Genomic Architecture of the Clownfish Hybrid *Amphiprion leucokranos*

Sarah Schmid [1,2], Diego A. Hartasánchez [1], Wan-Ting Huang [1], Ashton Gainsford [3,4], Geoffrey P. Jones [3,4], Nicolas Salamin [1,*]

[1]Department of Computational Biology, University of Lausanne, Lausanne 1015, Switzerland

[2]Ecosystems and Landscape Evolution, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland

[3]College of Science and Engineering, James Cook University, Townsville 4811, Australia

[4]ARC Center of Excellence for Coral Reef Studies, James Cook University, Townsville 4811, Australia

*Corresponding author: E-mail: nicolas.salamin@unil.ch.

Accepted: February 20, 2025

Abstract

Natural hybridization is increasingly recognized as playing a significant role in species diversification and adaptive evolution. *Amphiprion leucokranos*, the naturally occurring clownfish hybrid between *Amphiprion chrysopterus* and *Amphiprion sandaracinos*, is found within the hybrid zone of the two parental species. Based on whole-genome sequencing of parental and hybrid individuals sampled in Kimbe Bay, Papua New Guinea, we found that most of the hybrids collected were first-generation hybrids, a few were first- and second-generation backcrosses with *A. sandaracinos*, and the first evidence, to our knowledge, of both an early backcross with *A. chrysopterus* and a second-generation hybrid in the wild, highlighting the richness and diversity of genomic architectures in this hybrid zone. The frequent backcrossing with *A. sandaracinos* has led to higher levels of introgression from *A. chrysopterus* into the *A. sandaracinos* genomic background, potentially allowing for adaptive introgression. We have additionally identified morphological features which could potentially allow differentiating between first-generation hybrids and backcrosses. By comparing population genetic statistics of first-generation hybrids, backcrosses, parental populations within the hybrid zone, and parental allopatric populations, we provide the context to evaluate population differentiation and the consequences of ongoing hybridization. This study is the first whole-genome analysis of a clownfish hybrid population and builds upon the growing body of literature relative to the evolutionary outcomes of hybridization in the wild and its importance in evolution.

Key words: hybridization, whole-genome, introgression, admixture, hybrid zone.

Significance

Clownfish represent an important group of coral reef fishes engaged in mutualistic relationships with sea anemones. Recent studies have suggested that hybridization across the clownfish phylogeny has been important in the rapid diversification of clownfish species. In this study, we focus on one of the two known naturally occurring clownfish hybrids. By analyzing whole-genome sequences of *Amphiprion leucokranos* and its two parental species, we have investigated inheritance patterns, consistency of introgression among hybrids, and the impact of genetic exchanges on parental genomes. This research provides important insight into hybridization dynamics in clownfishes and will help to better understand the maintenance of species identity in the face of extensive gene flow between parental species in this iconic group.

Introduction

Hybrid zones are geographic areas where divergent lineages exchange genetic variants through the process of hybridization (Barton and Hewitt 1985; Gompert and Buerkle 2016). They are an ideal focal point to study the interaction between divergent genomic backgrounds, which might help elucidate the mechanisms underlying speciation and eventually lead to the identification of candidate genes contributing to reproductive isolation (Harrison and Larson 2014). Hybrid zones also experience ongoing evolutionary processes (Abbott et al. 2013). They can act as conduits for adaptive alleles and enable them to cross species boundaries (Pardo-Diaz et al. 2012), or increase reproductive isolation by coupling of incompatibilities (Barton and De Cara 2009) or reinforcement, in which a reduction in hybrid fitness entails the establishment of a premating barrier (Servedio and Noor 2003). We thus expect different evolutionary outcomes in hybrid zones. Extreme scenarios could either result in the completion of the speciation process by reinforcement or in genetic swamping and substitution of local parental genotypes by the hybrid ones (Allendorf et al. 2001). Alternatively, hybrid zones could lead to adaptive introgression (Hanemaaijer et al. 2018), formation of a new hybrid species (i.e. hybrid speciation; Lamichhaney et al. 2018), or persistence of hybrids beyond the initial generations through backcrossing and interbreeding (i.e. hybrid swarm; Abbott et al. 2013; Harrison and Larson 2014). Finally, hybrid zones might persist over time and be sustained by the balance between selection against hybrids and continuous mating among the parental species coming from adjacent areas (i.e. tension zone model; Barton and Hewitt 1989). Consequently, the fate of a given hybrid zone relies on its ecological and genomic contexts, but our understanding of the influence of each factor is still in its infancy.

Until recently, knowledge about the mechanisms preventing populations from fully mixing and the nature of barriers to genetic exchanges relied on a few genetic markers and phenotypic variation across and within hybrid zones (e.g. Taylor et al. 2006). The advent of high-throughput sequencing has allowed us to build on this fundamental knowledge and to help characterize the underlying determinants of the observed introgression patterns and outcomes. In particular, these genomic approaches have provided new insight into the mosaicism of hybrid genomes and into how the genomes of the parental species are blended to form hybrid individuals (e.g. Zhang et al. 2023). Hybrid genomes are characterized by the presence of admixture tracts—the genetic blocks inherited from the parental species—whose size decreases with successive generations due to recombination events. With ongoing hybrid generations, some tracts can become fixed and others eliminated by the combination of genetic drift and selection, with selection against incompatibilities likely playing an important role (Buerkle and Rieseberg 2008). Eventually, the interruption of genetic exchanges between the parental species and the hybrid can lead to genome stabilization of the hybrid taxon (Buerkle and Rieseberg 2008).

During the stabilization process, the type of selective pressure acting on the hybrid genome highly influences its architecture (Runemark et al. 2019). When hybrids mainly face ecological pressure, selection favors reproductive isolation from the parental species and adaptation to a novel niche. The genome then stabilizes through the purging of incompatibilities and the maintenance of co-adapted gene complexes and pathways, resulting in a hybrid genome with both parents' ancestry tracts (Schumer et al. 2016). In the case of numerous incompatibilities, selection tends to remove the alleles of the minor parent—the species from which hybrids derive less of their genome—in gene-rich regions through backcrossing. Such a mechanism will result in a stabilized genome with a few adaptive admixed regions (Pardo-Diaz et al. 2012; Hanemaaijer et al. 2018). The proportion of the genome inherited by the hybrids from each parent might thus vary substantially within species after subsequent hybrid generations. The hybrids can exhibit mosaic genomes with an equal contribution from both parental species (e.g. Italian sparrow; Elgvin et al. 2017), a highly unbalanced contribution (*Heliconius* butterflies; Jiggins et al. 2008), or even genomes with an ancestry almost completely biased toward a single donor (e.g. *Anopheles* mosquito; Hanemaaijer et al. 2018). Moreover, some hybrids also exhibit high variation in parental genome proportions among individuals, as shown in swordtail fishes (Schumer et al. 2016). As puzzling as such patterns might be, the causes and evolutionary consequences of hybridization outcomes are still poorly understood.

Here, we focus on a precise location (i.e. Kimbe Bay, Papua New Guinea) of the mosaic hybrid zone (as defined by Rand and Harrison 1989) between two genetically and morphologically distinct clownfish species—*Amphiprion sandaracinos* and *Amphiprion chrysopterus*. Like all clownfish species, *A. sandaracinos* and *A. chrysopterus* exhibit a mutualistic interaction with specific host sea anemones (Marcionetti et al. 2019), and have an overlap in host preference. They have a size-based hierarchy, where the largest individual in the sea anemone is the dominant female, followed in size by the male and smaller nonbreeding subordinates (Buston 2003). Both species display mostly allopatric distributions but coexist in a narrow area between the Solomon Islands and the Halmahera Island (Indonesia), where they hybridize and form the hybrid *Amphiprion leucokranos* (Gainsford et al. 2015). Due to its distinctive color patterns, the hybrid was previously considered a nominal species, but genetic, morphological, and ecological data have confirmed its hybrid status (Gainsford

et al. 2015; He et al. 2019). Hybrids exhibit intermediary size, colors, and ecological traits compared to the parental species and, similar to the parents, use mainly *Stychodactyla mertensii* as the host sea anemone or *Heteractis crispa* as an alternative host (Gainsford et al. 2015). The size difference of parental species combined with their hierarchical social structure were suggested to be the main drivers of their mating patterns, with the larger *A. chrysopterus* being the female when reproducing with the smaller *A. sandaracinos* (Gainsford et al. 2015).

Three distinct locations (i.e. (1) Kimbe Bay, (2) Kavieng, in Papua New Guinea, and (3) Solomon Islands) within the hybrid zone between *A. sandaracinos* and *A. chrysopterus* were previously studied based on a few genetic markers and morphological data (Gainsford et al. 2015, 2020). The outcome of hybridization events in each of these locations is thought to be dependent on their specific ecological context, with the relative abundances and the distinct sizes of the parental species most likely explaining the hybridization patterns (Gainsford et al. 2015, 2020). Mixed-species assemblages frequently occur in the Solomon Islands—where *A. chrysopterus* is more abundant compared to *A. sandaracinos*. In addition, hybrids from both Kavieng and Solomon Islands exhibit balanced parental ancestry. In contrast, there is a bias toward *A. sandaracinos* ancestry in hybrids in Kimbe Bay, where both parental species have relatively similar frequency and mainly form conspecific assemblages (Gainsford et al. 2020). Indeed, Kimbe Bay comprises hybrids with equal contribution from both parents alongside putative *A. sandaracinos* backcrosses (BCs). However, BCs with *A. chrysopterus* are seldom observed (Gainsford et al. 2015, 2020). This pattern is surprising given that hybrids inherit their mitochondrial genome from the female *A. chrysopterus* and mounting evidence that incompatible mitonuclear interactions might result in hybrid fitness reduction (Runemark et al. 2018).

In this study, we took a step further into understanding the dynamics of the Kimbe Bay hybrid zone location using whole-genome sequencing data. We investigated the genomic outcomes of the recurrent hybridization among *A. sandaracinos* and *A. chrysopterus*, focusing on its impact on both the hybrids and the parental species. Given the presence of widespread hybridization in clownfishes (Litsios et al. 2014; Schmid et al. 2022; Marcionetti and Salamin 2023) and its potential role in their evolutionary history, we wanted to characterize the Kimbe Bay hybrid zone through an in-depth genomic analysis of *A. sandaracinos*, *A. chrysopterus* and hybrid (hereafter *A. leucokranos*) individuals and test if it is a tension zone (Barton and Hewitt 1985) with mainly first-generation (F1) hybrids or if it is a hybrid swarm (Grant 1981; Allendorf et al. 2001) with a mix of early and later hybrid generations. Furthermore, we analyzed the patterns of admixture among hybrid individuals and the impact of recurrent genetic exchanges on the parental genomes. Shedding light on these matters will improve current knowledge of the mechanisms maintaining species integrity and this hybrid's genomic architecture.

Results

Sequencing, Mapping, and Single Nucleotide Polymorphism (SNP) Calling Statistics

Our dataset consists of 23 *A. chrysopterus*, 24 *A. leucokranos*, 20 *A. sandaracinos*. All 24 *A. leucokranos*, as well as ten *A. chrysopterus* and ten *A. sandaracinos* were sampled in Kimbe Bay (Papua New Guinea). Within the species range overlap between *A. chrysopterus* and *A. sandaracinos* (He et al. 2019), we sampled seven *A. sandaracinos* and eight *A. chrysopterus* from the Solomon Islands, and one *A. chrysopterus* from Palau. We additionally sampled parental species from allopatric populations in regions with no parental species overlap: three *A. sandaracinos* from Christmas Island and four *A. chrysopterus* from Fiji (Fig. 1 and supplementary table S1, Supplementary Material online). We achieved an average of ~5.91 billion raw paired reads across 67 samples, with the number of raw reads per sample ranging from ~40 to ~191 million (supplementary table S2, Supplementary Material online). After trimming low-quality regions and removing low-quality reads, we ended up with ~36 to 173 million paired reads per sample, corresponding to an estimated coverage between 6.1× and 28.9× (supplementary table S2, Supplementary Material online).

We mapped the trimmed reads to the *Amphiprion clarkii* (Moore et al. 2023) genome and 90.2% to 96.6% of the reads mapped properly to the reference (i.e. with pairs having the correct orientation and insert-size; supplementary table S3, Supplementary Material online). We further filtered the data by removing duplicated reads and reads mapped with low-confidence to obtain a final average coverage ranging between 3× and 24.3× (supplementary table S3, Supplementary Material online). Following these steps, three individuals with low depth were removed and we proceeded to call variants and filtered the resulting variant call format (VCF) file to obtain two final data set consisting of (1) 3,304,054 single nucleotide polymorphisms (SNPs) for 64 individuals (*A. leucokranos*, *A. chrysopterus*, and *A. sandaracinos* data set; after removal of three individuals with low average read depth—see "Materials and methods") and (2) 9,456,356 SNPs for 74 individuals (full dataset including 10 *A. clarkii* (Schmid et al. 2024) used as outgroup—see "Materials and methods"; supplementary table S4, Supplementary Material online).

