The Easter Egg Weevil (*Pachyrhynchus*) genome reveals synteny in Coleoptera across 200 million years of evolution

Matthew H. Van Dam[1,3,*], Analyn Anzano Cabras[2], James B. Henderson[3], Cynthia Pérez Estrada[4], Arina D. Omer[4], Olga Dudchenko[4], Erez Lieberman Aiden[4], 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.*

[4]*The Center for Genome Architecture, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA*

[*]*Correspondence to be sent to: Matthew H. Van Dam,*
E-mail: ***matthewhvandam@gmail.com***

The Easter Egg Weevil (*Pachyrhynchus*) genome reveals synteny in Coleoptera across 200 million years of evolution

Abstract

Patterns of genomic architecture across insects remain largely undocumented or decoupled from a broader phylogenetic context. For instance, it is unknown whether translocation rates differ between insect orders? We address broad scale patterns of genome architecture across Insecta by examining synteny in a phylogenetic framework from open source insect genomes. To accomplish this, we add a chromosome level genome to a crucial lineage, Coleoptera. Our assembly of the *Pachyrhynchus sulphureomaculatus* genome is the first chromosome scale genome for the hyperdiverse Phytophaga lineage and currently the largest insect genome assembled to this scale. The genome is significantly larger than those of other weevils, and this increase in size is caused by repetitive elements. Our results also indicate that, among beetles, there are instances of long-lasting (>200 Ma) localization of genes to a particular chromosome with few translocation events. While some chromosomes have a paucity of translocations, intra-chromosomal synteny was almost absent, with gene order thoroughly shuffled along a chromosome. To place our findings in an evolutionary context, we compared syntenic patterns across Insecta. We find that synteny largely scales with clade age, with younger clades, such as Lepidoptera, having especially high synteny. However, we do find subtle differences in the maintenance of synteny and its rate of decay among the insect orders.

INTRODUCTION

Beetles represent one of the most diverse groups of metazoans, with ~400,000 described species (Hammond 1992), and estimates of total diversity up to 0.9–2.1 million species (Stork et al. 2015). Among beetles, weevils (Coleoptera: Curculionidae) are one of the most diverse insect groups (>60,000 species (Oberprieler et al. 2007)), encompassing a huge range of life history strategies and occupying every conceivable niche in a terrestrial ecosystem. With morphological forms specialized to ecological habits, such as feeding on fungi, seeds, pollen, wood, roots, and even kangaroo dung, weevils make an excellent system in which to study the evolution of different ecomorphologies (Zimmerman 1994, Oberprieler et al. 2007). Weevils belong to the group Phytophaga whose members comprise lineages that specialize on and have co-diversified with many plant lineages (McKenna et al. 2009, Seppey et al. 2019). Given their vast diversity and economic importance as pollinators and crop pests, knowing more about the genomic architecture of beetles should be of broad applicability. However, to date, there are only four available genomes resolved to chromosome level for Coleoptera and none for weevils or the hyperdiverse beetle lineage Phytophaga (Van Belleghem et al. 2018, Fallon et al 2018, Zhang et al. 2020, Herndon et al. 2020). Here we present the first genome resolved to chromosome level for the Phytophaga beetle lineage *Pachyrhynchus sulphureomaculatus* Schultze, 1922.

Figure 1. *Pachyrhynchus sulphureomaculatus*, lateral habitus. (photo by A. Cabras)

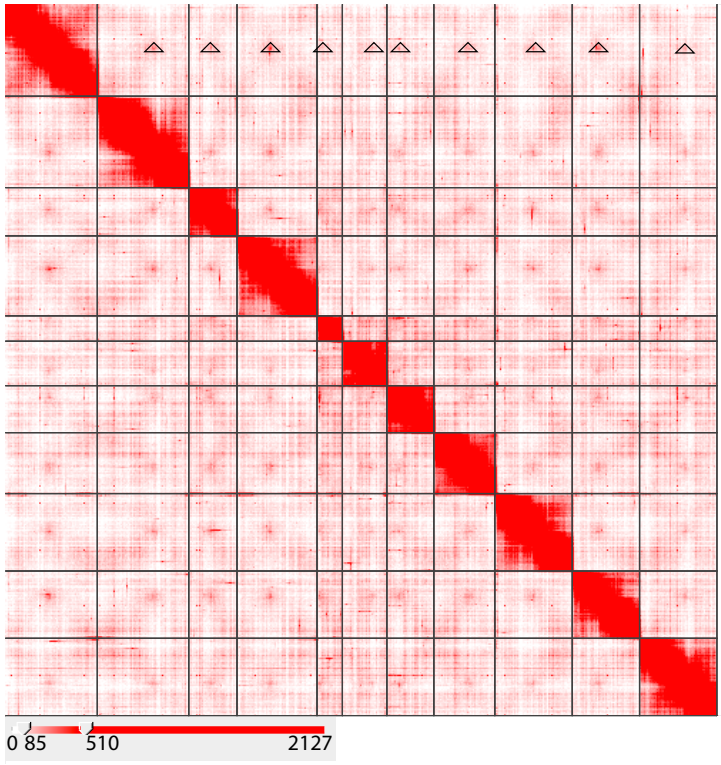


Figure 2. Hi-C contact map heatmap of *Pachyrhynchus sulphureomaculatus* Schultze, 1922. Eleven chromosome boundaries are indicated by black lines. Heatmap scale lower left, range in counts of mapped Hi-C reads per megabase squared. Rabl-like pattern highlighted along chromosome 1, top row, open triangles indicate contact between centromere regions. X-like pattern between adjacent off diagonal regions indicative of contact between distal portions of chromosomes.

Recent advances in genome assembly techniques, such as in situ high throughput conformation capture technology (Hi-C) (Lieberman-Aiden et al. 2009), have substantially enhanced our knowledge of genome architecture (Rao et al. 2014, Dudchenko et al. 2017, Eagen et al. 2017). Increasing the accuracy and contiguity of genome assemblies has also been aided by using long-read sequencing technology in combination with in situ Hi-C (Ghurye et al. 2017, Matthews et al. 2018, Song et al. 2018, Kingan et al. 2019, Scheffer et al. 2020). These innovations have allowed researchers to not only reconstruct genomes to chromosome scale but also to do so

relatively quickly and cheaply (Dudchenko et al. 2018). In addition, in situ Hi-C technology has shown that the 3D conformation of genomes is not random and that this conformation can influence gene expression and linkage (Sanborn et al. 2015). The result of these new sequencing techniques has increased the number of high quality genomes for non-model insect species, including beetles (Matthews et al. 2018, Hill et al. 2019, Lu et al. 2019, Liu et al. 2019, Biello et al. 2020, Herndon et al. 2020, Zhang et al. 2020). Because in situ Hi-C orders scaffolds and corrects misjoins, we can study synteny between organisms with more confidence (Dudchenko et al. 2017, Ghurye et al. 2019).

With the influx of new chromosome-level genomes, we can now begin to explore patterns of genome architecture within and between major insect lineages. For example, in Lepidoptera (butterflies), genome architecture has been characterized as relatively stable with few (6%) orthologous loci being translocated (Ahola et al. 2014, Davey et al. 2016, Hill et al. 2019, Wan et al. 2019). Holocentric chromosomes observed throughout Lepidoptera are implicated in reducing hybridization limitations, (Marec et al. 2001, Lukhtanov et al. 2018, Edelman et al. 2019, Hill et al. 2019) suggesting that genome architecture plays a significant role in their biology. In *Heliconius* butterflies, the inversion and rearrangements that do occur do not seem to hinder hybridization (Edelman et al. 2019). In contrast to Lepidoptera, *Drosophila* species have many more translocations and rearrangements (Renschler et al. 2019). In beetles, however, even a basic understanding of genomic architecture remains undocumented. The basic blueprints as revealed by in situ Hi-C maps of how a genome is organized (e.g. – with a Rabl-like conformation (Rabl 1885, Csink and Henikoff 1998), holocentric chromosomes, chromosome domain territories, compartments, and topological associated domain loops) remain non-existent

and therefore unplaced in a phylogenetic context. A general synthesis across insects linking these genomic architectural patterns to their function and potential influence on speciation remains incomplete. For example, do different insect orders have distinct rates of genomic rearrangements (the breakage of synteny between genes), or are the patterns we observe merely due to clade age? Are there aspects of a lineages' genomic architecture that contribute to their observed syntenic patterns? Here we address these questions and provide a new chromosome-level genome for Coleoptera.

