/j.tree.2014.02.003

Santos, M. E., Braasch, I., Boileau, N., Meyer, B. S., Sauteur, L., Böhne, A., Belting, H.-G., Affolter, M., & Salzburger, W. (2014). The evolution of cichlid fish egg-spots is linked with a cis-regulatory change. *Nature Communications*, *5*, 5149. https://doi.org/10.1038/ncomms6149

Schiffels, S., & Durbin, R. (2014). Inferring human population size and separation history from multiple genome sequences. *Nature Genetics*, *46*(8), 919–925. https://doi.org/10.1038/ng.3015

Schiffels, S., & Wang, K. (2020). MSMC and MSMC2: The multiple sequentially Markovian coalescent. *Methods in Molecular Biology*, *2090*, 147–166. https://doi.org/10.1007/978-1-0716-0199-0_7

Schlötterer, C., Tobler, R., Kofler, R., & Nolte, V. (2014). Sequencing pools of individuals-mining genome-wide polymorphism data without big funding. *Nature Reviews Genetics*, *15*(11), 749–763. https://doi.org/10.1038/nrg3803

Schmidt, J. M., Good, R. T., Appleton, B., Sherrard, J., Raymant, G. C., Bogwitz, M. R., Martin, J., Daborn, P. J., Goddard, M. E., Batterham, P., & Robin, C. (2010). Copy number variation and transposable elements feature in recent, ongoing adaptation at the Cyp6g1 locus. *PLoS Genetics*, *6*(6), e1000998. https://doi.org/10.1371/journal.pgen.1000998

Schrader, L., Kim, J. W., Ence, D., Zimin, A., Klein, A., Wyschetzki, K., Weichselgartner, T., Kemena, C., Stökl, J., Schultner, E., Wurm, Y., Smith, C. D., Yandell, M., Heinze, J., Gadau, J., & Oettler, J. (2014). Transposable element islands facilitate adaptation to novel environments in an invasive species. *Nature Communications*, *5*(1), 5495. https://doi.org/10.1038/ncomms6495

Schrader, L., & Schmitz, J. (2019). The impact of transposable elements in adaptive evolution. *Molecular Ecology*, *28*(6), 1537–1549. https://doi.org/10.1111/mec.14794

Sharakhov, I. V., White, B. J., Sharakhova, M. V., Kayondo, J., Lobo, N. F., Santolamazza, F., della Torre, A., Simard, F., Collins, F. H., & Besansky, N. J. (2006). Breakpoint structure reveals the unique origin of an interspecific chromosomal inversion (2La) in the *Anopheles gambiae* complex. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(16), 6258–6262. https://doi.org/10.1073/pnas.0509683103

Shropshire, J. D., Leigh, B., & Bordenstein, S. R. (2020). Symbiont-mediated cytoplasmic incompatibility: What have we learned in 50 years? *Elife*, *9*, 1–36. https://doi.org/10.7554/ELIFE.61989

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Sniegowski, P. D., Gerrish, P. J., Johnson, T., & Shaver, A. (2000). The evolution of mutation rates: Separating causes from consequences. *BioEssays*, *22*(12), 1057–1066. https://doi.org/10.1002/1521-1878(200012)22:12<1057:AID-BIES3>3.0.CO;2-W.

Sprenger, P. P., & Menzel, F. (2020). Cuticular hydrocarbons in ants (Hymenoptera: Formicidae) and other insects: How and why they differ among individuals, colonies, and species. *Myrmecological News*, *30*, 1–26. https://doi.org/10.25849/myrmecol.news_030:001

Stanke, M., Schöffmann, O., Morgenstern, B., & Waack, S. (2006). Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics*, *7*(1), 62. https://doi.org/10.1186/1471-2105-7-62

Stanke, M., & Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics*, *19*(Suppl 2), 215–225. https://doi.org/10.1093/bioinformatics/btg1080

Stapley, J., Santure, A. W., & Dennis, S. R. (2015). Transposable elements as agents of rapid adaptation may explain the genetic paradox of invasive species. *Molecular Ecology*, *24*(9), 2241–2252. https://doi.org/10.1111/mec.13089

Steinegger, M., & Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology*, *35*, 1026–1028. https://doi.org/10.1038/nbt.3988

Tajima, F. (1989a). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*, 585–595. https://doi.org/10.1093/genetics/123.3.585

Tajima, F. (1989b). The effect of change in population size on DNA polymorphism. *Genetics*, *123*(3), 597–601. https://doi.org/10.1093/genetics/123.3.597