Principal Component Analysis, Admixture, and Mitochondrial Phylogenetic Reconstruction

To describe the relationship between the two parental species and the hybrid, we first performed a principal

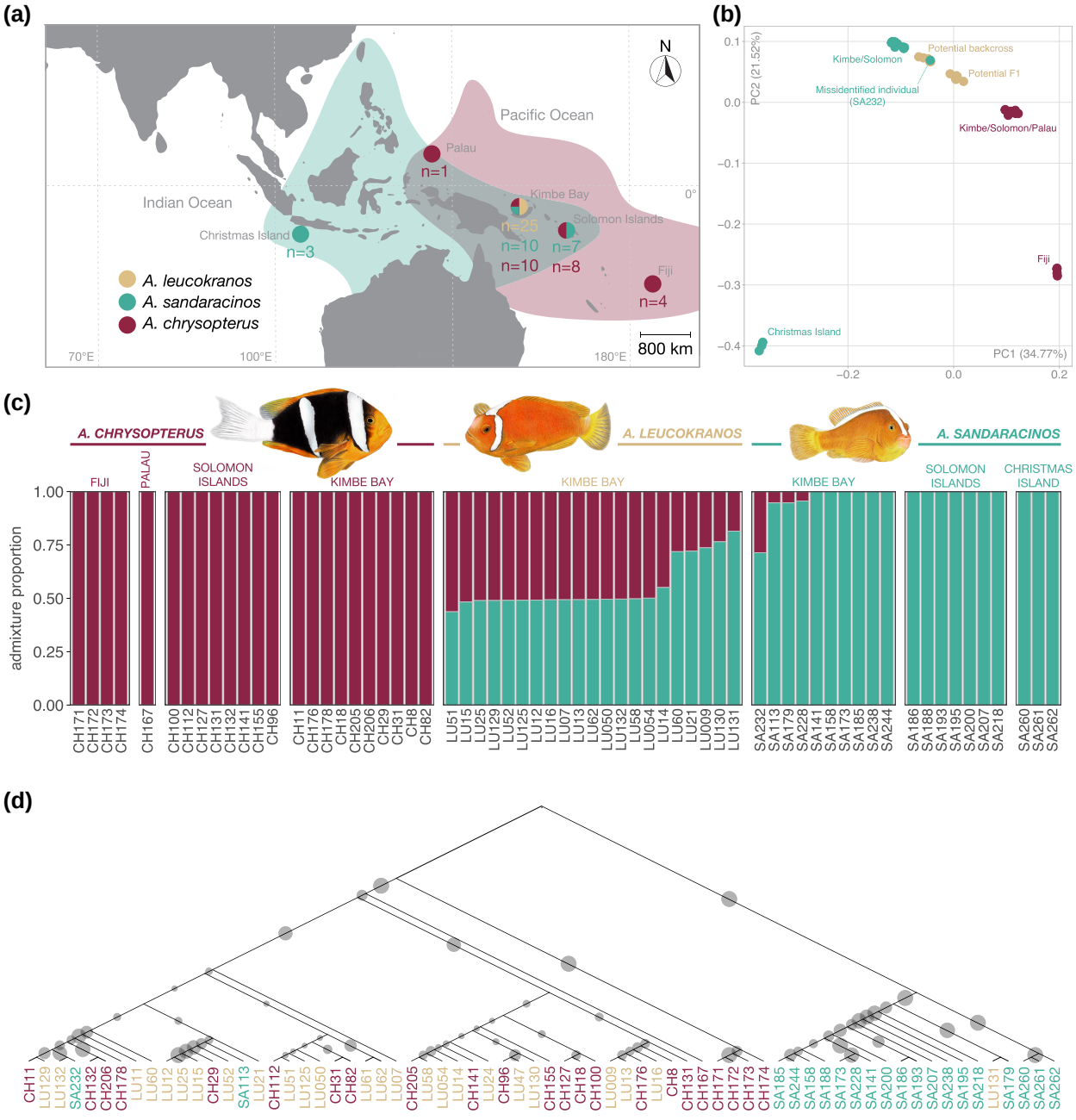

Fig. 1. Sampling and population genetic structure of *A. chrysopterus*, *A. leucokranos* (hybrid), and *A. sandaracinos*. a) Colored circles represent sampling sites for the study and the two color patches correspond to the geographic range of the two parental species (*A. chrysopterus* and *A. sandaracinos*) following He et al. (2019). The numbers below correspond to the number of samples for each site and species. b) PCA showing a separation of the two parental species (*A. chrysopterus* and *A. sandaracinos*) on the first axis (PC1), with the hybrid species (*A. leucokranos*) in the middle, and separation of individuals within the Kimbe Bay hybrid zone with allopatric populations *A. chrysopterus* from Fiji and *A. sandaracinos* from Christmas Island on the second axis (PC2). c) Admixture bar plot for $K = 2$. Each bar represents the ancestry proportion of each individual for both parental species (*A. chrysopterus* or *A. sandaracinos*). d) Mitochondrial midpoint-rooted phylogenetic reconstruction with IQ-tree V2.2.2. The phylogeny is based on the 67 mitochondrial genomes reconstructed with MITObim using *A. percula* reference mtDNA genome. The size of the dots on the branches corresponds to the bootstrap support (ranging from 5 to 100) based on 1,000 ultrafast bootstraps.

component analysis (PCA) on the covariance matrix between all individuals (Fig. 1b, showing the first two principal components, and supplementary fig. S1, Supplementary Material online showing the second and third principal components). We found a clear separation between *A. chrysopterus* and *A. sandaracinos*—the parental

species—along the first axis of the PCA, which explains 34.77% of the variance. We can additionally identify two clusters of *A. leucokranos* individuals: One equally distant from both parental species clusters and the second one closer to the *A. sandaracinos* cluster. Among the latter, there is an *A. sandaracinos* individual (SA232) potentially representing a misidentified *A. leucokranos*. The second axis of the PCA explains 21.52% of the variance and separates *A. chrysopterus* individuals from Fiji as well, as the *A. sandaracinos* individuals from Christmas Island, from the main cluster. These two clusters representing the allopatric populations are separated from the rest of individuals which are within the parental species overlapping range, potentially representing that variants are shared among all these individuals through hybridization and backcrossing events within the hybrid zone.

We quantified the contribution of each parental species to the hybrid genome and potential introgression events in parental populations by calculating the admixture proportion of each individual (Fig. 1c and supplementary fig. S2, Supplementary Material online). Cross-validation values for $K = 2$ (Fig. 1c), $K = 4$, and $K = 5$ were the lowest and highly similar (~0.32), followed by $K = 3$ (0.34), $K = 6$ (0.36), and $K = 7$ (0.38). At $K = 3$, each species (hybrid and parental species) formed a separate cluster, at $K = 4$ individuals from Fiji (allopatric *A. chrysopterus* population) formed an additional cluster, and at $K = 5$, the *A. sandaracinos* samples from Christmas Island (allopatric population) formed the fifth cluster (see supplementary fig. S2, Supplementary Material online for results from $K = 3$ to $K = 7$). We found that 14 *A. leucokranos* individuals had half of each parental species ancestry (considered as F1-hybrids or "F1s" in subsequent analyses), four individuals (LU60, LU21, LU009, and LU130) had a ~0.75 proportion of *A. sandaracinos* ancestry (considered as first-generation *A. sandaracinos*-BCs or "BCs", that is, resulting from an F1 × *A. sandaracinos* cross, in subsequent analyses). One individual (LU51) had slightly below half of *A. chrysopterus* ancestry, one individual LU14 had slightly above half, and another individual (LU131) had slightly above 0.75. Four *A. sandaracinos* individuals from Kimbe Bay displayed introgression from *A. chrysopterus*. The one with the highest level of introgression (SA232) was probably a misidentified *A. leucokranos* (in particular, an *A. sandaracinos*-BC), while the other three individuals (SA113, SA179, and SA228) had much lower levels of *A. chrysopterus* ancestry.

Finally, we determined the parental inheritance of the mitochondrial genome in hybrids by reconstructing a maximum-likelihood tree based on the whole mitochondrial sequence (Fig. 1d). All *A. leucokranos* clustered with *A. sandaracinos*, except for a single individual (LU131). We also found that two *A. sandaracinos* individuals grouped with *A. chrysopterus* (SA113 and SA232), one of them (SA232) being the potentially misidentified *A. leucokranos*.

Ancestry Tracks in Hybrids

To try to determine the origin of individuals not labeled as F1s or *A. sandaracinos*-BCs in the previous analysis, we estimated allele dosage scores at each SNP using the two-layer hidden Markov model implemented in ELAI (Guan 2014) for 25 individuals with mixed ancestry according to the admixture analysis. Among the 21 *A. leucokranos* individuals and consistent with the admixture analysis, 14 had more than 98% heterozygous genomes consistent with our previous labeling of them as F1s (Fig. 2a and supplementary fig. S3, Supplementary Material online). One individual (LU51) had ~13% of homozygous *A. chrysopterus* variants, consistent with being the descendant of a first-generation *A. sandaracinos*-BC × *A. chrysopterus* cross (Fig. 2a and supplementary fig. S3, Supplementary Material online). One individual (LU14) exhibited ~25% homozygous *A. sandaracinos* and 14% homozygous *A. chrysopterus*, a pattern difficult to interpret but which could be consistent with a second-generation hybrid (F2) with a reduced *A. chrysopterus* homozygous component (Fig. 2a and supplementary fig. S3, Supplementary Material online). Six individuals (LU009, LU21, LU60, LU130, LU131, and SA232) displayed between 42.6% and 57.5% of homozygous *A. sandaracinos* tracks consistent with them being BCs, and including the misidentified individual (Fig. 2a and supplementary fig. S3, Supplementary Material online). These six individuals form a separate cluster from most other *A. leucokranos* in the PCA (Fig. 1b) and are consistent with being BCs. Among these, one individual (LU131) had the highest (57.5%) homozygous *A. sandaracinos* component and clustered among the *A. sandaracinos* in the mitochondrial phylogeny, suggesting it is a BC but with an *A. sandaracinos* mother. Two individuals (SA113 and SA179) had ~87.5 homozygous *A. sandaracinos* tracks which would be consistent with being second-generation BCs (i.e. descendants of BC × *A. sandaracinos* crosses). However, among these, SA113 clusters with *A. chrysopterus* in the mitochondrial phylogeny, which would be difficult to reconcile with it being a second-generation *A. sandaracinos*-BC. Finally, SA228 appears to be a pure *A. sandaracinos*, inconsistent with the admixture analysis. We additionally analyzed the genomic tracks with homozygous *A. sandaracinos* ancestry in five *A. sandaracinos*-BCs (LU009, LU21, LU60, LU130, and LU131) to check for the presence of regions being consistently inherited from either the *A. sandaracinos* or the *A. chysopterus* grandparent. The distribution of tracks with homozygous *A. sandaracinos* ancestry in BC individuals was scattered across the genome. Most chromosomes displayed *A. sandaracinos* homozygous tracks in at least one BC except for chromosome 8

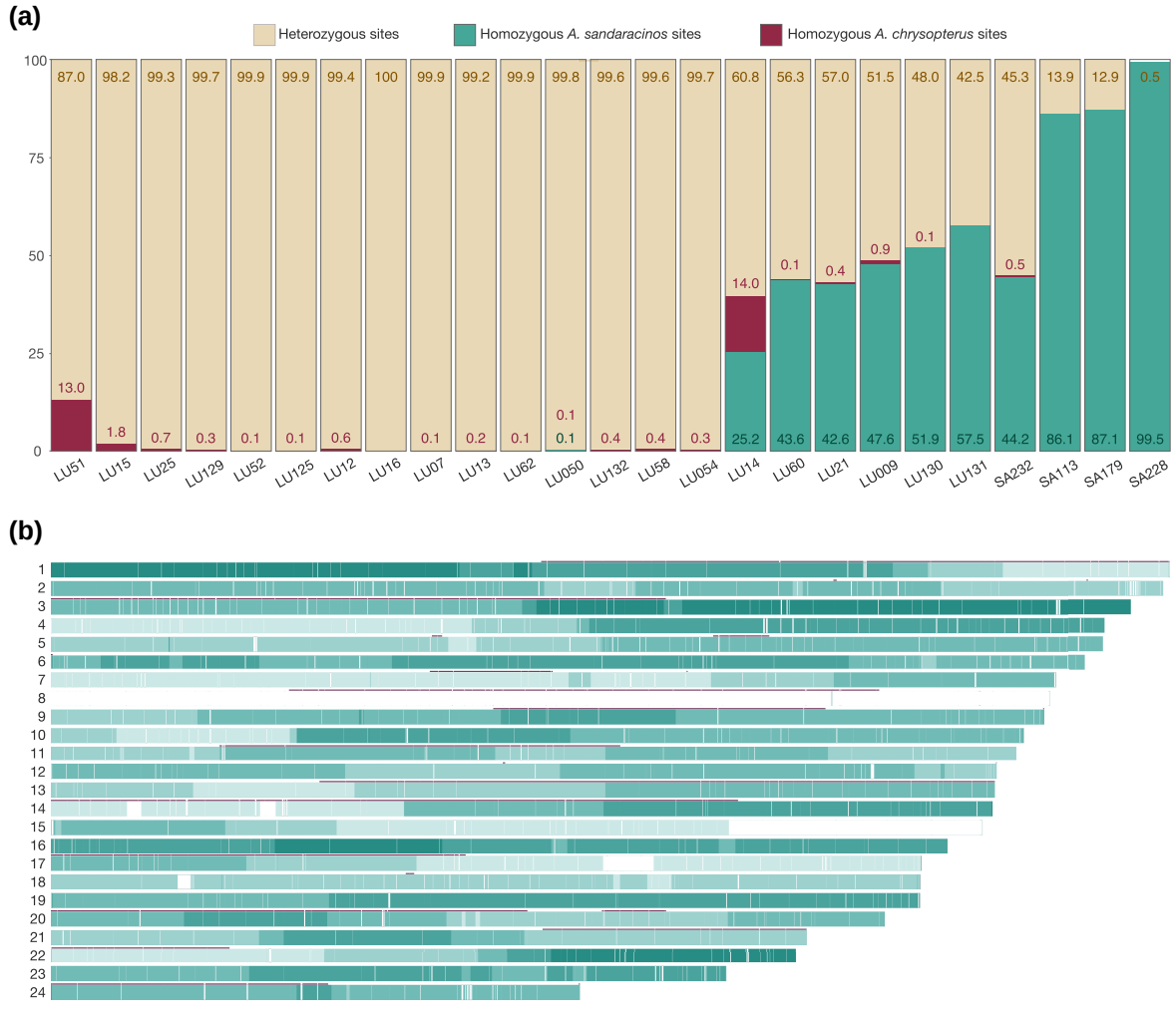

Fig. 2. Local ancestry inference with ELAI (Guan 2014). a) Percentage of heterozygous sites, homozygous sites for *A. sandaracinos*, and homozygous sites for *A. chrysopterus* for each hybrid *A. leucokranos* (LU individuals) and *A. sandaracinos* (SA individuals) displaying introgression. Percentages are also indicated at the top of each bar if not equal to zero. The distribution of heterozygous and homozygous sites across chromosomes for each individual is available in supplementary fig. S3, Supplementary Material online. b) Karyotype of the parental ancestry estimated with ELAI (Guan 2014) across the 24 chromosomes of *A. sandaracinos*-BCs (LU60, LU21, LU009, LU130, and LU131). The number of BCs with homozygous *A. sandaracinos* ancestry ranges from zero to five individuals and is shown with increasing color shades. The line above each chromosome highlights regions with *A. chrysopterus* homozygous ancestry in LU51.