RESULTS

Sequencing and assembly results

From our PacBio library we sequenced a total of 87,452,300,317 base pairs (bp) with an N50 read length of 31,404 bp (see Supplemental Information Table 1 for full report). From our in situ Hi-C library (we refer to the in situ Hi-C library or reads as "Hi-C" throughout), we sequenced a total of 228,169,567 paired reads after cleaning. Only 2.53% of our Hi-C reads were

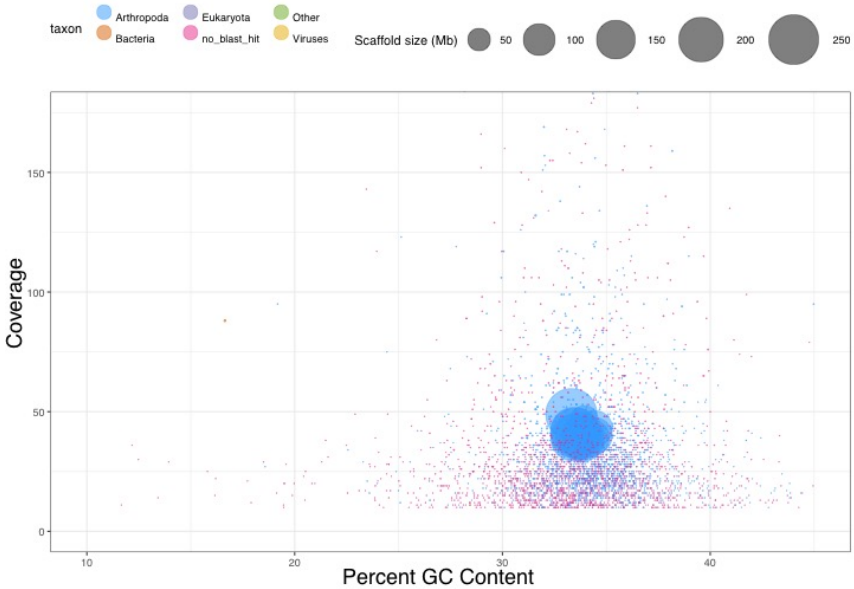


Figure 3. *P. sulphureomaculatus* scaffold bubble plot of coverage versus GC content. Scaffolds included are from the unfiltered assembly. Taxonomic annotation provided via *blastn* alignment to the NCBI nt database.

unmapped, and we had a total of 80,652,881 Hi-C contacts. For a list of the intra-/inter-chromosomal contacts and long/short range Hi-C contacts, see Table 1.

Our initial assembly after 3X polishing in RACON (Vaser 2017) consisted of 18,240 scaffolds and was 2,982,578,979 bp in total length. After removing duplicate haplotigs with *Purge Haplotigs* (Roach et al. 2018), 9,751 scaffolds and 2,052,097,903 bp remained (Fig. 2). Our initial Hi-C assembly resulted in 4,111 and 2,057,226,403 bp total. The size increase is due to 500 bp insertions of Ns (the 3D-DNA default), between scaffolds merged into super-scaffolds. Running *Pilon* (v. 1.23) (Walker et al. 2014) in "--fix bases" mode and removal of mitochondrial and contaminant scaffolds (virus or bacteria) resulted in 4,093 scaffolds and 2,051,389,195 bp. A bubble plot of scaffolds by taxon category, and a table of the chromosomes scaffold N50s are shown in Fig. 3 and Table 2, respectively. The identity of other scaffolds not included in the main chromosomes are ambiguous (14 potential viruses and 31 potential bacteria). We retained these but did remove any with bacteria or virus as their best blast score those previously mentioned. The full summary statistics of our final assembly are shown in Table 2. From the different versions of BUSCO (Felipe et al. 2015) Insecta gene sets (1658 BUSCOs version 2, 1367 version 4 and 5-beta), the percentage of complete genes varied (90.8% V2, 87.6% V4 and 91.1% V5 (Fig. 4)), indicating a relatively complete assembly. Compared to other chromosome-level beetle genomes, we found a comparable number of complete BUSCO genes. However, the results vary somewhat depending on which version of BUSCO and which genes were used (Fig. 4). We found a relatively low duplication rate compared to that found in two other beetle (*Photinus* firefly and *Propylea* ladybeetle) genomes that used primarily long-read and Hi-C sequencing in their assembly.

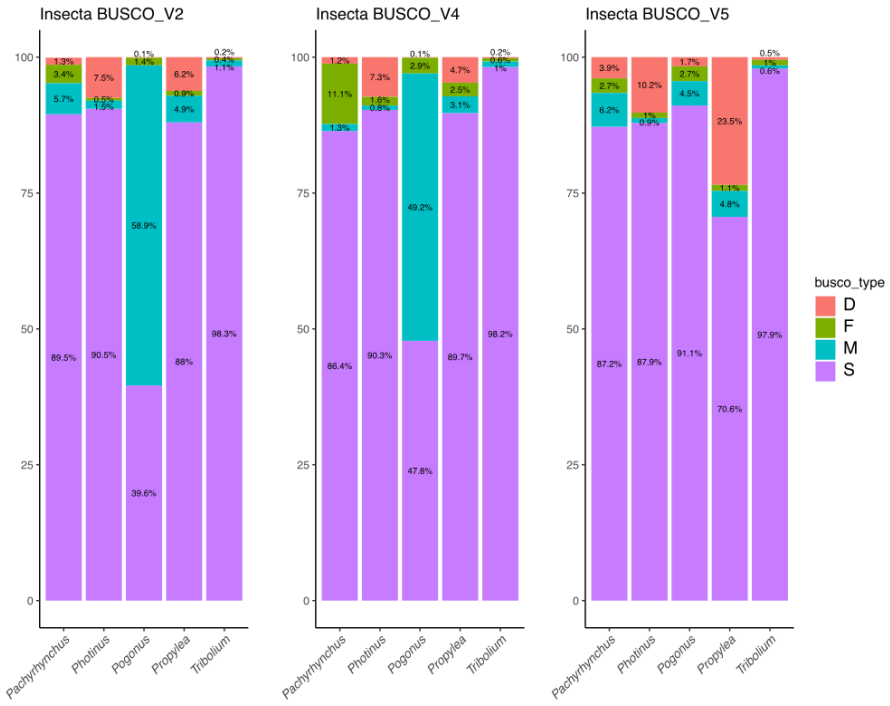
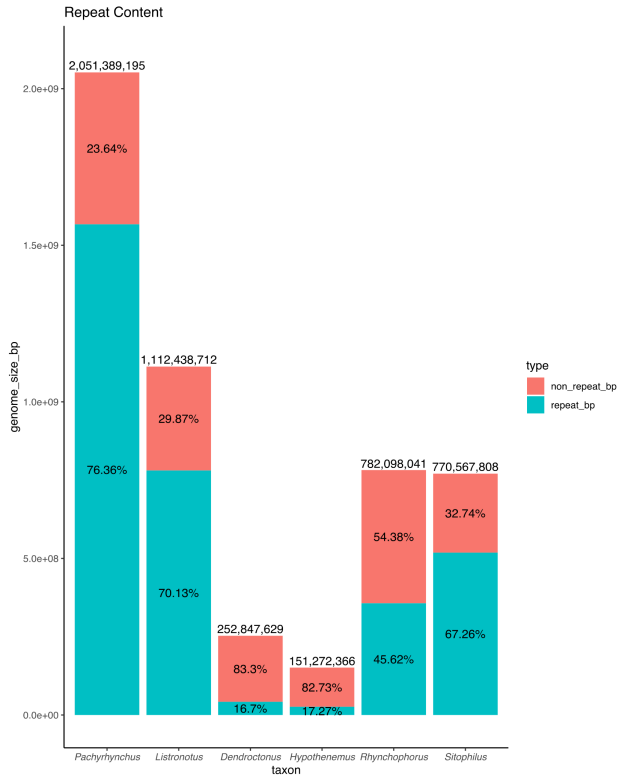


Figure 4. Stacked bar plot of Insecta BUSCO gene sets by category for chromosome-level beetle genomes. Y-axis is the percent of BUSCO genes, X-axis labels are the genus names. The abbreviations in the legend are: D=duplicated, F=fragmented, M=missing and S=single.

