






ORIGINAL ARTICLE OPEN ACCESS

Reproductive Barriers and Genomic Hotspots of Adaptation During Allopatric Species Divergence

Riddhi Deshmukh^{1,2}  | Saurav Baral^{1,3}  | Muktai Kuwalekar^{1,4}  | Athulya Girish Kizhakke¹  | Krushnamegh Kunte¹ 

¹National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bengaluru, India | ²Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland | ³Department of Zoology, Stockholm University, Stockholm, Sweden | ⁴Institute of Biotechnology, HiLIFE, University of Helsinki, Helsinki, Finland

Correspondence: Krushnamegh Kunte (krushnamegh@ncbs.res.in)

Received: 8 January 2024 | **Revised:** 15 July 2024 | **Accepted:** 18 September 2024

Handling Editor: Brent Emerson

Funding: This work was partially funded by a Ramanujan Fellowship from the Department of Science and Technology, Government of India, and an NCBS Research Grant to K.K., support from the Department of Atomic Energy, Government of India, under project nos. 12-R&D-TFR-5.04-0800 and 12-R&D-TFR-5.04-0900, a CSIR Shyama Prasad Mukherjee Fellowship to R.D., and NCBS Student Fellowships to S.B. and A.G.K. Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung grant (31003A_173189) provides financial support to R.D. at the University of Lausanne.

Keywords: assortative mating | postzygotic reproductive barriers | prezygotic reproductive barriers | selective sweeps | signatures of selection | speciation genomics

ABSTRACT

Theory predicts that in allopatric populations, genomic divergence and reproductive barriers may be driven by random genetic drift and thereby evolve slowly in large populations. However, local adaptation and divergence under selection may also play important roles, which remain poorly characterised. Here, we address three key questions in young allopatric species: (a) How widespread are genomic signatures of adaptive divergence?, (b) What is the functional space along which young sister species show divergence at the genomic level? and (c) How quickly might prezygotic and postzygotic reproductive barriers evolve? Analysis of 82 re-sequenced genomes of the Oriental *Papilio polytes* species group revealed surprisingly widespread hotspots of intense selection and selective sweeps at hundreds of genes, spanning all chromosomes, rather than divergence only in a few genomic islands. These genes are involved in diverse ecologically important adaptive functions such as wing development, colour patterning, courtship behaviour, mimicry, pheromone synthesis and olfaction, and host plant use and digestion of secondary metabolites, that could contribute to local adaptation and subsequent reproductive isolation. Divergence at such functional genes appeared to have evolved in conjunction with reproductive consequences: behavioural and hybridisation experiments revealed strong assortative mate preference (prezygotic barriers) as well as postzygotic barriers to hybridisation in timespans as short as 1.5 my, indicating that speciation was already complete rather than incipient. Our study thus demonstrates an underappreciated role of intense selection and potential local adaptation in creating genome-wide hotspots of rapid molecular evolution and divergence during differentiation and speciation in young allopatric species.

1 | Introduction

Speciation and morphological diversification have shaped the remarkable diversity of life forms. The causes and consequences of selective agents and processes that drive this diversification are critical in structuring biological diversity. Research in this area

was historically disconnected to a degree. Traits within a species were studied largely through the lens of selection and local adaptation (Darwin 1859; Gould and Lewontin 1979; Sheppard 1958; Wallace 1897). On the other hand, speciation was studied largely through the lens of geographic isolation, which dominated ideas and taxonomic practices related to speciation in the 20th century

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2025 The Author(s). *Molecular Ecology* published by John Wiley & Sons Ltd.

(Mayr 1942). The long-lasting consensus was that allopatric divergence is the predominant mode of speciation, in which drift is expected to play a prominent role in genomic divergence that could lead to reproductive isolation (Mayr 1942). However, little is known about how populations diverge in allopatry: to what degree drift and/or local selection contribute to divergence, how many traits may be involved in divergence, whether reproductive isolation largely evolves through neutral processes, and how widespread signatures of selection are expected to be in genomic divergence between allopatric versus sympatric species pairs. This is unfortunate because adaptive divergence and ecological, physiological and inherent (e.g., genetic as now recognised) bases of reproductive isolation were very much at the core of early thinking about speciation (Darwin 1859; Dobzhansky 1937; Schluter 2009; Wallace 1897). The two views focusing on trait divergence and geographic basis began to coalesce with the renewed interest in the process of adaptive divergence and speciation and the availability of genetic and genomic data (Coyne and Orr 2004; Schluter 2009).

The recent shift towards understanding the biological aspects, rather than geographic aspects, of divergence and speciation has generated a robust understanding of how ecological pressures create disruptive natural and sexual selection. Disruptive selection fuels adaptive divergence along morphological, behavioural, sensory and physiological axes across the population-to-species continuum. Some of the prominent examples of adaptive population divergence along this continuum are: (a) sympatric host race formation in pea aphids (Peccoud et al. 2009) and *Timema* stick insects (Nosil et al. 2002), and in divergent male courtship signals and female preference in bushcrickets (*Ephippiger diurnus*) (Barbosa et al. 2016), and frogs (*Engystomops petersi*) (Guerra and Ron 2008); (b) at early stages of reproductive isolation and speciation, polygenic basis of reproductive isolation involving hybrid male sterility and genomic regions with strong differentiation in *Oryctolagus cuniculus* rabbits (Carneiro et al. 2014); (c) incipient ecological speciation primarily due to local adaptation to water availability influencing several morphological traits in subspecies of *Boechera stricta*, a perennial mustard (Lee and Mitchell-Olds 2013); and at the other end, (d) reproductive isolation in sister species, for example, assortative mating by wing colouration among *Heliconius cydno* and *H. pacheus*, governed by a single locus that controls both wing colour pattern and mate preference (Kronforst et al. 2006). Similarly, pollinators mediate prezygotic isolation in *Mimulus lewisii* and *M. cardinalis*, the underlying flower colour being controlled by a major quantitative trait locus (Yuan et al. 2013). Work in this area has so far focused largely on the tempo and mode of graded divergence, gene flow and genetics of ecological adaptation and reproductive isolation along the population-to-species continuum (Nosil and Feder 2012; Roux et al. 2016). However, most of this knowledge comes from instances of sympatric speciation, while few comparisons have been made in allopatric species (McGaugh and Noor 2012; Orr 2005) about broad genome-wide signatures of genetic incompatibilities, adaptive divergence and the roles of selection and drift in shaping this genomic divergence among species.

There is value in understanding the interaction between ecological pressures and geographic isolation, and the population genetic signatures they may leave on adaptation and genomic divergence during the process of speciation. In allopatric populations, one might expect neutral processes such as genetic

drift to act in concert with selection; however, there may be no specific selection for divergence itself (Barton 1996; Mayr 1942). Therefore, populations may diverge slowly, producing few hotspots of adaptive divergence (except under some local adaptation) and slower evolution of reproductive isolation (Coyne and Orr 1989; Matute and Cooper 2021; Turelli et al. 2014). On the other hand, genetic/genomic divergence and reproductive isolation may be relatively accelerated in sympatric populations evolving under diversifying selection and adaptive divergence, particularly under reinforcement, resulting in distinct genomic signatures of positive and directional selection or, in extreme cases, selective sweeps (Coyne and Orr 1989; Matute and Cooper 2021; Turelli et al. 2014). How frequently each of these scenarios is observed and whether genetic signatures of selection accompany the evolution of genomic divergence remain largely unexplored, especially in non-model organisms. Indeed, certain ecological adaptations appear to be primed for rapid divergence and local adaptation in the process of population differentiation and speciation, perhaps because they often experience intense selection and have a malleable genetic basis that responds to selection rapidly. Known examples include colour polymorphisms, mimicry, armature and other traits used in predator avoidance (Barrett et al. 2008; Jiggins et al. 2001; Mallet et al. 2007; Nadeau et al. 2014), diet specialisation (Ballhorn et al. 2016; Mikheyev et al. 2013; Peccoud et al. 2009) and sexually selected adaptations such as courtship behaviours, sexual weapons and ornaments (Emlen et al. 2007; Mérot et al. 2015; Shirangi et al. 2009). It would be interesting to find out whether these traits contribute to rapid divergence more frequently in sympatric settings due to the selection for divergence than in allopatric settings.

Here we focus on the tempo and genomic consequences of divergence between young species. We explore this framework by analysing genomes of the *Papilio polytes* species group. The *P. polytes* group has recently emerged as a model system in evolutionary and developmental genetics because of its iconic mimetic polymorphism (Clarke and Sheppard 1960; Deshmukh et al. 2018, 2020; Fisher 1958; Ford 1945; Iijima et al. 2018, 2019; Kunte et al. 2014; Nishikawa et al. 2015; Sheppard 1975; VanKuren et al. 2023; Zhang et al. 2017). *Papilio* are edible prey but females of many species have evolved independently multiple times into Batesian mimics of the aposematic *Pachliopta* butterflies to avoid predators (Kunte 2009; Uesugi 1995, 1996). The female forms and their corresponding geographic distributions are illustrated in Figure 1D. Across a large part of the range of these species, female forms *cyrus* and *polytes* co-occur across, while f. *romulus* is restricted to Peninsular India and Sri Lanka, and f. *theseus* is endemic to the islands of SE Asia (Clarke and Sheppard 1972). Historically, the *P. polytes* group was considered to have only three species—*P. polytes*, *P. phestus* and *P. ambrax* (Figure 1A). Of these, *P. polytes* was believed to be the most widespread, occurring from Sri Lanka and India to the Lesser Sunda Islands, Sulawesi, the Philippines and southern Japanese islands in the east, splintered into at least 20 subspecies structured in the mainland-island mosaic and exhibiting small variation in wing colouration and the presence/absence of tails (Clarke and Sheppard 1972; D'Abbrera 1982; Kunte et al. 2023). However, there has been some evidence indicating reproductive incompatibility between a few subspecies. For instance, Clarke and Sheppard had different degrees of success

