How the Easter Egg Weevils Got Their Spots: Phylogenomics reveals Müllerian Mimicry in *Pachyrhynchus* (Coleoptera, Curculionidae).

Matthew H. Van Dam*[1,3], Analyn Anzano Cabras[2], Athena W. Lam[3]

[1]Entomology Department, Institute for Biodiversity Science and Sustainability, California Academy of Sciences, 55 Music Concourse Dr., San Francisco, CA 94118, USA

[2]Coleoptera Research Center, Institute for Biodiversity and Environment, University of Mindanao, Matina, Davao City, 8000, Philippines

[3]Center for Comparative Genomics, Institute for Biodiversity Science and Sustainability, California Academy of Sciences, 55 Music Concourse Dr., San Francisco, CA 94118, USA

*correspondence sent to: mvandam@calacademy.org

ABSTRACT

The evolutionary origins of mimicry in the Easter Egg weevil, *Pachyrhynchus*, have fascinated researchers since first noted more than a century ago by Alfred Russel Wallace. Müllerian mimicry, or mimicry in which two or more distasteful species look similar, is widespread throughout the animal kingdom. Given the varied but discrete color patterns in *Pachyrhynchus*, this genus presents one of the best opportunities to study the evolution of both perfect and imperfect mimicry. We analyzed more than 10,000 UCE loci using a novel partitioning strategy to resolve the relationships of closely related species in the genus. Our results indicate that many of the mimetic color patterns observed in sympatric species are due to convergent evolution. We suggest that this convergence is driven by frequency-dependent selection.

INTRODUCTION

Mimicry as a driver of species diversification

Mimicry is central to evolutionary biology and has important implications for various evolutionary processes, such as adaptive radiation[1], coevolution[2], niche partitioning[3,4], population structuring [5], adaptation of chemical defense[6]. In particular, Müllerian mimicry is hypothesized to be a primary driver of diversification resulting from mutualistic evolution[7–10]. This type of mimicry describes an antipredator strategy for groups of unpalatable species—by sharing similar color and pattern, each individual prey experiences a lower probability of being mistaken as a food source by predators[7]. Müllerian mimicry and aposematism have been the subject of many fascinating studies in a wide range of taxa: butterflies [11–13], net-winged beetles[14], velvet ants [15], spiny plants [16], dart frogs[17,18], vipers[19], coral snakes[20], fish[21], and possibly toxic birds[22].

Pachyrhynchus is a diverse and charismatic group of beetles with elaborate patterning and iridescent colors (Figs. 1-2). Beetles in this genus have hard cuticles with fused elytra, which presents a line of defense against predation[23,24] , and they are hypothesized to be distasteful to their predators, which consist of birds, lizards, and frogs[23,25]. Their hard cuticle is derived in part through the bacteria endosymbiont *Nardonella*, which produces all precursors to tyrosine, a key amino acid for cuticular hardening[26]. *Pachyrhynchus* distinctive color patterns are structural colors produced by their scales' inner nanostructure which scatters incident light[27].

Both Batesian and Müllerian mimicry are associated with and found within *Pachyrhynchus*. The first record of mimicry among Pachyrhynchini was noted by

Wallace[28] when he observed sympatric species with the same colors and elytral patterns. It was also noted by Schultze where he provided a list of 19 sympatric species of *Pachyrhynchus, Metapocyrtus* (also in the Pachyrhynchini), and *Doliops* (Cerambycidae), all sharing the same coloration and patterns between genera[25,29]. He reported 14 additional sympatric species of *Metapocyrtus* exhibiting very similar elytral patterns, only distinguishable from each other by close inspection of diagnostic characters of the rostrum[29]. Their distinctive patterns are also observed in other unrelated weevils (e.g., *Polycatus*, *Eupyrgops*, *Neopyrgops*, *Alcidodes, Coptorhynchus, Calidiopsis*). Following the typical Batesian model, long-horned beetles (e.g., *Doliops*, *Paradoliops*) and even a cricket mimic *Pachyrhynchus*' aposematic signals are thought to aid in avoiding predation[28]. In this study, we explore patterns of Müllerian mimicry in *Pachyrhynchus* through detailed study of the group's phylogenetics, biogeography, and ancestral state reconstructions.

Pachyrhynchus' natural history and taxonomy

The genus *Pachyrhynchus* was established based on the species *Pachyrhynchus moniliferus* from Luzon Island, Philippines, and placed in the tribe Pachyrhynchini (Germar). The tribe is entirely flightless with 18 described genera and more than 500 described species. *Pachyrhynchus* is endemic to oceanic islands but the entire genus is flightless. It is distributed in the Philippines (excluding Palawan), Ryukyu Island, Green and Orchid Island in Taiwan, and Talaud and Moluccas Island in Indonesia[25,30,31], with the highest diversity in the Philippines. There are approximately

145 species of *Pachyrhynchus*, 93% of which are Philippine endemics[25,32]. Currently, despite the recent interest in the group[31–35], *Pachyrhyhnchus* taxonomy and the relationships between different species remains poorly understood. Although naturalists and scientists have been collecting and describing *Pachyrhynchus* for centuries, there remains no concerted effort to study their phylogenetic relationships in a robust manner. While some species have been well quantified using morphometrics combined with DNA characters, the taxonomic scope of these studies are limited to only a few species found on Green and Orchid Islands of Taiwan[33]. The foundations of *Pachyrhynchus* taxonomy remains problematic, Heller[36] and Schultze[25] tried grouping species of *Pachyrhynchus* according to elytral color patterns. This was flawed because a heterogeneous set of unrelated species may lead to polyphyletic classifications. For example, classifications based on color pattern alone may cause over splitting of species because of intraspecific variation, or, conversely, over lumping of species because of convergent coloration and patterns. *Pachyrhynchus* species concepts remain largely untested by more complete morphological data and genetic data as well. We hope to provide a basic structure for the relationships within the genus to advance towards a more meaningful discussion on this group's mimicry and biology.

Biogeographic model

The Pleistocene Aggregate Island Complex (PAIC) which has been used as a biogeographic framework in the Philippines for decades, refers groups of islands in the Philippines with less than -120 m isobath separation that coalesced in the Pleistocene

Epoch due to lowering of sea levels during glacial cycles[37–42] . These groups of islands are centers of biological endemism and share a high percentage of common floral and faunal elements. There are nine biogeographic subregions in the Philippines such as Batanes PAIC, Babuyanes PAIC, Luzon PAIC, Mindoro PAIC, Romblon PAIC, Palawan PAIC, Mindanao PAIC and Sulu PAIC[40] of which the three largest biogeographically significant sub-provinces are Greater Luzon PAIC, Mindanao PAIC, and Greater Negros–Panay[38]. Studies on the distribution of mammals, birds, and herpetofauna have shown the consistency of PAIC boundaries on species distribution[38,40]. The connectedness of these islands allowed for the exchange of flora and fauna amongst the islands. The Philippines which has a unique geologic history including; paleo-island accretion, late Pleistocene sea-level fluctuations, and landmass connectivity has led to the isolation of lineages and restricted distributional patterns such as in the case of *Pachyrhynchus*.

Here, we use both dry pinned and newly collected specimens from the last 30 years and more than 10,000 UCE loci to gain a more thorough understanding of the *Pachyrhynchus* mimetic system. Our results are the first well-supported phylogeny of this charismatic and threatened beetle genus.

We explore the following study questions:

1) Müllerian Mimicry - We aim to identify how sympatric *Pachyrhynchus* species with similar patterns acquired their coloration. We tested if their similarity is due

to inheritance alone or is the color pattern independently evolved in one or more sympatric lineages.

2) Intraspecific polymorphism - Some species of *Pachyrhynchus* display a striking polymorphism where some individuals possess solid maculations while others exhibit the same pattern but is not “filled” (i.e. same outline but lacking scales in the middle). We tested whether this trait is constrained to a single lineage within *Pachyrhynchus* or if it is more widespread throughout the genus.
3) Biogeography - Does *Pachyrhynchus* follow the Pleistocene Aggregate Island Complex (PAIC) hypothesis[37,43,44] , where speciation follows the interglacial periods when the land masses of the Philippines were at their most isolate in the last ~3MYA?

METHODS

DNA isolation and Quality Check

We used 71 pinned and 16 ethanol preserved specimens for this study. For a complete list of voucher specimens, see Supplemental Table S1. Voucher specimens are stored in two museum collections: the Coleoptera Research Center, University of Mindanao, Philippines (CRC) and the California Academy of Sciences, Entomology Department, USA (CASENT). To minimize contaminants, specimens were carefully dissected, excluding exoskeleton and beetle guts. In some specimens, contaminants

146 are apparent upon examination (e.g., yeast, fungus and mites are observed in the body
147 cavity), in those cases legs were removed and punctured to allow enzymatic
148 (proteinase-K) digestion of soft tissue. DNA was extracted from the resulting tissues
149 using the QIAamp micro kit (Qiagen, Germany) following the manufacturer's protocol.
150 We assessed the quantity of all isolated DNA using a Qubit 2.0 Fluorometer (Invitrogen,
151 USA DNA quality (i.e., fragment size distributions) was determined using 1% agarose
152 gel in newly collected specimens and historic samples with low starting concentrations,
153 with a 2100 BioAnalyzer (Agilent Technologies, USA). We found that our starting DNA
154 quality and quantity varied significantly among specimens: from 4.8 ng–5000 ng and
155 200 bp–50 kbp. This is not surprising as collection dates and storage conditions of each
156 specimen varied widely.
157 High molecular weight (HMW) DNA for 10X Genomics library construction was
158 extracted from a single newly collected, dry frozen, *P. miltoni* specimen, using
159 MagAttract HMW DNA Kit (Qiagen, Germany) and following manufacturer's protocol.
160 DNA fragment size was quantified by pulsed-field capillary electrophoresis using the
161 Femto Pulse System (Agilent Technologies, USA).
162
163 *Library Construction*
164
165 We obtained paired-end Illumina data from 85 whole genome libraries, and an
166 additional three from another studies (Van Dam et al. 2021, in prep). When necessary,
167 DNA was sheared using a Covaris M220 (Covaris Inc., USA). Libraries were
168 constructed with the NEBNext® Ultra™II DNA Library Preparation kit (New England

Biolabs Inc, USA) following the manufacturer's protocol. To minimize PCR replicates in the final dataset, titration of the number of PCR cycles was performed as described in Belton et al. 2012[45]. Average sizes of the final libraries were around 250 bp (for libraries constructed from historic samples) to 400 bp (for libraries constructed using newly collected material).

The 10X Genomics linked-read library for Pachyrhynchus miltoni was prepared at QB3 Genomics at the University of California, Berkeley

Low Coverage Genome Assembly

First, we trimmed adapters and low quality bases from the ends of our reads with *fastp* version-0.20.0 using the "detect_adapter_for_pe" setting[46]. Because we used two lanes of sequencing for some of our libraries, we concatenated the forward and reverse reads of these two lanes. We then concatenated unpaired reads into a single file for each species. Next, we used *SPAdes-3.11.1*[47] to assemble the reads into scaffolds with k-mer values of 21, 33, 55, 77, 99 and 127 in the *SPAdes* assembly pipeline, using default settings for everything other than the memory ("-m 800") and cpu threads ("-t 32").

Genome Assembly of Pachyrhynchus miltoni

We used 594M illumina 2x150 paired reads which had been barcoded by the 10X Genomics Chromium instrument. The 10X Genomics linked-read assembly was constructed with Supernova v2.0.1[48] with default settings.

Ultraconserved Element (UCE) Marker Design

We designed a custom ultraconserved element (UCE) probe set using the PHYLUCE pipeline[49,50]. To maximize the effectiveness of the probes, we selected individuals that spanned the phylogenetic diversity of the Pachyrhynchini. For the base taxon, we used the chromosome level genome assembly of *Pachyrhynchus sulphureomaculatus* Schultze, 1922[52]. This taxon was used to select the initial bates because: 1) preliminary investigation of *P. sulphureomaculatus* indicated that it is not recovered at distal portions of the *Pachyrhynchus* phylogeny, which has been demonstrated to increase the number of UCE loci captured[52], 2) the genome is complete and soft masked for repetitive elements, and 3) perhaps most important, the genome is free of contamination, which can lead to off-target loci capture (Van Dam et al. 2021, in prep). Soft masked genome is critical because the PHYLUCE pipeline only screens for repetitive DNA if the base genome is soft masked. If the genome is not soft masked, probes may be designed from paralogous loci. To diminish, as much as possible, the possibility of designing loci from non-target species in the probes themselves, we used a stringent screening method for our probes, selecting only loci that were recovered in our base taxon (Van Dam et al. 2021, in prep). We included eight *Pachyrhynchus* species from different species groups (Supplemental Figure S1).

We selected two outgroup species, close relatives to *Pachyrhynchus*: *Coptorhynchus* and *Oribius* (Celeuthetini)[53], to ensure that the probe set is relatively universal across the phylogenetic diversity of our target species, for a total dataset of ten species.