Repeat content analyses

At 2.05 Gbp, the *Pachyrhynchus sulphureomaculatus* genome is roughly 1.8 times as large as the next largest weevil (Curculionoidea) genome published to date, the 1.11 Gbp *Listronotus bonariensis*, the Argentine Stem Weevil (Harrop et al. 2020), and 2.6 times the next largest, the 782 Mbp Red Palm Weevil, *Rhynchophorus ferrugineus* (Hazzouri et al. 2020) genome. To help explain the size difference, we categorized the repeat content of *P. sulphureomaculatus*. The repeat content analyses from *RepeatMasker* shows that the genome of *P. sulphureomaculatus* consists of more than three quarters (76.36%) repetitive DNA, similar to the repeat percentage of *Listronotus*, which is the closest relative to *Pachyrhynchus*. Compared to other weevil genomes (Fig. 5), *P. sulphureomaculatus* has roughly the same percentage of non-repetitive DNA as

136 *Listronotus* and *Sitophilus*. However, the genomes of the two bark beetles of the subfamily
137 Scolytinae (*Dendroctonus* and *Hypothenemus*), are ~1/12 the size of *P. sulphureomaculatus* and consist of only ~17% repetitive content. The *P. sulphureomaculatus* genome consisted of 73.1% interspersed repeats, with SINEs being 0.1%, LINEs 20.8%, LTR elements 2.6%, DNA elements 33% and unclassified repeats 16.6%. A sliding window analysis suggests that repetitive content tends to be found in a higher percentage towards the ends of the chromosomes in *P. sulphureomaculatus*, except in chromosome 5 (Fig. 6).



149 **Figure 5.** Histogram of repeat content for weevil genomes.

150

151 ***Genome annotation***

152 After removing low quality reads from our transcriptome library, a total of 20,551,938
153 paired reads remained. Our initial 3 transcriptome assemblies, *Trinity* de novo, *Trinity* genome
154 guided assembly and *rnaSPAdes*, resulted in fairly similar assemblies, with each having a high
155 number (~90%) of the BUSCO v.2 Arthropoda genes (see Supplemental Information Table 2 for
156 details).

157 As the nuclei of cells between different species generally do not interact (except for
158 viruses), and because Hi-C mapping will remove any non-*Pachyrhynchus* DNA from the
159 chromosomes, we only annotated genes found within the 11 chromosomes comprising

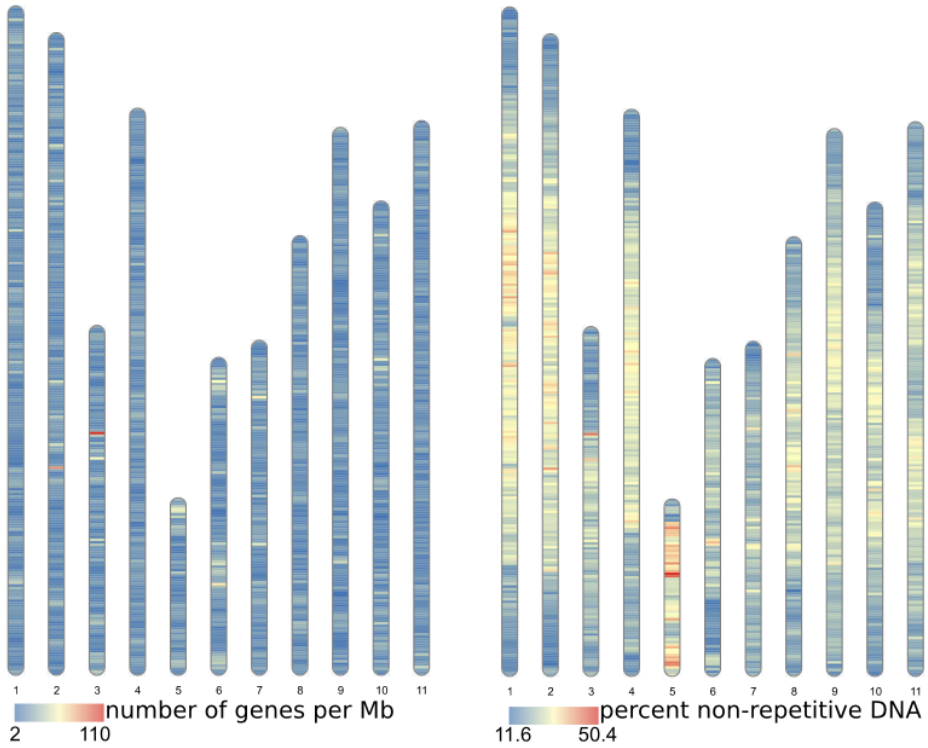
2,000,581,858 bp. The *EVidenceModeler* analysis found that the *P. sulphureomaculatus* contained, as percentages of length, 26.00% genes, 1.46% exons and 24.54% introns, with 30,175 genes, totaling 520,120,665 bp with a mean length of 17,236.81. There are a total of 10,009 single exon genes (33.17% of genes), a total of 120,454 exons with a mean of 3.99 exons per gene, a length of 242.1 bp and per gene length of 966.41bp. There are a total of 90,279 introns with a mean of 2.99 per gene, a length of 5,438.24 bp and a per gene length of 16,270.4 bp. The distribution of gene, exon and intron sizes can be found in the Supplemental Information file "Size_of_Genes_Exons_Introns". The gff and tRNA annotations are also in the Supplemental Information. Chromosome gene distribution is relatively even, with only a few regions enriched with genes (Fig. 6).

Synteny across coleopteran chromosome-level genomes

We found BUSCO v.2 loci (1658 Insecta gene set) had a low level of translocations between chromosomes (Fig. 7). Our UCE set resulted in 295 loci among taxa and recovered an identical topology and similar dates as in McKenna et al. (2019). We also found a similar synteny pattern between BUSCO genes and those from our UCE set (Supplemental Information file "BUSCO_UCE_chromosome_Tcas_Psulph"). Results show that within a chromosome, the order of BUSCO genes is not conserved (Fig. 7), with few long segments of synteny within a chromosome. Synteny is greatest between *P. sulphureomaculatus* and the three Polyphaga beetles, and least between Adephaga (*Pogonus*) and *P. sulphureomaculatus*. Interestingly, there is more synteny between *P. sulphureomaculatus* and *Photinus pyralis* (firefly) (Fallon et al. 2018) than between *P. sulphureomaculatus* and *Propylea japonica* (ladybird beetle), the closer

182 relative of the two, indicating that the lineage leading to *Propylea* has undergone many more
183 chromosomal translocation events (Fig. 7 and 8).

184 We computed the Ensemble Gene Order Conservations (GOC) scores (ref) across all
185 pairwise comparisons for our 34 taxa; results are in Supplemental Information file
186 "GOC_results_matrix.txt". We recovered 1356 BUSCO Genes for the 70% complete matrix,
187 totaling 546,311 amino acids in length. The phylogeny recovered the same clades as in Misof et
188 al. (2014). The GOC scores tended to vary with phylogenetic distance between taxa. For
189 example, the hemipterans *Triatoam rubrofasciata* and *Rhodnius prolixus* (Reduviidae) scored
190 0.73, and between *R. prolixus* and the pea aphid *Acyrthosiphon pisum*, 0.01. Results from the
191 Mantel test between phylogenetic distance and GOC score found that the two variables were
192 correlated (P≤0.001).



193

194 **Figure 6.** Heat map of gene density and non-repetitive DNA per 1 Mb sliding window. The 11 chromosomes
195 are in the same order as in the Hi-C heat map (Fig. 1) and fasta file of the genome.

196

197 **DISCUSSION**

198

199 The combination of long-read DNA and Hi-C sequencing was successful in resolving a
200 large and highly repetitive insect genome. To date, this is the largest insect genome and one of
201 the largest arthropod genomes assembled to chromosome scale, the horseshoe crab's (*Tachypleus*
202 *tridentatus*) being only slightly larger (2.06 Gb vs 2.05 Gb) (Zhou et al. 2020). This is
203 remarkable because the assembly of relatively large and highly repetitive insect genomes into
204 highly contiguous ones such as this was previously unattainable (Li et al. 2020). Those efforts
205 were hindered by repetitive contents breaking scaffolds or misjoining them (Dudchenko et al.
206 2017, Hill et al. 2019, Li et al. 2020). The unusually large size of the *Pachyrhynchus* genome is
207 mostly due to the inflated proportion of repetitive content, 76.4% of the genome (Fig. 5). Again,
208 highlighting the need for long sequencing reads to span the repetitive content. Here we used a
209 single individual to create both our Hi-C and PacBio libraries. The main advantage over using
210 multiple individuals is little loss of Hi-C reads mapped to the scaffolds; it also eliminates the
211 need for isogenic lines to be established before sequencing. In our previous attempts to assemble
212 a genome for *Pachyrhynchus*, we were greatly hindered by the loss of mappable reads when
213 using multiple individuals. As long read sequencing improves in its capabilities of using a small
214 amount (5–50 ng) of DNA, capitalizing on this combination of Hi-C and long-read sequencing
215 will make it feasible to assemble chromosome scale genomes from single, very small insect
216 specimens (Kingan et al. 2019, Schneider et al. 2020).