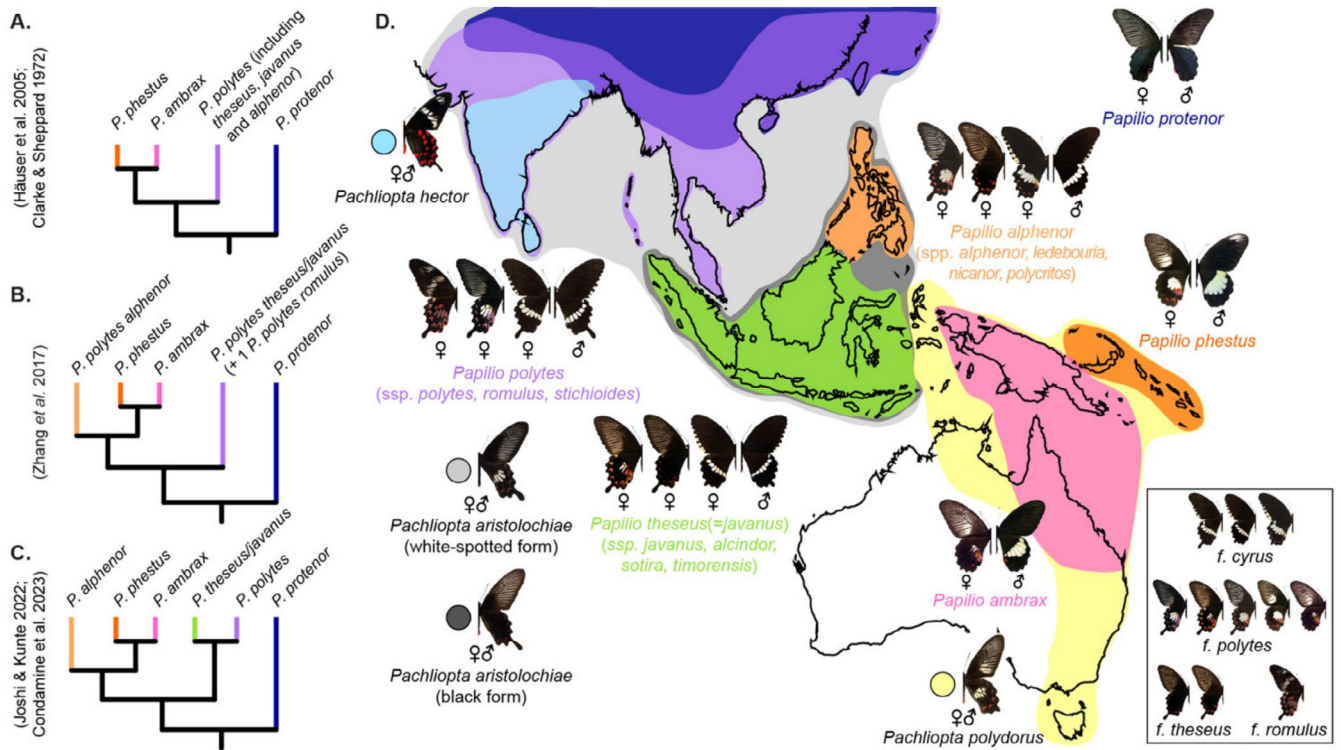


FIGURE 1 | Putative species relationships, geographical distribution and various Batesian mimetic forms in the *Papilio polytes* species group. (A–C) Different interpretations of species relationships in the *polytes* species group. (D) Geographical ranges of each species, their sexual dimorphism, female polymorphisms and Batesian mimetic relationships. Names of all aposematic models are in black, whereas names of all the *polytes* group species are colour coded with their geographic ranges as per latest taxonomic proposals (this work). Sexual dimorphism, polymorphism and mimicry phenotypes are also illustrated for each taxon.

in crossing several putative subspecies of *P. polytes* to identify the genetic bases of mimetic wing polymorphism (Clarke and Sheppard 1972). Conflicting phylogenetic evidence has also recently emerged. Reconstruction of the evolutionary history of mimetic polymorphism in this group with genome-wide markers showed *P. polytes* to be either a polyphyletic or a paraphyletic species because *alphenor*, traditionally considered to be the Philippine subspecies of *P. polytes*, was sister to *P. phestus* and *P. ambrax* outside of *P. polytes* (Zhang et al. 2017) (Figure 1B). Further, mito-nuclear phylogenies of the *polytes* species group (including many of the subspecies) suggested that *P. polytes* is not one but a complex of three geographically well-separated (allopatric) species—*P. polytes*, *P. theseus* and *P. alphenor*—each with multiple island subspecies included in them (note that *P. theseus* was earlier listed as *P. javanus*, including by us, but the oldest available name for this species is *Papilio theseus* Cramer, 1777, not *Papilio javanus* C. Felder, 1862) (Condamine et al. 2023; Joshi and Kunte 2022) (Figure 1A–C). The time-calibrated phylogeny and node age estimates show the group to have evolved approx. 4 mya, with the split between the youngest *P. polytes* and *P. theseus* species pair being 1.5 mya (Condamine et al. 2023). Thus, the systematics of this complex polytypic group is beginning to be finally resolved, making it feasible to perform a detailed genome-level analysis of selection, divergence and differentiation, along with an understanding of reproductive barriers.

Here, we first characterise species in the *P. polytes* species group with phylogenetic, genomic and behavioural analyses. We then use this characterisation to address the following questions: (a)

Is *P. polytes* a polyphyletic, paraphyletic or monophyletic species? (b) What forms of reproductive barriers maintain these species? (c) What are the genome-wide signatures of selection that might provide insights into the roles of adaptation and drift during allopatric divergence? Are signatures of selective divergence concentrated in a few genomic islands, or is divergence more widespread at the genomic level? and (d) What is the functional space of genes that show signatures of selective divergence? Are these shared between species? We address these questions with detailed population genomic analysis of selection and divergence using 82 genomes across the entire species group.

2 | Results

2.1 | *Papilio polytes* Is a Cluster of Three Well-Supported, Monophyletic, Geographically Structured Species

2.1.1 | Phylogenetic Structure

We generated Bayesian phylogenies of the *polytes* species group, including multiple subspecies, with four marker datasets (Figure 2). A species tree reconstructed using only the mitochondrial markers was unresolved (Figure 2D), showing that barcodes and other mitochondrial markers alone are not adequate to resolve relationships in this species group. However, the three trees reconstructed using the nuclear and mito-nuclear datasets produced an identical, well-supported

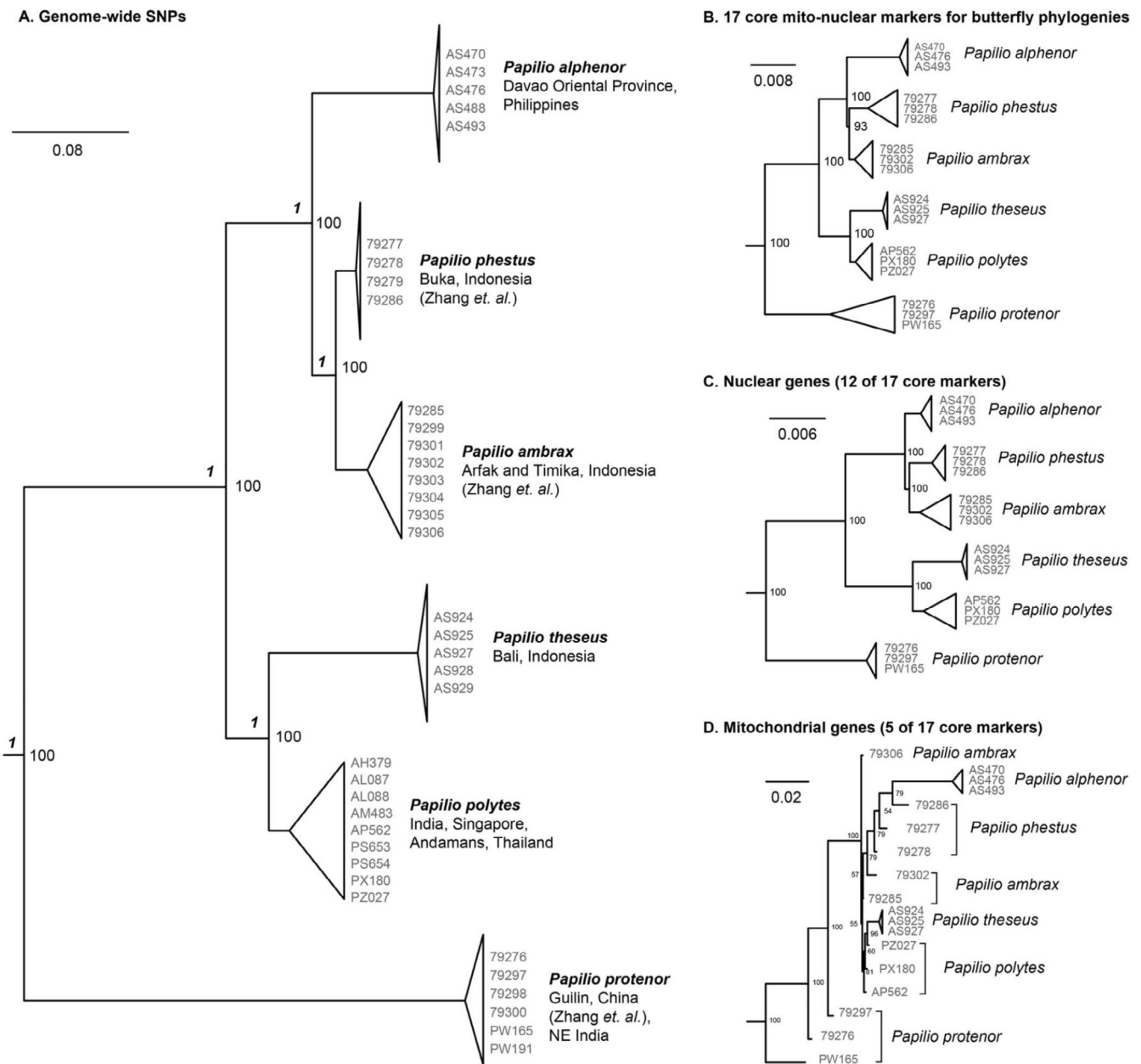


FIGURE 2 | Species phylogenies of the *polytes* species group using different sets of genetic markers including concatenated genome-wide SNPs (A), 17 widely used mito-nuclear markers (B), only nuclear genes (C), and only mitochondrial genes (D). Numbers adjacent to collapsed branches represent specimen codes or genome accession numbers (sample details in Table S1). Support values for all splits are shown at the nodes. Italicised numbers in (A) indicate mPTP support for species divergence events. 30,760 SNPs were used to construct tree A (details of SNP calling and filtering are provided in the Methods). The use of only mitochondrial markers resulted in an unresolved tree topology (D).

topology showing the following species relationships: (*protenor*,((*alphenor*,(*phestus*,*ambrax*)),(*polytes*,*theseus*))) (Figure 2A–C).

Thus, *P. polytes* as traditionally viewed came out as a polyphyletic species. The three distinct, strongly supported monophyletic groupings that resulted, which do not even show sister species relationships with respect to *alphenor*, can be treated as three distinct species, as *P. polytes*, *P. theseus* and *P. alphenor*. The tree topology, node age estimates, mainland-island distributions and allopatry of these three monophyletic groups all indicate that the three species have diverged in allopatry

over the past 4 million years, that is, since early Pliocene (Condamine *et al.* 2023).

2.1.2 | Species Delimitation and PCA

We performed species delimitation analysis with SNAPP using the 30,760 genome-wide SNP dataset (dataset 1). We tested five models with different combinations of species partitions (Table 1) against two null models. The first null model assumed that *P. polytes* and *P. theseus* are distinct species (based on Figures 1C and 2A–C). None of the alternative

TABLE 1 | Comparison of alternative species models with SNAPP.

Model	No. of species	MLE	MLE rank	Bayes factor (null model 1)	Bayes factor (null model 2)
A (<i>polytes</i> species group has 6 species; Figures 1C and 2A–C)	6	−119,050	1	—	36,985.730
B (<i>polytes</i> and <i>theseus</i> are one species; Figure 1B)	5	−137,543	5	−36,986.268	—
C (<i>polytes</i> , <i>theseus</i> and <i>alphenor</i> are one species; Figure 1A)	4	−195,363	6	−152,626.3	−115,640.3
D (<i>ambrax</i> and <i>phestus</i> are one species; Figures S3 and S4)	5	−127,348	3	−16,596.529	20,389.476
E (<i>polytes</i> split into two species, one occurring in South Asia (India) and the other in Indo-China (Thailand))	7	−121,271	2	−4441.1624	32,544.838
F (Thailand <i>polytes</i> + <i>theseus</i>)	6	−137,311	4	−36,522.321	463.679

Note: Marginal Likelihood Estimates (MLE) and Bayes factors for different models of species delimitation are shown. The two null models for calculation of Bayes factor assumed either that *P. polytes* and *P. theseus* were distinct species (model A in the table, based on Figure 1C), or that *P. polytes* and *P. theseus* were a single species (model B in the table, based on Figure 1B). In model F, we included the Thailand population of *P. polytes* (which appears as most basal in the *P. polytes* clade in Figure S1) with *P. theseus*, to check whether combining geographically closer populations of different species affected the model support. $N = 3$ individuals for all species.

models obtained sufficient support: models B–F were rejected (Bayes factor < 0) in favour of Null Model 1, which had the highest MLE rank (Table 1; also see Methods). The second null model assumed that *P. polytes* and *P. theseus* are the same species (based on Zhang et al. 2017; Figure 1B). This analysis rejected Null Model 2, with model A having the strongest support (Table 1).