Extracting UCE Loci and Alignment Construction

To match and extract our probe set from the *Pachyrhynchus* scaffolds and then align our loci we followed the PHYLUCE pipeline[49] using default settings unless otherwise noted. We used the PHYLUCE script "phyluce_probe_run_multiple_lastzs_sqlite" with an "identity" of 60. After matching probes to scaffolds, we used "phyluce_probe_slice_sequence_from_genomes" to extract the flanking 500 bases around our probes. After the initial alignment step using *mafft*[54] we used "phyluce_align_get_trimal_trimmed_alignments_from_untrimmed" to internally trim our matrices. This step uses *trimAl*[55] to help trim ambiguously aligned sites in the alignments. We used "phyluce_align_get_only_loci_with_min_taxa" to select loci with a minimum of 50% complete matrices. Lastly, we used "phyluce_align_format_nexus_files_for_raxm" to produce the final concatenated matrix.

UCE Phylogenomics

We used two different types of analyses for phylogenetic reconstruction: (1) a concatenated analysis using RAxML-NG v1.0.0[56], and (2) a summary species tree analyses using ASTRAL-MP v5.7.4[57,58].

237

238 *Concatenated Phylogenetic Analyses*

239

240 We used the General Time Reversible + gamma (GTRGAMMA) site rate
241 substitution model across our alignment. We used 10 independent parsimony-based
242 starting trees for our maximum-likelihood (ML) searches in RAxML-NG. Non-parametric
243 bootstrap replicates (BS) were done using the autoMRE option, with a maximum of 200
244 replicates to optimize the number of bootstrap replicates for this large dataset. Lastly,
245 we mapped the bootstrap replicate values onto the best-scoring ML tree.

246

247 *Species Tree Analyses*

248

249 *Identifying UCE loci within genomic windows:* We used methodologies that are similar to
250 Van Dam et al. 2020[51] for combining UCE loci. First, we identified UCE loci that are
251 within 25 kb non-overlapping windows, as similar window sizes were used in Edelman
252 et al. 2019[59]. To accomplish this, we divided up the base genome's chromosomes into
253 25 kb non-overlapping windows using the "*tileGenome*" function in the *GenomicRanges*
254 package[60] in *R*. Next, we extracted the genomic coordinates from these windows and
255 the coordinates produced in PHYLUCE for the probe set, and then we identified
256 overlapping ranges using the "*GRanges*" function in the *GenomicRanges* package. If
257 there was a UCE that overlapped two windows, it was combined with the adjoining
258 UCEs to maximize the number of UCEs within a window.

259

Partitioning procedures: Data preperation before partitioning and partitioning procedures were carried out following Van Dam et al. 2017[62]. Before we partitioned UCE loci, we first used the R package *ips*[61] with an *R* script from Van Dam et al. 2017[62]. We removed any columns composed exclusively of "-","n" and/or "?" using the "deleteEmptyCells" function, followed by removing any ragged ends of the matrix with the "*trimEnds*" function, with a minimum of four taxa present in the alignment. Below we expand on partitioning procedures to consider the potential site rate heterogeneity found in UCE loci as well as partition loci found in the same genomic window or potentially the same gene[51]. We use an alternative partitioning scheme from Van Dam et al. 2017 and Tagliacollo and Lanfear 2018[62,63]. Although the central core regions of UCEs tend to be more conserved than the flanking regions [49,62–64], this observed amount of variability is greatly reduced due to internal trimming procedures in the PHYLUCE pipeline, which is necessary to accurately align loci. Internal trimming reduces the variability of informative sites to present a relatively even distribution across these loci[64], essentially taking the U-shaped distribution and reducing it to a more or less flat line[64]. With this in mind, we unlinked the character sets of the flanking regions and used *PartitionFinder2* v2.1.1[65] to group these separate partitions based on the best fitting model and substitution rates for the data. First, we divided the loci into a central core region of 80 bp because some loci are reduced by internal trimming to be less than the original 160 bp of the probes. Next, we divided each of the flanking regions into 5 separate character sets based on their proportion of sequence length (5 to the left of the central core and 5 to the right). For the loci that shared genomic windows, we combined character sets; thus, if two loci were concatenated, there would be 22-character sets total. These character sets were

283 then input into *PartitionFinder2* for partitioning and model selection. Because RAxML-
284 NG uses a more diverse set of nucleotide substitution models than previous versions of
285 RAxML[56], we tested for the best model fit from a total of 39 different models
286 (Supplemental material, model list).
287
288 *Gene tree and species tree reconstruction:* We used ten independent parsimony-based
289 starting trees for our maximum likelihood (ML) searches in RAxML-NG. We then
290 performed 100 non-parametric bootstrap replicates, followed by mapping the bootstrap
291 replicate values onto the best-scoring ML tree. Next, we collapsed/contracted branches
292 in the gene trees with BS ≤20 using newickutils[66]; which has been demonstrated to
293 have a strong positive impact on the accuracy of species tree reconstruction[67,68]. These
294 resulting trees were used in species tree reconstruction. We used ASTRAL-MP with the
295 default settings to reconstruct the species tree and annotate the tree with support
296 values calculated for the normalized quartet support (NQS) and local posterior
297 probability (LPP)[69].
298
299 *Divergence Dating Biogeography*
300
301 We used MCMCTREE[70] to perform our divergence dating analyses. We used the
302 topology from our ASTRAL analyses as the starting topology for the MCMCTREE. No
303 fossils exist for our ingroup or near relatives, so we used a geological calibration for the
304 maximum age of the Philippine Islands. We used 25–30 Ma as the root node of the
305 Pachyrhynchini, an approximate date for the emergence of the Philippine proto-islands

306 proposed by Hall 2002. We used MCMCtreeR[71] to estimate a normal distribution around
307 the maximum age of our calibration point as well as to format the tree file for
308 MCMCTREE. Next, we used the aforementioned tree to obtain a rough estimate of the
309 substitution rate using basml[72]. To help accomplish this, we randomly selected 300 loci,
310 using many more will prevent the analysis from completing, as well as using more data
311 is not necessary to approximate the uncertainty of the divergence dates[72,73]. Finally, we
312 estimated the gradient and Hessian of the branch lengths[74] to assist in the final
313 estimation of our divergence dates.

314 To reconstruct the broad scale biogeographical patterns of *Pachyrhynchus*, we
315 used BioGeoBEARS v1.1.2[75]. Because biogeographic model selection is sensitive to
316 duplicate taxa, we removed all potential duplicates from the same metapopulation
317 lineage/species[75]. We wanted to examine if the biogeography followed the Pleistocene
318 Aggregate Island Complex (PAIC) hypothesis[37,43,44], and thus we defined our areas
319 according to the PAIC scheme. Many of the defined areas are congruent with the
320 various geological histories of the major island groups[42,76–78] (Fig. 4). We followed this
321 scheme except for the island of Marinduque because it contains a number of unique
322 species which do not have any obvious close relatives in the Luzon PAIC. For this
323 island, we were curious to see its colonization history separate from the Luzon PAIC.
324 We initially examined three different biogeographic models, DEC[79], DIVA-like[80],
325 BAYAREA-like[81] (see Table 3), and also included the "+J" parameter for founder-
326 event/jump speciation at cladogenesis events[75]. This parameter has been demonstrated
327 to greatly improve model fit for island- and island-like systems[75,82,83]. We used the

328 Akaike information criterion corrected for sample size (AICc) to identify which model
329 best explained our data given the number of free parameters.
330
331 *Ancestral State Reconstruction of Color Patterns*
332
333 To identify instances of mimicry and/or convergence in the color pattern of
334 *Pachyrhynchus*, we used the R package *phytools* v0.7.70 "fitMk" [84] to reconstruct
335 ancestral character states. We categorized the color pattern into 11 different states (Fig.
336 5). We define 11 different color patterns on their elytra as follows: 1) "Black", species
337 whose integument is black or almost entirely black. 2) "Rainbow" elongate linear bands
338 or elongate maculations that nearly touch composed of yellow/orange color at distal and
339 apical ends gradually transitioning to blue in the center. 3) "Vertical bars" elongate linear
340 bands or elongate maculations that nearly touch composed of a single color running the
341 length of the elytra. 4) "Filled Moroccan tile", large central patches of color with three
342 apical and three distal areas not covered by colored scales. 5) "Open bands" three pairs
343 of lines that circumscribe an ovoid shape or run the width of an elytra. 6) "Moroccan tile"
344 net-like pattern of lines. 7) "Checker" two vertical lines near apex of elytra with a
345 horizontal central line and two vertical lines near anterior end of elytra. 8) "Spots" at
346 least three pairs of circular maculations. 9) "Irregular grid" two–four pairs of short
347 convergent or parallel lines near apex, a central horizontal band, and two–four pairs of
348 short convergent lines at the anterior end of the elytra. 10) "Vertical lines" narrow lines
349 running the length of the elytra. 11) "Filled bands" three pairs of broad ovoid
350 maculations running roughly the width of an elytra.

We then selected the best fitting model (equal rates ER, symmetric backward and forward rates SYM, or all-rates-different for transitions ARD) using the AIC weights as our selection criteria. All models treated character transitions as unordered. Next, we simulated 1000 different character histories using *phytools*' "make.simmap" function using the "mcmc" option to estimate the transition rate matrix, under our best fitting model. Lastly, because some *Pachyrhynchus* species possess discrete polymorphisms, we wanted to examine the history of this trait as well. The polymorphic trait in *Pachyrhynchus* presents itself as the color pattern being "filled" (solid center) or "open" (outlined) maculations (Fig. 4). The addition of this polymorphic state to the previous 11 character states (Fig. 4) resulted in a transition matrix that was much too large to analyze. To examine how widespread open or filled maculations are across the phylogeny, we coded it as two different polymorphic states (filled, not filled, or species that have both states). We performed the same model selection and ancestral state reconstruction methods as above but using *phytools*' "*fitpolyMk*" and also included the model "*transient*", where the rates of gaining or losing polymorphic states differ. We combined this information with their biogeography to identify whether color patterns were due to inheritance (related species look alike via shared ancestry of trait), convergence (species look alike but independently evolve a trait in allopatry), or mimetic convergence (species look alike, and of those that are similar, at least one lineage independently evolved the trait in sympatry).

RESULTS

374 ***Genome Assembly of Pachyrhynchus miltoni***

375 The 10X Genomics linked-read assembly resulted in a total length of
376 2,204,292,973 bp (roughly the same size as the *P. sulphureomaculatus* base
377 genome[51]), with an N50 of 10,763 bp and 368,426 scaffolds. To use the assembly in
378 PHYLUCE, we modified the headers of the scaffolds to resemble those from SPAdes,
379 giving each a unique name and matching sequence length of the scaffold.

380

381 ***Probe Design, Sequencing Results, and UCE Loci Recovery***

382

383 Our probe design selected a total of 12,522 UCE loci shared by all 11 taxa. Our
384 sequencing results were largely successful with only one DNA library producing less
385 than a million reads, and this sample was removed from all subsequent analyses.
386 Filtering loci that contain ≥ 50% complete matrices resulted in 10,108 of 12,323 total loci
387 recovered. We sequenced an average of 77,990,397 ±11,627,118 95% CI, paired reads
388 per sample (Supplementary Fig. S2), with a final 50% complete matrix with a mean of
389 8,756 ± 698 95% CI per sample (Supplementary Fig. S2). These statistics exclude loci
390 recovered from the long read or pseudo long read assemblies, which recovered an
391 average of ~11,000 UCEs (Supplementary Table S1). We found a total of 2,995 UCEs
392 sets in 25 kb windows, with a total of 1,332 windows containing multiple UCEs and
393 7,113 UCEs not in sets (Supplementary Fig. S2).

394

395

396

397

Length of UCE sets	2	3	4	5	6	7
Number of sets by length	1087	184	44	11	4	2

Table 1. Count of UCEs by the number of UCEs concatenated in a "set". Upper row gives the number of UCEs concatenated in a 25 kb non-overlapping window, and the lower row is the count by category.

Loci type	ABS_mean	ABS_median	ABS_95%_CI	Number of loci
Sets	55.4	58	0.67796	1332
Singles	48.1	51.9	0.41227	7113

Table 2. Average bootstrap support by locus type. "Sets" are concatenated UCEs found in a 25kb non-overlapping sliding window bin, and "Singles" are those where only a single UCE occupies a 25kb bin in the *P. sulphureomaculatus* base genome.

Phylogenetic Analyses

Concatenated analysis: Our alignment totaled 6,353,380 bp in length. The analysis completed 200 non-parametric bootstrap replicates, and we provided the phylogenetic hypothesis for the genus in Supplemental Fig 1S. BS support was relatively high throughout the tree (Supplementary Fig. S1), with strong support along the backbone (BS =100%). The genus *Pachyrhynchus* was recovered as monophyletic, but the genus *Metapocyrtus* was not. The genus *Pantorhytes* is sister to the Philippine members of the Pachyrhynchini, and the Celuthenini taxa are sister to Pachyrhynchini.