217

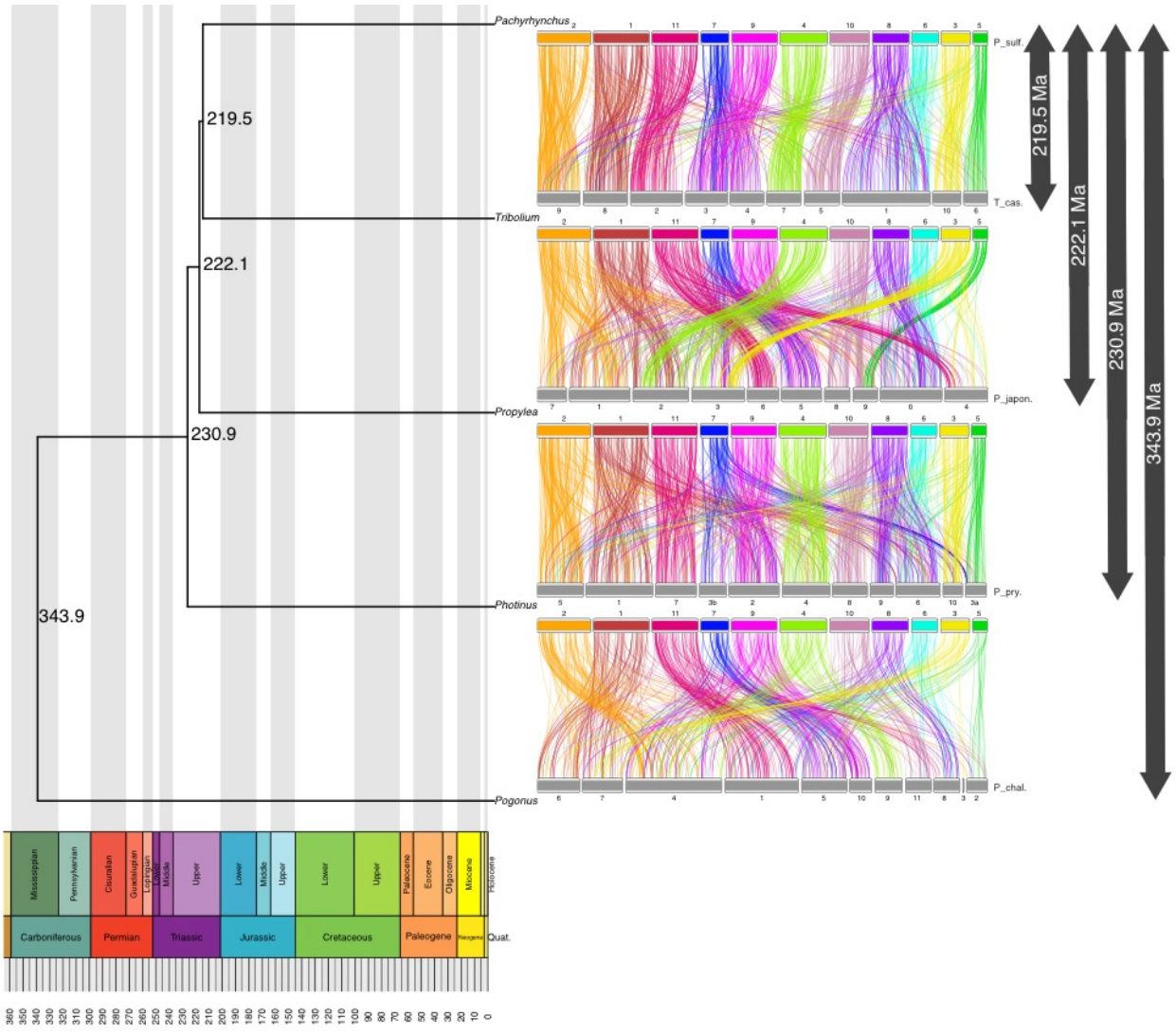


Figure 7. Chronogram and ideograms of the 5 beetle genomes which have chromosome level assemblies.

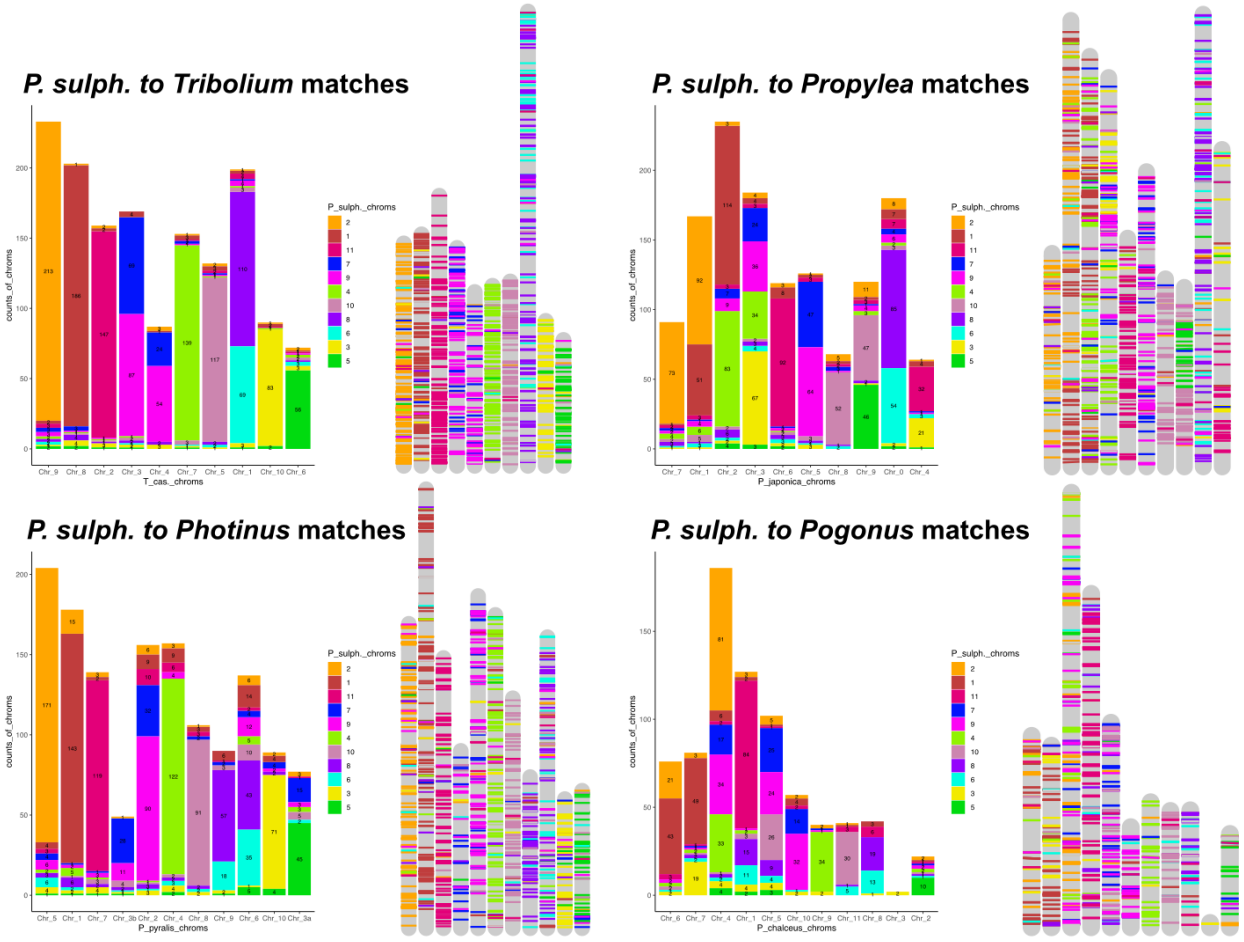


Figure 8. Stacked bar plots and chromosome mappings of BUSCO genes' placements. Colors correspond to *P. sulphureomaculatus* chromosomes. The numbering scheme of chromosomes matches the names found in the genome's fasta file.