We ascertained the most supported species partitions using mPTP (multi-rate Poisson Tree Processes model) analysis, which models speciation rate using sequence substitutions. This analysis also strongly supported the topology in Figure 2A–C, that *P. polytes*, *P. theseus* and *P. alphenor* are three distinct, monophyletic species (mPTP support values in italics in Figure 2A).

We also performed a PCA with our dataset using genome-wide SNPs and obtained a similar pattern to the phylogenetic and mPTP analysis. Sub-clade 1, containing *P. polytes* and *P. theseus*, and sub-clade 2, containing *P. alphenor*, *P. phestus* and *P. ambrax*, separated from the basal *P. protenor* along PC1 and PC2, and the sister species *P. polytes* and *P. theseus*, and *P. alphenor*, *P. phestus* and *P. ambrax*, separated along PC3 and PC4 (statistical significance represented by 95% confidence limits; Figure S2). PC axes 1–4 explained 79.05% of the total variation in the dataset.

2.1.3 | Population Structure

We used ADMIXTURE to determine population structure in the *polytes* species group using biallelic genome-wide SNPs. The most supported result with the smallest cross-validation error split the group into six clusters ($K = 6$; Figure S3). *Papilio polytes*, *P. theseus* and *P. alphenor* formed separate population clusters, with *P. polytes* being split into two distinct clusters over its geographic range. Such a clear population structure in *P. polytes* appears to be a result of considerable genomic variation across its

large population, possibly owing to genetic drift or local adaptation over its very broad geographic range.

ADMIXTURE analysis treated *P. phestus* and *P. ambrax* together as a single population cluster in the best-supported clustering ($K = 5$; Figure S3). It is possible that highly unequal sample sizes of species biased the result (*P. polytes*: 40 individuals, *P. phestus*: 4 individuals), or that ADMIXTURE analysis pooled *P. phestus* and *P. ambrax* because of the relatively small genomic differences among them compared with genomic differences among other species. To account for unequal sample sizes, we reran the ADMIXTURE analysis with randomly chosen seven samples each from the focal *P. polytes*, *P. theseus* and *P. alphenor*, which removed the large disparity in the sample sizes among the species (Figure S4). This analysis, too, pooled *P. phestus* and *P. ambrax* as a single population ($K = 5$; Figure S4). Since all the other analyses that we performed (phylogenetic reconstruction, species delimitation, SNAPP and PCA) supported *P. phestus* and *P. ambrax* as distinct species, we suggest that the incongruent pooling of *P. phestus* and *P. ambrax* in the ADMIXTURE analyses may be due to relatively small differences between them, and not due to highly unequal sample sizes in our first analysis (Figure 2A–C, Figure S2). Indeed, *P. phestus* and *P. ambrax* have the shortest branch lengths separating sister species in the *P. polytes* species group (Figure 1, Figure S1).

2.1.4 | Genome-Wide Divergence

These species show considerable percentage divergence between the genome-wide variable sites (30,760 SNPs from dataset 1; range: 3.6%–17% divergence; Figure 2A, Figure S1). We further estimated pairwise genomic divergence (F_{st}) and absolute divergence (d_{xy}) in 10 kb windows using the full dataset of 27 million genome-wide variants across the five mimetic sister species (Figure S5). The average genome-wide F_{st} was considerably high, ranging from 0.4 to 0.8, revealing high divergence

among all the species (average F_{st} values: *P. polytes*–*P. theseus*: 0.43; *P. polytes*–*P. alphenor*: 0.64; *P. theseus*–*P. alphenor*: 0.82; *P. phestus*–*P. ambrax*: 0.59; Figure S5). The d_{xy} estimates follow a similar trend to F_{st} for pairwise comparisons of genome-wide divergence, although average d_{xy} values were considerably lower than F_{st} values (Figure S5). This is interesting since high F_{st} and low d_{xy} values may be indicative of either rapid recent divergence or selective sweeps after divergence (Henderson and Brelsford 2020). The genomic spread and functional aspects of selection and selective sweeps are addressed in sections 3 and 4 below. Moreover, the majority of F_{st} and d_{xy} values are within a relatively narrow range for each species pairwise comparison, with few genomic regions showing high or low levels of divergence (Figure S5). These narrow ranges of genome-wide divergence may indicate that the genomes of each species have diverged largely at a constant pace, presumably under drift, as expected.

Interestingly, despite *P. polytes* and *P. theseus* being phylogenetically equally distant from *P. alphenor*, there is considerable difference between average F_{st} estimates between them (0.64 and 0.82 for *P. polytes*–*P. alphenor* and *P. theseus*–*P. alphenor*). Island demographic effects and perhaps local adaptation may be responsible for these differences since both *P. theseus* and *P. alphenor* are island species and may have thus diverged disproportionately more compared with *P. polytes*. Additionally, although effective population size may influence the pattern of genomic divergence observed, the three species have

relatively large and widespread populations across their geographical ranges, and they occur in allopatry with no known hybrid zones (Condamine et al. 2023; Joshi and Kunte 2022; Zhang et al. 2017). These aspects need further attention in the future.

2.2 | *Papilio polytes*, *P. theseus* and *P. alphenor* Are Reproductively Isolated by Prezygotic and Postzygotic Barriers

We tested mate preference (prezygotic barriers) and post-mating reproductive success (largely postzygotic barriers) using naturally occurring and hand-paired intraspecific and interspecific matings in the three taxa (*polytes*, *theseus* and *alphenor*) that are still widely treated as subspecies of the same species. In the first experiment in a mixed population of freshly eclosed individuals (see Methods), all three species mated assortatively (Pearson's Chi-squared test, $X^2=100.44$, $df=4$, $p<0.0001$; Figure 3A, Tables S3 and S5) even when they had no prior experience of their own species over other species, showing that the assortative mate preference was instinctive. In the second set of experiments, we tested reproductive success by setting up intraspecific and interspecific hand-paired matings. This experiment removed the behavioural choice of mates but was able to reveal post-mating selection, reproductive compatibilities and hybrid fitness. We compared several measures of reproductive success: (1) the duration of mating,

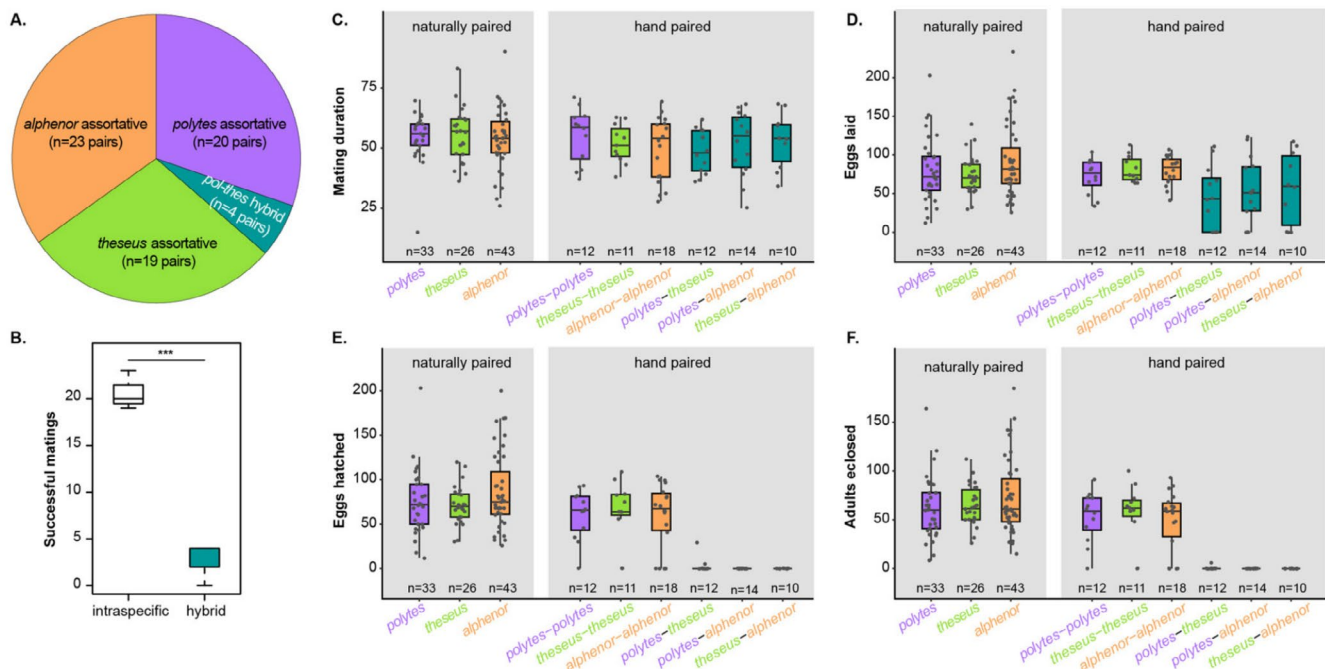


FIGURE 3 | Reproductive success and isolation among *P. polytes*, *P. theseus* and *P. alphenor*. (A, B) Test of prezygotic barriers to reproduction. Assortative mating in a mixed population containing equal numbers of males and females of *P. polytes*, *P. theseus* and *P. alphenor*. ($p<0.001$ represented by '***'). (C–F) Tests of postzygotic barriers to reproduction. Mating duration (C), number of eggs laid (D), hatching success (E) and eclosion success (F) across the three species and their hybrids in naturally paired and hand-paired matings. $p<0.001$, Tukey's test with correction for multiple comparisons, for all pairwise tests between intraspecific and interspecific matings in (E) and (F). Details of the statistical tests are provided in Table S6. Intraspecific matings are colour coded, interspecific matings are dark green. n indicates sample sizes for each comparison. Hybrid mating pair sample sizes in panel (A): $N=3$ for *P. polytes* males \times *P. theseus* females and $n=1$ for *P. theseus* male \times *P. polytes* female. (C–F) $N=6$ each for *P. polytes* males \times *P. theseus* females and *P. theseus* males \times *P. polytes* females; $n=5$ for *P. polytes* males \times *P. alphenor* females and $n=9$ for *P. alphenor* males \times *P. polytes* females; $n=3$ for *P. theseus* males \times *P. alphenor* females and $n=7$ for *P. alphenor* males \times *P. theseus* females.

(2) no. of eggs laid, (3) no. of eggs that hatched successfully and (4) no. of adults that eclosed successfully. We compared the reproductive success in hand-paired intraspecific matings with those of naturally occurring intraspecific matings. Naturally paired and hand-paired intraspecific matings showed similar trends in all the traits measured (Tukey's test with correction for multiple comparisons, $p > 0.08$ for all pairwise comparisons; Figure 3B,C, Tables S2 and S4–S6), showing that hand-pairings were successful. Likewise, mating duration as well as the number of eggs laid between intraspecific and interspecific matings were similar (Tukey's test with correction for multiple comparisons for all pairwise tests: mating duration: $p > 0.8$, except *polytes-alphenor* hand-paired mating where $p = 0.04$; no. of eggs laid: $p > 0.1$), further showing that hand-pairings resulted in normal mating and oviposition behaviours. This result also suggests that the prezygotic barriers between these species do not appear to affect the maturation and oviposition of eggs. However, the hatching and eclosion success in hand-paired interspecific matings was very low (Tukey's test with correction for multiple comparisons, $p < 0.001$ for all pairwise tests; Figure 3D,E, Tables S2 and S4–S6), showing that the three species have genetic/developmental incompatibilities that contribute to postzygotic barriers to hybridisation.