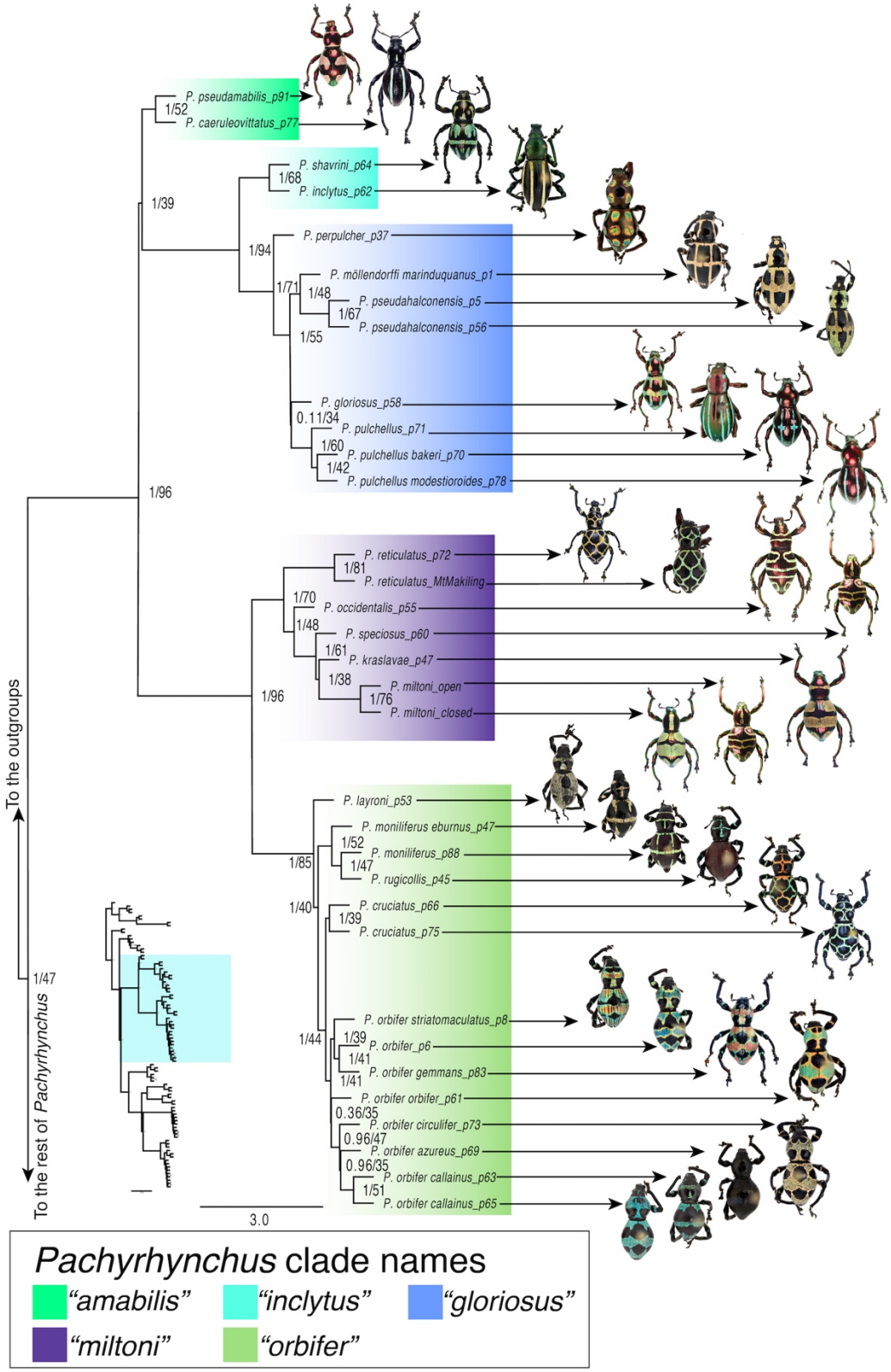
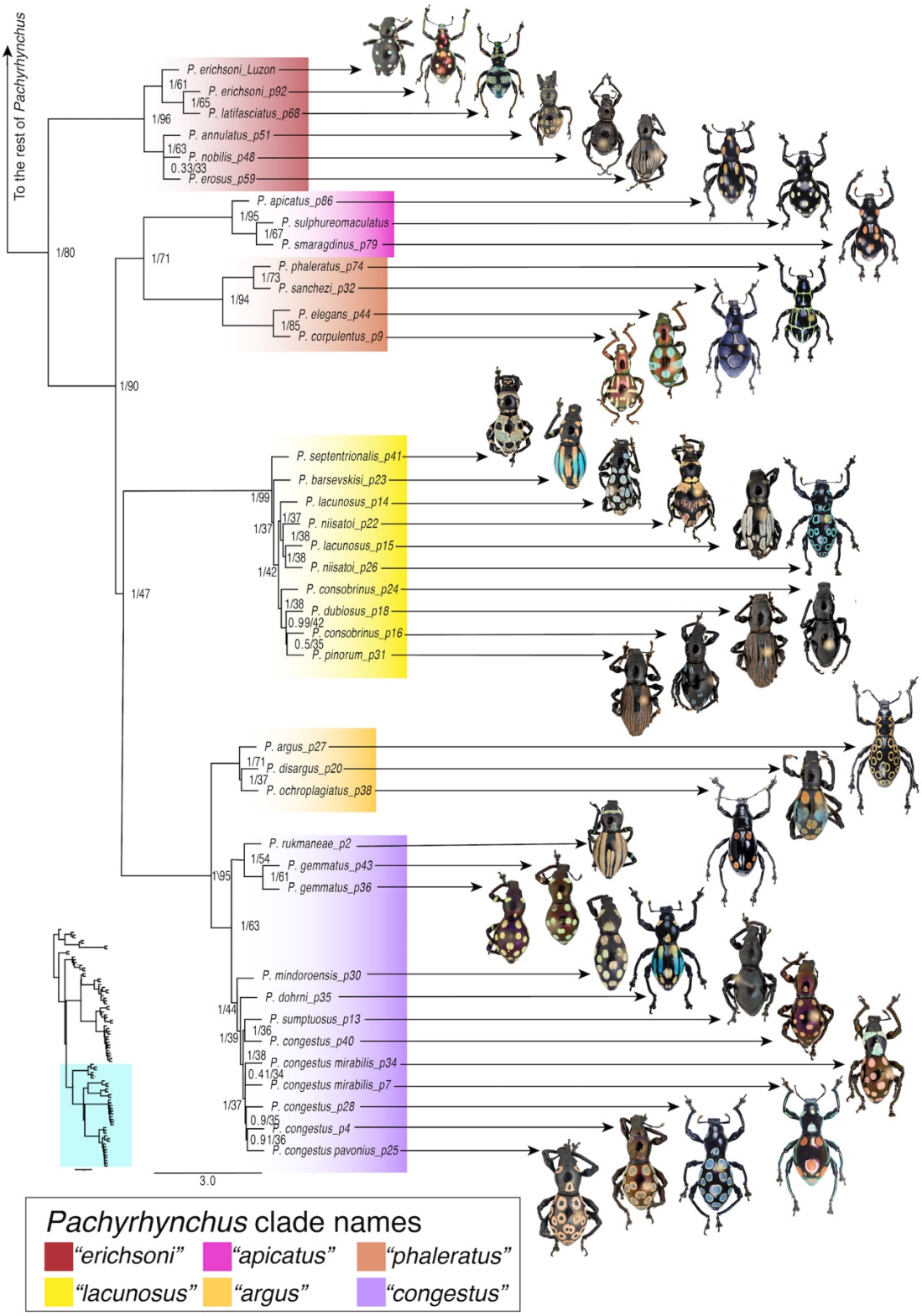


Figure 1. Phylogeny of the *Pachyrhynchus* ASTRAL species tree, part one. The other half of the tree is found in Fig. 2. Node labels correspond to local posterior probability (LPP) and normalized quartet support (NQS).

421



422

423 **Figure 2. Phylogeny of the *Pachyrhynchus* ASTRAL species tree, part two.** The other half of the tree
424 is found in Fig. 1. Node labels correspond to local posterior probability (LPP) and normalized quartet
425 support (NQS).

Species tree analyses: We present the first genus wide phylogeny, and recovered 11 major lineages (Figs. 1–2). We found surprising patterns that contradicts the idea that pattern alone is a sufficient character to group *Pachyrhynchus* species. For example, *P. speciosus*, previously described as a single species, was composed of several unrelated lineages[85]. In addition, *P. reticulatus* and *P. cruciatus* were actually rather distantly related despite similar "Moroccan tile" like patterns (Fig. 1).

The partitioning results of the gene trees showed that most UCEs require more than three partitions; with the mean number of partitions in single UCEs being 3.6 ± 0.04, and the mean number of partitions in UCE sets 17.8 ± 0.65. These results indicate that UCE loci have different rates and models of nucleotide substitutions as well, not just a central core and symmetrically variable flanking regions. In addition, the substitution models selected were varied, with the traditional GTR+G selected second-most frequently (Fig.3). The topology of the species tree reconstructed with ASTRAL had an identical backbone to the tree from the concatenated analysis and only a few differences at nodes near the tips (Supplementary Fig. S3), but all species remained in the same clades (Figs. 1–2) in both trees. The local posterior probabilities (LPP)[69] were similar to the BS values of the RAxML-concatenated analysis, with high support along the backbone of the tree and lower support near the tips, where conflict between topologies was typically due to short internode distances (Supplementary Fig. S3). Normalized quartet support (NQS) values were similar to the other support values (Figs. 1–2) with most of the gene trees congruent with the species tree topology along the backbone, as well as at nodes separated by relatively long internode distances. However, nodes separated by very short distances had more conflict among the gene

tree topologies, with only ~35–40% of the gene tree quartets in agreement. The number of loci and our computer resources impaired the completion of the 100 BS replicates. However, the statistics we used are a more reliable measure of gene tree conflict and support for species trees[69].

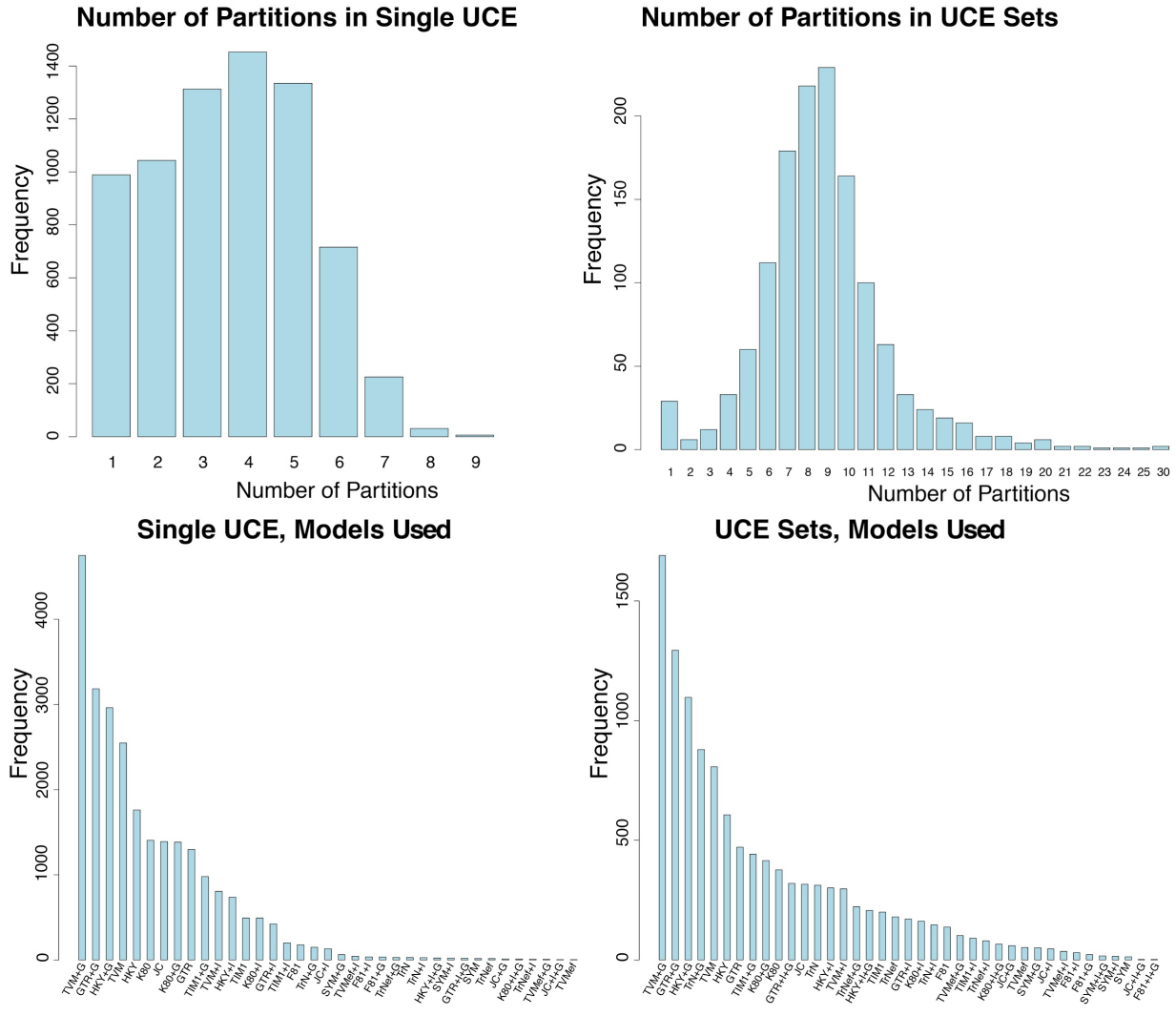


Figure 3. Barplot of UCE loci partitioning results. UCE "Sets" are concatenated UCEs which occupy a 25 kb non-overlapping sliding window bin, and "Singles" are those where only a single UCE occupies a 25 kb bin in the *P. sulphureomaculatus* base genome. Upper panels, barplot of the number of partitions found in a UCE locus. Lower panels, barplot of the number of times a particular model of nucleotide substitution was used in a partition.