(Fig. 2b). A percentage of 15.5 of the genome displayed *A. sandaracinos* homozygous ancestry in at least four out of the five BC individuals.

Genomic and Phenotypic Characterization of *A. leucokranos*

To explore the genotype–phenotype relationship among *A. leucokranos* individuals, we performed two additional PCAs on the 20 *A. leucokranos* for which we have a picture (see "Materials and methods"). We computed the first PCA using the whole-genome SNPs dataset (Fig. 3a) and the second using the nine binary phenotypic traits characterizing each hybrid (Fig. 3b; see supplementary table S5, Supplementary Material online for the matrix of phenotypes). The phenotypic traits used were based on white bands asymmetry (A), the extent of the white dorsal band (*A. sandaracinos* characteristic), and the number of vertical white stripes (*A. chrysopterus* characteristic). The first axis of the genomic PCA (Fig. 3a) split the *A. sandaracinos*-BC individuals from the rest of the individuals with the single second-generation *A. chrysopterus*–*A. sandaracinos*-BC individual (LU51) located in the opposite end. PC1 explained 8.14% of the variance, with the

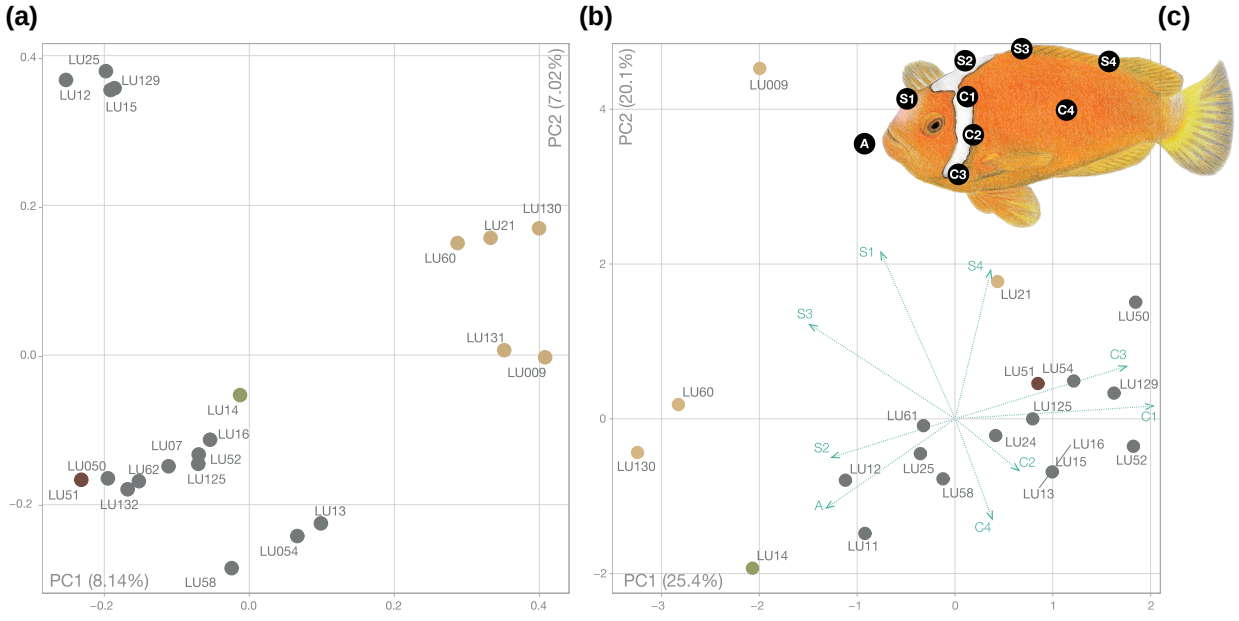

Fig. 3. PCA based on genetic and phenotypic data considering only *A. leucokranos* individuals. a) PCA based on whole-genome SNPs for *A. leucokranos* only. Labels correspond to the sample IDs. b) PCA based on phenotype matrix (supplementary table S5, Supplementary Material online). Arrows indicate loadings for each trait. The color code is the same as in a). Labels correspond to the sample IDs. c) Illustration of the nine phenotypic characteristics used for the PCA.

only putative F2 hybrid (LU14) in-between them. The second axis explained 7.02% of the variance and split the F1-hybrids into three clusters. The first and second axes of the PCA based on phenotypic data (Fig. 3b), respectively, explained 25.4% and 20.1% of the variance. F1-hybrids formed a cluster, whereas *A. sandaracinos*-BCs and the F2 hybrid were spread across the left and top parts of the PCA space and exhibited *A. sandaracinos*-like phenotypic traits.

Population Genomic Statistics Between Hybrids and Parental Species

To investigate the genome-wide patterns of differentiation, divergence, and diversity among the hybrids and the different parental species populations, we averaged the differentiation (F_{ST}), the absolute divergence (d_{xy}), and the nucleotide diversity (π) values across 50-kb genomic windows (Table 1 and supplementary table S6, Supplementary Material online). Pairwise d_{xy} values between hybrids and parental species were stable across the three parental populations. Comparatively, pairwise F_{ST} values were more variable across parental populations potentially due to differences in sample sizes. Finally, nucleotide diversity in *A. sandaracinos* was around half of that of *A. chrysopterus* across corresponding populations; and 4.5 and 3.5 higher than *A. chrysopterus* in Kimbe Bay for F1s and BCs, respectively.

Table 1 Mean F_{ST}, d_{xy}, and nucleotide diversity (π) between the hybrids and the parental species

Parameter	Species	Parental population origins		
		Kimbe Bay	Solomon Islands	Allopatric
F_{ST}	CHP-BC	0.425 ± 0.162	0.496 ± 0.163	0.476 ± 0.150
	CHP-F1	0.225 ± 0.039	0.207 ± 0.033	0.190 ± 0.031
	SAN-BC	0.091 ± 0.083	0.109 ± 0.086	0.189 ± 0.118
	SAN-F1	0.187 ± 0.031	0.206 ± 0.036	0.139 ± 0.031
	BC-F1	0.031 ± 0.049	-	-
d_{xy}	CHP-BC	0.346 ± 0.065	0.346 ± 0.065	0.362 ± 0.055
	CHP-F1	0.252 ± 0.029	0.252 ± 0.029	0.295 ± 0.029
	SAN-BC	0.135 ± 0.060	0.137 ± 0.060	0.173 ± 0.057
	SAN-F1	0.238 ± 0.032	0.240 ± 0.032	0.260 ± 0.033
	BC-F1	0.241 ± 0.030	-	-
π	CHP	0.056 ± 0.035	0.054 ± 0.035	0.062 ± 0.033
	SAN	0.028 ± 0.037	0.027 ± 0.034	0.039 ± 0.037
	BC	0.193 ± 0.067	-	-
	F1	0.252 ± 0.030	-	-

Allopatric corresponds to the *A. chrysopterus* Fiji population and *A. sandaracinos* Christmas Island population. All comparisons and population statistics are in supplementary table S6, Supplementary Material online. CHP, *A. chrysopterus*; SAN, *A. sandaracinos* (admixed individuals removed); BC, backcross with *A. sandaracinos*; F1, first-generation hybrids.

Characterization of BC's Genomic Architecture

Following the identification of homozygous tracts of *A. sandaracinos* ancestry in BCs, we considered "BC windows" all of those windows which had homozygous *A. sandaracinos* tracts in four out of the five BCs analyzed.

Table 2 Mean population genomics statistics and number of putative barrier loci in BC windows compared to the genomic background

Parameter	Species	BC windows	Background	Significance	Effect size
F_{ST}	CHP-BC	0.603 ± 0.118	0.468 ± 0.159	[d]	Large
	CHP-F1	0.220 ± 0.036	0.226 ± 0.038	[d]	Negligible
	CHP-SAN	0.801 ± 0.114	0.815 ± 0.121	[d]	Negligible
	SAN-BC	0.023 ± 0.031	0.109 ± 0.081	[d]	Large
	SAN-F1	0.186 ± 0.027	0.188 ± 0.030	[d]	Negligible
	BC-F1	0.071 ± 0.030	0.022 ± 0.048	[d]	Large
d_{xy}	CHP-BC	0.390 ± 0.045	0.335 ± 0.064	[d]	Large
	CHP-F1	0.255 ± 0.024	0.251 ± 0.030	[d]	Negligible
	CHP-SAN	0.444 ± 0.050	0.444 ± 0.059	[c]	Negligible
	SAN-BC	0.086 ± 0.035	0.146 ± 0.059	[d]	Large
	SAN-F1	0.236 ± 0.028	0.237 ± 0.032	[d]	Negligible
	BC-F1	0.239 ± 0.026	0.241 ± 0.030	[d]	Negligible
π	CHP	0.063 ± 0.035	0.054 ± 0.034	[d]	Small
	SAN	0.025 ± 0.032	0.027 ± 0.036	[a]	Negligible
	BC	0.140 ± 0.043	0.205 ± 0.065	[d]	Large
	F1	0.252 ± 0.025	0.251 ± 0.031	[b]	Negligible
f_{dM}	SAN-SAN-CHP	0.006 ± 0.032	0.007 ± 0.036	[b]	Negligible
	CHP-CHP-SAN	0.002 ± 0.022	0.002 ± 0.022	NS	-
Barrier loci	…	1	41	[b]	-

For the comparison with the parental species, we only focused on *A. sandaracinos* and *A. chrysopterus* from the studied hybrid zone in Kimbe Bay (Papua New Guinea). Significance: nonsignificant (NS), [a]*P*-value < 0.05. [b]*P*-value < 0.01. [c]*P*-value < 0.001. [d]*P*-value < 0.0001. Effect size (ES) calculated with Cohen's *d* statistic: $d < 0.2$ (negligible), $d < 0.5$ (small), $d < 0.8$ (moderate), $d > 0.8$ (large). CHP, *A. chrysopterus*; SAN, *A. sandaracinos* (admixed individuals removed); BC, first-generation backcross with *A. sandaracinos*; F1, first-generation hybrids. f_{dM} SAN-SAN-CHP corresponds to the topology. P1, *A. sandaracinos* (Christmas Island); P2, *A. sandaracinos* (Kimbe Bay); P3, *A. chrysopterus* (Kimbe Bay) and f_{dM} CHP-CHP-SAN corresponds to the topology; P1, *A. chrysopterus* (Fiji); P2, *A. chrysopterus* (Kimbe Bay); P3, *A. sandaracinos* (Kimbe Bay).

We then calculated F_{ST}, d_{xy}, and π for these BC windows and compared the values obtained against the background (i.e. the rest of the genome). All values were significantly different between the BC and background genomic windows (Table 2). However, the effect size was negligible or small for most comparisons. The only exceptions were the comparisons involving the *A. sandaracinos*-BCs themselves, for which we detected a large effect size for the differences in nucleotide diversity and both divergence and differentiation with respect to the two parental species (Table 2). Furthermore, the observed frequency of putative barrier loci (see "Materials and methods" for definition) was significantly different from the expected frequency ($\chi^2 = 7$, df = 1, *P*-value = 0.006; Table 2).

Genomic Landscape of F1-hybrids

The admixture and local ancestry analyses highlighted that 14 *A. leucokranos* were F1-hybrids between *A. sandaracinos* and *A. chrysopterus*. To investigate the relationship between the F1-hybrids and the parental species in sympatry (Kimbe Bay), we compared the level of differentiation (F_{ST}) between the hybrids and both parents to identify highly divergent regions of the genome (see "Materials and methods" for definition). The distribution of the F_{ST} difference across the genome was significantly biased toward higher differentiation between the F1-hybrids and *A. chrysopterus* (mean difference significantly different from 0; *P*-value < 2E-16; Fig. 4a). The number of highly divergent windows per chromosome was not statistically different from the expected distribution generated with 10,000 permutations for most chromosomes, except for chromosomes 2, 10, 15, and 22, which exhibited an inflated number of outliers (i.e. respectively 19, 29, 33, and 15 windows; Fig. 4a).

Although all population genomic statistics were statistically different between the background and highly divergent windows (except for *A. sandaracinos*-F1 F_{ST} in highly divergent *A. chrysopterus* windows), the effect size was moderate or large in only part of the statistics (Table 3). Furthermore, the two f_{dM} statistics were higher in the outlier windows compared to the background windows (Table 3). The f_{dM} statistics based on the two different topologies ($P1 = CHP_{Fiji}$, $P2 = CHP_{Kimbe}$, $P3 = SAN_{Kimbe}$ and $P1 = SAN_{Chr.Isl}$, $P2 = SAN_{Kimbe}$, and $P3 = CHP_{Kimbe}$) were also statistically different in the background windows and highly divergent *A. chrysopterus* windows, with negligible and moderate effect size, respectively (Fig. 4b; *P*-value = 7.7e−78; $d = -0.16$ and *P*-value = 1.7e−35; $d = -0.60$).