2.3 | All *P. polytes* Group Species Show Widespread Genomic Signatures of Strong Selection and Divergence

We predicted that if the *P. polytes* group species have largely drifted apart in allopatry, their genomes would likely show little signatures of divergence under selection. On the other hand, if their divergence has occurred under some local selection, then we may observe genomic signatures of divergence under selection and adaptation, which have rarely been characterised in allopatric taxa. To test these two scenarios, we characterised the genomic regions that have experienced intense selection since the species started evolving independently after initial geographic isolation. We used Raised Accuracy in Sweep Detection (RAiSD) to identify selective sweeps, and a composite measure of selection (de-correlated composite of multiple signals (DCMS), including H1, H12, Fst, Tajima's D and nucleotide diversity) to identify regions showing signatures of selection across the genomes. We particularly chose signatures of selective sweeps as a clear indication of intense selection and its genetic manifestation, whereas the composite measure of selection may indicate a variable intensity of selection, from weak to strong. Together, these two approaches

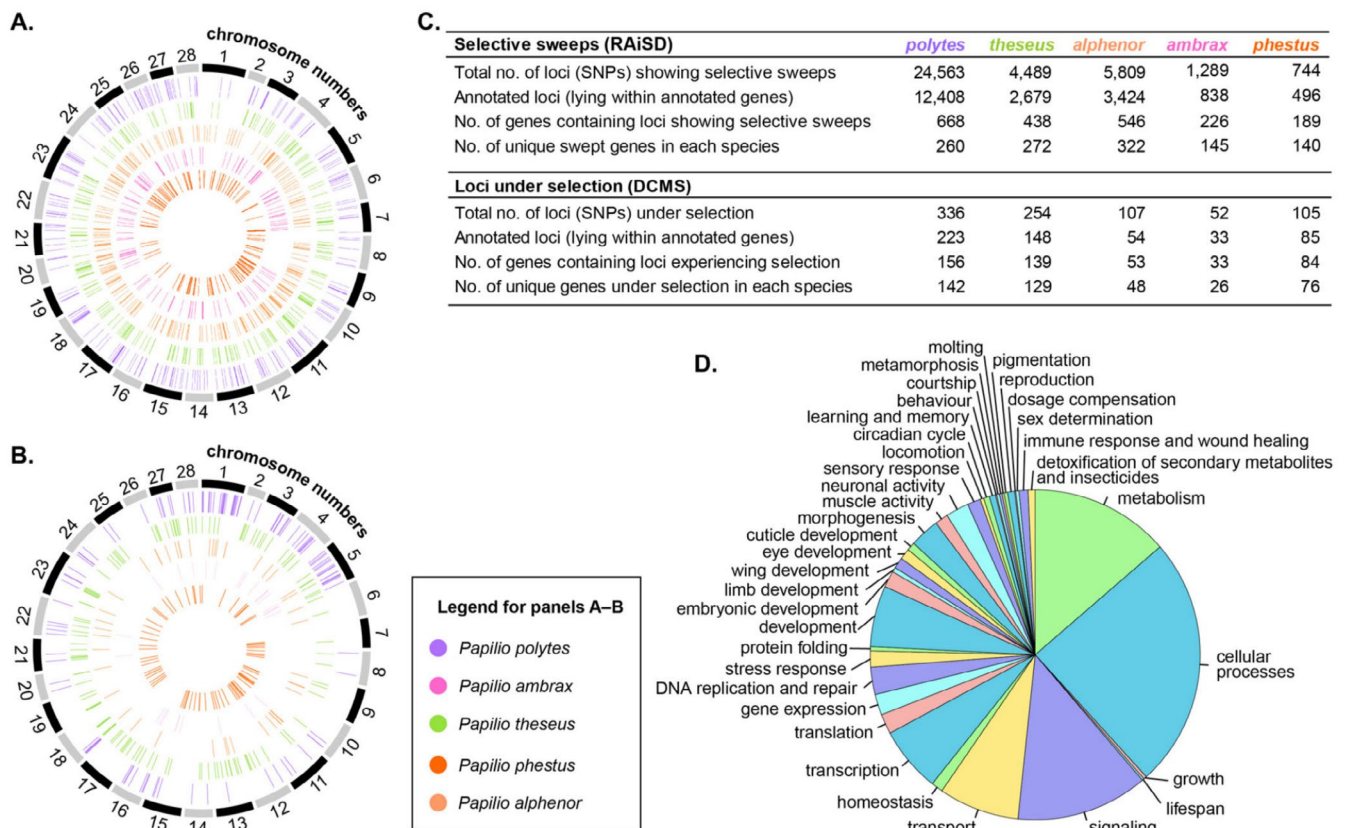


FIGURE 4 | Loci under selection during genomic divergence of species in the *P. polytes* species group. (A) Annotated genes showing selective sweeps in the five species, identified using RAiSD. The alternately coloured outer bands indicate chromosomes as defined in the reference *Bombyx mori* genome, whereas colour-coded lines inside represent genes under selection unique to each species. (B) Annotated genes showing signatures of selection, identified using DCMS. Colour coding as in panel (A). (C) Summary of the number of loci (SNPs and the genes they lie in) experiencing intense selection in the *P. polytes* species group. (D) Range of functions performed by genes containing sites under selection in the five *Papilio* species. Details of gene functions are provided in Table S10. Sample sizes: *P. polytes* $n = 19$, *P. theseus* $n = 17$, *P. alphenor* $n = 7$, *P. phestus* $n = 4$, *P. ambrax* $n = 8$, *P. protenor* $n = 6$.

revealed that: (a) hundreds of genes have experienced selective sweeps or diversifying selection in each species, (b) a large proportion of genes were under selection in individual species of the group and (c) few genes were under selection in more than two or three species (Figure 4A; Tables S7 and S8). Moreover, the sets of genes under selection as identified by RAiSD (selective sweeps) and DCMS (selection) were completely non-overlapping (Tables S7 and S8). Only the mimicry locus (*doublesex*) (Kunte et al. 2014; Nishikawa et al. 2015) showed signatures of selective sweeps in all five mimetic species, whereas DCMS did not identify any gene under selection in all species. The genes experiencing selective sweeps/selection were not concentrated in a few genomic islands but spread widely in numerous hotspots across all the annotated chromosomes in the genome (Figure 4A–C). Such a broad signature of genome-wide divergence under selection is particularly striking in light of the relatively recent splits between these species: the oldest split in this group was approx. 4 mya, between (*polytes,theseus*) and (*alphenor,(ambrax,phestus)*) and the most recent splits were only approx. 1.5 mya between (*polytes,theseus*) and between (*ambrax,phestus*) (Condamine et al. 2023). Moreover, the fact that the genome of each of these young allopatric species shows such widespread signatures of selective sweeps itself is startling.

2.4 | *Papilio polytes* Group Species Have Diverged Under Selection on Hundreds of Genes Involved in Local Adaptation, Sexual Selection and Housekeeping

Species of the *P. polytes* group show two striking patterns: (a) all the species have non-overlapping geographic distributions over different mainland/island groups (Figure 1D) and (b) a large number of genes are under selection in individual species (Figure 4C, Table S9). These observations together imply that these allopatric species have diverged with prominent episodes of strong selection, as shown by the widespread hotspots of genomic divergence, with hundreds of sites showing selective sweeps. What is the functional space of genes that show divergence under selection in these species? We mapped all the SNPs experiencing selection to annotated genes with known functions to characterise the functional space that these genes occupy. These genes belong to the following major functional categories: (a) local adaptation, such as host plant use, digestion of secondary metabolites and insecticide resistance, (b) mate recognition and reproductive isolation, such as wing patterning, eye development, pheromone production, odorant receptors and courtship behaviour, (c) embryonic and organismal development and gene expression and (d) cellular housekeeping genes involved in metabolism, transport and signalling (Figure 4D, Tables S7, S8 and S10). Of these, the first two functional categories have clear implications for ecological and sexual selection, and local adaptation and divergence. Indeed, the pre- and postzygotic barriers to hybridisation that we uncovered above, such as assortative mating and low reproductive success in hand-paired interspecific crosses, may have direct links to functional isolation along sensory axes, facilitated by mate choice, and by physiological and developmental breakdown in hybrid zygotes. A closer look at the biology of these species and additional hybridisation experiments may shed further light on this.

Additionally, annotation and functional association of SNPs that lay within intergenic regions could identify candidates involved in local adaptation and mate choice via gene regulation.

Of particular interest is the gene *doublesex*, which shows signatures of selective sweeps in all five mimetic species of the group but not in the basal non-mimetic *P. protenor*. It is important not only in mimetic polymorphism and wing patterning in the *polytes* species group but, as a key developmental regulator of sexual dimorphisms, it may be involved in other sex-specific adaptations in these species. The last two functional categories are important for a swath of critical functions at an organismal level, but it is unclear how selection on these genes might contribute to local adaptation and divergence. This needs to be explored further.

3 | Discussion

Isolation, divergence and speciation in the unusually fragmented biodiversity hotspots of the Oriental Region have led to high levels of species diversity and endemism (Braby and Pierce 2007; Brown and Siler 2014; Condamine et al. 2012; Müller et al. 2012; Welton et al. 2013). Unfortunately, the primary focus on geographical isolation, rather than on causes and consequences of adaptation and divergence in this biogeographically complex landscape, has historically misled taxonomists and evolutionary biologists in understanding and defining species, especially in polytypic taxa (de Queiroz 2007; Joshi and Kunte 2022; Mallet 2004; Mayr 1942, 1965). A unified approach to understanding species and application of genomic data and detailed phylogenetic analyses are now leading to improved understanding of species diversity and speciation patterns in relation to the long and complex biogeographic history of the fragmented land/seascape of the Oriental Region. Our study should be viewed in this light, where we were able to leverage a large genomic dataset, multiple lines of evidence using phylogenetic methods and population genetic tests, and data on mate choice and reproductive success. With these approaches, we were able to characterise taxonomic diversity and the potential role of selection and local adaptation in the genomic diversification in this clade. Our analyses reinforced the notion that *P. polytes*, as traditionally classified, is not a single super-widespread species (including different subspecies such as *P. polytes theseus/javanus* and *P. polytes alphenor*) but a complex of three monophyletic species that exhibit hallmarks of being distinct species (Figure 1). These hallmarks include not only a strong phylogenetic and population genetic structure (Figure 2, Figures S1 and S2) but also considerable genome-level differentiation under selection (Figure 4, Figure S5), and pre- and postzygotic barriers to hybridisation such as strong assortative mate preference and low success of interspecific crosses (Figure 3). These findings should be a lesson that polytypic taxa in the biogeographically complex Indo-Australian Region need to be reinvestigated carefully with respect to their systematic and taxonomic status based on detailed evolutionary and genetic studies. In this specific case, our findings fundamentally change the understanding of the evolutionary history of this species group that has become a model for testing theories about the evolution and genetics of adaptation. The systematic and genomic clarity provided above should particularly inform future work on the remarkable female-limited

polymorphic mimicry displayed in this clade—a classic example of natural selection.

Our phylogenetic analysis using genome-wide SNPs produced a different species tree topology (Figure 2A–C) compared with a topology shown recently (Zhang et al. 2017). This is because the Zhang et al. study did not adequately sample *P. polytes* (as delineated in our study), treating *P. theseus* as part of *P. polytes* instead and including only a single sample of the true *P. polytes*. The multiple lines of evidence based on broad sampling in the range of *P. polytes* from which we demonstrated that *P. polytes* and *P. theseus* are indeed different species should help resolve this discrepancy.