Divergence dating and Biogeography

The results of the MCMCTREE analysis estimated the root age from the 95% highest posterior density (HPD) of *Pachyrhynchus* to be between 25.1–17.9 Ma. Most *Pachyrhynchus* species have diversified within the last 5 Ma, and many species' HPDs straddle the Pliocene-Pleistocene boundary (Fig. 4). The complete results of the MCMCTREE analyses are shown in Supplementary Fig. S4.

The results of the BioGeoBEARS analyses are shown in Table 3 and Fig. 4. The BioGeoBEARS analyses gave the greatest AICc model weight to the DEC+J model (Table 3). All models favored adding the founder event speciation parameter to the model, indicating that this mode of dispersal was significant in forming the broad scale biogeographic pattern of the genus. The root node of *Pachyrhynchus* was reconstructed as a joint range between Mindanao and Luzon. However, the descendant nodes were reconstructed as most likely Luzon. The subsequent radiations of lineages on Mindanao all descended from Luzon lineages. These separate Mindanao lineages represent five independent colonization events. The lineages of Mindoro, however, represent a mix of biogeographic lineages, three from Luzon and two from Mindanao. The single taxon from Panay represents a founder event from Luzon. Marinduque, although part of the Luzon PAIC, had two separate founder events one from Luzon and one from Mindoro. There was no evidence of back-colonization.

Biogeographic model	LnL	numparams	d	e	j	AICc	AICc_wt
DEC	-80.65	2	0.0069	1.00E-12	0	165.5	2.60E-10
DEC+J	-58.12	3	1.00E-12	1.00E-12	0.038	122.7	0.53
DIVALIKE	-74.52	2	0.0089	1.00E-12	0	153.2	1.20E-07
DIVALIKE+J	-58.85	3	1.00E-12	1.00E-12	0.039	124.1	0.25
BAYAREALIKE	-100.1	2	0.0058	0.058	0	204.5	9.00E-19
BAYAREALIKE+J	-59.01	3	1.00E-07	1.00E-12	0.038	124.4	0.22

491 **Table3. BioGeoBEARS model results table.**

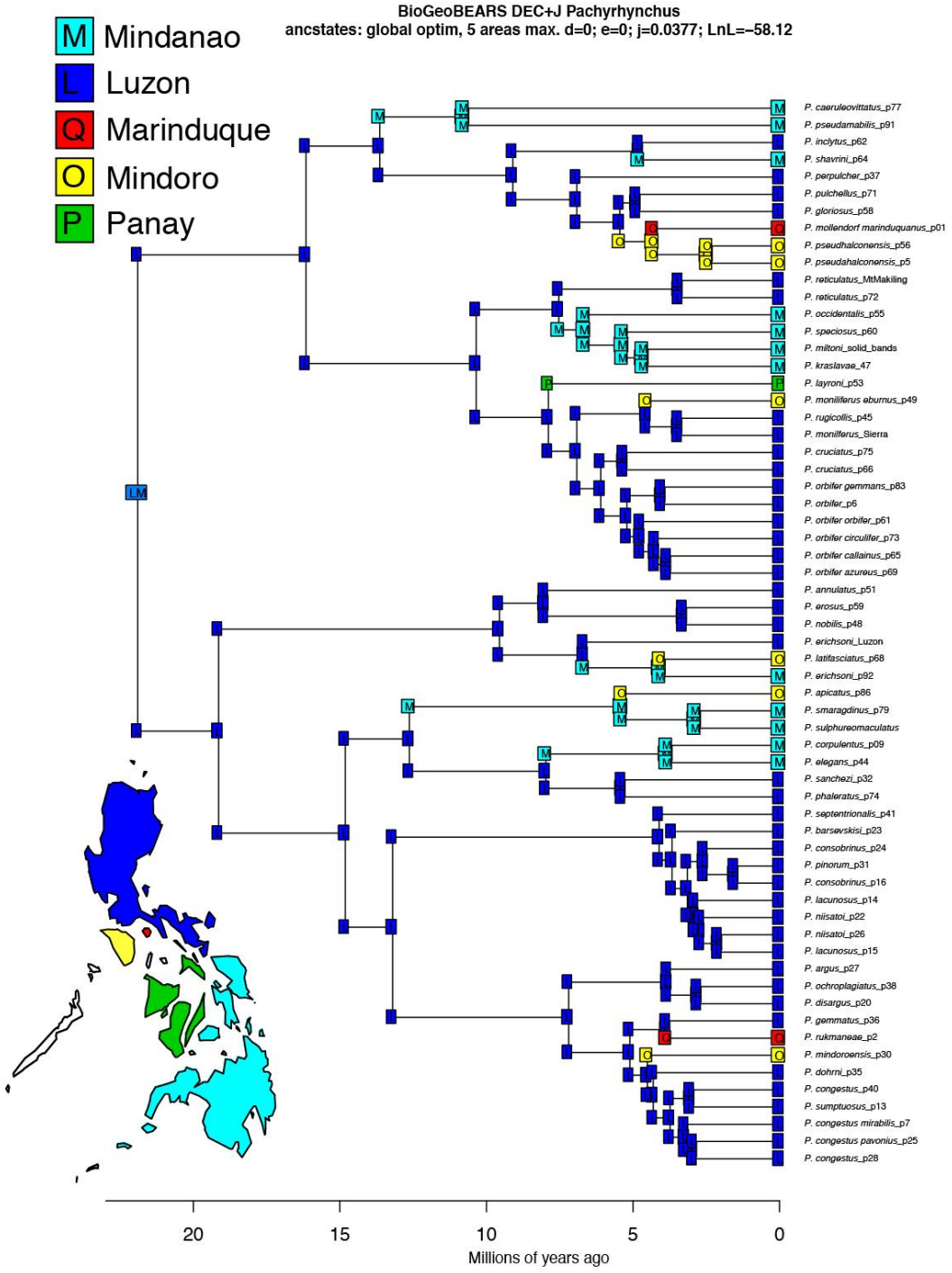


Figure 4. Biogeographic ancestral area reconstruction using *BioGeoBEARS* "*DEC+J*" model for *Pachyrhynchus*. Color codes correspond to Pleistocene Aggregate Island Complex (PAIC). Lower left color-coded map of the Philippines.

498 *Ancestral State Reconstruction*

499

500 The equal rates (ER) model was selected as the one that best fit our data given
501 the AIC weights (Table 4). The results of the ancestral state reconstructions of color
502 pattern show that few clades were restricted to a single color pattern. The root of the
503 tree was reconstructed as the spotted color pattern. The "*erichsoni"* and *"congestus"*
504 clades, approximately one half of the taxa, were predominantly reconstructed as
505 spotted. The changes from a spotted pattern to another state occurred along the
506 terminal branches in this clade (Fig. 5). The other half of the taxa showed changes at
507 deeper nodes, and their probabilities were not overwhelming given to one character
508 state. One of the most striking observations from the reconstruction was that patterns,
509 such as the net-like/Moroccan-tile pattern of *P. reticulatus*, occur independently in
510 several different places in the tree. The unique rainbow color pattern (blues and
511 yellows) also occurred in multiple clades (Figs. 2, 5).

512 For the analyses using the polymorphic Mk model framework, the "*transient*"
513 model of transitions between different character states was selected as having the best
514 fit to the data given the AIC weights. Here, the predominant pattern as well as the root
515 node were reconstructed as having the filled bands character state (Fig. 5). There were
516 six separate transitions from the filled state to the filled+open bands character state, two
517 of these occurring along terminal branches. Polymorphic species occur in all major
518 clades (Figs. 1–2, 5), except in the *"orbifer"* and *"gloriosus"* clades (Figs. 1, 5). In the
519 "*orbifer*" clade, there was a transition from the polymorphic state to the open state at the
520 node leading to *P. cruciatus*.

521

Mk model	df	AIC	AIC_wt
SYM	55	322.052	0
ER	1	249.767	1
ARD	110	410.323	0
Poly.Mk model	**df**	**AIC**	**AIC_wt**
SYM	2	95.562	0.217
ER	1	99.936	0.024
ARD	4	96.457	0.139
Transient	2	93.460	0.620

Table 4. Color pattern model selection results.

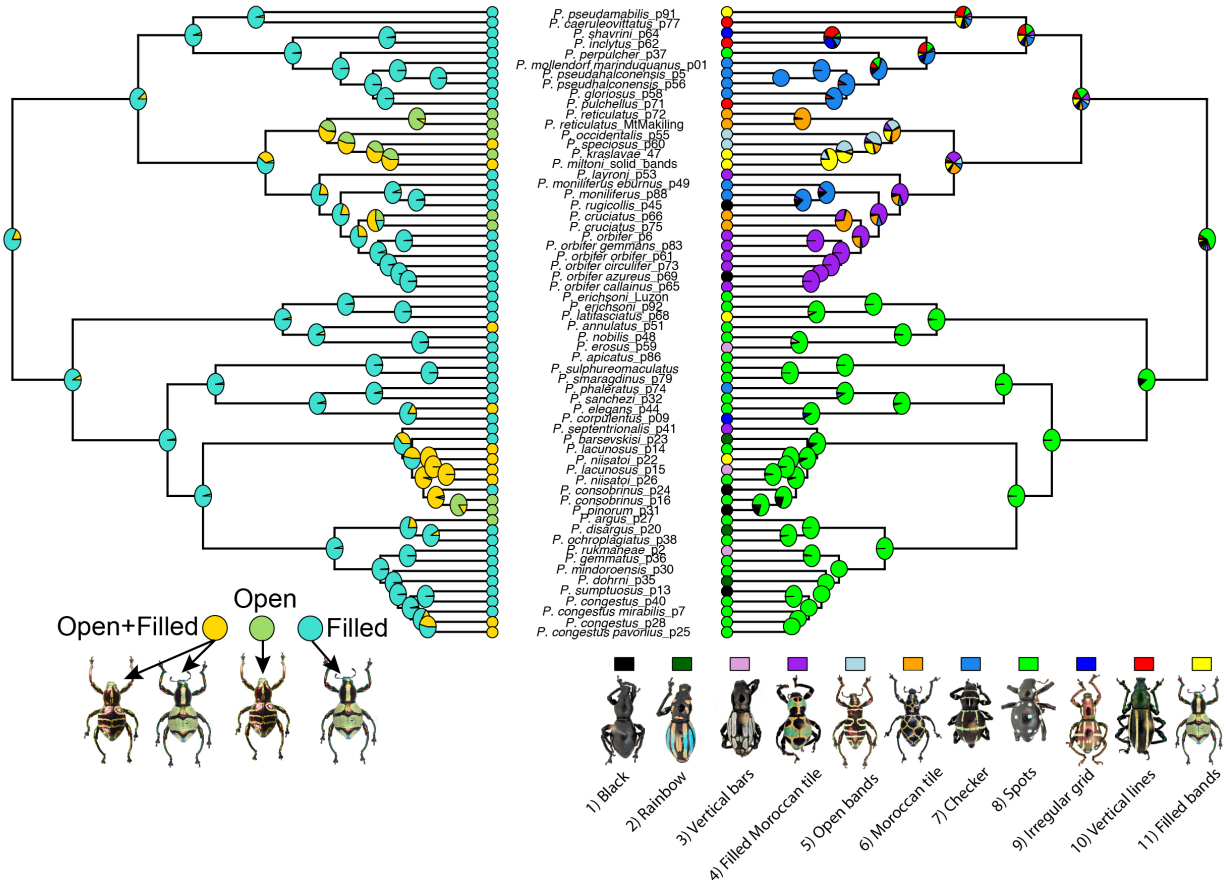


Figure 5. Ancestral state reconstruction of color patterns. Left, ancestral state reconstruction of polymorphic species states using *phytools* "*fitpolyMk*" function with the "*transient*" model. The character states were coded as "Open", "Filled", and "Open+Filled" where both states occur in a species. The polymorphic state can be associated with different color pattern states depending on the species (see lower right), for example "filled or open spots", but never two different patterns within the same species e.g. "filled *bands* and open *spots*". **Right**, ancestral state reconstruction using the "*fitMk*" function using the "*ER*" equal rates model. The 11 major color patterns observed in *Pachyrhynchus* are denoted in the lower right, colors correspond to those in the tree.