Population Genomics of the Parental Species in the Hybrid Zone and Outside

Having confirmed the occurrence of early-generation *A. leucokranos* hybrids (i.e. F1s, F2s, and BCs), we conducted population genomic analyses of the two parental species in Kimbe Bay to explore the impact of ongoing

Table 3 Mean population genomics statistics for highly divergent windows between F1-hybrids and parental species

Parameter	Species	CHP windows	SAN windows	Background	CHP *p* sign.	SAN *p* sign.	CHP ES	SAN ES
F_{ST}	CHP-BC	0.672 ± 0.227	0.265 ± 0.152	0.496 ± 0.157	[d]	[d]	Large	Large
	CHP-F1	0.242 ± 0.031	0.147 ± 0.062	0.225 ± 0.036	[d]	[d]	Small	Large
	CHP-SAN	0.829 ± 0.132	0.601 ± 0.214	0.814 ± 0.116	[d]	[d]	Negligible	Large
	SAN-BC	0.0254 ± 0.078	0.127 ± 0.114	0.092 ± 0.081	[d]	[d]	Large	Small
	SAN-F1	0.178 ± 0.045	0.157 ± 0.068	0.188 ± 0.029	NS	[d]	Small	Moderate
	BC-F1	0.096 ± 0.083	0.008 ± 0.042	0.032 ± 0.048	[d]	[d]	Large	Moderate
d_{xy}	CHP-BC	0.406 ± 0.090	0.257 ± 0.073	0.346 ± 0.063	[d]	[d]	Moderate	Large
	CHP-F1	0.255 ± 0.032	0.215 ± 0.052	0.252 ± 0.028	[a]	[d]	Negligible	Large
	CHP-SAN	0.457 ± 0.066	0.365 ± 0.091	0.445 ± 0.056	[d]	[d]	Negligible	Large
	SAN-BC	0.091 ± 0.083	0.190 ± 0.062	0.134 ± 0.060	[d]	[d]	Moderate	Large
	SAN-F1	0.246 ± 0.035	0.233 ± 0.050	0.237 ± 0.031	[d]	[b]	Small	Negligible
	BC-F1	0.247 ± 0.034	0.225 ± 0.043	0.240 ± 0.029	[d]	[d]	Small	Small
Π	CHP	0.040 ± 0.028	0.085 ± 0.048	0.056 ± 0.034	[d]	[d]	Small	Moderate
	SAN	0.041 ± 0.055	0.081 ± 0.084	0.026 ± 0.034	[d]	[d]	Negligible	Large
	BC	0.121 ± 0.105	0.216 ± 0.061	0.192 ± 0.066	[d]	[d]	Large	Negligible
	F1	0.258 ± 0.034	0.226 ± 0.043	0.251 ± 0.029	[d]	[d]	Small	Large
f_{dM}	SAN-SAN-CHP	0.019 ± 0.052	0.073 ± 0.122	0.007 ± 0.035	[d]	[d]	Small	Moderate
	CHP-CHP-SAN	0.008 ± 0.032	0.028 ± 0.071	0.002 ± 0.022	[d]	[d]	Small	Small
Barrier loci	…	0	0	42	NS	NS	-	-

For the comparison with the parental species, we only focused on *A. sandaracinos* and *A. chrysopterus* from the studied hybrid zone in Kimbe Bay (Papua New Guinea). CHP windows correspond to the windows with values above the 99th percentile (F_{ST} bias toward *A. chrysopterus*, meaning that F_{ST} is higher in *A. chrysopterus* compared to *A. sandaracinos*). SAN windows correspond to the windows with values below the 1st percentile (F_{ST} bias toward *A. sandaracinos*, meaning that F_{ST} is higher in *A. sandaracinos* compared to *A. chrysopterus*). Significance: nonsignificant (NS), [a]*P*-value < 0.05. [b]*P*-value < 0.01. [d]*P*-value < 0.0001. ES calculated with Cohen's *d* statistic: $d < 0.2$ (negligible), $d < 0.5$ (small), $d < 0.8$ (moderate), $d > 0.8$ (large). CHP, *A. chrysopterus*; SAN, *A. sandaracinos* (admixed individuals removed); BC, backcross with *A. sandaracinos*; F1, first-generation hybrids. f_{dM} SAN-SAN-CHP corresponds to the topology. P1, *A. sandaracinos* (Christmas Island); P2, *A. sandaracinos* (Kimbe Bay); P3, *A. chrysopterus* (Kimbe Bay) and f_{dM} CHP-CHP-SAN corresponds to the topology. P1, *A. chrysopterus* (Fiji); P2, *A. chrysopterus* (Kimbe Bay); P3, *A. sandaracinos* (Kimbe Bay).

genetic exchanges and recurrent backcrossing in *A. sandaracinos* (Fig. 5a). Mean differentiation and divergence across the genome between *A. sandaracinos* and *A. chrysopterus* were, respectively, 0.811 (±0.124) and 0.445 (±0.059). The mean nucleotide diversity was higher in *A. chrysopterus* (0.056 ± 0.035) compared to *A. sandaracinos* (0.028 ± 0.037). Finally, although the mean f_{dM} values across the genome were close to zero (0.007 for $P1 = SAN_{Chr.lsl}$, $P2 = SAN_{Kimbe}$, $P3 = CHP_{Kimbe}$ and 0.002 for $P1 = CHP_{Fiji}$, $P2 = CHP_{Kimbe}$, $P3 = SAN_{Kimbe}$), we found some positive peaks in chromosomes 2, 6, 10, 11, and 15 (Fig. 5a).

We also compared the density distribution of the same population genomics statistics to other *A. sandaracinos*–*A. chrysopterus* pairwise populations either in sympatry (Solomon Islands) or in allopatry (Fiji and Christmas Island; Fig. 5b). F_{ST} and d_{xy} distributions were similar among the three pairwise population comparisons, with the distribution of values for the allopatric populations slightly shifting toward smaller values compared to the two sympatric populations. Finally, nucleotide diversity had highly similar distributions for both species in Kimbe Bay and the Solomon Islands and allopatric populations of both species displayed higher nucleotidic diversity compared to the two sympatric populations (Fig. 5b, see supplementary fig. S4, Supplementary Material online for all pairwise comparisons).

Discussion

Hybrid zones are robust systems for studying the interaction between diverging gene pools and thus provide the basis for understanding the mechanisms leading to speciation (Barton and Hewitt 1985; Gompert and Buerkle 2016). Besides their importance in studying evolutionary processes, hybrid zones are puzzling phenomena displaying their own evolutionary history, as their fate ranges from population collapse to generation of wide admixture zones (Abbott et al. 2013). Here, we conducted whole-genome analyses of the Kimbe Bay hybrid zone between two clownfish species, *A. chrysopterus* and *A. sandaracinos*. We aimed to investigate the impact of hybridization on the integrity of the parental species and characterize the genomic architecture of the resulting hybrid *A. leucokranos.* Our analyses showed that besides the parental species, the hybrid zone comprised F1-hybrids and early BC generations with *A. sandaracinos*, with later generation hybrids and BCs with *A. chrysopterus* being also present but scarce. An in-depth investigation of the genomic architecture of the BC individuals revealed that the introgression of *A. sandaracinos* into the hybrid genome was unbalanced between chromosomes. The genomic examination of the parental species integrity highlighted high-level introgression concentrated in some genomic regions and a potential

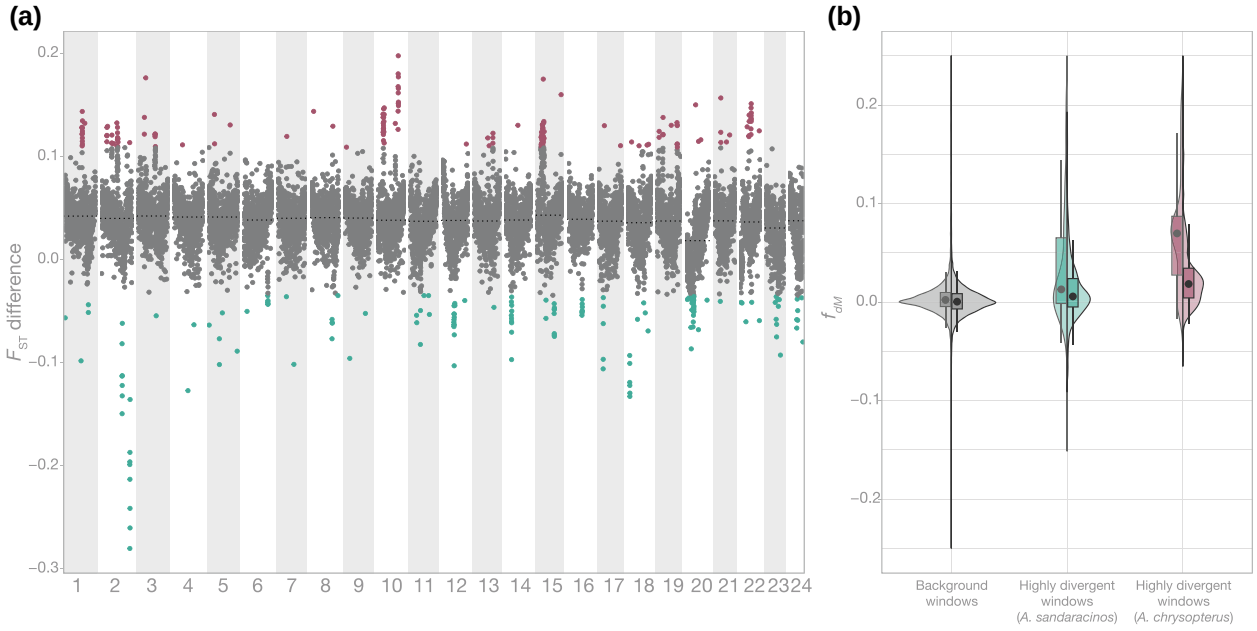

Fig. 4. Characterization of the F1-hybrids genomic landscape and biased introgression between parental species. a) F_{ST} difference between F1-hybrids and both parental species from Kimbe Bay (CHP-F1 F_{ST}—SAN-F1 F_{ST}) calculated for consecutive 50-kb windows. Color points correspond to values above the 99th percentile (high F_{ST} for *A. chrysopterus*-F1 and low F_{ST} for *A. sandaracinos*-F1) or below the 1st percentile (low F_{ST} for *A. chrysopterus*-F1 and high F_{ST} for *A. sandaracinos*-F1) and were considered highly divergent windows. The mean F_{ST} difference for each chromosome is indicated with a dotted line. b) Distribution of f_{dM} across all background windows and across the highly divergent windows biased toward *A. sandaracinos* or toward *A. chrysopterus*. The left side of the violin plot represents the f_{dM} distribution for the following topology: P1, *A. sandaracinos* (Christmas Island); P2, *A. sandaracinos* (Kimbe); P3, *A. chrysopterus* (Kimbe); Outgroup, *A. clarkii*. The right side of the violin plot represents the f_{dM} distribution for the following topology: P1, *A. chrysopterus* (Fiji); P2, *A. chrysopterus* (Kimbe); P3, *A. sandaracinos* (Kimbe); Outgroup, *A. clarkii*. Dots correspond to the median f_{dM} value. Mean difference between pairwise f_{dM} values was significant for the background windows with a negligible effect size (*P*-value = 7.7e−78; $d = -0.16$) and for the *A. chrysopterus* divergent windows with a moderate effect size (*P*-value = 1.7e−35; $d = -0.60$).

flow of *A. chrysopterus* genome into *A. sandaracinos* through the hybrids. The Kimbe Bay hybrid zone might thus act as a conduit for transferring potentially adaptive alleles from one parental species to the other.

Hybrid Zone Population Structure

Based on whole-genome analysis of the Kimbe Bay hybrid zone, we found that the two parental species were mainly co-occurring with F1-hybrids and early-generation *A. sandaracinos*-BCs (Figs. 1 and 2). Such hybrid composition does not match the typical hybrid zone categorizations suggested by Jiggins and Mallet (2000), which aims at evaluating the level of reproductive isolation among the parental lineages. This characterization typically relies on the relative composition of hybrids and parents. Bimodal hybrid zones mainly constituted by the parental forms involve strong reproductive isolation between the parents, whereas unimodal hybrid zones—consisting in a mix between various hybrid generations and the parents (i.e. hybrid swarm)—are caused by weak reproductive barriers (Jiggins and Mallet 2000). Similar to pufferfishes (Takahashi et al. 2017) and North Atlantic eels hybrid zones (Pujolar et al. 2014), the Kimbe Bay hybrid zone coincides more with a trimodal hybrid zone; parents and early-generation hybrids coexist, even though the presence of early-generation *A. sandaracinos*-BCs is unexpected under this model. This asymmetry in introgression was previously attributed to various factors, including differences in generation time (Barton 1986), mating behavior (Lamb and Avise 1986), fitness (Ostberg et al. 2013), or relative abundances of parental species (Lepais et al. 2009). In the Kimbe Bay hybrid zone, the abundance and generation time of parental species are comparable (Gainsford et al. 2020) and are unlikely to generate the observed pattern. However, as Gainsford et al. (2020) suggested, this could be a consequence of the size-based hierarchy of the clownfish. Indeed, similar to numerous fish species, reproductive success is biased toward dominant individuals, which are usually bigger and more aggressive, and thus benefit from a priority for reproduction and other limiting resources (Ang and Manica 2010). As a result, the intermediary-sized hybrids might