The conserved inter-chromosomal synteny (few chromosome translocations) between the beetle genomes is surprising given the divergence times of the different lineages. For example, we recovered chromosomes that have remained 80–92% intact for more than 200 Ma (Figs. 7–8). By contrast, the order of the BUSCO genes inside of the chromosomes are highly rearranged, such as chromosomes 8 and 6 in *Pachyrhynchus* and chromosome 1 in *Tribolium* (Figs. 7–8). This initial finding prompted us to examine whether similar patterns are observed across other insect orders. A characteristic of Lepidoptera is having a high level of synteny across different families

(Hill et al. 2019, Wan et al. 2019). When we examine the ages or relative branch lengths between clades, we find that much of the synteny is correlated with clade age (Fig. 9), as well as from the results of the Mantel test ($P \leq 0.001$). Although Lepidoptera tends to exhibit more synteny given the phylogenetic distance between taxa, there are a few examples other than *Drosophila* in this same time period to compare against (Fig. 9). Currently, chromosome-level genomes are not available for Trichoptera (caddisflies, the sister lineage to Lepidoptera) or early diverging lineages of Lepidoptera. With the addition of these lineages, we could determine whether the observed pattern of synteny conservation is found only in Lepidopteran crown groups or whether it is more widely dispersed across the entire Lepidopteran lineage. Another order with a somewhat similar level of synteny as in Lepidoptera given their divergence times are Hemiptera (Fig. 9), although there are only 2 comparisons with the same level of divergence (Fig. 9). Taxa with much less synteny given their divergence times are Hymenoptera and Diptera (Fig. 9). The finding that synteny tends to decay with age is not surprising; however, there are some insect orders that are more or less syntenic than expected given their age. For example, in *Drosophila*, there is less synteny between members of this genus (~40 Ma) than across all of Lepidoptera or the Aphidae that we examined. These results of gene order conservation are consistent with research of *Drosophila* topological associated domains (TADs) that showed synteny break points at approximately every 6th gene between *D. melanogaster, D. virilis* and *D. busckii*, which have a similar level of divergence as the *Drosophila* taxa we examined, about 40 Ma of divergence (Renschler et al. 2019). In addition, the chromosomal rearrangement across *Drosophila* tends to occur at TAD boundaries, not inside the loops (Renschler et al. 2019, Liao et al. 2020). In *Anopheles* mosquitos, the TAD structures seem to be associated with cytological structures as well (Lukyanchikova et al. 2020).

Recent studies in Diptera have demonstrated that syntenic breakpoints tend to occur at the boundaries between TADs (Renschler et al. 2019, Liao et al. 2020, Lukyanchikova et al. 2020). Despite having many breakpoints, with relatively few chromosome translocations, the dipteran chromosomes largely remain intact (Bracewell et al. 2019). In Coleoptera we find a somewhat similar syntenic pattern, in that the chromosomes remain intact while also being highly shuffled (Figs. 7–8). Given the divergence time between our taxa, when translocations do occur, their initial positions are lost due to a high level of reorganization producing a pattern of interwoven segments. For example, chromosomes 8 and 9 in *Pachyrhynchus* and the large chromosome 1 in *Tribolium* (Fig. 8), there are no large syntenic runs of genes or obvious places of translocation. In contrast, chromosome 9 of *Propylea* and chromosome 5 of *Pachyrhynchus* are still largely intact, with the homologous segment of chromosome 5 inserted into roughly the middle of *Propylea*'s chromosome 9. Across the Coleoptera we examined, this was the only fusion event that could be roughly placed. Given the relative amount of reshuffling along other parts of this chromosome, the ability to place the insertion indicates that this was a relatively recent event. This large amount of reshuffling within chromosomes with few inter-chromosomal events contrasts with patterns seen in mammals in which the chromosomes tend to exchange larger blocks of material more readily (Chowdhary et al. 1998, Kemkemer et al. 2009, Deakin 2018, Simison et al. 2020).

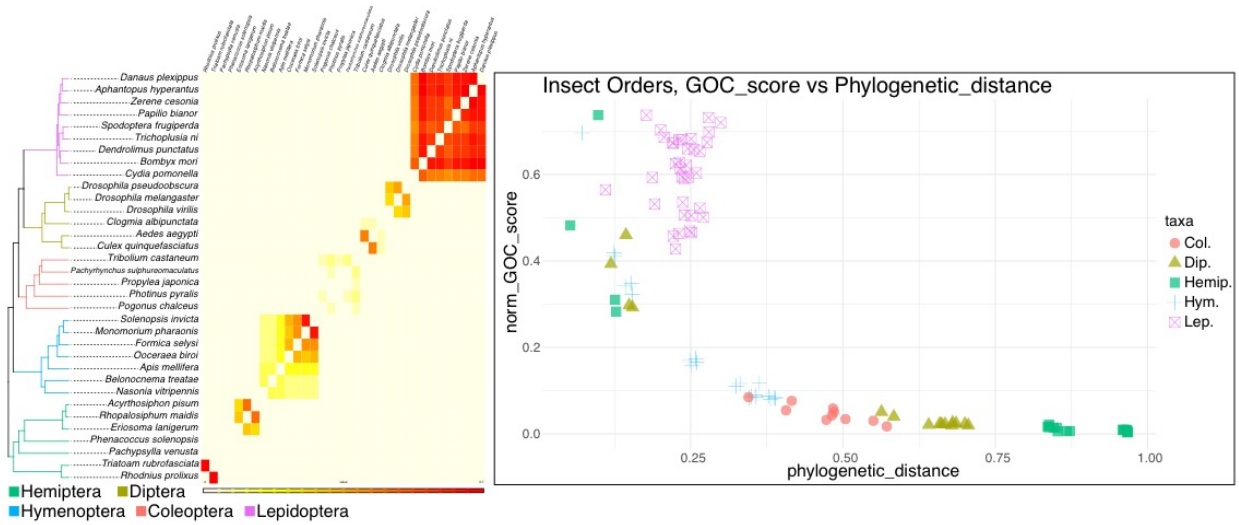


Figure 9. Insecta, gene order conservation score (GOC) plot. Left panel, chronogram, branches colored by order. Heat map of normalized pairwise GOC scores, redder boxes indicate more synteny between pairs. Right panel, within order normalized pairwise GOC score versus phylogenetic distance.

Another architectural feature of *Pachyrhynchus*' genome above the chromosome level includes the Rabl-like configuration of chromosomes, where centromeres and telomeres cluster at opposite/different regions of the cell. These features are important to note because they may serve an important evolutionary function, such as reducing chromosomal entanglements during interphase as well as regulating chromosomal compartmentalization (Mizuguchi et al. 2015, Pouokam et al. 2019). Both major lineages of Diptera, the Nematocera (e.g. mosquitoes and Psychodidae) and Schizophora (e.g. *Drosophila*), have cells with a Rabl-like configuration (Csink and Henikoff 1998, Dudchenko et al. 2017, Matthews et al. 2018). These taxa span much of the phylogenetic distance across the dipteran lineage, and thus this pattern of chromosomal organization may be characteristic of Diptera. We also observe the Rabl-like configuration in *Pachyrhynchus* as well as in the Hi-C map of *Tribolium* (DNAZoo Consortium et al. 2020). Hi-C map observations for the other taxa do not indicate any other obvious cases of the Rabl-like

292 configuration within the Insecta. However, improving the quality of existing Hi-C maps would
293 provide more evidence for this observation because a lack of valid Hi-C reads can obscure this
294 type of chromosomal architecture.

295 The Rabl-like configuration is not restricted to beetles and flies; it is also found in the
296 yeast genome (Jin et al. 1998, Goto et al. 2001, Mizuguchi et al. 2015, Kim et al. 2017) as well
297 as in wheat, barley and *Brassica* (Santos and Shaw 2004, Mascher et al. 2017, Concia et al.
298 2020, Wang et al. 2019), and was originally described from salamander cells (Rabl 1885). It is
299 unclear how widespread the Rabl-like configuration is in Coleoptera. The Hi-C maps of the other
300 beetle genomes do not display this formation and are from similar tissue types to what we used
301 (Fallon et al. 2018). It could be that this configuration is only in the Tenebrionoidea and
302 Phytophaga lineages, where it is presently observed. It is assumed that the Rabl-like
303 configuration is found in all life stages, as appears to be the case in Diptera (Dudchenko et al.
304 2017, Matthews et al. 2018, Lukyanchikova et al. 2020). Additionally, while the Rabl-like
305 configuration is found in different life stages (egg, larva, adult) of *Drosophila*, it is found
306 intermittently in cells, e.g. found in *Drosophila* larvae early G1 and not in late G1 interphase of
307 mitosis (Csink and Henikoff, 1998). Moreover, some organisms' cells possess a Rabl-like
308 configuration more often during mitosis as compared to other organisms (Idziak et al. 2015).
309 Therefore, what is visualized on the Hi-C map is not that all cells possess the Rabl-like
310 configuration all the time; instead, it is an average of the occurrence in a particular organism's
311 cells. While the Rabl-like configuration is the predominant chromosomal arrangement observed
312 thus far in Diptera and some Coleoptera, its evolutionary significance remains unclear. Genomic
313 architecture's influence on diversity, if any, is hindered by the sparse, haphazard sampling of
314 insect genomes. It may be tempting to ascribe patterns to clades (such as Lepidoptera being

highly conserved in their genome architecture), but such patterns fade in a broader phylogenetic context or remain to be fully tested. Rather than one to one comparisons, it is more meaningful to describe patterns for a clade in a broader phylogenetic context.