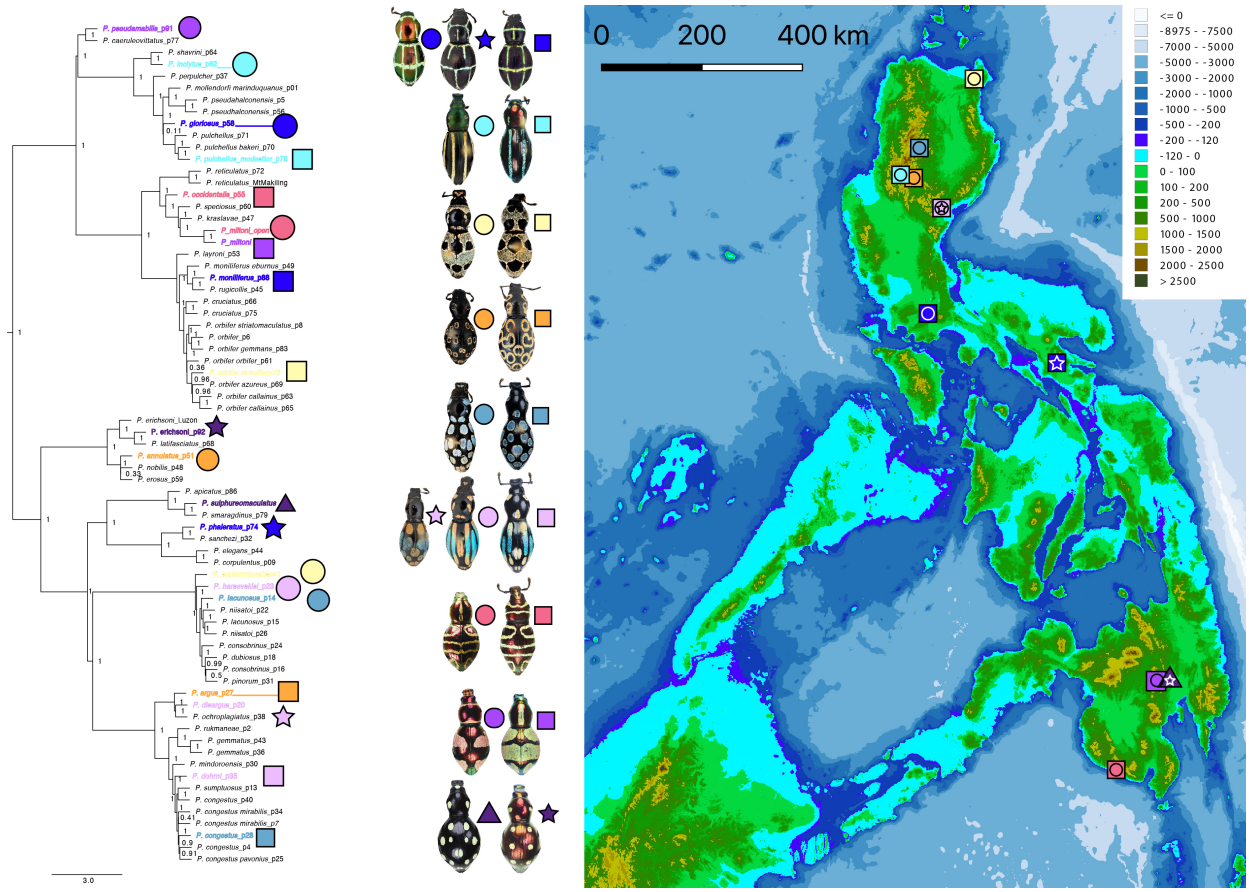


Figure 6. ASTRAL species tree of *Pachyrhynchus* with sympatric mimetic species. Colored symbols on phylogeny correspond to those in the central column and map. Node labels correspond to the LPP. **Central column**, mimetic species of *Pachyrhynchus*. **Right**, map of the Philippines, turquoise blue color denotes the -120m isobath, indicating land connection during Pleistocene glacial cycles.

DISCUSSION

We present the first phylogeny of the genus *Pachyrhynchus* based on 71 pinned and 16 ethanol preserved specimens using 10,108 UCE loci. Our results from both the concatenated and species tree analyses are highly supported with largely concordant topologies, demonstrating the benefit of Next Generation Sequencing and the value of historical museum specimens. With the phylogeny, biogeographic analysis, and ancestral state reconstructions of color patterns, we addressed how the wide array of

554 phenotypes originated and broad scale evolutionary trends in this Müllerian mimicry
555 system.
556
557 *Müllerian Mimicry in Pachyrhynchus*
558
559 For this study, we functionally define Müllerian mimetic groups as two or more
560 species of armored[23,24] *Pachyrhynchus* weevils that share similar color patterns and are
561 sympatric but do not share a most recent common ancestor. From our dataset, we
562 identified 11 color patterns and at least nine distinctive mimetic groups in this system
563 (Fig. 6). Members of each mimetic group closely resemble each other in their elytral
564 patterns, and they also occur in sympatry, presumably sharing the same populations of
565 predators (birds, lizards and frogs[23,25]). We found that most similar color patterns arose
566 independently in distantly related taxa (Fig. 6), with a divergence time up to 20 MY
567 between these taxa. For example, although *P. moniliferus* and *P. phaleratus* look
568 superficially similar, they are not each other's closest relatives, and are in fact at other
569 ends of the tree (Fig. 6, top row, circle and square). As well, some color patterns do
570 share common origins. For example, the *P. absurdus* species group and *P. speciosus*
571 species group, co-occur throughout the greater Mindanao PAIC (Fig. 6, sixth row). This
572 indicates that Müllerian mimetic evolution in *Pachyrhynchus* is mostly driven by
573 convergence in conjunction with rare examples of shared ancestry. The mixed evolution
574 of patterns has also been observed in *Heliconius* butterflies[12,86].
575 While some members of the mimetic groups strongly resemble each other,
576 others have only minor differences in pattern and/or color that are detectable by human

577 eyes. Recent studies suggest that Müllerian mimicry might not be as clear-cut as
578 originally proposed, and imperfect mimics and polymorphisms exist and may be
579 common in nature[14,87–90]. These studies suggest that varied forms of mimics are able to
580 persist due to limited cognitive capabilities of predators[91–93] and/or predator avoidance
581 of imperfect mimics due to the high cost of error [89,92,94]. Imperfect Müllerian mimicry
582 adds additional complexity to the diversification process and phenotypic variation in
583 *Pachyrhynchus*. For example, sympatric species in southern Mindanao, such as *P.*
584 *erichsoni*, *P. miltoni*, and *P. pseudamabilis* all have a metallic red coloration to their
585 cuticle, but differ in their maculations, spots or filled bands. (Fig. 6, bottom two rows).
586
587 Biogeography of the Pleistocene Aggregate Island Complex (PAIC)
588
589 We find that all of the deeper lineages of the genus originated well before the
590 PAIC was formed (Fig. 4). Additionally, within the major clades (Figs. 1–2), we find that
591 most of their diversification events occurred during the Pliocene with some at the
592 beginning of the Pleistocene (Fig. 4). This is significant because late Pliocene and early
593 Pleistocene periods consisted of glacial/interglacial cycles[95–97]. These interglacial time
594 periods should have promoted isolation between populations of *Pachyrhynchus*.
595 Another period of higher sea level is the mid-Pliocene when it was 2–3°C hotter than
596 today[98], the interglacial periods would have caused significant isolation and perhaps
597 promoted separate insular color morphs that would not have been compatible with
598 neighboring larger PAIC islands when they were connected during lower sea levels. We
599 see highly different color patterns today on such islands as Marinduque once connected

to Luzon during the Pleistocene. The colonization history of such nearshore islands also contributes to differences from the larger PAIC islands. The divergence date for many species largely coincides with the mid-Pliocene time period, and as many *Pachyrhynchus* species are confined to higher elevations this may have influenced divergence on the larger PAIC islands (e.g., Luzon and Mindanao) as well. However, as there is a wide 95 HPD around these divergence times, we cannot entirely rule out early Pleistocene cycles as well. One striking pattern about the biogeography of *Pachyrhynchus* is that there is no back colonization of Luzon given its relative proximity and size to other major PAIC island groups. One factor that may have promoted this is that species colonizing an already inhabited landscape with other *Pachyrhynchus* species would not have had color patterns that matched the local fauna. However, this also could be a random effect, or an artifact of our sampling. Lastly, more precise geological data for when different islands emerged above sea level would greatly benefit Philippine biogeography.

Insight into Pachyrhynchus Phylogeny and the Evolution of Color Patterns

Our phylogeny suggests that solely focusing on external morphology and color pattern may have driven inaccurate conclusions in previous taxonomic literature. For instance, *P. reticulatus* and *P. cruciatus* species share the superficial similarity of color pattern, while our phylogeny shows, with strong support, that they are distantly related, indicating convergence in color patterns. Although outside of the scope of this paper, our phylogeny could provide a great resource for taxonomic revision and species delimitation.

In addition to convergence, discrete polymorphism of patterns occurs frequently in populations and can complicate species identification if descriptions are based solely on morphology. Our phylogenetic results indicate that individuals with varied patterns should not be treated as separate species. For example, *P. miltoni* individuals may look different based on color pattern but are nearly identical genetically. We suggest that genetic data should be considered into future taxonomy and systematics of Pachyrhynchini beetles that have such complex mimetic evolution and biogeographic history. By integrating phylogenetic information to the taxonomy and systematics, the risk of creating synonyms or erroneous taxa will be diminished.

The type of binary polyphenism mentioned above is unusual in mimetic systems, with few examples in Coleoptera. One similar example occurs in the *Harmonia* ladybird beetles, but unlike *Pachyrhynchus*, the patterning is more of a continuum[99]. Additionally, in the ladybird beetle system, the color patterns are derived from pigments, and in *Pachyrhynchus* they are structural colors[27]. In Lepidoptera, discrete color polymorphism is more widespread and is known to occur in *Arctia plantaginis* tiger moths[100], between the sexes of *Neophasia terlooii* butterflies, males are white and females are orange[101], as well as in *Heliconius*[12].

We suggest that selection of color patterns is frequency-dependent. This hypothesis is supported by casual observations of *Pachyrhynchus* specimens in the CASENT collection, we counted the relative abundances of beetles with particular patterns in some well-studied localities. For instance, *P. moniliferus* is a widespread, relatively common species found in southern Luzon. It co-occurs with several other species with more restricted ranges. For example, in Iriga, Camarines, Luzon, 69 *P.*

moniliferus specimens (CASENT) were collected, but only three of the larger *P. phaleratus* (Fig. 6 top row, star and square) were collected at the same location and time period (May–Oct. 1931). In addition, at Mt. Makiling, Luzon, (May–Dec. 1930–31), 38 specimens of *P. gloriosus* were collected compared to 225 *P. moniliferus* (Fig. 6 top row, circle and square). Because all *Pachyrhynchus* species have a hard cuticle (often bending pins during preparation)[23], but have many different color patterns, frequency dependent selection is likely acting on the color pattern and not cuticular hardness. In the case of polymorphic species, frequency dependent selection may also be a cause of the observed frequencies. For example, in *P. miltoni*, found in the southern Davao City Province of Mindanao, the ratio of the filled band morphs to the open band morphs is 35:2, but in neighboring populations near Mt. Apo, the open-banded morph is most abundant (CRC and CASENT collections). Although more methodical studies should be done to control for artifacts caused by uneven sampling, different years of collection, seasonality, etc.; including information about polymorphism helps to elucidate how particular patterns evolved. For instance, *P. reticulatus* species with open bands were descended from polymorphic populations of *P. reticulatu*s inferred from our ancestral state reconstructions (Fig. 6). Because these beetles are flightless with patchy distributions, (some only restricted to a single mountain), the relative frequencies of these color morphs are likely to be fixed in isolated populations and is a possible cause for speciation. This could perhaps explain why we find convergent color patterns between allopatric species that are not each other's closest relatives.