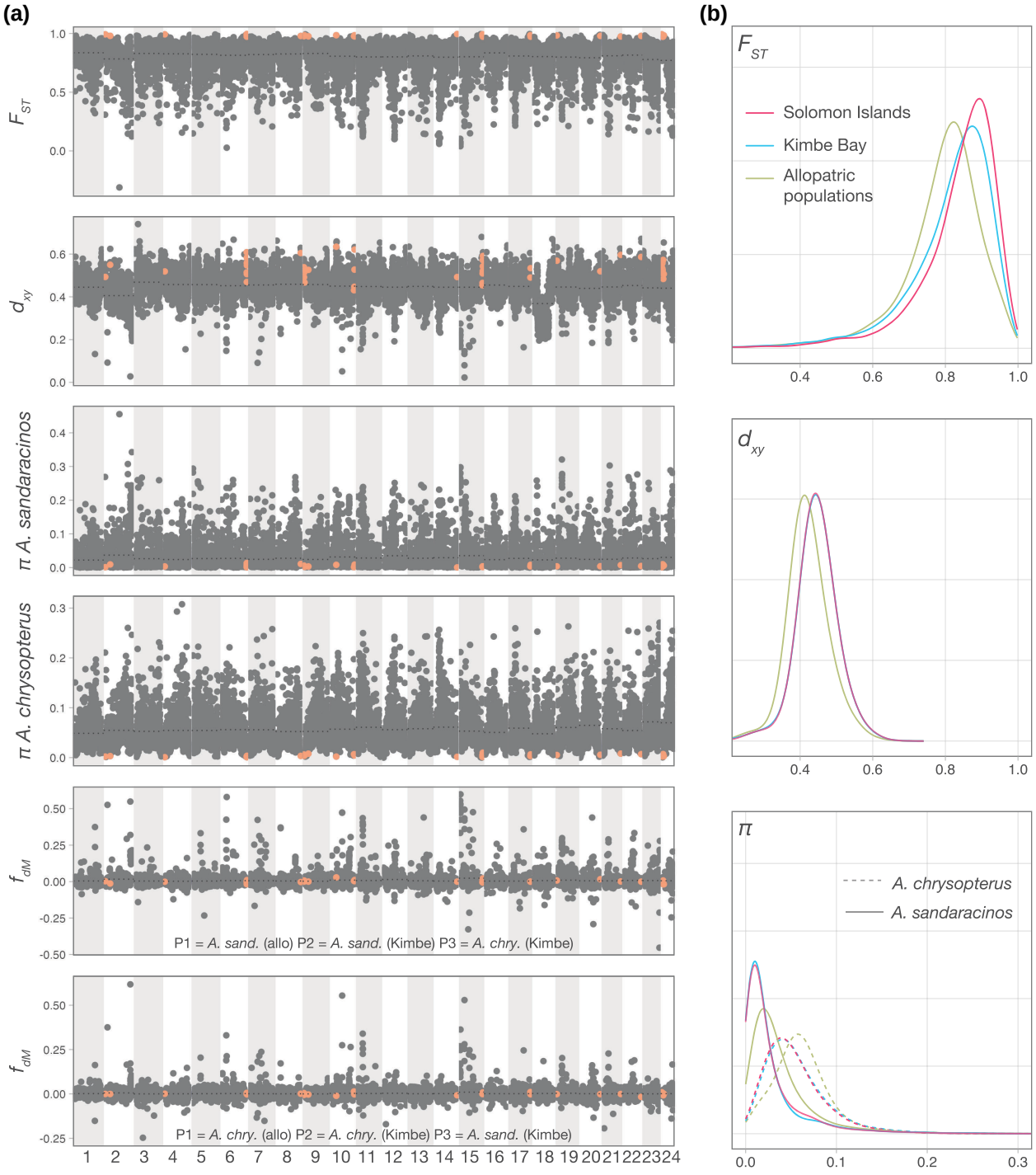

Fig. 5. Genomic landscape of parental species in the hybrid zone and compared to other populations. a) F_{ST}, d_{xy}, π, and f_{dM} in 50-kb windows across the genome of *A. sandaracinos* and *A. chrysopterus* in Kimbe Bay. For the two f_{dM} plots, the topology used for the calculation is indicated at the bottom of the plot. In both cases, the outgroup was *A. clarkii*. Color points correspond to potential barrier loci identified using outlier F_{ST} values among the three parental populations. Mean values for each chromosome are indicated with a dotted line. b) Density plot of F_{ST}, d_{xy}, and π for Kimbe Bay, Solomon Islands and for two populations (Fiji and Christmas Island) where the two parental species are in allopatry. F_{ST} and d_{xy} were computed between *A. sandaracinos* and *A. chrysopterus* occurring in Kimbe Bay, between *A. sandaracinos* and *A. chrysopterus* occurring in Solomon Islands and between *A. sandaracinos* from Christmas Island and *A. chrysopterus* from Fiji to compare populations where the two species are in allopatry. *A. chry.*, *A. chrysopterus*; *A. sand.*, *A. sandaracinos*.

choose to reproduce with the smaller-sized *A. sandaracinos*, since the hybrids are usually ignored by the bigger *A. chrysopterus*, which favors mating with conspecific individuals (Gainsford et al. 2020). Such mechanisms are further supported by the observation that heterospecific assemblages of *A. sandaracinos*-hybrids were more common in Kimbe Bay compared to *A. chrysopterus*-hybrid or only-hybrid groups (Gainsford et al. 2020).

We also found that the hybrid zone was characterized by a paucity of late-generation hybrids (i.e. *A. leucokranos* × *A. leucokranos*). Indeed, the frequency of F1-hybrids and the prevalence of backcrossing with *A. sandaracinos* confirm that the hybrids are fertile and viable. In addition, the mosaic structure of the Kimbe Bay hybrid zone—which displays various combinations of sea anemone host species, depth, and substrate—should be well suited for the formation of a hybrid swarm, which is defined as a mix between early and late hybrid generations (Grant 1981; Allendorf et al. 2001). The lack of late-generation hybrids might therefore be a consequence of their reduced fitness, as previously observed in numerous studies (e.g. Orr 1995; Christie and Strauss 2018). Different mechanisms could explain this pattern. One prominent explanation is the selection against hybrids due to Dobzhansky–Muller incompatibilities (DMIs; Orr 1995). Indeed, DMIs are increasingly revealed after the first F1-hybrid generation due to the combined effect of recombination and independent assortment, which breaks down gene complexes that coevolved over thousands of generations (Edmands and Burton 1999). This process—known as "hybrid breakdown"—usually has a more severe impact on F2 hybrids compared to F1-hybrids and was highlighted in various hybridizing fish species (e.g. Schumer et al. 2014; Carlon et al. 2021). In addition, the intermediate phenotypes of hybrids can be suboptimal in an environment where both parental species display phenotypic optimum for their specific niche (Thompson 2020). In the clownfish-specific case, this can be translated into the availability of host sea anemone, which is a limiting resource (Litsios et al. 2014). The hybrids do not display transgressive phenotypes enabling them to colonize distinct host anemone species compared to the parental species (Gainsford et al. 2015). Thus, they rarely form conspecific assemblages and predominantly share the host sea anemone with one of the parental species (Gainsford et al. 2020), which might prevent them from mating with other hybrids.

The phenotypically identified hybrids were mainly F1s or first-generation BCs with *A. sandaracinos*. However, some individuals identified as *A. sandaracinos* based on their phenotype displayed a substantial amount of *A. chrysopterus* variants (Fig. 2a) and might consist of descendants of second-generation *A. sandaracinos* BCs. Thus, only a couple of generations of backcrossing appear to erase *A. chrysopterus* phenotypic characteristics, resulting in BC individuals identical to the pure parental species, a pattern described in a sparrow hybrid zone (Walsh et al. 2015). If such a mechanism is at play, it might have biased our sampling, which relies on the individuals' phenotype. Indeed, by selecting putative hybrids based on their phenotype, we probably missed BC individuals which are phenotypically totally similar to *A. sandaracinos*, thus biasing our sampling toward F1-hybrids. Therefore, analysis of additional *A. sandaracinos* coming from the Kimbe Bay hybrid zone would be required to evaluate the true extent of backcrossing.

Hybrid Zone Genomic Architecture and Potential Outcomes

We detected considerable variation in admixed regions across the genome of BC individuals (Fig. 2b). Indeed, some regions were consistently inherited from *A. sandaracinos* (notably in chromosomes 1 and 3), whereas other chromosome segments displayed a total lack of *A. sandaracinos* components (such as chromosome 8). Such results are expected under the semipermeable view of species boundaries, which translate into differential admixture patterns across the genome (Harrison and Larson 2014). However, the admixed genomic windows did not present dissimilar differentiation, divergence, and nucleotide diversity values compared to the background windows (Table 2), which might have hinted at the potential selection forces at play. Such a lack of difference might result from the length of the admixed genomic tracts in the BCs. Indeed, the BC individuals in our dataset correspond to early BC generations and thus exhibit large tracts of DNA inherited by the parental species. The low number of generations implies that large admixed tracts have not yet been broken down by recombination into smaller genomic regions, thus preventing the detection of loci that may have been selected and fixed (Navascués et al. 2021). Despite this lack of resolution to detect selection signals and the limited number of BC samples, the apparent absence of *A. sandaracinos* components across all five individuals in chromosome 8 is intriguing. Loci under divergent selection (and potentially linked regions) associated with population or species divergence usually present particularly reduced admixture rates (Barton and Hewitt 1985; Harrison and Larson 2014). Interestingly *casz1* and various *hoxc* genes are located on this chromosome and were strongly associated with color patterns in the hamlet reef fishes (*Hypoplecturus* spp; Hench et al. 2022). This chromosome might thus have an essential role in determining the hybrid color pattern. However, such a conclusion is highly speculative, and further investigations are required to link the absence of admixture and the presence of coloration genes.

Although we did not detect any evolutionary relevant signal in the admixed tracts of BC individuals, features

linked with the genomic landscape have likely contributed to the pattern of introgression in clownfish species. Indeed, an increasing amount of studies demonstrated that introgression rate is variable across the genome of many species (e.g. Janoušek et al. 2015; Hagberg et al. 2022), including fish (e.g. Schaefer et al. 2016). The well-known case of sex chromosomes subject to reduced introgression (Burton and Barreto 2012) is not particularly relevant for species displaying a socially controlled sex change such as clownfish. Also, being sequential hermaphrodites, hybrid clownfish do not represent a case to evaluate Haldane's Rule, which posits that in the cases of hybridization for which one sex is absent, rare, or sterile in the offspring, this sex is the heterogametic sex (Haldane 1922). Although characterization of differential loss of male versus female function has been studied in hybrid hermaphrodites (Schilthuizen et al. 2011), this type of analysis remains beyond the scope of our study.

Furthermore, some genomic patterns common to multiple independent hybridization events start to emerge. Notably, genomic regions exhibiting a reduced density of coding genes or conserved elements experience an increased level of introgression (Brandvain et al. 2014; Martin et al. 2019). In addition, introgression is favored in the genomic portions subject to an elevated recombination rate (Sankararaman et al. 2014; Janoušek et al. 2015). This correlation results from recombination breaking down the association between neutral or adaptive alleles and deleterious alleles, eventually decreasing the strength of selection against the introgressed tracts (Martin et al. 2019; Veller et al. 2023). Genomic architecture can thus impact the level of introgression and partially predict the outcome of hybridization events (Brandvain et al. 2014; Martin et al. 2019). Although we are missing crucial information such as the recombination rate across the genome to predict the Kimbe Bay hybrid zone outcomes accurately, we can still formulate some hypotheses based on the observed patterns of introgression and population structure of the hybrid zone.

The coexistence of early-generation *A. sandaracinos*-BCs with *A. sandaracinos* individuals exhibiting low levels of admixture (Figs. 1c and 2a) suggests that hybrid genome stabilization could occur through multiple BC generations and subsequent reduction of the minor parent genome (*A. chrysopterus*). The resulting stabilized hybrid genome would consist of a few potentially adaptive tracts of *A. chrysopterus* embedded in an *A. sandaracinos* background (Runemark et al. 2019), similar to the stabilized hybrid genomes of the *Anopheles* mosquito (Hanemaaijer et al. 2018) or the monkeyflowers (Brandvain et al. 2014). This hypothetical outcome is further supported by the few peaks of introgression we detected between the *A. sandaracinos* and *A. chrysopterus* in the hybrid zone, suggesting that past genetic exchanges occurred between the parental species (Fig. 5). However, the time to reach genome stabilization is difficult to predict since high variability exists among species. For instance, fixation of ancestry blocks was swift in experimental *Helianthus* sunflower hybrids (Rieseberg et al. 1996) and introgressed Neanderthal regions became fixed in the human genome ca. 2,000 generations following hybridization (Sankararaman et al. 2014).

Another hybrid zone outcome highlighted in various species is hybrid speciation (Mavárez et al. 2006). Homoploid hybrid speciation requires reproductive isolation of the hybrid against both parental species, which can arise through reproductive barriers acting before or after fertilization (Abbott et al. 2013). Assortative mating between hybrids (Lamichhaney et al. 2018), genomic structural disparities (Rieseberg 2001), sorting of parental incompatibilities (Hermansen et al. 2014), or adaptation to novel ecological niches through transgressive phenotypes (Schwarzbach et al. 2001) are among the multiple barriers that could lead to homoploid hybrid speciation. However, despite one putative F2 hybrid in our data set, we have found no evidence of ongoing mating between hybrids, suggesting that homoploid hybrid speciation is an unlikely scenario. The Kimbe Bay hybrid zone might thus promote the exchange of adaptive alleles among the parental species through the hybrid individuals.

Parental Species: Ongoing Genetic Exchanges, Impact and Comparison with Other Populations

We detected high differentiation and divergence throughout the genome of the parental species *A. sandaracinos* and *A. chrysopterus* (Fig. 5). Such divergence was expected since the two species shared a common ancestor ca. 6.2 million years ago, a considerable evolutionary time considering that the clownfish radiation started ca. 10 million years ago (Gaboriau et al. 2024). Nevertheless, the occurrence of first-generation and early BC hybrids in the Kimbe Bay hybrid zone (Figs. 1c and 2a) indicates ongoing genetic exchanges between the two parental species and confirms that a high level of genetic divergence is a poor predictor of the ability to hybridize (Jiggins and Mallet 2000). The ongoing genetic exchange between the two parental species is not surprising since clownfishes are known for recurrent past hybridization (Schmid et al. 2022; Marcionetti and Salamin 2023) and frequently hybridize when raised in aquaria (Fautin and Allen 1997). Furthermore, the shared use of resources—here, the host sea anemone species shared by hybridizing individuals—might also favor hybridization, as observed for other reef fish species (Montanari et al. 2016). Reproductive isolation is thus not complete yet, raising the question of the barriers enabling the two parental species to maintain their integrity in the face of genetic exchange. Conflict in the genomes of

the parental species (i.e. genetic incompatibilities) appears to maintain species boundaries in a swordfish hybrid zone despite frequent interbreeding and overlapping environments (Schumer et al. 2014). In addition, mating behavior and reduced fitness of second-generation hybrids due to the accumulation of DMIs were suggested to preserve intact species boundaries in parrotfishes (Carlon et al. 2021). Beyond these examples in fishes, other species exhibit various prezygotic (Sobel et al. 2019) and postzygotic barriers (Pulido-Santacruz et al. 2018) which maintain parental species integrity if sufficiently strong. However, the genetic exchange might be able to dissolve those barriers provided the latter are weak enough (Xiong and Mallet 2022).

In the Kimbe Bay hybrid zone, genetic exchange seems to impact the integrity of one of the parental species. Indeed, the difference in pairwise F_{ST} among F1-hybrids and both parental lineages exhibit an unexpected biased distribution, with most genomic windows revealing a lower differentiation between the F1-hybrids and *A. sandaracinos* compared to *A. chrysopterus* (Fig. 4a). Since the parental species are highly differentiated (Fig. 5), a large proportion of SNPs should be alternatively fixed in each lineage (i.e. homozygous for the reference allele in one parent and homozygous for the alternative allele in the other). The remaining SNPs are thus expected to be heterozygous in some individuals. They should be randomly distributed across the genomes of the parental species, generating an expected distribution of F_{ST} differences between the F1-hybrid and both parents centered around zero. The unexpected pattern we observed is most likely a consequence of the genomic features of the parental species rather than a process impacting F1-hybrids directly since the latter should be shielded against intrinsic incompatibilities because hybrid breakdown usually emerges in the second generation (Burton 1990). A potential explanation could be that genetic exchange among the parental species alters the *A. sandaracinos* genomic integrity. The recurrent unidirectional genetic exchange due to backcrossing might impact *A. sandaracinos* population allelic frequencies, which might start to be more similar to the ones from *A. chrysopterus*. Consequently, F1-hybrids—which are expected to be totally heterozygous—would display a reduced whole-genome differentiation against *A. sandaracinos* compared to the differentiation against *A. chrysopterus*, leading to the observed bias. As a consequence of the unilateral introgression via hybrids due to size dimorphism in parental species, we might thus witness the diffusion of *A. chrysopterus* variants into *A. sandaracinos* in the Kimbe Bay hybrid zone.