Tsutsui, N. D., & Suarez, A. V. (2003). The colony structure and population biology of invasive ants. *Conservation Biology*, *17*(1), 48–58. https://doi.org/10.1046/j.1523-1739.2003.02018.x

Tsutsui, N. D., Suarez, A. V., Holway, D. A., & Case, T. J. (2000). Reduced genetic variation and the success of an invasive species. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(11), 5948–5953. https://doi.org/10.1073/pnas.100110397

Ün, Ç., Schultner, E., Manzano-Marín, A., Flórez, L. V., Seifert, B., Heinze, J., & Oettler, J. (2021). Cytoplasmic incompatibility between Old and New World populations of a tramp ant. *Evolution*, *75*(7), 1775–1791. https://doi.org/10.1111/evo.14261

Valencia-Montoya, W. A., Elfekih, S., North, H. L., Meier, J. I., Warren, I. A., Tay, W. T., Gordon, K. H. J., Specht, A., Paula-Moraes, S. V., Rane, R., Walsh, T. K., & Jiggins, C. D. (2020). Adaptive introgression across semipermeable species boundaries between local *Helicoverpa zea* and invasive *Helicoverpa armigera* moths. *Molecular Biology and Evolution*, *37*(9), 2568–2583. https://doi.org/10.1093/molbev/msaa108

van Boheemen, L. A., Lombaert, E., Nurkowski, K. A., Gauffre, B., Rieseberg, L. H., & Hodgins, K. A. (2017). Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia. *Molecular Ecology*, *26*(20), 5421–5434. https://doi.org/10.1111/mec.14293

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Current Protocols in Bioinformatics*, *43*(Suppl 43), 11.10.1–11.10.33. https://doi.org/10.1002/0471250953.bi1110s43

Verhoeven, K. J. F., Macel, M., Wolfe, L. M., & Biere, A. (2011). Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1702), 2–8. https://doi.org/10.1098/rspb.2010.1272

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., & Earl, A. M. (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*, *9*(11), e112963. https://doi.org/10.1371/journal.pone.0112963

Warren, R. L., Coombe, L., Mohamadi, H., Zhang, J., Jaquish, B., Isabel, N., Jones, S. J. M., Bousquet, J., Bohlmann, J., & Birol, I. (2019). ntEdit: Scalable genome sequence polishing. *Bioinformatics*, *35*(21), 4430–4432. https://doi.org/10.1093/bioinformatics/btz400

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, *7*(2), 256–276. https://doi.org/10.1016/0040-5809(75)90020-9

Wilfert, L., Gadau, J., & Schmid-Hempel, P. (2007). Variation in genomic recombination rates among animal taxa and the case of social insects. *Heredity*, *98*(4), 189–197. https://doi.org/10.1038/sj.hdy.6800950

Woronik, A., Tunström, K., Perry, M. W., Neethiraj, R., Stefanescu, C., Celorio-Mancera, M. D. L. P., Brattström, O., Hill, J., Lehmann, P., Käkelä, R., & Wheat, C. W. (2019). A transposable element insertion is associated with an alternative life history strategy. *Nature Communications*, *10*(1), 5757. https://doi.org/10.1038/s41467-019-13596-2

Wu, C., & Lu, J. (2019). Diversification of transposable elements in arthropods and its impact on genome evolution. *Genes*, *10*(5), 338. https://doi.org/10.3390/genes10050338

Yan, H., Bombarely, A., & Li, S. (2020). DeepTE: A computational method for de novo classification of transposons with convolutional neural network. *Bioinformatics*, *36*(15), 4269–4275. https://doi.org/10.1093/bioinformatics/btaa519

Yang, S., Wang, L., Huang, J. U., Zhang, X., Yuan, Y., Chen, J.-Q., Hurst, L. D., & Tian, D. (2015). Parent-progeny sequencing indicates higher mutation rates in heterozygotes. *Nature*, *523*(7561), 463–467. https://doi.org/10.1038/nature14649

Zhu, Y., Bergland, A. O., González, J., & Petrov, D. A. (2012). Empirical validation of pooled whole genome population re-sequencing in *Drosophila melanogaster*. *PLoS One*, *7*(7), 1–7. https://doi.org/10.1371/journal.pone.0041901

SUPPORTING INFORMATION

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

How to cite this article: Errbii, M., Keilwagen, J., Hoff, K. J., Steffen, R., Altmüller, J., Oettler, J., & Schrader, L. (2021). Transposable elements and introgression introduce genetic variation in the invasive ant *Cardiocondyla obscurior*. *Molecular Ecology*, 30, 6211–6228. https://doi.org/10.1111/mec.16099