In summation, we have reconstructed one of the largest and most repetitive arthropod genomes. With the combination of Hi-C reads and PacBio long-read sequencing data, we were able to resolve a highly contiguous, chromosome-level genome. We find patterns of genomic architecture, specifically, synteny across Insecta, largely scales with clade age, with some groups, such as Lepidoptera and Diptera, showing subtly different patterns.

METHODS

Taxon selection and natural history

Pachyrhynchus, from the entirely flightless tribe Pachyrhynchini, is found from the Philippines to Papua New Guinea, Australia, Taiwan, Japan, and Indonesia (Schultze, 1923; Alonso-Zarazaga & Lyal, 1999). They are known for their bright, iridescent and unique elytral markings, which they use as an aposematic signal to warn predators of their unpalatability (Tseng et al., 2014). Members of other weevil groups (e.g. *Polycatus*, *Eupyrgops*, *Neopyrgops*, *Alcidodes*) and long-horned beetles (e.g. *Doliops*, *Paradoliops*) mimic *Pachyrhynchus*' aposematic signals to ward off predators. Currently, the Pachyrynchini has 17 known genera, with the majority found exclusively in the Philippines (Schultze, 1923; Yap & Gapud, 2007; Yoshitake, 2013, 2018).

Pachyrhynchus Germar, 1824 has the widest geographic range among Pachyrynchini. There are presently 145 species in the genus, of which 93% of are endemic to the Philippines (Rukmane, 2018), with the majority of species having a narrow geographic range, limited to a mountain range, island, or Pleistocene Aggregate Island Complex (PAIC) (Inger 1954, Heaney 1985, Brown and Siler 2014). The general diagnostic characters of *Pachyrhynchus* Germar, 1824 include a head lacking a distinct transverse groove or distinct basal border, entire episternal suture, and antennal scape not reaching the hind eye (Schultze, 1923; Yoshitake, 2012). *P. sulphureomaculatus* Schultze, 1922, is only recorded from Mindanao Island (Schultze, 1922; Cabras et al., 2017; Rukmane, 2018). This species was described from material collected in South Cotabato but has recently been recorded (personal observations of A. Cabras) in other areas of Mindanao (e.g. Marilog, Davao City, Arakan, Cotabato, Mt. Kiamo, Bukidnon). This species belongs to the *P. venustus* group, conspicuous for their large size, prothorax with two dorsolateral spots in the middle a large, oblong spot at the lateral margins, and elytra with oval or oblong spots (Schultze, 1923).

Collection and extraction of DNA

Specimens were collected near the edge of the road in a secondary forest (HWY 81, Arakan, Cotabato, Philippines [N7.487059, E125.248795]). One individual was used for both in situ Hi-C and high molecular weight DNA libraries. A second individual was used for transcriptome sequencing. Individuals were collected live, then frozen and stored at -80°C until library preparation.

Beetle tissues were dissected carefully to avoid inclusion of contaminants from guts and impurities from chitinous cuticles. Half of the resulting tissues were used for Phenol Chloroform

(PCI) based high molecular weight (HMW) DNA extraction for PacBio sequencing (the other half of the material was used as starting material for Hi-C library preparation, see below). Tissues were homogenized on ice using a sterile razor blade. ATL buffer (140 µl) and Proteinase K (60 µl) were then added to the homogenized material and incubated at 65°C for 1 hr. The 200 µl of resulting lysate was used as starting material for the PCI extraction following a PacBio recommended protocol. (https://www.pacb.com/wp-content/uploads/2015/09/SharedProtocol-Extracting-DNA-usinig-Phenol-Chloroform.pdf). Two additional rounds of PCI clean-up were performed to eliminate impurities such as chitin to meet the DNA requirement for PacBio sequencing. In particular, to achieve OD ratios of 1.8–2.0. DNA concentration was determined with the Qubit™ dsDNA HS Assay Kit (Invitrogen corp., Carlsbad, CA), and high molecular weight content was confirmed by running a Femto Pulse (Agilent, Santa Clara, USA).

In situ Hi-C library preparation

Tissues from the same sample were homogenized using a sterile razor blade on ice. An in situ Hi-C library was prepared as described in Rao et al. (2014) with a few modifications. Briefly, after the Streptavidin Pull-down step, the biotinylated Hi-C products underwent end repair, ligation and enrichment using the NEBNext® Ultra™II DNA Library Preparation kit (New England Biolabs Inc, Ipswich, MA). Furthermore, titration of the number of PCR cycles was performed as described in Belton et al. (2012).

Transcriptome library preparation

RNA extraction was performed using tissues from a frozen sample. Tissue was extracted from the prothorax and abdomen with the digestive tract removed. The Monarch Total RNA

Miniprep kit (New England Biolabs Inc, Ipswich, MA) was used for extraction. The manufacturer's protocol for total RNA purification from tissue was followed (cite) (https://www.neb.com/protocols/2017/11/08/total-rna-purification-from-tissues-and-leukocytes-using-the-monarch-total-rna-miniprep-kit-neb-t2010). RNA concentration was determined using the Qubit® RNA HS Assay Kit (Invitrogen corp., Carlsbad, CA), and intact RNA content was confirmed by running a Bioanalyzer High Sensitivity RNA Analysis (Agilent, Santa Clara, USA). The resulting RNA was sent to Novogene Inc. for library preparation and sequencing, from which 12.5 Gbp of data were obtained.

Genome sequencing and assembly

First, we performed an initial quality control of the in situ Hi-C library using the CPU version of *Juicer* v 1.5.7 (Durand et al. 2016) to determine if enough ligation motifs were present in the sample. To accomplish this, we first cleaned our reads with *fastp* (Chen et al. 2018) to remove sequencing adapters and low quality reads with default settings except for the more sensitive '--detect_adapter_for_pe' setting on. After passing the quality control of having >30% ligation motifs present, we proceeded to sequence the full library at higher coverage. We only considered ligation motifs as this was a de novo assembly without a closely related reference genome to align to the Hi-C reads. The full Hi-C library was sequenced on a paired-end (2x150 bp) lane on an Illumina HiSeq4000. High molecular weight DNA was sent to the QB3 Genomics facility at the University of California Berkeley for sequencing on a Pacific Biosciences Sequel II platform, sequencing one cell with CLR version 2 chemistry (PacBio, Menlo Park, CA, USA).

We used *PacBio Assembly Tool Suite pb-assembly* v 0.0.8 (which includes the FALCON assembly pipeline) to assemble the primary scaffolds. Next, we polished the primary assembly using 3 rounds of mapping the raw fastq reads using *minimap2* (Li 2018) followed by using

RACON (Vaser 2017) to help error correct the initial assembly. This was followed by running the *Purge_Haplotigs* (Roach et al. 2018) pipeline to eliminate haplotigs (alternative haplotype contigs) in the assembly. Next, using the CPU version of *Juicer* v 1.5.7, we created a site positions file for the restriction enzyme MboI using *Juicer*'s *generate_site_positions.py* script, followed by running *Juicer* until it creates the mapping stats file and a "*merged_nodups*" file. Then we used the 3D-DNA (Dudchenko et al. 2017) pipeline with default settings to correct misjoins and place scaffolds into chromosome groups. After generating a Hi-C heat map, we corrected any assembly errors manually via *Juicebox Assembly Tools* v 1.11.08 (Durand et al. 2016, Dudchenko et al. 2018). After, (Fig 1.) we ran 3D-DNA's *run-asm-pipeline-post-review.sh* to produce a final assembly file and fasta. To polish our final assembly further, we aligned our Hi-C reads to our scaffolds using *bwa mem* followed by *SAMclip* and *SAMtools* 'view' (Li et al. 2009) with options '-S -b -f 2 -q 1 -F 1536'. After grouping scaffolds into chromosomes, we divided each into a separate fasta (due to memory constraints) and used *Pilon* (v. 1.23) (Walker et al. 2014) in "--fix bases" mode as to not break our scaffolds and to fix any homopolymer repeat errors. The resulting assembly was used in all subsequent analyses.

Removal of mitochondrial/contaminant DNA

To identify scaffolds that contained mitochondrial cytochrome oxidase subunit 1 (COI) DNA, we used BLAT v. 35 (Kent 2002; 2012) using a reference sequence from *Pachyrhynchus smaragdinus* (Supplemental Information file "P79_coI.fasta") to query our scaffolds. Once identified, these scaffolds were removed. We also used *blast* (Camacho et al. 2008) with the nt database and default settings to identify contaminant (non-arthropod or undetermined) sequences and then removed these from the final assembly. These represented only a handful of sequences.