Furthermore, the significant difference in f_{dM} values calculated based on different topologies—particularly when examining outlier windows biased toward *A. chrysopterus*—suggest a scenario of asymmetric introgression from *A. chrysopterus* into *A. sandaracinos* (Fig. 4b). This pattern likely arises as a consequence of previous genetic exchange between the parental species, resulting in the fixation of introgressed alleles from *A. chrysopterus* within the *A. sandaracinos* genome. The shaping of parental species genomes through past and ongoing genetic exchange has also been documented across numerous fish hybrid zones (e.g. Takahashi et al. 2017; Barth et al. 2020).

We investigated other impacts on the parental species based on the comparison with other *A. sandaracinos* and *A. chrysopterus* populations either occurring in sympatry (Solomon Islands) or allopatry (respectively, Christmas Island and Fiji; Fig. 5b). Nucleotide diversity varied among populations and was higher in both allopatric populations, and its level was generally lower in *A. sandaracinos* compared to *A. chrysopterus*, consistent with what was previously highlighted using microsatellite markers (Gainsford et al. 2015). Patterns of F_{ST} were highly similar among the three pairwise comparisons, but genetic divergence and differentiation were unexpectedly lower when comparing more geographically distant parental populations. Due to the clownfishes' propensity to hybridize (Schmid et al. 2022; Marcionetti and Salamin 2023), we cannot exclude the possibility that past hybridization events involving other clownfish species occurred in the allopatric *A. chrysopterus* and *A. sandaracinos* parental populations. The highest nucleotide diversity in the allopatric populations, as well as the PC2 values for the two populations, also point toward this conclusion. Parallels can be drawn with the ancestry patterns previously observed in the skunk clownfish complex—whose evolutionary history was marked by ancestral genetic exchange (Marcionetti et al. 2024). A sister species to *A. sandaracinos*—*Amphiprion perideraion*—occurs in both allopatric locations and has a history of past hybridization with other clownfish species, notably with *A. sandaracinos* (Marcionetti et al. 2024). We thus carefully suggest that the reduced divergence between the allopatric parental populations might be a consequence of past genetic exchange with the clownfish *A. perideraion* and is not linked to the dynamics of the Kimbe Bay hybrid zone. The subsequent split in the admixture plots of the Fiji population from the other *A. chrysopterus* populations (at $K = 4$) followed by the separation of the Christmas Island *A. sandaracinos* from other conspecific populations (at $K = 5$; supplementary fig. S2, Supplementary Material online) further supports that distinct evolutionary processes might have shaped the genome of the two parental allopatric populations compared to the sympatric ones.

Future Directions

The recurrent backcrossing of the F1-hybrids with *A. sandaracinos* combined with the scarcity of second-generation hybrids and the genomic signature of past genetic exchange among the parental species suggest that the

hybridization events occurring in the Kimbe Bay hybrid zone might lead to adaptive introgression. The resulting stabilized hybrid genome would thus consist of a few potentially *A. chrysopterus* adaptive loci introgressing into the *A. sandaracinos* genomic background. However, we expect distinct outcomes at other locations of the *A. chrysopterus–A. sandaracinos* hybrid zone—which ranges from the Solomon Islands to the Halmahera Island. Indeed, BCs were detected in Kimbe Bay but not in other locations (i.e. Solomon Islands and Kavieng), which consisted mainly of equally admixed individuals (Gainsford et al. 2020). Dissimilar hybridization outcomes also occurred in the rainbow trout hybrid zone, where some locations only consisted of advanced BC hybrids, while others were constituted of a mix between F1-hybrids and BCs (Mandeville et al. 2019). Geographical heterogeneity in hybridization outcomes was also highlighted in Catostomus suckers (Mandeville et al. 2017) or in the hybrid zone between the rainbow and the westslope trout (Muhlfeld et al. 2017). This spatial variation of introgression patterns and hybrid frequency is a consequence of the variability among locations in multiple processes such as reproductive isolation (Mandeville et al. 2017), context-specific selection, or drift (Ottenburghs et al. 2017). We thus expect a high variability in the hybridization outcomes within the *A. chrysopterus–A. sandaracinos* hybrid zone, which we might be able to address with the genomic analysis of more locations leading to a deeper understanding of the genomic architecture of *A. leucokranos* and ultimately, of the mechanisms linked with clownfish speciation and reproductive isolation.

Materials and Methods

Sampling, Library Preparation, and Sequencing

Between 2011 and 2014, we collected tissue samples consisting of 1-cm-long dorsal fin clips and took pictures from 23 *A. chrysopterus*, 24 *A. leucokranos*, and 20 *A. sandaracinos*. All *A. leucokranos* (hybrid species) came from the hybrid zone in Kimbe Bay (Papua New Guinea). For the parental species, we collected *A. chrysopterus* and *A. sandaracinos* in the hybrid zone and in allopatric populations (Fig. 1a and supplementary table S1, Supplementary Material online). We used an outgroup, the raw DNA reads of ten *A. clarkii* collected in Indonesia sequenced in a previous study (Schmid et al. 2024). For the newly collected samples, we extracted the genomic DNA of each fin clip following the DNeasy Blood and Tissue kit standard procedure and performed the final elution twice in 100 µl of Buffer AE (QIAGEN, Hombrechtikon, Switzerland). We quantified the extracted DNA using a Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) and evaluated the integrity by electrophoresis. We followed the TruSeq Nano DNA library prep standard protocol to generate libraries with a 350-bp insert-size for whole-genome paired-end sequencing (Illumina, San Diego, CA, USA). We validated the fragment length distribution of the libraries with a fragment analyzer (Agilent Technologies, Santa Clara, CA, USA). We pooled all samples together, and the Genomic Technologies Facility of the University of Lausanne performed the sequencing on ten lanes of Illumina 4000 HiSeq to reach a ~10× coverage for each sample.

Sequenced Data Processing, Mapping and Genotyping

We trimmed the generated reads for adapter sequences, filtered reads shorter than 36 bp, and cut nucleotides off the start and end of a read if the quality was below three using Trimmomatic V0.392.3 (Bolger et al. 2014). We assessed read quality before and after processing with FastQC V0.12.1 (Andrews 2010). We mapped the processed reads to *A. clarkii* reference genome (GenBank accession number JALBFV000000000.1; Moore et al. 2023) using BWA-MEM V0.7.17 (Li et al. 2009). Divergence times of *A. sandaracinos* and *A. chrysopterus* from *A. clarkii* are ~5.1 and ~6.2 million years, respectively (Gaboriau et al. 2024). We subsequently sorted, indexed, and filtered them according to mapping quality (>30) using SAMtools V1.17 (Li et al. 2009). Then, we assigned all the reads to read-groups using Picard Tools V2.9.0 (http://broadinstitute.github.io/picard), fixed paired reads, removed secondary alignments and sorted the SAM file with SAMtools V1.17, marked duplicates with Picard Tools V2.9.0, and validated the mapping output using various statistics generated with BamTools V2.5.2 (Barnett et al. 2011).

After mapping, we called haplotypes for each sample and each chromosome using GATK V4.4.0.0 (Van der Auwera and O'Connor 2020), merged the generated gVCF files with Picard Tools V2.9.0 and proceeded to joint genotyping with GATK V4.4.0.0. At this step, we removed three individuals with low-quality data (LU11, LU61, and LU24). Then, we generated two different VCF files before subsequent filtering: (i) one with the two parental species (*A. chrysopterus* and *A. sandaracinos*) and the hybrid species *A. leucokranos*, and (ii) one including also the outgroup species *A. clarkii* in addition of the hybrid and parental species. Then, we proceeded to hard-filtering the two VCF files following the GATK recommendations, since no well-curated training resources are available for base recalibration (QD < 2.0, MQ < 40.0, FS > 60.0, SOR > 3.0, MQRankSum < −12.5, ReadPosRankSum < −8.0). We performed additional filtering using VCFtools V0.1.16 (Danecek et al. 2011) and kept only biallelic SNPs with a quality above 30 (–minQ 30), which were informative for at least 85% of the individuals (–max-missing 0.85), with a minimum depth of 7 (–minDP 7), a maximum depth of

40 (–maxDP 40) and a minor allele count greater than or equal to 3 (–mac 3).

PCA and Admixture

To select only independent SNPs for the PCA, we performed linkage disequilibrium (LD) pruning using PLINK V1.9 (Purcell et al. 2007). We removed variants exhibiting high LD using a sliding window approach with the following parameters: A window size of 50 kb, a step size of 10 bp, and an LD threshold of $r^2 = 0.2$. For each window, pairs of variants with $r^2 = 0.2$ values exceeding the threshold were identified, and one variant from each pair was excluded. Then, using the pruned dataset, we performed a PCA with PLINK V1.9 (Purcell et al. 2007). Then, we estimated individual admixture proportions using ADMIXTURE V1.3.0 (Alexander et al. 2009) with *K*-values ranging between 2 and 7. The *K*-value that balances minimizing cross-validation error and parsimony was chosen for the final analysis. To further investigate the relationship between the hybrids, we performed a second PCA analysis, including only the *A. leucokranos* individuals.

Mitochondrial Genome Reconstruction and Phylogeny

To explore the genetic relationship between individuals and determine the parental inheritance of mitochondrial DNA in hybrids, we first reconstructed the mitochondrial genome of the 64 individuals (hybrids and parental species) using the baiting and iterative mapping approach implemented in MITObim V1.9.1 (Hahn et al. 2013). We first merged the paired reads using FLASh V1.2.11 (Magoč and Salzberg 2011). Then, we used the complete mitochondrial genome of *Amphiprion percula* (GenBank accession: KJ174497.1; Tao et al. 2016) as reference to reconstruct the mitochondrial genomes of the 64 samples. We circularized the 64 resulting genomes with MARS (Ayad and Pissis 2017) and aligned them using MAFFT V7.505 (Katoh and Standley 2013). We then used the multiple sequences alignment as input to build a maximum-likelihood tree with IQ-tree V2.2.2 (Minh et al. 2020) and selected the substitution model (TN + F + I + I + R3) that minimizes the Bayesian information criterion with ModelFinder (Kalyaanamoorthy et al. 2017). We assessed branch support by implementing the single-branch test (Shimodaira-Hasegawa-like approximate likelihood ratio test [SH-like aLRT]; Guindon et al. 2010) and the ultrafast bootstrap approximation (UFBoot; Hoang et al. 2018) with 1,000 replicates and eventually visualize the resulting phylogeny with iTOL V7 (Letunic and Bork 2021).

Local Ancestry Inference

We used the two-layer hidden Markov model implemented in ELAI (Guan 2014) to infer the local ancestry of the hybrid individuals. This approach estimates linkage disequilibrium between predefined parental groups to calculate the allele dosage scores at each SNP. The values range from 0 to 2 in a two-way admixture and indicate the site's most likely ancestry proportion (0 or 2 indicates homozygous states, whereas a value of 1 corresponds to a heterozygous state). We ran ELAI on the complete SNPs dataset and considered the Kimbe Bay populations of *A. chrysopterus* and *A. sandaracinos* as the two parental groups. We set the number of upper-layer clusters (-C) to 2 (representing *A. sandaracinos* and *A. chrysopterus*) and the lower-layer clusters (-c) to 10 (five times the upper-layer cluster, following ELAI recommendations). We applied six different admixture generation (-mg) parameters (1, 2, 3, 5, 10, and 20) and performed ELAI analysis five times with 30 expectation maximization runs (-s) for each generation. We averaged the 30 independent runs and considered the sites with allele dosage between 0.8 and 1.2 as heterozygous, sites with allele dosage above 1.8 as homozygous for *A. sandaracinos* ancestry and sites below 0.2 as homozygous for *A. chrysopterus* ancestry, similar to what Seixas et al. (2018) suggested. We regarded values in-between those ranges as noninformative due to the high level of uncertainty. To summarize the information, we calculated the percentage of homozygous and heterozygous sites for each hybrid individual.

Hybrid Color Patterns Matrix and PCA

The parental species differ in their phenotype in terms of white bands and background color patterns: *A. chrysopterus* displays two white lateral bands and dark-orange background coloration, whereas *A. sandaracinos* has a single dorsal white stripe on a light-orange background (Fig. 1c). The *A. leucokranos* individuals exhibit various intermediate traits between the two parental species. We built a color pattern matrix consisting in nine binary traits (either absent or present in the hybrid): Four traits describing the extent of the dorsal white band (S1–S4), four traits depicting the extent of the lateral white band (C1–C4), and one trait describing the asymmetry of pattern between the right and left side (A; Fig. 3c). We thus characterized the 20 hybrid individuals with available pictures (no pictures were available for individuals LU131, LU132, and LU07) according to the presence or absence of those nine traits. Finally, we performed a PCA on the resulting matrix to summarize the color information in a 2D space and explore the relationship between the hybrids.