UCE Partitioning

We have demonstrated that the universal use of a single model of nucleotide substitution is inadequate to accommodate site rate heterogeneity across UCE sets. The preferred use of multiple modes within a UCE locus has also been observed when constructing gene trees with MrBayes[62]. Additionally, we found that using more than three partitions was selected as optimal. While UCE loci have been treated as varying symmetrically around the central core[62,63,102], we find that often this is not the case as a variable, even number of partitions was selected. This suggests that flanking regions tend to be highly variable and that treating them as symmetrical units is a suboptimal partitioning design. More analyses are required to compare the design of Tagliacollo and Lanfear 2018[63] to the design we propose here to make a decisive decision of which scheme is optimal for UCE partitioning. The partition method we employ is the only one that accommodates the combination of neighboring or co-genic UCEs[51] (the method of Tagliacollo and Lanfear 2018[63] could be easily updated). The asymmetry in variation away from the UCE core is perhaps due to the way in which UCE loci are treated in the alignment process, mainly that difficult to align regions are trimmed in a less symmetric manner around the core of the locus.

Future Directions

Several unique features add intriguing complexity to our study system and, thus, inspire further studies. Color patterns in *Pachyrhynchus* are formed by the arrangement of scales, and the different scales' colors result from light reflectance on photonic crystals coupled with the background color of the elytra. The colors are structural[27],

differing from the pigment-based colors of most other Coleoptera[14] or *Heliconius* butterflies. This indicates a different genetic pathway that underlies their evolution and diversification. Our phylogeny provides a robust basis for further research to uncover the genetic mechanism controlling structural coloration in an evolutionary framework.

The armored exoskeleton is essential for a weevil's survival. Thick cuticle formation relies on the generation of its precursor amino acid, tyrosine, produced by obligate bacterial endosymbionts of the *Nardonella* lineage[26]. *Nardonella* has a much-reduced genome of 0.2 Mb and lacks genes responsible for most metabolic pathways; they rely on the beetle's metabolic output for survival[26]. Interestingly, *Pachyrhynchus*' elytra and cuticle are initially soft and easily deformed when they are teneral (soft bodied) adults, but the color patterns are apparent as soon as weevils emerge from the pupae[26] (personal obs. A Cabras). Their bold patterning likely provides protection from predators that have learned to avoid the aposematic signal in older adults. Predator avoidance during vulnerable stages is likely essential for survival until reproduction. More ecological studies on population densities of sympatric *Pachyrhynchus* species and predator preferences can also help to elucidate color pattern evolution on a landscape scale.

In summation, we clearly demonstrated that many of the mimetic color patterns observed in sympatry are due to convergent evolution and not simply due to inheritance. Based on our observations in natural history collections, we hypothesize that convergence between these independent color pattern forms is likely driven by frequency dependent selection. Lastly, the use of a UCE design specific to the tribe Pachyrhynchini was highly successful in resolving the relationships between our taxon.

716 Our study presents an interesting system of Müllerian mimicry and provides a
717 framework for an integrative approach to study other similar systems.
718

719 **Acknowledgments**

720 This research was funded in part by a generous gift from to the C.A.S. from Will and
721 Margaret Hearst for supporting the 2011 Filipino-American Hearst Biodiversity
722 expedition to Luzon. We were also funded in part through NSF:DEB award number
723 1856402 made to MHVD. We would like to thank the Ruth Tawan-tawan, Ceso II of the
724 Philippines' Department of Environment and Natural Resources Region XI for help with
725 the Gratuitous and export permits. We would also like to thank the University of
726 Mindanao for the mobility support, and Milton N. Medina and Chrestine Torrejos of U.M.
727 for help collecting specimens. We would also like to thank Sarah Crews and Alejandra
728 Hernandez-Agreda of C.A.S. for help with the manuscript text. We would like to thank
729 Jim Henderson C.A.S. for help with the 10X assembly.
730

731 **References**

732 1. Jiggins, C. D. & McMillan, W. O. The genetic basis of an adaptive radiation:
733 Warning colour in two Heliconius species. *Proc. R. Soc. B Biol. Sci.* **264**, 1167–
734 1175 (1997).
735 2. Brower, A. V. Z. Parallel race formation and the evolution of mimicry in Heliconius
736 butterflies: A phylogenetic hypothesis from mitochondrial DNA sequences.
737 *Evolution (N. Y).* **50**, 195–221 (1996).
738 3. Mallet, J. & Gilbert, L. E. Why are there so many mimicry rings? Correlations
739 between habitat, behaviour and mimicry in Heliconius butterflies. *Biol. J. Linn.*
740 *Soc.* **55**, 159–180 (1995).
741 4. Beccaloni, G. W. Vertical stratification of ithomiine butterfly (Nymphalidae:
742 Ithomiinae) mimicry complexes: the relationship between adult flight height and
743 larval host-plant height. *Biol. J. Linn. Soc.* **62**, 313–341 (1997).
744 5. Kronforst, M. R. & Gilbert, L. E. The population genetics of mimetic diversity in
745 Heliconius butterflies. *Proc. R. Soc. B Biol. Sci.* **275**, 493–500 (2008).
746 6. Dettner, K. & Liepert, C. Chemical Mimicry and Camouflage. *Annu. Rev. Entomol.*
747 **39**, 129–154 (1994).

7. Muller, F. Ituna and Thyridia: a remarkable case of mimicry in butterflies. *R. Ent. Soc. London Trans* 12–29 (1879).
8. Mallet, J. & Joron, M. Evolution of Diversity in Warning Color and Mimicry: Polymorphisms, Shifting Balance, and Speciation. *Annu. Rev. Ecol. Syst.* **30**, 201–233 (1999).
9. Sherratt, T. N. The evolution of Müllerian mimicry. *Naturwissenschaften* **95**, 681–695 (2008).
10. Chouteau, M., Arias, M. & Joron, M. Warning signals are under positive frequencydependent selection in nature. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 2164–2169 (2016).
11. Brower, A. V. Z. Rapid morphological radiation and convergence among races of the butterfly Heliconius erato inferred from patterns of mitochondrial DNA evolution. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 6491–6495 (1994).
12. Kapan, D. D. Three-butterfly system provides a field test of müllerian mimicry. *Nature* **409**, 338–340 (2001).
13. Ritland, D. B. & Brower, L. P. The viceroy butterfly is not a batesian mimic. *Nature* **350**, 497–498 (1991).
14. Motyka, M., Kampova, L. & Bocak, L. Phylogeny and evolution of Müllerian mimicry in aposematic Dilophotes: Evidence for advergence and size-constraints in evolution of mimetic sexual dimorphism. *Sci. Rep.* **8**, (2018).
15. Wilson, J. S., Williams, K. A., Forister, M. L., Von Dohlen, C. D. & Pitts, J. P. Repeated evolution in overlapping mimicry rings among North American velvet ants. *Nat. Commun.* **3**, 1–7 (2012).
16. Lev-Yadun, S. Müllerian mimicry in aposematic spiny plants. *Plant Signal. Behav.* **4**, 482–483 (2009).
17. Symula, R., Schulte, R. & Summers, K. Molecular phylogenetic evidence for a mimetic radiation in Peruvian poison frogs supports a Müllerian mimicry hypothesis. *Proc. R. Soc. B Biol. Sci.* **268**, 2415–2421 (2001).
18. Chiari, Y. *et al.* New evidence for parallel evolution of colour patterns in Malagasy poison frogs (Mantella). *Mol. Ecol.* **13**, 3763–3774 (2004).
19. Sanders, K. L., Malhotra, A. & Thorpe, R. S. Evidence for a Müllerian mimetic radiation in Asian pitvipers. *Proc. R. Soc. B Biol. Sci.* **273**, 1135–1141 (2006).
20. Moore, T. Y., Danforth, S. M., Larson, J. G. & Davis Rabosky, A. R. A Quantitative Analysis of Micrurus Coral Snakes Reveals Unexpected Variation in Stereotyped Anti-Predator Displays Within a Mimicry System. *Integr. Org. Biol.* **2**, (2020).
21. JE Randall. A review of mimicry in marine fishes. *Zool. Stud.* **44**, 299–328 (2005).
22. Dumbacher, J. P. & Fleischer, R. C. Phylogenetic evidence for colour pattern convergence in toxic pitohuis: Müllerian mimicry in birds? *Proc. R. Soc. B Biol. Sci.* **268**, 1971–1976 (2001).
23. Tseng, H.-Y., Lin, C.-P., Hsu, J.-Y., Pike, D. A. & Huang, W.-S. The Functional Significance of Aposematic Signals: Geographic Variation in the Responses of Widespread Lizard Predators to Colourful Invertebrate Prey. *PLoS One* **9**, e91777 (2014).
24. Wang, L. Y., Rajabi, H., Ghoroubi, N., Lin, C. P. & Gorb, S. N. Biomechanical strategies underlying the robust body armour of an aposematic weevil. *Front. Physiol.* **9**, 1410 (2018).

25. Schultze, W. A monograph of the pachyrrhynchid group of the Brachyderinae, Curculionidae: Part I. The genus Pachyrrhynchus Germar. *Philipp. J. Sci.* **23**, 609–673 (1923).
26. Anbutsu, H. *et al.* Small genome symbiont underlies cuticle hardness in beetles. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E8382–E8391 (2017).
27. Wilts, B. D. & Saranathan, V. A Literal Elytral Rainbow: Tunable Structural Colors Using Single Diamond Biophotonic Crystals in *Pachyrrhynchus congestus* Weevils. *Small* **14**, 1802328 (2018).
28. Wallace, A. *Darwinism: an exposition of the theory of natural selection with some of its applications*. (Macmillan and Company, 1889).
29. Schultze, W. A monograph of the pachyrrhynchid group of the Brachyderinae, Curculionidae: Part III. The genera Apocyrtidius Heller and Metapocyrtus Heller. *Philipp. J. Sci.* **26**, 131–310 (1925).
30. Hsu, C.-F., Tseng, H.-Y., Hsiao, Y. & Ko, C.-C. First record of the host plant and larvae of *Pachyrhynchus sonani* (Coleoptera: Curculionidae) on Lanyu Island, Taiwan. *Entomol. Sci.* **20**, 288–291 (2017).
31. Yoshitake, H. A New Genus and Two New Species of the Tribe Pachyrhynchini (Coleoptera: Curculionidae) from Palawan Island, the Philippines. *Elytra* **8**, 4–15 (2018).
32. Rukmane A. An annotated checklist of genus Pachyrhynchus (Coleoptera: Curculionidae: Pachyrhynchini). in **18 (1):63-68**, (Acta Biol. Univ. Daugavp, 2018).
33. Chen, Y.-T. *et al.* Integrated species delimitation and conservation implications of an endangered weevil Pachyrhynchus sonani (Coleoptera: Curculionidae) in Green and Orchid Islands of Taiwan. *Syst. Entomol.* **42**, 796–813 (2017).
34. Bollino, M., Sandel, F. & Rukmane, A. New species of the genus Pachyrhynchus Germar, 1824 (Coleoptera: Curculionidae) from Mindanao, Philippines. *Balt. J. Coleopterol.* **17**, 189–204 (2017).
35. Rukmane-Bārbale, A. Short contribution to distribution and appearance of Pachyrhynchus decussatus Waterhouse, 1841 (Entimine: Pachyrhynchini) with description of one new taxon from Catanduanes Island, Philippines. *Balt. J. Coleopterol.* **20**, 81–85 (2020).
36. Heller, K. . New Philippine Zygopinae, Calandrinae, and Cryptoderminae (Curculionidae, Coleóptera). *Philipp. J. Sci.* **25**, 287–309 (1924).
37. Heaney, L. Zoogeographic evidence for middle and late Pleistocene land bridges to the Philippine Islands. *Mod Quatern Res SE Asia* **9**, 127–144 (1985).
38. Heaney, L. R. Biogeography of mammals in SE Asia: estimates of rates of colonization, extinction and speciation. *Biol. J. Linn. Soc.* **28**, 127–165 (1986).
39. Heaney, L. R. & Regalado, J. C. . J. Vanishing treasures of the Philippine rain forest. *Vanish. treasures Philipp. rain For.* (1998).
40. Brown, R. & AC Diesmos. Application of lineage-based species concepts to oceanic island frog populations: the effects of differing taxonomic philosophies on the estimation of Philippine biodiversity. *Silliman J.* **41**, 133–162 (2002).
41. Sinha, C. & Heaney, L. R. *Philippine Biodiversity: Principles and Practice*. (2006).
42. Steppan, S. J., Zawadzki, C. & Heaney, L. R. Molecular phylogeny of the endemic Philippine rodent Apomys (Muridae) and the dynamics of diversification in an oceanic archipelago. *Biol. J. Linn. Soc.* **80**, 699–715 (2003).