It is notable that we did not find evidence in these young species for Haldane's Rule—the observation that reproductive incompatibilities among different populations/species disproportionately hamper the production of individuals of the heterogametic sex in hybrid progeny (Schilthuizen et al. 2011; Wu et al. 1996). Due to the low success of hybridisation and poor performance of hybrid progeny (Tables S5 and S6), we did not have enough samples to effectively determine the sex ratio of hybrid offspring. It is possible that either these species have become sufficiently reproductively incompatible not to show effects of Haldane's Rule or, more likely, that Haldane's Rule needs to be specifically tested with respect to sex and species with further crossing experiments in this species group. Recent studies in allopatric pairs of *Heliconius pardalinus* subspecies (Rosser et al. 2022), as well as *Papilio bianor* and *Papilio dehaanii* (Xiong et al. 2023) have characterised hybrid incompatibility via Haldane's rule and demonstrated a complex polygenic basis of incompatibility involving the Z chromosome. Both these allopatric pairs show a high level of genetic divergence that is comparable to pairwise comparisons within the *polytes* species group. Perhaps a comparison of genetic divergence and its correlation with reproductive success and hybrid survival could be made among *Papilio* species occurring in allopatry.

With respect to speciation, when should we expect few local genomic islands versus several widespread hotspots of adaptation and adaptive divergence? Theory predicts that a few genomic islands of adaptive divergence would be responsible for reproductive isolation among populations that are diverging in sympatry (Coyne and Orr 2004; Feder et al. 2013). For speciation to occur, this divergence is expected to be rapid. Otherwise, ongoing genetic exchange would prevent speciation, unless the fitness consequences of ecological selection on co-adapted traits/gene complexes are strong. Yet well-studied cases of sympatric speciation reveal genome-wide divergence even in young species pairs (Kautt et al. 2020; Michel et al. 2010), countering the idea of a very small number of genomic islands of differentiation and speciation.

In allopatric populations, a neutral theory would predict divergence across multiple unlinked traits and loci, but the tempo of this genomic divergence leading up to reproductive isolation and speciation to be rather slow, without producing hotspots of divergence under selection. However, we above showed genome-wide differentiation in allopatric populations, this differentiation showing unexpectedly widespread signatures of strong selection producing selective sweeps in hundreds of genes across the

genome, in timespans as short as 1.5 my. Since genome-level divergence was not correlated with phylogenetic distance, and especially since so many genes with ecological and sexual functions showed signatures of selective sweeps in each species, it is clear that selection has played a prominent role in divergence among these allopatric species. We have not yet tested whether these genes (via changes in the coding sequences or their regulation) are directly responsible for reproductive incompatibility and local adaptation for traits mentioned above, or they are in linkage with other loci involved in these adaptations. Similarly, we have not yet specifically tested the relative strengths and signatures of genetic drift versus selection in the genomes of these species. Nonetheless, these findings offer many new hypotheses for future work on adaptive diversification and speciation in this species group. Moreover, they suggest that the degree of selection for local adaptation and underlying genome-level divergence during both sympatric and allopatric speciation may have been underestimated. Perhaps a future study comparing the levels of genomic divergence and signatures of selection within and between populations of allopatric species may help contextualise the patterns we reported above, and determine whether these interspecific patterns are outliers or common in other species groups, too.

Although several recent studies have looked at the genetic basis of adaptation and divergence that pull populations and species apart, genome-level characterisation of adaptive differentiation in sister species is rarely attempted. As a result, there are few benchmarks available to compare the tempo and extent of genomic divergence under drift versus selection at advanced stages of isolation and divergence. Taking advantage of the natural experiment of diversification in these allopatric species, our work offers such a benchmark for future studies. How factors such as speciation in sympatry, secondary contact and reinforcement and adaptive introgression change the genome-wide signatures of selection and divergence is just beginning to be explored (Arnold and Kunte 2017; Edelman et al. 2019; Foote 2018; Kautt et al. 2020; Michel et al. 2010; Richards et al. 2019). Our study offers a contrast in allopatric divergence and might serve as a further impetus for investigations in this direction.

4 | Methods

4.1 | Sample Collection, Sequencing and SNP Calling

We performed whole-genome sequencing of wild-caught samples preserved in 100% ethanol, of four species in the *polytes* group (*P. polytes*, *P. alphenor*, *P. theseus* and *P. protenor*), supplementing these with previously published genome sequences (Zhang et al. 2017) from the SRA database for our analysis (sample details in Table S1). We obtained *P. alphenor* and *P. theseus* from wild populations and native commercial butterfly breeding facilities in the Philippines and Java, respectively. We extracted DNA from thoracic muscles of each butterfly using the QIAGEN DNeasy blood and tissue kit. We quantified extracted DNA using Qubit fluorometric quantification and prepared libraries using the Illumina TruSeq DNA PCR-free library preparation kit. We sequenced them on the Illumina HiSeq 2500 using 2×100 PE runs. We checked the quality of sequences and downloaded

genomes using FastQC, trimmed Illumina adapters using Trimmomatic (v0.35) (Bolger et al. 2014) and aligned them to the reference *P. polytes* reference genome from NCBI (Ppol_1.0, genome version 1.0, annotation version 1.0) (Nishikawa et al. 2015) using the BWA mem aligner (Li and Durbin 2009). We marked and removed duplicates from aligned files, merged files from the same sample, performed indel realignment and called SNPs with the haplotype caller in the GATK pipeline (McKenna et al. 2010). We filtered the resulting SNPs with the recommended parameters (QD < 2.0, SOR > 3.0, FS > 60.0, MQ < 40.0, MQRankSum < -12.5, ReadPosRankSum < -8.0), retained samples with an average coverage of 9X across the genome, and obtained a set of 27,965,662 variants for the *polytes* species group. We calculated average heterozygosity and inbreeding coefficient (F) for each species using vcfTools 0.1.13 (Danecek et al. 2011). We observed large variation between individuals of the same species (*P. polytes* $n = 40$, $F = 0.284$; *P. theseus* $n = 17$, $F = 0.368$; *P. alphenor* $n = 7$, $F = 0.141$; *P. phestus* $n = 4$, $F = 0.445$; *P. ambrax* $n = 8$, $F = 0.582$; *P. protenor* $n = 6$, $F = 0.140$). Surprisingly, some of the individuals obtained from breeders showed higher heterozygosity than some wild-caught individuals (Zhang et al. 2017). We filtered this SNP dataset to remove monoallelic sites and sampled SNPs every 5000 bases to obtain a set of 30,760 SNPs that could be used for phylogenetic analysis with moderate run times.

4.2 | PCA, Phylogenetic Reconstruction and Species Delimitation

We resampled sites every 10 kb from the SNP dataset of 27 million SNPs to obtain unlinked sites across the genome to perform a PCA using plink2. We represented PCs 1 to 4 encompassing 79.05% of the variation between the six species (Figure S2). We reconstructed Bayesian phylogenies with MrBayes 3.2.7a (Ronquist et al. 2012) using four datasets: (1) 30,760 genome-wide SNPs, (2) 17 standard phylogenetic markers including both nuclear (*thiolase*, *carbamoyl-phosphate synthetase 2*, *aspartate transcarbamylase and dihydroorotase (CAD)*, *catalase (CAT)*, *dopa decarboxylase (DDC)*, *glyceraldehyde 3-phosphate dehydrogenase (GAPDH)*, *isocitrate dehydrogenase (IDH)*, *malate dehydrogenase (MDH)*, *ribosomal protein S2 (RPS2)*, *ribosomal protein S5 (RPS5)*, *hairy cell leukaemia protein 1 (HCL)*, *elongation factor-I alpha (EF1-a)* and *wingless*) and mitochondrial genes (*cytochrome c oxidase-I*, *tRNA leucine*, *cytochrome c oxidase II*, *16S* and *ND5*) (Wahlberg and Wheat 2008), (3) a subset of the 17 markers including only the 12 nuclear genes, and (4) a subset of the 17 markers including only the five mitochondrial genes. The four datasets were used to compare the results between conventionally used markers for phylogenetic analyses for systematics and those from whole-genome datasets. Our expectation was that we would obtain better resolution with the whole-genome dataset compared to commonly used nuclear and mitochondrial markers. We extracted DNA sequence data for datasets 2–4 from whole-genome sequences of 29 samples using bcftools-1.6 mpileup (Li 2011): three samples each for the six species in the *polytes* species group. We compiled the genome-wide SNP dataset from 48 samples: 37 samples from the *polytes* species group—five samples of *P. alphenor*, eight samples of *P. ambrax*, four samples of *P. phestus*, five samples of *P. theseus*, nine samples of *P. polytes* and six samples of *P. protenor*—and 11

outgroup species. We aligned DNA sequence data in datasets 2–4 using the MUSCLE algorithm on MEGA X (Edgar 2004) and concatenated them to get a final data block of 17,450 bp for dataset 2, 12,087 bp for dataset 3 and 5363 bp for dataset 4. Since most of these markers are housekeeping genes and therefore well-represented, we had adequate coverage for the SNPs called. We used PartitionFinder 2.1.1 (Lanfear et al. 2016) to choose the best partitioning scheme and models of evolution under Bayesian information criterion (BIC) for all four datasets. PartitionFinder suggested a single partition for dataset 1 (GTR), 12 Partitions for dataset 2 (GTR+I, HKY+I, HKY+G, K80, HKY+G, K80+I, GTR+I, HKY+I+G, GTR+I+G, HKY+I+G, GTR+G and HKY+I+G), six partitions for dataset 3 (GTR+I, HKY+I, HKY+G, K80, HKY+G and K80+I) and six partitions for dataset 4 (GTR+I+G, HKY+I for 2, HKY+G, HKY+I, GTR+G and HKY+I+G) (Dryad DOI for tree files will be submitted in the final submission after the initial acceptance of the manuscript). We ran the MrBayes analysis for one million generations for dataset 1 and two million generations for datasets 2–4. All four datasets had two independent runs with four chains each, with sampling done every 1000 generations. We used a split frequency below 0.01 and visual inspection of the run parameters on Tracer v1.7.1 (Rambaut et al. 2018) to ensure stationarity. We built consensus trees after discarding 25% of the samples as burn-in. We used FigTree v1.4.3 (Rambaut 2014) to view and edit all the phylogenies.

We used coalescent-based SNAPP (SNP and AFLP Package for Phylogenetic analysis) to compare alternative species models (Bryant et al. 2012), and mPTP (multi-rate Poisson Tree Processes) to determine the most supported species partition scheme by modelling branching events based on the number of mutations (Zhang et al. 2013). We performed SNAPP analysis with SNAPP v 1.5.0 (Bryant et al. 2012) on Beast2 v 2.6.2 (Bouckaert et al. 2014) using the dataset of 30,760 genome-wide SNPs. We used custom scripts in R to convert SNP data to the SNAPP binary format. Due to computational limitations, we used a reduced dataset of three randomly selected samples per species to reconstruct a phylogeny using whole-genome SNPs.

We tested six alternative species models (hypotheses) as shown in Table 1. For every species model, we set default model parameters: a gamma distribution with $\alpha = 2$ and $\beta = 200$ for the speciation rate prior of the Yule model (λ) and a gamma distribution with $\alpha = 1$ and $\beta = 150$ for the rate prior using Beauti V 2.6.2 (Bouckaert et al. 2014). We ran each model for 20 million generations, sampling every 2000 generations. For the path sampling analysis, we used 48 steps, an alpha of 0.3 and a burn-in of 0.25. We calculated Bayes Factor values from the Marginal Likelihood Estimates (MLE) produced by the path sampler to compare each model against the null hypothesis. We carried out both ML and MCMC analysis on mPTP (v.0.2.4) using the multi-rate option. The MCMC analysis was run for 100 million generations, sampling every 10,000 generations in 10 independent runs. The first 2 million samples were discarded as burn-in. We performed both the ML start and random start options. We examined convergence by assessing the plot of log-likelihood against MCMC iteration. Since the Bayes factor is calculated by comparing the null model with alternative models, the highest positive value is the best-supported model (null model 1 in our analyses; Table 1), whereas negative values

for alternative models do not reject the null model (Bryant et al. 2012; Leaché et al. 2014).