Repeat content analyses

To address what is making the genome of *Pachyrhynchus sulphureomaculatus* so large relative to other complete weevil genomes (>85% Benchmarking Universal Single-Copy Orthologs BUSCO Insecta genes), we compared the repeat content of *P. sulphureomaculatus* to 5 other weevil genomes from NCBI (Supplemental Information file "NCBI_numbers_for_Weevils_used_in_repeatmasker"). We used the de novo *RepeatModeler* v. open-1.0.11 (Smit et al. 2015) repeat set combined with all repbase recs to first model for repeat content. Next, we used *RepeatMasker* v. 4.1.0 (Smit et al. 2015) to annotate and soft mask repeat content. For *Listronotus*, we downloaded the results from Harrop et al. (2020), who used comparable methodologies. We also calculated the percentage of repetitive content (bases soft masked) in a 1 Mb sliding window across the chromosomes in *R* using a custom script.

Genome annotation

We first cleaned our reads with *fastp* and concatenated the unpaired cleaned reads. We performed 3 different initial reconstructions of the transcriptome: 1) *Trinity* v. 2.11.0 (Grabherr et al. 2013; Haas et al. 2014) de novo assembly using default settings, 2) *Trinity* genome guided assembly, where we first aligned our reads with *tophat* v. 2.1.1 (Kim et al. 2013), 3) *rnaSPAdes* (Bushmanova et al. 2019) de novo assembly. Selecting the *rnaSPAdes* assembly, because it had the most single copy *BUSCO V2* Arthropoda genes (Felipe et al. 2015), we mapped our reads to this soft masked assembly using *HISAT2* v. 2.2.0 (Kim et al. 2019), and formatted a bam file using *SAMtools* 'view -b -f 3 -F 256 -q 10'. Next, we used *BRAKER v. 2.1.5* (Hoff et al. 2019) to create an annotated gff. This process used the bam file from *HISAT2* and

453 results from a *BUSCO* search as 'seeding' genes to make the resulting gff. In addition, we used
454 the *PASA* pipeline (Campbell et al. 2006; Haas 2008) which used our *rnaSPAdes* transcripts
455 aligned to the genome assembly with *BLAT* (Kent 2002) and *gmap* (Wu and Watanabe 2005).
456 Lastly, we used *EVidenceModeler* (Haas et al. 2008) to evaluate our different annotations using
457 the developers' recommended weights for each assembly type to produce the final gene model
458 gff.
459
460 ***Synteny across coleopteran and Insecta chromosome-level genomes***
461 To examine the gene synteny between other Coleoptera genomes, we downloaded chromosome-
462 level genomes from NCBI or supplied form the journal or authors website (Supplemental
463 Information file "NCBI_all_taxa_genomes_list") (Fallon et al. 2018; Zhang et al. 2020; Herndon
464 et al. 2020; Van Belleghem 2018). We also used the unpublished genome assemblies (*Tribolium*
465 *castaneum* [GCF_000002335.3], *Bombyx mori* [GCA_000151625.1], *Clogmia albipunctata*
466 [clogmia.6], *Culex quinquefasciatus* [CpipJ3], and *Rhodnius prolixus* [Rhodnius_prolixus-
467 3.0.3]), generated by the DNA Zoo Consortium (dnazoo.org). The assemblies were based on the
468 whole genome sequencing data from (Herndon et al., 2020; *Tribolium* Genome Sequencing
469 Consortium 2008, International Silkworm Genome Consortium 2008, Arensburger et al. 2010,
470 Mesquita et al. 2015) as well as Hi-C data generated by the DNA Zoo Consortium and
471 assembled using 3D-DNA (Dudchenko et al., 2017) and Juicebox Assembly Tools
472 (Dudchenko et al., 2018). Next, we identified the BUSCO v.2 loci, (1658 Insecta gene set) and
473 extracted their coordinates for the single and fragmented loci. We then compared the coordinates
474 of *Pachyrhynchus sulphureomaculatus* to the other Coleoptera genomes. Following, we
475 calculated the number of loci found in *P. sulphureomaculatus* chromosomes and those in the

476 other Coleoptera and calculated the percent conserved within a chromosome. To visualize the
477 shared synteny, we plotted the different pairs using the R package *RIdeogram* (Hao et al. 2020).
478 To help visualize the relationship between the different taxa, we generated an ultraconserved
479 elements (UCE) dataset between the taxa using the PHYLUCE pipeline (Faircloth 2012). We
480 used the loci to help reconstruct a concatenated phylogeny in RAxML (Stamatakis 2014) and
481 calculate branch lengths to render the tree ultrametric. We dated the tree using dates (95%
482 highest posterior density interval HPD) from McKenna et al. (2019) using the R package *ape*
483 v.5.4 'makeChronosCalib' function (Paradis and Schliep 2019) (see Supplemental Information
484 file "Insecta_Claibrations_table" for dates).
485 Next, we investigated whether the observed synteny was distinctive within Coleoptera
486 relative to other orders of insects, such as Lepidoptera, in which high levels of synteny between
487 taxa have been recorded (Hill et al. 2019, Ahola et al. 2014). We used all insect genomes (with
488 some exceptions) available from NCBI that were marked as "chromosome" level. (See
489 Supporting Information for a complete list.) We tried to sample evenly across insect orders. For
490 example, we excluded the many *Drosophila* genomes as they are all phylogenetically close
491 relatives, and this would cause over-representation (i.e. we want patterns of chromosomal
492 evolution across Diptera, not just *Drosophila*). Instead, we sampled individual species across the
493 phylogenetic breadth of the genus. In addition, we also gathered genomes from the literature.
494 (see Supplemental Information file "NCBI_all_taxa_genomes_list".) Next, we identified all
495 BUSCO version 5-beta loci that were single copy and calculated the gene order conservation
496 (GOC) score (see https://m.ensembl.org/) using a custom script (Supplemental Information files
497 "1make_scaff_order_busco_tsv.sh", and "2busco_GOC.sh"). First, we ordered the BUSCO v5-
498 beta genes by scaffold and position and then identified two genes upstream and downstream

from a particular gene. Next, to determine if a set of 4 genes are in the same order in our target genome, they receive a score of 1, 0.75, 0.5, 0.25 or 0 based on whether 4, 3, 2, 1 or 0 genes are in the same order, respectively. Missing genes between the two genomes are discarded from comparisons. This process is repeated along the length of the two genomes. We then summed the scores for the four categories 0-100% and added these categories together (e.g. if 8 matched *sets* were found at 25% and 1 at 100%, the total score would be 5). These total scores were normalized by dividing by the minimum number of genes present in the comparisons. We computed the total GOC scores for all pairwise comparisons among the 34 taxa. Next, to consider the effect of the phylogenetic relationships, we reconstructed the relationship among our taxa using the BUSCO gene sets' amino acids. We used custom scripts to identify a 70% complete matrix and used *mafft* with 1000 iterations and the "localpair" settings to align the sequences. Next, we used *trimAI* (Capella-Gutierrez et al. 2009) with "automated1" settings to remove ambiguously aligned positions. RAxML-ng with the LG+G8+F site rate substitution model was used to reconstruct the phylogeny for our exemplar taxa across Insecta. We calibrated our tree using the same methods for the beetle tree (above), and calibration points can be found in the Supplemental Information file "Insecta_Claibrations_table", from Misof et al 2014, Obbard et al. 2012 and Mckenna et al. 2019). This calibration was done to help visualize the data as the subsequent Mantel test did not require an ultrametric tree. Lastly, to test if pairwise phylogenetic distance covaries with pairwise synteny values, we conducted a Mantel test.

Data availability

Assembly available at https://www.dnazoo.org/

Supplemental Information available on Dryad [link pending submission]

522 Sequencing data from this project are archived on the SRA under accession #######
523 [embargoed until 2021-11-01 or publication].
524

525 **Acknowledgments**

526 We would like to thank the Ruth Tawan-tawan, Ceso II of the Philippines' Department of
527 Environment and Natural Resources Region XI for help with the Gratuitous and export permits.
528 We would also like to thank the University of Mindanao for the mobility support, and Milton N.
529 Medina and Chrestine Torrejos of U.M. for help collecting specimens. We would like to thank
530 Zane Colaric of B.C.M., for the help loading the QC library runs. We would also like to thank
531 Sarah Crews of C.A.S. for help with the manuscript text. We were funded in part through
532 NSF:DEB award number 1856402 made to MHVD.
533

534 **Figure Legends**
535

536 **Figure 1.** *Pachyrhynchus sulphureomaculatus*, lateral habitus. (Photo by A. Cabras)
537

538 **Figure 2.** Hi-C contact map heatmap of *Pachyrhynchus sulphureomaculatus* Schultze, 1922.
539 Eleven chromosome boundaries are indicated by black lines. Heatmap scale lower left, range in
540 counts of mapped Hi-C reads per megabase squared. Rabl-like pattern highlighted along
541 chromosome 1, top row, open triangles indicate contact between centromere regions. X-like
542 pattern between adjacent off diagonal regions indicative of contact between distal portions of
543 chromosomes.
544

Figure 3. *P. sulphureomaculatus* scaffold bubble plot of coverage versus GC content. Scaffolds included are from the unfiltered assembly. Taxonomic annotation provided via *blastn* alignment to the NCBI nt database.