43. Inger. Systematics and zoogeography of Philippine Amphibia. *Fieldiana* **33**, 182–531 (1954).
44. Brown, R. M. & Siler, C. D. Spotted stream frog diversification at the Australasian faunal zone interface, mainland versus island comparisons, and a test of the Philippine 'dual-umbilicus' hypothesis. *J. Biogeogr.* **41**, 182–195 (2014).
45. Belton, J. M. *et al.* Hi-C: A comprehensive technique to capture the conformation of genomes. *Methods* **58**, 268–276 (2012).
46. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. in *Bioinformatics* **34**, i884–i890 (Oxford University Press, 2018).
47. Bankevich, A. *et al.* SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
48. Weisenfeld, N. I., Kumar, V., Shah, P., Church, D. M. & Jaffe, D. B. Direct determination of diploid genome sequences. *Genome Res.* **27**, 757–767 (2017).
49. Faircloth, B. C. *et al.* Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst. Biol.* **61**, 717–726 (2012).
50. Faircloth, B. C. Identifying conserved genomic elements and designing universal bait sets to enrich them. *Methods Ecol. Evol.* **8**, 1103–1112 (2017).
51. Van Dam, M. H. *et al.* The Easter Egg Weevil (Pachyrhynchus) genome reveals synteny in Coleoptera across 200 million years of evolution. *bioRxiv* 2020.12.18.422986 (2020). doi:10.1101/2020.12.18.422986
52. Gustafson, G. T. *et al.* Ultraconserved element (UCE) probe set design: Base genome and initial design parameters critical for optimization. *Ecol. Evol.* **9**, 6933–6948 (2019).
53. Toussaint, E. F. A., Tänzler, R., Rahmadi, C., Balke, M. & Riedel, A. Biogeography of Australasian flightless weevils (Curculionidae, Celeuthetini) suggests permeability of Lydekker's and Wallace's Lines. *Zool. Scr.* **44**, 632–644 (2015).
54. Katoh, K., Misawa, K., Kuma, K. I. & Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
55. Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
56. Kozlov, A. M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453–4455 (2019).
57. Yin, J., Zhang, C. & Mirarab, S. ASTRAL-MP: scaling ASTRAL to very large datasets using randomization and parallelization. *Bioinformatics* **35**, 3961–3969 (2019).
58. Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**, 153 (2018).
59. Edelman, N. B. *et al.* Genomic architecture and introgression shape a butterfly radiation. *Science (80-.).* **366**, 594–599 (2019).
60. Lawrence, M. *et al.* Software for Computing and Annotating Genomic Ranges.

PLoS Comput. Biol. **9**, e1003118 (2013).

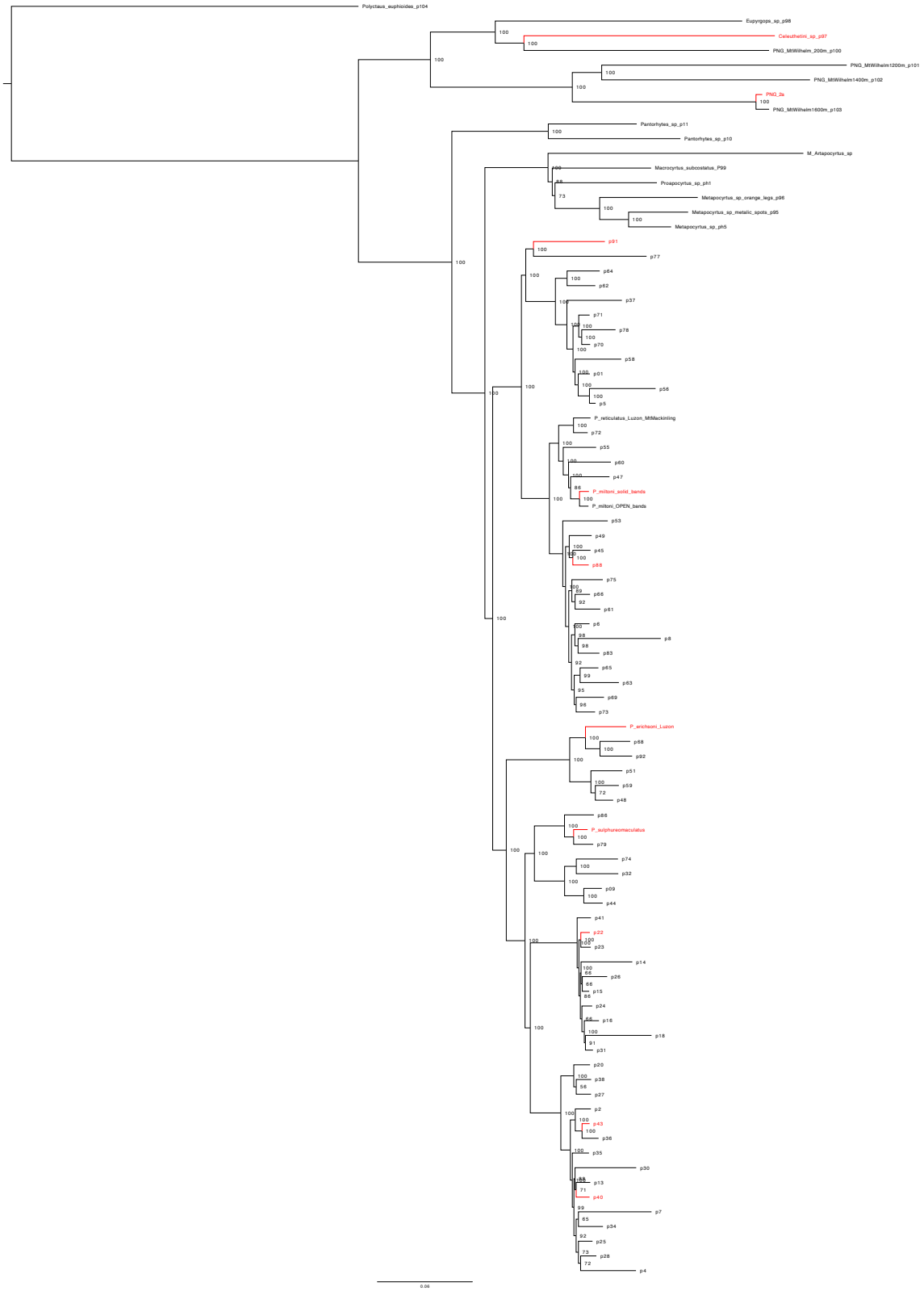
61. Heibl, C. PHYLOCH: R language tree plotting tools and interfaces to diverse phylogenetic software packages. (2008). Available at: http://www.christophheibl.de/Rpackages.html. (Accessed: 17th January 2021)
62. Van Dam, M. H. *et al.* Ultraconserved elements (UCEs) resolve the phylogeny of Australasian smurf-weevils. *PLoS One* **12**, e0188044 (2017).
63. Tagliacollo, V. A. & Lanfear, R. Estimating improved partitioning schemes for ultraconserved elements. *Mol. Biol. Evol.* **35**, 1798–1811 (2018).
64. Van Dam, M. H., Trautwein, M., Spicer, G. S. & Esposito, L. Advancing mite phylogenomics: Designing ultraconserved elements for Acari phylogeny. *Mol. Ecol. Resour.* **19**, 465–475 (2019).
65. Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T. & Calcott, B. Partitionfinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* **34**, 772–773 (2017).
66. Junier, T. & Zdobnov, E. M. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* **26**, 1669–1670 (2010).
67. Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**, 153 (2018).
68. Mirarab, S. Species Tree Estimation Using ASTRAL: Practical Considerations. *arXiv* (2019).
69. Sayyari, E. & Mirarab, S. Fast Coalescent-Based Computation of Local Branch Support from Quartet Frequencies. *Mol. Biol. Evol.* **33**, 1654–1668 (2016).
70. Yang, Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
71. Puttick, M. N. MCMCtreeR: functions to prepare MCMCtree analyses and visualize posterior ages on trees. *Bioinformatics* **35**, 5321–5322 (2019).
72. Dos Reis, M. & Yang, Z. The unbearable uncertainty of Bayesian divergence time estimation. *J. Syst. Evol.* **51**, 30–43 (2013).
73. Zhu, T., Reis, M. Dos & Yang, Z. Characterization of the uncertainty of divergence time estimation under relaxed molecular clock models using multiple loci. *Syst. Biol.* **64**, 267–280 (2015).
74. Reis, M. Dos & Yang, Z. Approximate likelihood calculation on a phylogeny for Bayesian Estimation of Divergence Times. *Mol. Biol. Evol.* **28**, 2161–2172 (2011).
75. Matzke, N. J. Model Selection in Historical Biogeography Reveals that Founder-Event Speciation Is a Crucial Process in Island Clades. *Syst. Biol.* **63**, 951–970 (2014).
76. Tangunan, Fernando & Arcilla, C. Notes on the occurrence of Late Cretaceous nannofossils from Codon Point, San Andres, Catanduanes. *J. Geol. Soc. Philipp.* **64**, 35–47 (2013).
77. Vallejo, B. The Biogeography of Luzon Island Biogeography of Luzon Island , Philippines. **2**, (2014).
78. Knittel, U. *et al.* Diverse protolith ages for the Mindoro and Romblon Metamorphics (Philippines): Evidence from single zircon U-Pb dating. *Isl. Arc* **26**, e12160 (2017).

79. Ree, R. H. & Smith, S. A. Maximum Likelihood Inference of Geographic Range Evolution by Dispersal, Local Extinction, and Cladogenesis. *Syst. Biol.* **57**, 4–14 (2008).
80. Ronquist, F. Dispersal-Vicariance Analysis: A New Approach to the Quantification of Historical Biogeography. *Syst. Biol.* **46**, 195 (1997).
81. Landis, M. J., Matzke, N. J., Moore, B. R. & Huelsenbeck, J. P. Bayesian Analysis of Biogeography when the Number of Areas is Large. *Syst. Biol.* **62**, 789–804 (2013).
82. Brown, R. M. *et al.* Phylogeny of the island archipelago frog genus Sanguirana: Another endemic Philippine radiation that diversified 'Out-of-Palawan'. *Mol. Phylogenet. Evol.* **94**, 531–536 (2016).
83. Van Dam, M. H. & Matzke, N. J. Evaluating the influence of connectivity and distance on biogeographical patterns in the south-western deserts of North America. *J. Biogeogr.* **43**, 1514–1532 (2016).
84. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).
85. Waterhouse, G. Descriptions of the Species of the Curculionideous Genus Pachyrhynchus, Sch., collected hi/ H. Cuming, Esq.,in the Philippine Islands. *Trans. Entomol. Soc. London* **3**, 310–327 (1841).
86. Hines, H. M. *et al.* Wing patterning gene redefines the mimetic history of Heliconius butterflies. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 19666–19671 (2011).
87. Sherratt, T. N. The evolution of imperfect mimicry. *Behav. Ecol.* **13**, 821–826 (2002).
88. Penney, H. D., Hassall, C., Skevington, J. H., Abbott, K. R. & Sherratt, T. N. A comparative analysis of the evolution of imperfect mimicry. *Nature* **483**, 461–464 (2012).
89. Kikuchi, D. W. & Pfennig, D. W. Imperfect mimicry and the limits of natural selection. *Q. Rev. Biol.* **88**, 297–315 (2013).
90. Brown, K. S. & Benson, W. W. Adaptive Polymorphism Associated with Multiple Mullerian Mimicry in Heliconius numata (Lepid. Nymph.). *Biotropica* **6**, 205 (1974).
91. Dittrich, W., Gilbert, F., Green, P., Mcgregor, P. & Grewcock, D. Imperfect mimicry: a pigeon's perspective. *Proc. R. Soc. London. Ser. B Biol. Sci.* **251**, 195–200 (1993).
92. Edmunds, M. Why are there good and poor mimics? *Biol. J. Linn. Soc.* **70**, 459–466 (2000).
93. Richards-Zawacki, C. L., Yeager, J. & Bart, H. P. S. No evidence for differential survival or predation between sympatric color morphs of an aposematic poison frog. *Evol. Ecol.* **27**, 783–795 (2013).
94. Duncan, C. J. & Sheppard, P. M. Sensory Discrimination and Its Role in the Evolution of Batesian Mimicry. *Behaviour* **24**, 269–282 (1965).
95. Haq, B. U., Hardenbol, J. & Vail, P. R. Chronology of fluctuating sea levels since the Triassic. *Science (80-.).* **235**, 1156–1167 (1987).
96. Miller, K. G. *et al.* The phanerozoic record of global sea-level change. *Science* **310**, 1293–1298 (2005).
97. Oaks, J. R., Siler, C. D. & Brown, R. M. The comparative biogeography of Philippine geckos challenges predictions from a paradigm of climate-driven