In this framework, we tested two null models: Null Model 1, in which *polytes* and *theseus* were treated as two distinct species (based on Figures 1C and 2A–C) and Null Model 2, in which *polytes* and *theseus* were treated as a single species (based on Zhang et al. 2017; Figure 1B).

4.3 | Genetic Divergence Between the Species of *Polytes* Group

To find genome-wide genetic divergence between *polytes* group species, we grouped the 27 million genome-wide SNPs by species and calculated Weir-Cockerham's F_{st} and d_{xy} (in 10 kb windows) for pairwise comparisons between *P. polytes*, *P. theseus*, *P. alphenor*, *P. phestus* and *P. ambrax* using vcfTools 0.1.13 and pixy 1.0.0 (Danecek et al. 2011; Korunes and Samuk 2021). We used ADMIXTURE to estimate ancestry and population structure in the *polytes* group species using biallelic sites from the 27 million SNPs (Alexander and Lange 2011). We tested ADMIXTURE runs with ancestral populations (K) ranging from 2 to 13. We plot here results of three runs with the lowest cross-validation error ($K=5$, $K=6$ and $K=7$; Figure S3). We downsampled some species to rerun ADMIXTURE analysis with roughly equal sample sizes (seven per species for *P. polytes*, *P. theseus* and *P. ambrax*). We again plot results of the four runs with the lowest cross-validation error (Figure S4).

4.4 | Mate Choice and Hybridisation Success Experiments

We maintained pure-breeding populations of *polytes*, *theseus* and *alphenor* in separate cages. These lab-bred populations were originally started with multiple wild-caught individuals, done as a standard practice in non-model organisms. We allowed males and females from each population to mate freely and raised the progeny in the common population-specific cages. We monitored a subset of naturally occurring intra-population matings to study mating success within the population.

We set up an experiment to study interpolation mate choice and mating success in a separate cage, in which individually marked unmated males and females from each population were allowed to choose between potential mates from the three populations. These males and females were maintained in this cage from eclosion, that is, they had no prior experience of individuals of their own population over individuals from other populations. We thus excluded the possibility of early exposure to olfactory and other signals of any particular population before they could choose between potential mates, which could otherwise bias their subsequent mate choice. This ensured that individuals were choosing mates based purely on instinctive mate preference, and not early exposure. We set up this mixed population with an equal number of males and females from each pure-breeding population.

Separately, we set up experimental hand-paired matings, in which we individually selected 3- to 5-day-old unmated males

and females. We paired them by gently squeezing the abdominal end and opening up the claspers of a male and aligning its genital opening with that of the female, and slowly releasing the male while his claspers closed onto the female's genital opening. Successful hand-pairings remove the opportunity for butterflies to court and behaviourally choose mates, and only measure the post-mating success of a pair. Such hand-pairings are possible in *Papilio*; they are often successful when done by experienced people and have been used in hybridisation experiments in the past (Clarke and Sheppard 1956). We noted down the duration of mating, the number of eggs laid and the stages of metamorphosis until which the progeny of the hand-paired butterflies developed successfully. All statistical tests were performed using R. We set up hand-pairings of individuals from within and across populations to compare the success of hand-paired matings. We maintained each mated female in a separate small cage in which she was allowed to lay eggs, and we raised her brood separately.

We did not include in the analysis short-lived butterflies, that is, those that died before reaching maturity to mate (usually within the 1–3 days of eclosion) or those females that died before being able to lay eggs (usually within a day after mating). Because of the low numbers, we did not attempt to study species \times sex effects on the success of hybrid matings or the mating and breeding success of hybrid progeny. Sample sizes for mate-choice and hand pairing assays, including species of origin for males and females from the mating pair, are provided in Figure 3.

4.5 | Estimation of Signatures of Selection in the *P. polytes* Species Group

To identify loci showing signatures of selection in the genomic dataset, we used two methods that could potentially produce an overlapping set of SNPs that could be screened. We used a composite measure of selection (de-correlated composite of multiple signals, i.e., DCMS) (Ma et al. 2015; Yurchenko et al. 2018) and scanned the genomes for signatures of selection. We used the 27 million variant dataset (*P. polytes* $n=19$, *P. theseus* $n=17$, *P. alphenor* $n=7$, *P. phestus* $n=4$, *P. ambrax* $n=8$, *P. protenor* $n=6$) and calculated the μ statistic using Raised Accuracy in Sweep Detection (RAiSD, v.2.5 (Alachiotis and Pavlidis 2018)) to identify loci that have experienced selective sweeps. We calculated Tajima's D , nucleotide diversity and F_{st} using vcfTools 0.1.13 (Danecek et al. 2011) and H1 and H12 using published scripts (<https://github.com/ngarud/SelectionHapStats>) (Garud et al. 2015) for the *polytes* group species and estimated a composite metric DCMS (Yurchenko et al. 2018) to find SNPs showing signatures of selection. We used a conservative cut-off and considered SNPs in the 99.5 percentile for both our methods and annotated them. We obtained a large number of hits with RAiSD compared to DCMS, which is expected since RAiSD is known to have a high rate of false positives depending on background selection (Alachiotis and Pavlidis 2018), with no overlapping hits detected by both methods. To screen hits for false positives, we randomly resampled ~1000,000 SNPs from all species with different sample sizes to determine the FDR threshold ($\alpha=0.05$) for the μ statistic using RAiSD. The resulting threshold was lower than our cut-off. The SNPs experiencing sweeps identified by each of the methods lay within annotated genetic elements as well as intergenic regions that lack annotation features. Genes

with functional significance potentially associated with local adaptation in each species are summarised in Table S10. We mapped the genes under selection in each species and the genes common across the five mimetic species to the *Bombyx mori* genome (Lu et al. 2019) using BLAST+ to identify their chromosomal locations. We used Circos-0.69-9 (Krzywinski et al. 2009) to map these genes across chromosomes for each species, assuming synteny between *P. polytes* and *Bombyx mori* genomes. Additionally, we mapped the loci showing signatures of selection to the congeneric *Papilio bianor* genome. We observed that > 70% of hits retained synteny with *B. mori* (Table S11).

We were unable to annotate with reasonable confidence genomic elements such as enhancers and other intergenic regions that may play significant roles in ecological and local adaptation, since most of them remain unidentified in these butterflies. A large number of loci that showed signatures of intense selection fell in these intergenic regions. This factor is responsible for the discrepancy in the number of total positions under selection versus the number of such loci in annotated genes (Figure 4C).

4.6 | Characterising the Functional Space of Genes Under Selection

We cross-referenced gene function for all the genes identified in our analyses to be under selection and showing the signatures of selective sweeps with the UniProt database (Bateman et al. 2021). This database provides manually annotated and peer-reviewed information on protein sequence and gene function from all organisms, although not all individual gene functions have been verified in butterflies. It is possible that some of these genes have additional functions in local and morphological adaptation in butterflies, as recently discovered for the role of *doublesex*, *WntA* and other genes in polymorphic mimicry and wing patterning (Deshmukh et al. 2018; Kunte et al. 2014; Mazo-Vargas et al. 2017; Nishikawa et al. 2015).

Author Contributions

Riddhi Deshmukh and Saurav Baral generated and analysed genome and genotype sequence datasets. Riddhi Deshmukh and Muktai Kuwalekar performed selection analyses. Riddhi Deshmukh analysed mate choice and reproductive success data. Athulya Girish Kizhakke and Saurav Baral performed phylogenetic analyses. Athulya Girish Kizhakke performed species delimitation analyses. Krushnamegh Kunte conceived and directed the project, designed research, performed experiments for mate choice and reproductive success; and wrote the paper with contributions from Riddhi Deshmukh. All authors approved the final manuscript version.

Acknowledgements

We thank Deepa Agashe, Chris Wheat, Jim Mallet, Ilik Saccheri and Claudius Kratochwil for comments on the project and manuscript; Tarun Karmakar, Anuj Jain, Sarika Baidya, Arjan Basu Roy, Adam M. Cotton and Hariprasath for contributing samples; NCBS Sequencing Facility, AgriGenome and Genotypic for whole-genome sequencing; and Biodiversity Lab Research Collections for sample storage facilities. Wild-caught samples were obtained largely from the NCBS-TIFR campus/field stations and private lands, and from wildlife sanctuaries and national parks in India under the following research and collection permits issued by the state forest departments in Karnataka (permit

no. 227/2014-2015 dated 2015/04/16), West Bengal (permit no. 2115(9)/WL/4K-1/13/BL41, dated 2013/11/06; and permit no. 1107/42/2W-705/18, dated 2018/05/07), and Meghalaya (permit no. FWC/G/173/Pt-II/474-83, dated 2014/05/27), for which we thank the Principal Chief Conservator of Forest, Deputy Conservators of Forest, Wildlife Wardens and field officers of those states.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Raw data for mating experiments are given in Tables S2–S4. Raw genome sequence data generated in this study are available at the NLM-NCBI SRA database (PRJNA1166847). Sequence alignments, partition models and tree files for the phylogenetic analysis are available on Dryad (<https://doi.org/10.5061/dryad.wpzgmsbxx>).

References

- Alachiotis, N., and P. Pavlidis. 2018. “RAiSD Detects Positive Selection Based on Multiple Signatures of a Selective Sweep and SNP Vectors.” *Communications Biology* 1, no. 1: 79. <https://doi.org/10.1038/s42003-018-0085-8>.
- Alexander, D. H., and K. Lange. 2011. “Enhancements to the ADMIXTURE Algorithm for Individual Ancestry Estimation.” *BMC Bioinformatics* 12, no. 1: 246. <https://doi.org/10.1186/1471-2105-12-246>.
- Arnold, M. L., and K. Kunte. 2017. “Adaptive Genetic Exchange: A Tangled History of Admixture and Evolutionary Innovation.” *Trends in Ecology & Evolution* 32, no. 8: 601–611. <https://doi.org/10.1016/j.tree.2017.05.007>.
- Ballhorn, D. J., F. P. Rakotoarivelo, and S. Kautz. 2016. “Coevolution of Cyanogenic Bamboos and Bamboo Lemurs on Madagascar.” *PLoS One* 11, no. 8: e0158935. <https://doi.org/10.1371/journal.pone.0158935>.
- Barbosa, F., D. Rebar, and M. D. Greenfield. 2016. “Female Preference Functions Drive Interpopulation Divergence in Male Signalling: Call Diversity in the Bushcricket *Ephippiger diurnus*.” *Journal of Evolutionary Biology* 29, no. 11: 2219–2228. <https://doi.org/10.1111/jeb.12940>.
- Barrett, R. D. H., S. M. Rogers, and D. Schluter. 2008. “Natural Selection on a Major Armor Gene in Threespine Stickleback.” *Science* 322, no. 5899: 255–257. <https://doi.org/10.1126/science.1159978>.
- Barton, N. H. 1996. “Natural Selection and Random Genetic Drift as Causes of Evolution on Islands.” *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 351, no. 1341: 785–795. <https://doi.org/10.1098/rstb.1996.0073>.
- Bateman, A., M.-J. Martin, S. Orchard, et al. 2021. “UniProt: The Universal Protein Knowledgebase in 2021.” *Nucleic Acids Research* 49, no. D1: D480–D489. <https://doi.org/10.1093/nar/gkaa1100>.
- Bolger, A. M., M. Lohse, and B. Usadel. 2014. “Trimmomatic: A Flexible Trimmer for Illumina Sequence Data.” *Bioinformatics* 30, no. 15: 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Bouckaert, R., J. Heled, D. Kühnert, et al. 2014. “BEAST 2: A Software Platform for Bayesian Evolutionary Analysis.” *PLoS Computational Biology* 10, no. 4: e1003537. <https://doi.org/10.1371/journal.pcbi.1003537>.
- Braby, M. F., and N. E. Pierce. 2007. “Systematics, Biogeography and Diversification of the Indo-Australian Genus *Delias* Hübner (Lepidoptera: Pieridae): Phylogenetic Evidence Supports an ‘Out-of-Australia’ Origin.” *Systematic Entomology* 32, no. 1: 2–25. <https://doi.org/10.1111/j.1365-3113.2006.00349.x>.
- Brown, R. M., and C. D. Siler. 2014. “Spotted Stream Frog Diversification at the Australasian Faunal Zone Interface, Mainland Versus