Population Genomic Analyses in Sliding Windows

Based on the admixture analysis and the local ancestry estimation, we identified four groups of *A. leucokranos*: 14 first-generation hybrids (F1s), one second-generation hybrid (F2), five *A. sandaracinos*-BCs, and one first-generation *A. sandaracinos*-BC × *A. chrysopterus* cross

(see supplementary table S1, Supplementary Material online). We only considered pure F1s (14 individuals) and *A. sandaracinos*-BCs (five individuals) for subsequent analyses. To investigate the genome-wide patterns of differentiation, divergence, and diversity among the hybrids and the two parental populations, we estimated between-population differentiation (F_{ST}), between-population absolute divergence (d_{xy}), as well as population nucleotide diversity (π) in nonoverlapping 50-kb windows (containing at least 100 SNPs) using popgenWindows.py (https://github.com/simonhmartin/genomics_general). We considered as populations the eight following groups: (i) F1s, (ii) *A. sandaracinos*-BCs, (iii to v) *A. sandaracinos* from Solomon Islands, Kimbe Bay and Christmas Island, and (vi to viii) *A. chrysopterus* from Solomon Islands, Kimbe Bay, and Fiji (supplementary table S1, Supplementary Material online). Finally, we quantified the extent of genetic exchanges between the two parental lineages in the hybrid zone (Kimbe Bay) and calculated the f_{dM} statistic (Malinsky et al. 2015) in nonoverlapping 50-kb windows (containing at least 100 SNPs) using ABBABABAwindows.py. Like other *f*-statistics (also known as ABBA–BABA statistics), the f_{dM} statistic relies on patterns of ancestral and derived alleles in a group of species/populations of interest (P1, P2, and P3) and an outgroup to discern hybridization from incomplete lineage sorting. Compared to other similar statistics, f_{dM} has the additional advantage of accommodating genetic exchanges between both P2 and P3 (positive value) and P1 and P3 (negative values). Since we were interested in genetic exchanges between parental species in Kimbe Bay, we considered the two following topologies to calculate two f_{dM} values: (i) P1 = *A. sandaracinos* (Christmas Island), P2 = *A. sandaracinos* (Kimbe Bay), P3 = *A. chrysopterus* (Kimbe Bay), and outgroup = *A. clarkii*; and (ii) P1 = *A. chrysopterus* (Fiji), P2 = *A. chrysopterus* (Kimbe Bay), P3 = *A. sandaracinos* (Christmas Island), and outgroup = *A. clarkii*. We subsequently used those statistics to characterize the genomic landscapes of the F1-hybrids, the *sandaracinos*-BC hybrids, and the parental lineages based on the following analyses.

Identification of Potential Barrier Loci in Parental Populations

To identify potential barrier loci among the parental species *A. chrysopterus* and *A. sandaracinos*, we used the pairwise F_{ST} between the two species in sympatry (the Solomon Islands and Kimbe Bay) and in allopatry (*A. chrysopterus* from Fiji and *A. sandaracinos* from Christmas Island). Regions of the genome with elevated F_{ST} can point to fixed or nearly fixed allele frequency differences between populations (i.e. barrier loci/genomic islands of differentiation). We thus extracted the 50-kb genomic windows with F_{ST} values above the 99th percentile of the whole-genome distribution and in common between three pairwise comparisons. We subsequently considered those windows as potential barrier loci.

Identification of Regions of Elevated Divergence Between F1-hybrids and One of the Parental Species

We used F_{ST} disparities between lineages to identify genomic regions of inflated divergence between the F1-hybrids and either or both parental species (also referred in results and discussion as "outlier windows"). Thus, for each 50-kb window across the genome, we computed the difference in F_{ST} between the F1-hybrids and both parental species from Kimbe Bay (F1-parent divergence = CHP-F1 F_{ST}—SAN-F1 F_{ST}), and subsequently selected the F_{ST}-difference values above the 99th percentile (biased toward *A. chrysopterus*) and below the first percentile (biased toward *A. sandaracinos*). To detect if the number of highly divergent windows was biased toward some chromosomes, we generated a random distribution of highly divergent windows across the genome using 10,000 permutations. We then compared the observed mean value of divergent windows per chromosome to the expected mean value. We corrected the *P*-value for multiple testing using the Benjamini–Hochberg false-discovery rate method (Benjamini and Hochberg 1995). We performed a two-sample Wilcoxon test to evaluate differences in F_{ST}, d_{xy}, and f_{dM} values between the background genomic windows and the highly divergent windows. We performed the same test to determine if the two computed f_{dM} values averaged across the highly divergent windows were statistically different, which would indicate biased introgression from one parental species to the other. Since the statistical significance indicated by the *P*-value is highly influenced by the sample size and can be misleading when many observations are compared, we computed the effect size using Cohen's *d* statistic to assess the practical significance of the difference between the two groups. We also performed a χ^2 test to compare the number of barrier loci between background and divergent genomic windows. Finally, we performed a one-sample Wilcoxon rank-sum test to test whether the mean divergence value across chromosomes was statistically different from 0.

Characterization of Homozygous A. sandaracinos *Tracts in BCs*

Similar to the analyses we made for the divergent windows between F1-hybrids and the parental species, we used the population genomic statistics to characterize the homozygous *A. sandaracinos* genomic tracts identified with ELAI in the *A. sandaracinos*-BCs. We considered the windows with at least four out of the five *A. sandaracinos*-BCs displaying homozygous *A. sandaracinos* ancestry to be BC windows. To test for a bias in some chromosomes

compared to the rest of the genome, we generated a random distribution of BC windows across the genome using 10,000 permutations. We then compared the observed mean value of BC windows per chromosome to the expected mean value. We corrected the *P*-value for multiple testing using the Benjamini–Hochberg false-discovery rate method (Benjamini and Hochberg 1995). Then, we tested for differences in the F_{ST}, d_{xy}, and f_{dM} values between the background and BC genomic windows using a two-sample Wilcoxon test and computed the effect size using Cohen's *d* statistic. Finally, we performed a χ^2 test to compare the number of barrier loci between background and BC genomic windows.

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Acknowledgments

We thank the Mahonia Na Dari Research and Conservation Centre and Walindi Resort for supporting field work in Kimbe Bay. Collections were made in accordance with James Cook University Animal Ethics approval number A1633. We also would like to thank the Lausanne Genomic Technologies Facility for sequencing the samples and the DCSR infrastructure of the University of Lausanne for the computational resources. We thank Nicolas Rodrigues for the clownfish drawings. We are grateful to Anna Marcionetti, and the group of Nicolas Salamin in general, for useful feedback on the manuscript and insightful discussions.

Author Contributions

S.S., D.A.H., and N.S. planned the study and interpreted the results. A.G. and G.P.J. collected the samples. S.S. performed the lab work, analyzed the genomic data, and drafted the first version of the manuscript. W.-T.H. contributed to the genomic data analysis. D.A.H. analyzed the morphological data. S.S. and D.A.H. wrote the final version of the manuscript with contributions from all authors.

Funding

This work was supported by the Faculté de Biologie et de Médecine, Université de Lausanne, and the Swiss National Science Foundation (310030_185223).

Data Availability

Raw reads were deposited on NCBI (BioProject ID: PRJNA1078272, SRA submission SUB15001078). VCF files, scripts and images are deposited on Zenodo (https://doi.org/10.5281/zenodo.14644030). Analysis pipeline is also available on GitHub (https://github.com/phylolab/leucokranosHybrid).

Literature Cited

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R, et al. Hybridization and speciation. J Evol Biol. 2013:26(2):229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x.

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

Allendorf FW, Leary RF, Spruell P, Wenburg JK. The problems with hybrids: setting conservation guidelines. Trends Ecol Evol. 2001:16(11):613–622. https://doi.org/10.1016/S0169-5347(01)02290-X.

Andrews S. FastQC: a Quality Control Tool for High Throughput Sequence Data; 2010 [accessed 2020 Jul 6]. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Ang TZ, Manica A. Unavoidable limits on group size in a body size-based linear hierarchy. Behav Ecol. 2010:21(4):819–825. https://doi.org/10.1093/beheco/arq062.

Ayad LAK, Pissis SP. MARS: improving multiple circular sequence alignment using refined sequences. BMC Genom. 2017:18(1):86. https://doi.org/10.1186/s12864-016-3477-5.

Barnett DW, Garrison EK, Quinlan AR, Strömberg MP, Marth GT. BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics. 2011:27(12):1691–1692. https://doi.org/10.1093/bioinformatics/btr174.

Barth JMI, Gubili C, Matschiner M, Tørresen OK, Watanabe S, Egger B, Han Y-S, Feunteun E, Sommaruga R, Jehle R, et al. Stable species boundaries despite ten million years of hybridization in tropical eels. Nat Commun. 2020:11(1):1–13. https://doi.org/10.1038/s41467-020-15099-x.

Barton NH. The effects of linkage and density-dependent regulation on gene flow. Heredity (Edinb). 1986:57(3):415–426. https://doi.org/10.1038/hdy.1986.142.

Barton NH, de Cara MAR. The evolution of strong reproductive isolation. Evolution. 2009:63(5):1171–1190. https://doi.org/10.1111/j.1558-5646.2009.00622.x.

Barton NH, Hewitt GM. Analysis of hybrid zones. Annu Rev Ecol Syst. 1985:16(1):113–148. https://doi.org/10.1146/annurev.es.16.110185.000553.

Barton NH, Hewitt GM. Adaptation, speciation and hybrid zones. Nature. 1989:341(6242):497–503. https://doi.org/10.1038/341497a0.

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B. 1995:57(1):289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

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

Brandvain Y, Kenney AM, Flagel L, Coop G, Sweigart AL. Speciation and introgression between *Mimulus nasutus* and *Mimulus guttatus*. PLoS Genet. 2014:10(6):e1004410. https://doi.org/10.1371/journal.pgen.1004410.

Buerkle CA, Rieseberg LH. The rate of genome stabilization in homoploid hybrid species. Evolution. 2008:62(2):266–275. https://doi.org/10.1111/j.1558-5646.2007.00267.x.

Burton RS. Hybrid breakdown in physiological response: a mechanistic approach. Evolution. 1990:44(7):1806–1813. https://doi.org/10.2307/2409509.

Burton RS, Barreto FS. A disproportionate role for mtDNA in Dobzhansky–Muller incompatibilities? Mol Ecol. 2012:21(20): 4942–4957. https://doi.org/10.1111/mec.12006.

Buston P. Size and growth modification in clownfish. Nature. 2003:424(6945):145–146. https://doi.org/10.1038/424145a.

Carlon DB, Robertson DR, Barron RL, Choat JH, Anderson DJ, Schwartz SA, Sánchez-Ortiz CA. The origin of the parrotfish species *Scarus compressus* in the Tropical Eastern Pacific: region-wide hybridization between ancient species pairs. BMC Ecol Evol. 2021:21(1):7. https://doi.org/10.1186/s12862-020-01731-3.

Christie K, Strauss SY. Along the speciation continuum: quantifying intrinsic and extrinsic isolating barriers across five million years of evolutionary divergence in California jewelflowers. Evolution. 2018:72(5):1063–1079. https://doi.org/10.1111/evo.13477.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. The variant call format and VCFtools. Bioinformatics. 2011:27(15): 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

Edmands S, Burton RS. Cytochrome C oxidase activity in interpopulation hybrids of a marine copepod: a test for nuclear-nuclear or nuclear-cytoplasmic coadaptation. Evolution. 1999:53(6):1972–1978. https://doi.org/10.1111/j.1558-5646.1999.tb04578.x.

Elgvin TO, Trier CN, Tørresen OK, Hagen IJ, Lien S, Nederbragt AJ, Ravinet M, Jensen H, Sætre G-P. The genomic mosaicism of hybrid speciation. Sci Adv. 2017:3(6):e1602996. https://doi.org/10.1126/sciadv.1602996.

Fautin DG, Allen GR. Field guide to Anemone fishes and their host sea anemones. 2nd Revised ed. Perth (AU): Western Australian Museum; 1997.

Gaboriau T, Marcionetti A, Garcia Jimenez A, Schmid S, Fitzgerald LM, Micheli B, Titus B, Salamin N. Host-use drives convergent evolution in clownfish and disentangles the mystery of an iconic adaptive radiation. bioRxiv 602550. https://doi.org/10.1101/2024.07.08.602550, 11 July 2024, preprint: not peer reviewed.

Gainsford A, Jones GP, Hobbs JA, Heindler FM, Herwerden L. Species integrity, introgression, and genetic variation across a coral reef fish hybrid zone. Ecol Evol. 2020:10(21):11998–12014. https://doi.org/10.1002/ece3.6769.

Gainsford A, van Herwerden L, Jones GP. Hierarchical behaviour, habitat use and species size differences shape evolutionary outcomes of hybridization in a coral reef fish. J. Evol. Biol. 2015:28(1): 205–222. https://doi.org/10.1111/jeb.12557.

Gompert Z, Buerkle CA. What, if anything, are hybrids: enduring truths and challenges associated with population structure and gene flow. Evol Appl. 2016:9(7):909–923. https://doi.org/10.1111/eva.12380.

Grant V. Plant speciation. 2nd ed. New York (NY): Columbia University Press; 1981.

Guan Y. Detecting structure of haplotypes and local ancestry. Genetics. 2014:196(3):625–642. https://doi.org/10.1534/genetics.113.160697.

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010:59(3):307–321. https://doi.org/10.1093/sysbio/syq010.

Hagberg L, Celemín E, Irisarri I, Hawlitschek O, Bella JL, Mott T, Pereira RJ. Extensive introgression at late stages of species formation: insights from grasshopper hybrid zones. Mol Ecol. 2022:31(8): 2384–2399. https://doi.org/10.1111/mec.16406.

Hahn C, Bachmann L, Chevreux B. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads – a baiting and iterative mapping approach. Nucleic Acids Res. 2013:41(13):e129–e129. https://doi.org/10.1093/nar/gkt371.

Haldane JBS. Sex ratio and unisexual sterility in hybrid animals. J Genet. 1922:12(2):101–109. https://doi.org/10.1007/BF02983075.

Hanemaaijer MJ, Collier TC, Chang A, Shott CC, Houston PD, Schmidt H, Main BJ, Cornel AJ, Lee Y, Lanzaro GC. The fate of genes that cross species boundaries after a major hybridization event in a natural mosquito population. Mol Ecol. 2018:27(24):4978–4990. https://doi.org/10.1111/mec.14947.

Harrison RG, Larson EL. Hybridization, introgression, and the nature of species boundaries. J Hered. 2014:105:795–809. https://doi.org/10.1093/jhered/esu033.

He S, Planes S, Sinclair-Taylor TH, Berumen ML. Diagnostic nuclear markers for hybrid Nemos in Kimbe Bay, PNG-*Amphiprion chrysopterus* × *Amphiprion sandaracinos* hybrids. Mar Biodiv. 2019:49(3):1261–1269. https://doi.org/10.1007/s12526-018-0907-4.