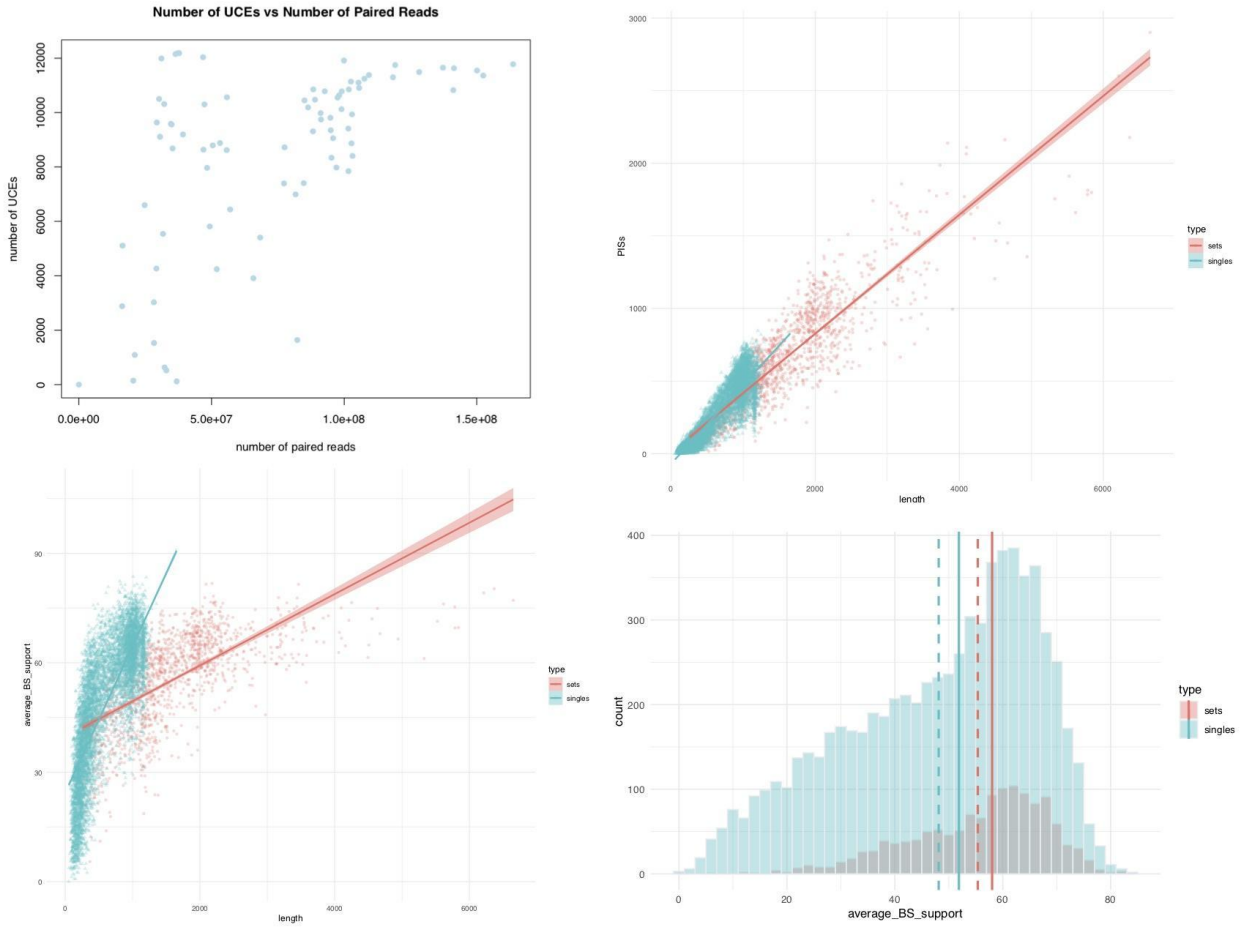
vicariant diversification across an island archipelago. *Evolution (N. Y).* **73**, 1151–1167 (2019).

98. Robinson, M. M., Dowsett, H. J. & Chandler, M. A. Pliocene Role in Assessing Future Climate Impacts. *Eos, Trans. Am. Geophys. Union* **89**, 501 (2008).
99. Gautier, M. *et al.* The Genomic Basis of Color Pattern Polymorphism in the Harlequin Ladybird. *Curr. Biol.* **28**, 3296-3302.e7 (2018).
100. Yen, E. C. *et al.* A haplotype-resolved, de novo genome assembly for the wood tiger moth (Arctia plantaginis) through trio binning. *Gigascience* **9**, 1–12 (2020).
101. Halbritter, D. A., Gordon, J. M., Keacher, K. L., Avery, M. L. & Daniels, J. C. Evaluating an alleged mimic of the monarch butterfly: Neophasia (lepidoptera: Pieridae) butterflies are palatable to avian predators. *Insects* **9**, (2018).
102. Freitas, F. V, Branstetter, M. G., Griswold, T. & Almeida, E. A. B. Partitioned Gene-Tree Analyses and Gene-Based Topology Testing Help Resolve Incongruence in a Phylogenomic Study of Host-Specialist Bees (Apidae: Eucerinae). *Mol. Biol. Evol.* (2020). doi:10.1093/molbev/msaa277

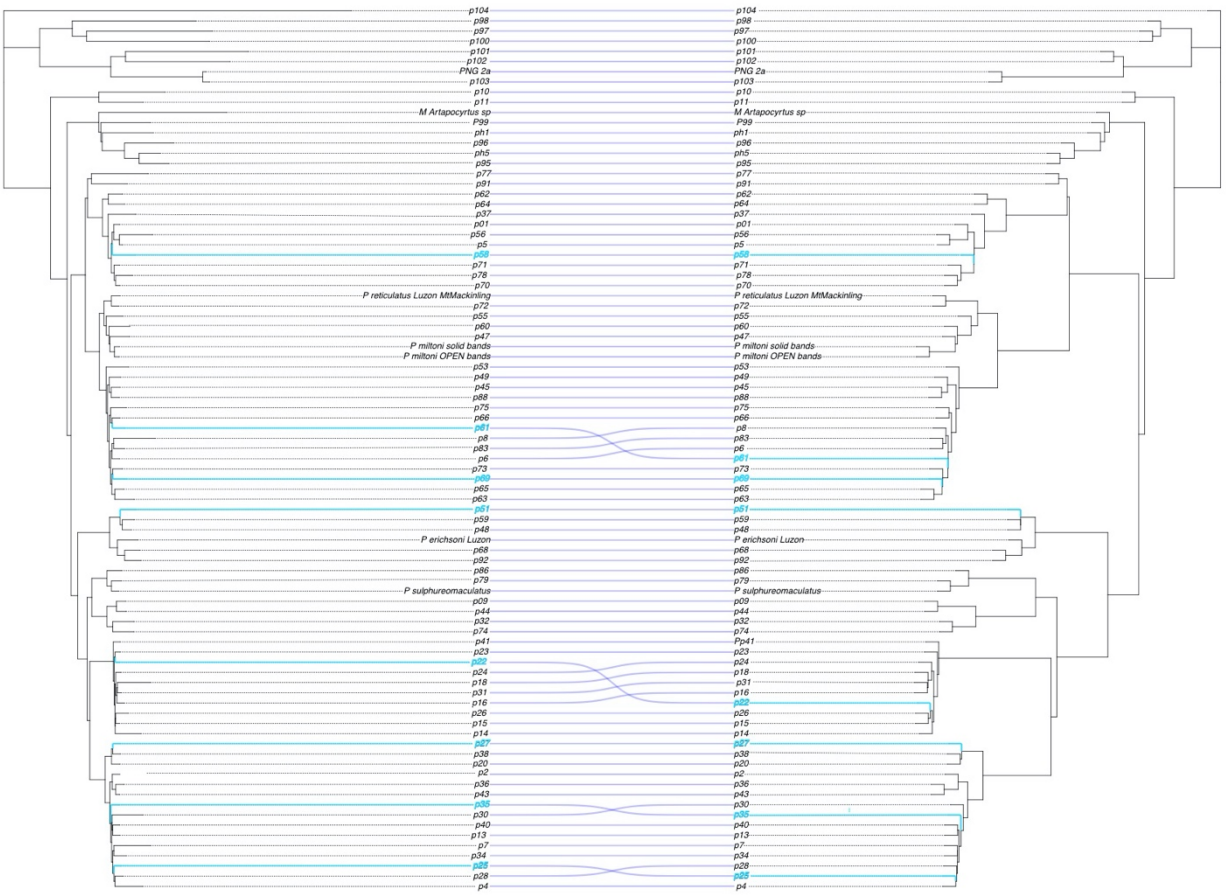
Supplemental Figures



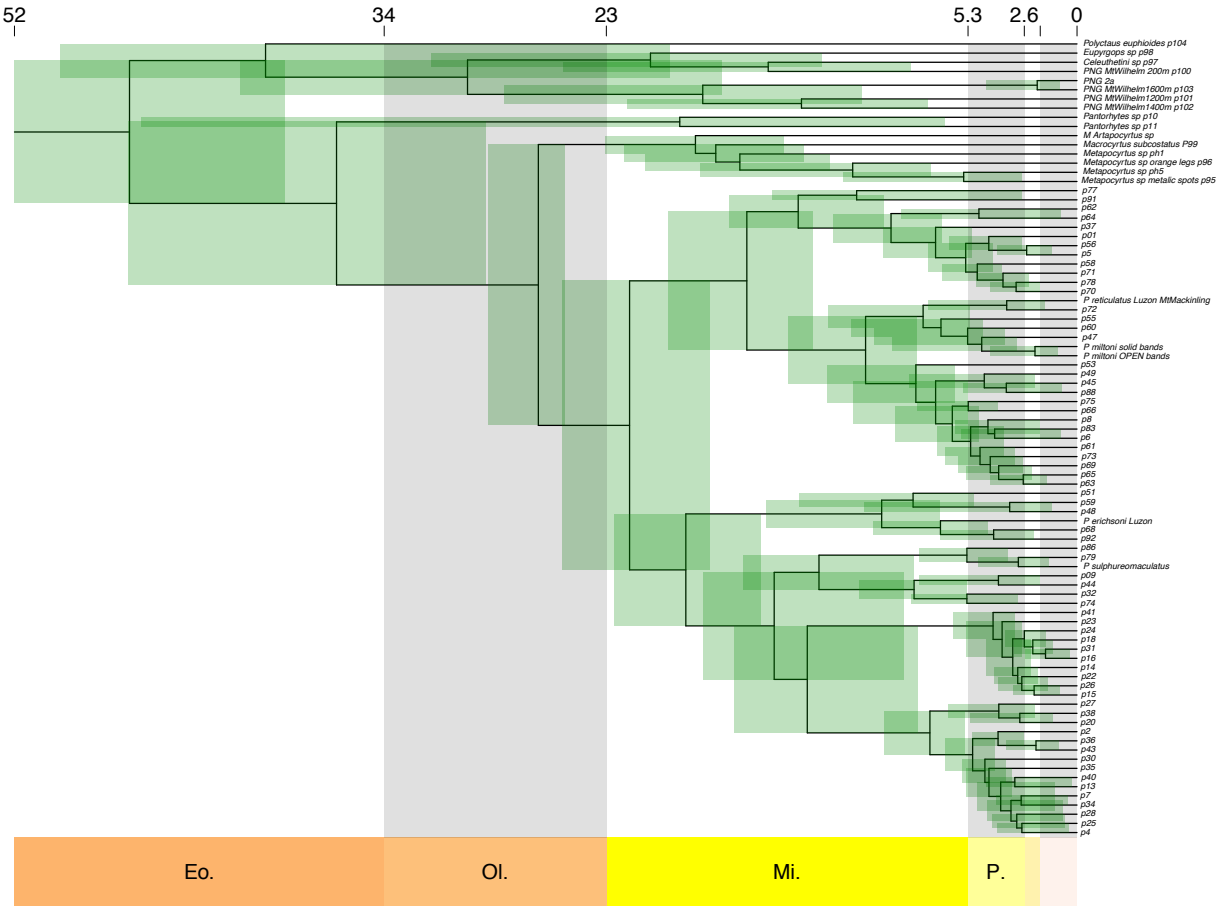
Supplementary Figure 1. *Pachyrhynchus* concatenated ML phylogeny constructed with RAxML-NG. Node labels are bootstrap support values. Branches in red are the taxa used in probe design.



Supplementary Figure 2. Upper left panel, Number of UCEs recovered vs number of paired reads. For the UCE type, the "sets" are concatenated UCEs found in a 25kb non-overlapping sliding window bin, and UCE "singles" are those where only a single UCE occupies a 25kb bin in the *P. sulphureomaculatus* base genome. **Upper right**, Phylogenetically informative sites vs length of UCE locus. **Lower left**, Average bootstrap support per locus by UCE type vs length of locus. **Lower right**, Average bootstrap support per locus by UCE type. Solid vertical lines are the mean and dashed vertical lines are the median by UCE type.



Supplementary Figure 3. RAxML concatenated phylogeny on left, ASTRAL species tree on right.



Supplementary Figure 4. *Pachyrhynchus* chronogram constructed using MCMCTREE.

Nucleotide Models Used: GTR+I, GTR+G, F81, F81+G, SYM, F81+I, JC, HKY, K80, HKY+I+G, K80+I, SYM+G, SYM+I, K80+G, GTR, HKY+G, SYM+I+G, HKY+I, F81+I+G, JC+G, GTR+I+G, JC+I, K80+I+G, JC+I+G, TIM, TVM, TVMef, TrN, TrNef, TIM+G, TVM+G, TVMef+G, TrN+G, TrNef+G, TIM+I, TVM+I, TVMef+I, TrN+I, TrNef+I, TIM1+I, TIM1,TIM1+G