- Island Comparisons, and a Test of the Philippine ‘Dual-Umbilicus’ Hypothesis.” *Journal of Biogeography* 41, no. 1: 182–195. <https://doi.org/10.1111/jbi.12192>.
- Bryant, D., R. Bouckaert, J. Felsenstein, N. A. Rosenberg, and A. RoyChoudhury. 2012. “Inferring Species Trees Directly From Biallelic Genetic Markers: Bypassing Gene Trees in a Full Coalescent Analysis.” *Molecular Biology and Evolution* 29, no. 8: 1917–1932. <https://doi.org/10.1093/molbev/mss086>.
- Carneiro, M., F. W. Albert, S. Afonso, et al. 2014. “The Genomic Architecture of Population Divergence Between Subspecies of the European Rabbit.” *PLoS Genetics* 10, no. 8: e1003519. <https://doi.org/10.1371/journal.pgen.1003519>.
- Clarke, C. A., and P. M. Sheppard. 1956. “Hand-Pairing of Butterflies.” *Lepidopterists’ News* 10: 47–53.
- Clarke, C. A., and P. M. Sheppard. 1960. “Super-Genes and Mimicry.” *Heredity* 14: 175–185. <https://doi.org/10.1038/hdy.1960.15>.
- Clarke, C. A., and P. M. Sheppard. 1972. “The Genetics of the Mimetic Butterfly *Papilio polytes* L.” *Philosophical Transactions of the Royal Society B* 263, no. 855: 431–458.
- Condamine, F. L., R. Allio, E. L. Reboud, et al. 2023. “A Comprehensive Phylogeny and Revised Taxonomy Illuminate the Origin and Diversification of the Global Radiation of *Papilio* (Lepidoptera: Papilionidae).” *Molecular Phylogenetics and Evolution* 183: 107758. <https://doi.org/10.1016/j.ympev.2023.107758>.
- Condamine, F. L., E. F. A. Toussaint, A. M. Cotton, G. S. Genson, F. A. H. Sperling, and G. J. Kergoat. 2012. “Fine-Scale Biogeographical and Temporal Diversification Processes of Peacock Swallowtails (*Papilio* Subgenus *Achillides*) in the Indo-Australian Archipelago.” *Cladistics* 1: 1–24. <https://doi.org/10.1111/j.1096-0031.2012.00412.x>.
- Coyne, J. A., and H. A. Orr. 1989. “Patterns of Speciation in *Drosophila*.” *Evolution* 43, no. 2: 362–381. <https://doi.org/10.1111/j.1558-5646.1989.tb04233.x>.
- Coyne, J. A., and H. A. Orr. 2004. *Speciation*. Sinauer Associates.
- D’Abrera, B. 1982. *Butterflies of the Oriental Region. Part I: Papilionidae, Pieridae & Danaidae*. Hill House Publishing.
- Danecek, P., A. Auton, G. Abecasis, et al. 2011. “The Variant Call Format and VCFtools.” *Bioinformatics* 27, no. 15: 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>.
- Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray.
- de Queiroz, K. 2007. “Species Concepts and Species Delimitation.” *Systematic Biology* 56, no. 6: 879–886. <https://doi.org/10.1080/10635150701701083>.
- Deshmukh, R., S. Baral, A. Gandhimathi, M. Kuwalekar, and K. Kunte. 2018. “Mimicry in Butterflies: Co-Option and a Bag of Magnificent Developmental Genetic Tricks.” *WIREs Developmental Biology* 7, no. 1: e291. <https://doi.org/10.1002/wdev.291>.
- Deshmukh, R., D. Lakhe, and K. Kunte. 2020. “Tissue-Specific Developmental Regulation and Isoform Usage Underlie the Role of Doublesex in Sex Differentiation and Mimicry in *Papilio* Swallowtails.” *Royal Society Open Science* 7, no. 9: 200792. <https://doi.org/10.1098/rsos.200792>.
- Dobzhansky, T. 1937. *Genetics and the Origin of Species*. Columbia University Press.
- Edelman, N. B., P. B. Frandsen, M. Miyagi, et al. 2019. “Genomic Architecture and Introgression Shape a Butterfly Radiation.” *Science* 366, no. 6465: 594–599. <https://doi.org/10.1126/science.aaw2090>.
- Edgar, R. C. 2004. “MUSCLE: Multiple Sequence Alignment With High Accuracy and High Throughput.” *Nucleic Acids Research* 32, no. 5: 1792–1797. <https://doi.org/10.1093/nar/gkh340>.
- Emlen, D. J., J. Marangelo, B. Ball, and C. W. Cunningham. 2007. “Diversity in the Weapons of Sexual Selection: Horn Evolution in the Beetle Genus *Onthophagus* (Coleoptera: Scarabaeidae).” *Evolution* 59, no. 5: 1060–1084.
- Feder, J. L., S. M. Flaxman, S. P. Egan, A. a. Comeault, and P. Nosil. 2013. “Geographic Mode of Speciation and Genomic Divergence.” *Annual Review of Ecology, Evolution, and Systematics* 44, no. 1: 73–97. <https://doi.org/10.1146/annurev-ecolsys-110512-135825>.
- Fisher, R. A. 1958. *The Genetical Theory of Natural Selection*. 2nd ed. Dover Publications.
- Foote, A. D. 2018. “Sympatric Speciation in the Genomic Era.” *Trends in Ecology & Evolution* 33, no. 2: 85–95. <https://doi.org/10.1016/j.tree.2017.11.003>.
- Ford, E. B. 1945. “Polymorphism.” *Biological Reviews* 20, no. 2: 73–88. <https://doi.org/10.1111/j.1469-185X.1945.tb00315.x>.
- Garud, N. R., P. W. Messer, E. O. Buzbas, and D. A. Petrov. 2015. “Recent Selective Sweeps in North American *Drosophila melanogaster* Show Signatures of Soft Sweeps.” *PLoS Genetics* 11, no. 2: e1005004. <https://doi.org/10.1371/journal.pgen.1005004>.
- Gould, S. J., and R. C. Lewontin. 1979. “The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme.” *Proceedings of the Royal Society B* 205, no. 1161: 581–598. <https://doi.org/10.1098/rspb.1979.0086>.
- Guerra, M. A., and S. R. Ron. 2008. “Mate Choice and Courtship Signal Differentiation Promotes Speciation in an Amazonian Frog.” *Behavioral Ecology* 19, no. 6: 1128–1135. <https://doi.org/10.1093/beheco/arn098>.
- Henderson, E. C., and A. Brelford. 2020. “Genomic Differentiation Across the Speciation Continuum in Three Hummingbird Species Pairs.” *BMC Evolutionary Biology* 20, no. 1: 113. <https://doi.org/10.1186/s12862-020-01674-9>.
- Iijima, T., R. Kajitani, S. Komata, et al. 2018. “Parallel Evolution of Batesian Mimicry Supergene in Two *Papilio* Butterflies, *P. polytes* and *P. memnon*.” *Science Advances* 4, no. 4: eaao5416. <https://doi.org/10.1126/sciadv.aao5416>.
- Iijima, T., S. Yoda, and H. Fujiwara. 2019. “The Mimetic Wing Pattern of *Papilio polytes* Butterflies Is Regulated by a Doublesex-Orchestrated Gene Network.” *Communications Biology* 2, no. 1: 257. <https://doi.org/10.1038/s42003-019-0510-7>.
- Jiggins, C. D., R. E. Naisbit, R. L. Coe, and J. Mallet. 2001. “Reproductive Isolation Caused by Colour Pattern Mimicry.” *Nature* 411: 302–305.
- Joshi, J., and K. Kunte. 2022. “Polytypy and Systematics: Diversification of *Papilio* Swallowtail Butterflies in the Biogeographically Complex Indo-Australian Region.” *bioRxiv*, 2022.03.23.485569. <https://doi.org/10.1101/2022.03.23.485569>.
- Kautt, A. F., C. F. Kratochwil, A. Nater, et al. 2020. “Contrasting Signatures of Genomic Divergence During Sympatric Speciation.” *Nature* 588, no. 7836: 106–111. <https://doi.org/10.1038/s41586-020-2845-0>.
- Korunes, K. L., and K. Samuk. 2021. “PIXY: Unbiased Estimation of Nucleotide Diversity and Divergence in the Presence of Missing Data.” *Molecular Ecology Resources* 21, no. 4: 1359–1368. <https://doi.org/10.1111/1755-0998.13326>.
- Kronforst, M. R., L. G. Young, D. D. Kapan, C. McNeely, R. J. O’Neill, and L. E. Gilbert. 2006. “Linkage of Butterfly Mate Preference and Wing Color Preference Cue at the Genomic Location of Wingless.” *Proceedings of the National Academy of Sciences of the United States of America* 103, no. 17: 6575–6580. <https://doi.org/10.1073/pnas.0509685103>.
- Krzywinski, M., J. Schein, I. Birol, et al. 2009. “Circos: An Information Aesthetic for Comparative Genomics.” *Genome Research* 19, no. 9: 1639–1645. <https://doi.org/10.1101/gr.092759.109>.