Hench K, Helmkampf M, McMillan WO, Puebla O. Rapid radiation in a highly diverse marine environment. Proc Natl Acad Sci U S A. 2022:119(4):e2020457119. https://doi.org/10.1073/pnas.2020457119.

Hermansen JS, Haas F, Trier CN, Bailey RI, Nederbragt AJ, Marzal A, Sætre G-P. Hybrid speciation through sorting of parental incompatibilities in Italian sparrows. Mol Ecol. 2014:23(23): 5831–5842. https://doi.org/10.1111/mec.12910.

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018:35(2):518–522. https://doi.org/10.1093/molbev/msx281.

Janoušek V, Munclinger P, Wang L, Teeter KC, Tucker PK. Functional organization of the genome may shape the species boundary in the house mouse. Mol Biol Evol. 2015:32(5):1208–1220. https://doi.org/10.1093/molbev/msv011.

Jiggins CD, Mallet J. Bimodal hybrid zones and speciation. Trends Ecol. 2000:15(6):250–255. https://doi.org/10.1016/S0169-5347(00)01873-5.

Jiggins CD, Salazar C, Linares M, Mavarez J. Hybrid trait speciation and *Heliconius* butterflies. Philos Trans R Soc Lond B Biol Sci. 2008: 363(1506):3047–3054. https://doi.org/10.1098/rstb.2008.0065.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017:14(6):587–589. https://doi.org/10.1038/nmeth.4285.

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013:30(4):772–780. https://doi.org/10.1093/molbev/mst010.

Lamb T, Avise JC. Directional introgression of mitochondrial DNA in a hybrid population of tree frogs: the influence of mating behavior. Proc Natl Acad Sci U S A. 1986:83(8):2526–2530. https://doi.org/10.1073/pnas.83.8.2526.

Lamichhaney S, Han F, Webster MT, Andersson L, Grant BR, Grant PR. Rapid hybrid speciation in Darwin's finches. Science. 2018:359(6372):224–228. https://doi.org/10.1126/science.aao4593.

Lepais O, Petit RJ, Guichoux E, Lavabre JE, Alberto F, Kremer A, Gerber S. Species relative abundance and direction of introgression in oaks. Mol Ecol. 2009:18(10):2228–2242. https://doi.org/10.1111/j.1365-294X.2009.04137.x.

Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021:49(W1):W293–W296. https://doi.org/10.1093/nar/gkab301.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. Bioinformatics. 2009:25(16):2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

Litsios G, Kostikova A, Salamin N. Host specialist clownfishes are environmental niche generalists. Proc Biol Sci. 2014:281(1795): 20133220. https://doi.org/10.1098/rspb.2013.3220.
Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 2011:27(21): 2957–2963. https://doi.org/10.1093/bioinformatics/btr507.
Malinsky M, Challis RJ, Tyers AM, Schiffels S, Terai Y, Ngatunga BP, Miska EA, Durbin R, Genner MJ, Turner GF. Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. Science. 2015:350(6267):1493–1498. https://doi.org/10.1126/science.aac9927.
Mandeville EG, Parchman TL, Thompson KG, Compton RI, Gelwicks KR, Song SJ, Buerkle CA. Inconsistent reproductive isolation revealed by interactions between *Catostomus* fish species. Evol Lett. 2017:1(5):255–268. https://doi.org/10.1002/evl3.29.
Mandeville EG, Walters AW, Nordberg BJ, Higgins KH, Burckhardt JC, Wagner CE. Variable hybridization outcomes in trout are predicted by historical fish stocking and environmental context. Mol Ecol. 2019:28(16):3738–3755. https://doi.org/10.1111/mec.15175.
Marcionetti A, Rossier V, Roux N, Salis P, Laudet V, Salamin N. Insights into the genomics of clownfish adaptive radiation: genetic basis of the mutualism with sea anemones. Genome Biol Evol. 2019:11(3): 869–882. https://doi.org/10.1093/gbe/evz042.
Marcionetti A, Bertrand JAM, Cortesi F, Donati GFA, Heim S, Huyghe F, Kochzius M, Pellissier L, Salamin N. Recurrent gene flow events occurred during the diversification of clownfishes of the skunk complex. Mol Ecol. 2024:33(11):e17347. https://doi.org/10.1111/mec.17347.
Marcionetti A, Salamin N. Insights into the genomics of clownfish adaptive radiation: the genomic substrate of the diversification. Genome Biol Evol. 2023:15(7):evad088. https://doi.org/10.1093/gbe/evad088.
Martin SH, Davey JW, Salazar C, Jiggins CD. Recombination rate variation shapes barriers to introgression across butterfly genomes. PLoS Biol. 2019:17(2):e2006288. https://doi.org/10.1371/journal.pbio.2006288.
Mavárez J, Salazar CA, Bermingham E, Salcedo C, Jiggins CD, LinaresM. Speciation by hybridization in *Heliconius* butterflies. Nature. 2006:441(7095):868. https://doi.org/10.1038/nature04738.
Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020:37(5):1530–1534. https://doi.org/10.1093/molbev/msaa015.
Montanari SR, Hobbs J-PA, Pratchett MS, van Herwerden L. The importance of ecological and behavioural data in studies of hybridisation among marine fishes. Rev Fish Biol Fisheries. 2016:26(2): 181–198. https://doi.org/10.1007/s11160-016-9420-7.
Moore B, Herrera M, Gairin E, Li C, Miura S, Jolly J, Mercader M, Izumiyama M, Kawai E, Ravasi T, et al. The chromosome-scale genome assembly of the yellowtail clownfish *Amphiprion clarkii* provides insights into the melanic pigmentation of anemonefish. G3 (Bethesda). 2023:13(3):jkad002. https://doi.org/10.1093/g3journal/jkad002.
Muhlfeld CC, Kovach RP, Al-Chokhachy R, Amish SJ, Kershner JL, Leary RF, Lowe WH, Luikart G, Matson P, Schmetterling DA, et al. Legacy introductions and climatic variation explain spatiotemporal patterns of invasive hybridization in a native trout. Glob. Change Biol. 2017:23(11):4663–4674. https://doi.org/10.1111/gcb.13681.
Navascués M, Becheler A, Gay L, Ronfort J, Loridon K, Vitalis R. Power and limits of selection genome scans on temporal data from a selfing population. Peer Community J. 2021:1:e37. https://doi.org/10.24072/pcjournal.47.
Orr HA. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics. 1995:139(4):1805–1813. https://doi.org/10.1093/genetics/139.4.1805.
Ostberg CO, Hauser L, Pritchard VL, Garza JC, Naish KA. Chromosome rearrangements, recombination suppression, and limited segregation distortion in hybrids between Yellowstone cutthroat trout (*Oncorhynchus clarkii bouvieri*) and rainbow trout (*O. mykiss*). BMC Genom. 2013:14(1):570. https://doi.org/10.1186/1471-2164-14-570.
Ottenburghs J, Kraus RHS, van Hooft P, van Wieren SE, Ydenberg RC, Prins HHT. Avian introgression in the genomic era. Avian Res. 2017:8(1):30. https://doi.org/10.1186/s40657-017-0088-z.
Pardo-Diaz C, Salazar C, Baxter SW, Merot C, Figueiredo-Ready W, Joron M, McMillan WO, Jiggins CD. Adaptive introgression across species boundaries in *Heliconius* butterflies. PLOS Genet. 2012:8: e1002752. https://doi.org/10.1371/journal.pgen.1002752.
Pujolar JM, Jacobsen MW, Als TD, Frydenberg J, Magnussen E, Jónsson B, Jiang X, Cheng L, Bekkevold D, Maes GE, et al. Assessing patterns of hybridization between North Atlantic eels using diagnostic single-nucleotide polymorphisms. Heredity (Edinb). 2014:112(6): 627–637. https://doi.org/10.1038/hdy.2013.145.
Pulido-Santacruz P, Aleixo A, Weir JT. Morphologically cryptic Amazonian bird species pairs exhibit strong postzygotic reproductive isolation. Proc Biol Sci. 2018:285(1874):20172081. https://doi.org/10.1098/rspb.2017.2081.
Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007:81(3):559–575. https://doi.org/10.1086/519795.
Rand DM, Harrison RG. Ecological genetics of a mosaic hybrid zone: mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. Evolution. 1989:43(2):432–449. https://doi.org/10.2307/2409218.
Rieseberg LH. Chromosomal rearrangements and speciation. Trends Ecol. 2001:16(7):351–358. https://doi.org/10.1016/S0169-5347(01)02187-5.
Rieseberg LH, Sinervo B, Linder CR, Ungerer MC, Arias DM. Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. Science. 1996:272(5262):741–745. https://doi.org/10.1126/science.272.5262.741.
Runemark A, Eroukhmanoff F, Nava-Bolaños A, Hermansen JS, Meier JI. Hybridization, sex-specific genomic architecture and local adaptation. Philos Trans R Soc Lond B Biol Sci. 2018:373(1757): 20170419. https://doi.org/10.1098/rstb.2017.0419.
Runemark A, Vallejo-Marin M, Meier JI. Eukaryote hybrid genomes. PLoS Genet. 2019:15(11):e1008404. https://doi.org/10.1371/journal.pgen.1008404.
Sankararaman S, Mallick S, Dannemann M, Prüfer K, Kelso J, Pääbo S, Patterson N, Reich D. The genomic landscape of Neanderthal ancestry in present-day humans. Nature. 2014:507(7492):354. https://doi.org/10.1038/nature12961.
Schaefer J, Duvernell D, Campbell DC. Hybridization and introgression in two ecologically dissimilar *Fundulus* hybrid zones. Evolution. 2016:70(5):1051–1063. https://doi.org/10.1111/evo.12920.
Schilthuizen M, Giesbers MC, Beukeboom LW. Haldane's rule in the 21st century. Heredity (Edinb). 2011:107(2):95–102. https://doi.org/10.1038/hdy.2010.170.
Schmid S, Bachmann Salvy M, Garcia Jimenez A, Bertrand JAM, Cortesi F, Heim S, Huyghe F, Litsios G, Marcionetti A, O'Donnell JL, et al. Gene flow throughout the evolutionary history of a colour

polymorphic and generalist clownfish. Mol Ecol. 2024:33(14): e17436. https://doi.org/10.1111/mec.17436.
Schmid S, Micheli B, Cortesi F, Donati G, Salamin N. Extensive hybridisation throughout clownfishes evolutionary history. bioRxiv 499304. https://doi.org/10.1101/2022.07.08.499304, 10 July 2022, preprint: not peer reviewed.
Schumer M, Cui R, Powell DL, Dresner R, Rosenthal GG, Andolfatto P. High-resolution mapping reveals hundreds of genetic incompatibilities in hybridizing fish species.McVean G, editor. eLife. 2014:3:e02535. https://doi.org/10.7554/eLife.02535.
Schumer M, Cui R, Powell DL, Rosenthal GG, Andolfatto P. Ancient hybridization and genomic stabilization in a swordtail fish. Mol Ecol. 2016:25(11):2661–2679. https://doi.org/10.1111/mec.13602.
Schwarzbach AE, Donovan LA, Rieseberg LH. Transgressive character expression in a hybrid sunflower species. Am J Bot. 2001:88(2): 270–277. https://doi.org/10.2307/2657018.
Seixas FA, Boursot P, Melo-Ferreira J. The genomic impact of historical hybridization with massive mitochondrial DNA introgression. Genome Biol. 2018:19(1):91. https://doi.org/10.1186/s13059-018-1471-8.
Servedio MR, Noor MAF. The role of reinforcement in speciation: theory and data. Annu Rev Ecol Syst. 2003:34(1):339–364. https://doi.org/10.1146/annurev.ecolsys.34.011802.132412.
Sobel JM, Stankowski S, Streisfeld MA. Variation in ecophysiological traits might contribute to ecogeographic isolation and divergence between parapatric ecotypes of *Mimulus aurantiacus*. J Evol Biol. 2019:32(6):604–618. https://doi.org/10.1111/jeb.13442.
Takahashi H, Toyoda A, Yamazaki T, Narita S, Mashiko T, Yamazaki Y. Asymmetric hybridization and introgression between sibling species of the pufferfish *Takifugu* that have undergone explosive speciation. Mar Biol. 2017:164(4):90. https://doi.org/10.1007/s00227-017-3120-2.
Tao Y, Li J-L, Liu M, Hu X-Y. Complete mitochondrial genome of the orange clownfish *Amphiprion percula* (Pisces: Perciformes, Pomacentridae). Mitochondrial DNA A. 2016:27(1):324–325. https://doi.org/10.3109/19401736.2014.892099.
Taylor EB, Boughman JW, Groenenboom M, Sniatynski M, Schluter D, Gow JL. Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. Mol Ecol. 2006:15(2):343–355. https://doi.org/10.1111/j.1365-294X.2005.02794.x.
Thompson KA. Experimental hybridization studies suggest that pleiotropic alleles commonly underlie adaptive divergence between natural populations. Am Nat. 2020:196(1):E16–E22. https://doi.org/10.1086/708722.
Van der Auwera GA, O'Connor BD. Genomics in the cloud: using Docker, GATK, and WDL in Terra. 1st ed. Sebastopol (CA): O'Reilly Media; 2020.
Veller C, Edelman NB, Muralidhar P, Nowak MA. Recombination and selection against introgressed DNA. Evolution. 2023:77(4): 1131–1144. https://doi.org/10.1093/evolut/qpad021.
Walsh J, Shriver WG, Olsen BJ, O'Brien KM, Kovach AI. Relationship of phenotypic variation and genetic admixture in the Saltmarsh–Nelson's sparrow hybrid zone. Auk. 2015:132(3):704–716. https://doi.org/10.1642/AUK-14-299.1.
Xiong T, Mallet J. On the impermanence of species: the collapse of genetic incompatibilities in hybridizing populations. Evolution. 2022:76(11):2498–2512. https://doi.org/10.1111/evo.14626.
Zhang B-L, Chen W, Wang Z, Pang W, Luo M-T, Wang S, Shao Y, He W-Q, Deng Y, Zhou L, et al. Comparative genomics reveals the hybrid origin of a macaque group. Sci Adv. 2023:9(22):eadd3580. https://doi.org/10.1126/sciadv.add3580.

Associate editor: Bonnie Fraser