- Kunte, K. 2009. "The Diversity and Evolution of Batesian Mimicry in *Papilio Swallowtail Butterflies*." *Evolution* 63, no. 10: 2707–2716. <https://doi.org/10.1111/j.1558-5646.2009.00752.x>.
- Kunte, K., S. Sondhi, and P. Roy. 2023. "Butterflies of India, v. 4.16. Published by the Indian Foundation for Butterflies and National Centre for Biological Sciences—TIFR." <https://www.ifoundbutterflies.org>.
- Kunte, K., W. Zhang, A. Tenger-Trolander, et al. 2014. "Doublesex Is a Mimicry Supergene." *Nature* 507, no. 7491: 229–232. <https://doi.org/10.1038/nature13112>.
- Lanfear, R., P. B. Frandsen, A. M. Wright, T. Senfeld, and B. Calcott. 2016. "PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses." *Molecular Biology and Evolution* 34, no. 3: 772–773. <https://doi.org/10.1093/molbev/msw260>.
- Leaché, A. D., M. K. Fujita, V. N. Minin, and R. R. Bouckaert. 2014. "Species Delimitation Using Genome-Wide SNP Data." *Systematic Biology* 63, no. 4: 534–542.
- Lee, C.-R., and T. Mitchell-Olds. 2013. "Complex Trait Divergence Contributes to Environmental Niche Differentiation in Ecological Speciation of *Boechera stricta*." *Molecular Ecology* 22, no. 8: 2204–2217. <https://doi.org/10.1111/mec.12250>.
- Li, H. 2011. "A Statistical Framework for SNP Calling, Mutation Discovery, Association Mapping and Population Genetical Parameter Estimation From Sequencing Data." *Bioinformatics* 27, no. 21: 2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>.
- Li, H., and R. Durbin. 2009. "Fast and Accurate Short Read Alignment With Burrows-Wheeler Transform." *Bioinformatics* 25, no. 14: 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- Lu, F., Z. Wei, Y. Luo, et al. 2019. "SilkDB 3.0: Visualizing and Exploring Multiple Levels of Data for Silkworm." *Nucleic Acids Research* 48: D749–D755. <https://doi.org/10.1093/nar/gkz919>.
- Ma, Y., X. Ding, S. Qanbari, S. Weigend, Q. Zhang, and H. Simianer. 2015. "Properties of Different Selection Signature Statistics and a New Strategy for Combining Them." *Heredity* 115, no. 5: 426–436. <https://doi.org/10.1038/hdy.2015.42>.
- Mallet, J. 2004. "Poulton, Wallace and Jordan: How Discoveries in Papilio Butterflies Led to a New Species Concept 100 Years Ago." *Systematics and Biodiversity* 1, no. 4: 441–452. <https://doi.org/10.1017/S1477200003001300>.
- Mallet, J., M. Beltran, W. Neukirchen, and M. Linares. 2007. "Natural Hybridization in Heliconiine Butterflies: The Species Boundary as a Continuum." *BMC Evolutionary Biology* 7: 28. <https://doi.org/10.1186/1471-2148-7-28>.
- Matute, D. R., and B. S. Cooper. 2021. "Comparative Studies on Speciation: 30 Years Since Coyne and Orr." *Evolution* 75, no. 4: 764–778. <https://doi.org/10.1111/evo.14181>.
- Mayr, E. 1942. *Systematics and the Origin of Species*. Columbia University Press.
- Mayr, E. 1965. *Animal Species and Evolution*. Harvard University Press.
- Mazo-Vargas, A., C. Concha, L. Livraghi, et al. 2017. "Macroevolutionary Shifts of WntA Function Potentiate Butterfly Wing-Pattern Diversity." *Proceedings of the National Academy of Sciences of the United States of America* 114, no. 40: 10701–10706. <https://doi.org/10.1073/pnas.1708149114>.
- McGaugh, S. E., and M. A. F. Noor. 2012. "Genomic Impacts of Chromosomal Inversions in Parapatric *Drosophila* Species." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 367, no. 1587: 422–429. <https://doi.org/10.1098/rstb.2011.0250>.
- McKenna, A., M. Hanna, E. Banks, et al. 2010. "The Genome Analysis Toolkit: A MapReduce Framework for Analyzing Next-Generation DNA Sequencing Data." *Genome Research* 20, no. 9: 1297–1303. <https://doi.org/10.1101/gr.107524.110>.
- Mérot, C., B. Frérot, E. Leppik, and M. Joron. 2015. "Beyond Magic Traits: Multimodal Mating Cues in *Heliconius* Butterflies." *Evolution* 69, no. 11: 2891–2904. <https://doi.org/10.1111/evo.12789>.
- Michel, A. P., S. Sim, T. H. Q. Powell, M. S. Taylor, P. Nosil, and J. L. Feder. 2010. "Widespread Genomic Divergence During Sympatric Speciation." *Proceedings of the National Academy of Sciences of the United States of America* 107, no. 21: 9724–9729. <https://doi.org/10.1073/pnas.1000939107>.
- Mikheyev, A. S., C. S. McBride, U. G. Mueller, et al. 2013. "Host-Associated Genomic Differentiation in Congeneric Butterflies: Now You See It, Now You Do Not." *Molecular Ecology* 22, no. 18: 4753–4766. <https://doi.org/10.1111/mec.12423>.
- Müller, C. J., P. F. Matos-Maraví, and L. B. Beheregaray. 2012. "Delving Into *Delias* Hübner (Lepidoptera: Pieridae): Fine-Scale Biogeography, Phylogenetics and Systematics of the World's Largest Butterfly Genus." *Journal of Biogeography* 40, no. 5: 1–13. <https://doi.org/10.1111/jbi.12040>.
- Nadeau, N. J., M. Ruiz, P. Salazar, et al. 2014. "Population Genomics of Parallel Hybrid Zones in the Mimetic Butterflies, *H. melpomene* and *H. erato*." *Genome Research* 24, no. 8: 1316–1333. <https://doi.org/10.1101/gr.169292.113>.
- Nishikawa, H., T. Iijima, R. Kajitani, et al. 2015. "A Genetic Mechanism for Female-Limited Batesian Mimicry in Papilio Butterfly." *Nature Genetics* 47 2014: 1–7. <https://doi.org/10.1038/ng.3241>.
- Nosil, P., B. J. Crespi, and C. P. Sandoval. 2002. "Host-Plant Adaptation Drives the Parallel Evolution of Reproductive Isolation." *Nature* 417, no. 6887: 440–443.
- Nosil, P., and J. L. Feder. 2012. "Genomic Divergence During Speciation: Causes and Consequences." *Philosophical Transactions of the Royal Society B* 367, no. 1587: 332–342. <https://doi.org/10.1098/rstb.2011.0263>.
- Orr, A. H. 2005. "The Genetic Basis of Reproductive Isolation: Insights From *Drosophila*." *Proceedings of the National Academy of Sciences of the United States of America* 102, no. Supplement 1: 6522–6526. <https://doi.org/10.1073/pnas.0501893102>.
- Peccoud, J., A. Ollivier, M. Plantegenest, and J.-C. Simon. 2009. "A Continuum of Genetic Divergence From Sympatric Host Races to Species in the Pea Aphid Complex." *Proceedings of the National Academy of Sciences of the United States of America* 106, no. 18: 7495–7500. <https://doi.org/10.1073/pnas.081117106>.
- Rambaut, A. 2014. *FigTree 1.4.2 Software*. Institute of Evolutionary Biology, University of Edinburgh.
- Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. "Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7." *Systematic Biology* 67, no. 5: 901–904. <https://doi.org/10.1093/sysbio/syy032>.
- Richards, E. J., M. R. Servedio, and C. H. Martin. 2019. "Searching for Sympatric Speciation in the Genomic Era." *BioEssays* 41, no. 7: 1900047. <https://doi.org/10.1002/bies.201900047>.
- Ronquist, F., M. Teslenko, P. van der Mark, et al. 2012. "MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space." *Systematic Biology* 61, no. 3: 539–542. <https://doi.org/10.1093/sysbio/sys029>.
- Rosser, N., N. B. Edelman, L. M. Queste, et al. 2022. "Complex Basis of Hybrid Female Sterility and Haldane's Rule in *Heliconius* Butterflies: Z-Linkage and Epistasis." *Molecular Ecology* 31, no. 3: 959–977. <https://doi.org/10.1111/mec.16272>.
- Roux, C., C. Fraisse, J. Romiguier, Y. Anciaux, N. Galtier, and N. Bierne. 2016. "Shedding Light on the Grey Zone of Speciation Along a Continuum of Genomic Divergence." *PLoS Biology* 14, no. 12: e2000234. <https://doi.org/10.1371/journal.pbio.2000234>.

- Schilthuizen, M., M. C. W. G. Giesbers, and L. W. Beukeboom. 2011. "Haldane's Rule in the 21st Century." *Heredity* 107, no. 2: 95–102. <https://doi.org/10.1038/hdy.2010.170>.
- Schluter, D. 2009. "Evidence for Ecological Speciation and Its Alternative." *Science* 323, no. 5915: 737–741. <https://doi.org/10.1126/science.1160006>.
- Sheppard, P. M. 1958. *Natural Selection and Heredity*. Hutchinson & Co.
- Sheppard, P. M. 1975. *Natural Selection and Heredity*. 4th ed. Hutchinson & Co.
- Shirangi, T. R., H. D. Dufour, T. M. Williams, and S. B. Carroll. 2009. "Rapid Evolution of Sex Pheromone-Producing Enzyme Expression in *Drosophila*." *PLoS Biology* 7, no. 8: e1000168. <https://doi.org/10.1371/journal.pbio.1000168>.
- Turelli, M., J. R. Lipkowitz, and Y. Brandvain. 2014. "On the Coyne and Orr-Igin of Species: Effects of Intrinsic Postzygotic Isolation, Ecological Differentiation, X Chromosome Size, and Sympatry on *Drosophila* Speciation." *Evolution* 68, no. 4: 1176–1187. <https://doi.org/10.1111/evo.12330>.
- Uesugi, K. 1995. "Mimicry in *Papilio polytes* and Its Ecological Meaning." In *Swallowtail Butterflies: Their Ecology & Evolutionary Biology*, edited by J. M. Scriber, Y. Tsubaki, and R. C. Lederhouse, 165–172. Scientific Publishers.
- Uesugi, K. 1996. "The Adaptive Significance of Batesian Mimicry in the Swallowtail Butterfly *Papilio polytes* (Insecta, Papilionidae): Associative Learning in a Predator." *Ethology* 102: 762–775.
- VanKuren, N. W., M. M. Doellman, S. I. Sheikh, D. H. Palmer Drogue, D. Massardo, and M. R. Kronforst. 2023. "Acute and Long-Term Consequences of Co-Opted Doublesex on the Development of Mimetic Butterfly Colour Patterns." *Molecular Biology and Evolution* 40: msad196. <https://doi.org/10.1093/molbev/msad196>.
- Wahlberg, N., and C. W. Wheat. 2008. "Genomic Outposts Serve the Phylogenomic Pioneers: Designing Novel Nuclear Markers for Genomic DNA Extractions of Lepidoptera." *Systematic Biology* 57, no. 2: 231–242. <https://doi.org/10.1080/10635150802033006>.
- Wallace, A. R. 1897. *Darwinism: An Exposition of the Theory of Natural Selection*. MacMillan.
- Welton, L. J., C. D. Siler, J. R. Oaks, A. C. Diesmos, and R. M. Brown. 2013. "Multilocus Phylogeny and Bayesian Estimates of Species Boundaries Reveal Hidden Evolutionary Relationships and Cryptic Diversity in Southeast Asian Monitor Lizards." *Molecular Ecology* 22, no. 13: 3495–3510. <https://doi.org/10.1111/mec.12324>.
- Wu, C.-I., N. A. Johnson, and M. F. Palopoli. 1996. "Haldane's Rule and Its Legacy: Why Are There So Many Sterile Males?" *Trends in Ecology & Evolution* 11, no. 7: 281–284. [https://doi.org/10.1016/0169-5347\(96\)10033-1](https://doi.org/10.1016/0169-5347(96)10033-1).
- Xiong, T., S. Tarikere, N. Rosser, X. Li, M. Yago, and J. Mallet. 2023. "A Polygenic Explanation for Haldane's Rule in Butterflies." *Proceedings of the National Academy of Sciences of the United States of America* 120, no. 44: e2300959120. <https://doi.org/10.1073/pnas.2300959120>.
- Yuan, Y.-W., J. M. Sagawa, R. C. Young, B. J. Christensen, and H. D. Bradshaw. 2013. "Genetic Dissection of a Major Anthocyanin QTL Contributing to Pollinator-Mediated Reproductive Isolation Between Sister Species of *Mimulus*." *Genetics* 194, no. 1: 255–263. <https://doi.org/10.1534/genetics.112.146852>.
- Yurchenko, A. A., H. D. Daetwyler, N. Yudin, et al. 2018. "Scans for Signatures of Selection in Russian Cattle Breed Genomes Reveal New Candidate Genes for Environmental Adaptation and Acclimation." *Scientific Reports* 8, no. 1: 12984. <https://doi.org/10.1038/s41598-018-31304-w>.
- Zhang, J., P. Kapli, P. Pavlidis, and A. Stamatakis. 2013. "A General Species Delimitation Method With Applications to Phylogenetic Placements." *Bioinformatics* 29, no. 22: 2869–2876. <https://doi.org/10.1093/bioinformatics/btt499>.
- Zhang, W., E. Westerman, E. Nitzany, S. Palmer, and M. R. Kronforst. 2017. "Tracing the Origin and Evolution of Supergene Mimicry in Butterflies." *Nature Communications* 8, no. 1: 1269. <https://doi.org/10.1038/s41467-017-01370-1>.

Supporting Information

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