Figure 4. Stacked bar plot of Insecta BUSCO gene sets by category for chromosome-level beetle genomes. Y-axis is the percent of BUSCO genes, X-axis labels are the genus names. The abbreviations in the legend are: D=duplicated, F=fragmented, M=missing and S=single.

Figure 5. Histogram of repeat content for weevil genomes.

Figure 6. Heat map of gene density and non-repetitive DNA per 1 Mb sliding window. The 11 chromosomes are in the same order as in the Hi-C heat map (Fig. 1) and fasta file.

Figure 7. Chronogram and ideograms of the 5 beetle genomes which have chromosome level assemblies.

Figure 8. Stacked bar plots and chromosome mappings of BUSCO genes' placements. Colors correspond to *P. sulphureomaculatus* chromosomes. The numbering scheme of chromosomes matches the names found in the genome's fasta file.

Figure 9. Insecta, gene order conservation score (GOC) plot. Left panel, chronogram, branches colored by order. Heat map of normalized pairwise GOC scores, redder boxes indicate more

567 synteny between pairs. Right panel, within order normalized pairwise GOC score versus
568 phylogenetic distance.

569
570
571
572
573 **TABLES**
574

Inter-chromosomal: 56,711,177 (24.85% / 38.63%)
Intra-chromosomal: 23,941,704 (10.49% / 16.31%)
Short Range (<20Kb): 18,227,329 (7.99% / 12.42%)
Long Range (>20Kb): 5,714,353 (2.50% / 3.89%)

575
576 **Table 1**. Summary of Hi-C reads mapped.
577
578
579
580
581
582

Number of scaffolds	4,093
Total size of scaffolds	2,051,389,195
Number of contigs	14,365
Number of contigs in scaffolds	10,283
Number of contigs not in scaffolds	4,082
Mean scaffold size	501,195
Median scaffold size	8,175
N50 scaffold length	215,921,627
scaffold %AT	33.11
scaffold %CG	16.87
scaffold %N	0.05
% bp of assembly in chromosomes	**97.52**

583
584 **Table 2**. Summary statistics for final assembly.
585

Chromosome	length bp	# of contigs	number of N's (runs of 100)	percent N's in chromo.	N50	N50 reached in # of contigs
Chr_1	263,832,947	1388	138,700	0.05%	287319	280
Chr_2	253,284,860	1222	122,100	0.04%	326447	246
Chr_3	137,890,936	683	68,200	0.04%	315991	127
Chr_4	223,502,247	1131	113,000	0.05%	297265	217
Chr_5	69,931,891	236	23,500	0.03%	452330	50
Chr_6	125,299,487	655	65,400	0.05%	304080	131
Chr_7	132,078,125	624	62,300	0.04%	368684	112
Chr_8	173,282,956	836	83,500	0.04%	335626	165
Chr_9	215,921,627	1221	122,000	0.05%	280522	237
Chr_10	186,927,849	988	98,700	0.05%	290475	197
Chr_11	218,628,933	1074	107,300	0.04%	316887	219

Table 3. Summary statistics for final assembly by chromosome.

REFERENCES

Ahola V, Lehtonen R, Somervuo P, Salmela L, Koskinen P, Rastas P, Välimäki N, Paulin L, Kvist J, Wahlberg N, et al. 2014. The Glanville fritillary genome retains an ancient karyotype and reveals selective chromosomal fusions in Lepidoptera. *Nature Communications* 2014;5 doi: 10.1038/ncomms5737. Article 4737

Alonso-Zarazaga MA and Lyal CHC. 1999. *A World Catalogue of Families and Genera of Curculionoidea (Insecta: Coleoptera). (Excepting Scolytidae and Platypodidae)*. 124 pp. Entomopraxis, Barcelona.

606 Anbutsu H, Moriyama M, Nikoh N, Hosokawa T, Futahashi R, Tanahashi M, Meng XY,
607 Kuriwada T, Mori N, Oshima K, Hattori M. 2017. Small genome symbiont
608 underlies cuticle hardness in beetles. *Proceedings of the National Academy of*
609 *Sciences* 114:E8382-E8391.

610 Arensburger P, Megy K, Waterhouse RM, Abrudan J, Amedeo P, Antelo B, Bartholomay L,
611 Bidwell S, Caler E, Camara F, et al. 2010. Sequencing of *Culex quinquefasciatus*
612 establishes a platform for mosquito comparative genomics. Science. 330(6000):86-8.
613 doi:10.1126/science.1191864.

614 Belton JM, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. 2012. Hi-C: a
615 comprehensive technique to capture the conformation of genomes. *Methods*
616 58:268-276.

617 Biello R, Singh A, Godfrey CJ, Fernández FF, Mugford ST, Powell G, Hogenhout SA,
618 Mathers TC. 2020. A chromosome level genome assembly of the woolly apple
619 aphid, *Eriosoma lanigerum* Hausmann (Hemiptera: Aphididae). *Molecular*
620 *Ecology Resources* 00: 1-11. https://doi.org/10.1111/1755-0998.13258

621 Bracewell R, Chatla K, Nalley MJ, Bachtrog D. 2019. Dynamic turnover of centromeres
622 drives karyotype evolution in *Drosophila*. *eLife* 2019;8:e49002.

623 Brown RM and Siler CD. 2014. Spotted stream frog diversification at the Australasian
624 faunal zone interface, mainland versus island comparisons, and a test of the
625 Philippine 'dual-umbilicus' hypothesis. *Journal of Biogeography*. 41:182-195.

626 Bushmanova E, Antipov D, Lapidus A, Prjibelski AD. 2019. rnaSPAdes: a *de novo*
627 transcriptome assembler and its application to RNA-Seq data. *GigaScience*.
628 8:giz100, https://doi.org/10.1093/gigascience/giz100
629
630 Campbell MA, Haas BJ, Hamilton JP, Mount SM, Buell CR. 2006. Comprehensive
631 analysis of alternative splicing in rice and comparative analyses with Arabidopsis.
632 *BMC Genomics* 2006, 7:327 http://www.biomedcentral.com/1471-2164/7/327

633 Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor.
634 *Bioinformatics* 34:884-890. doi: 10.1093/bioinformatics/bty560

635 Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. 2009. trimAl: a tool for automated
636 alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972-
637 1973.

638 Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL.
639 2008. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421.
640
641 Cabras A, Coritico F, Mohagan A, Rukname A. 2017. Pachyrynchini of Mt. Kiamo,
642 Malaybalay, Bukidnon, Mindanao Island. *Journal of Entomology and Zoology*
643 *Studies*. 5(3):979-983.
644

Csink AK and Henikoff S. 1998. Large-scale chromosomal movements during interphase progression in *Drosophila*. *J. Cell Biol.* 143:13-22.

Chowdhary BP, Raudsepp T, Frönicke L, Scherthan H. 1998. Emerging Patterns of Comparative Genome Organization in Some Mammalian Species as Revealed by Zoo-FISH. *Genome Res.* 8:577-589. doi:10.1101/gr.8.6.577

Ceja-Navarro JA, Karaoz U, Bill M, Hao Z, White RA 3rd, Arellano A, Ramanculova L, Filley TR, Berry TD, Conrad ME, et al. 2019. Gut anatomical development and microbial functional assembly promote lignocellulose deconstruction and colony subsistence of a wood-feeding beetle. *Nat. Microbiol.* 4:864-75.

Concia L, Veluchamy A, Ramirez-Prado JS, Martin-Ramirez A, Huang Y, Perez M, Domenichini S, Rodriguez Granados NY, Kim S, Blein T, Duncan S, et al. Wheat chromatin architecture is organized in genome territories and transcription factories. *Genome Biol.* 21(1):104. doi: 10.1186/s13059-020-01998-1.

Cowan CR, Carlton PM, Cande WZ. 2001. The Polar Arrangement of Telomeres in Interphase and Meiosis. Rabl Organization and the Bouquet. *Plant Physiology* 125(2):532-538. DOI: 10.1104/pp.125.2.532

Davey JW, Chouteau M, Barker SL, Maroja L, Baxter SW, Simpson F, Merrill RM, Joron M, Mallet J, Dasmahapatra KK, Jiggins CD. 2016. Major Improvements to the Heliconius melpomene Genome Assembly Used to Confirm 10 Chromosome Fusion Events in 6 Million Years of Butterfly Evolution. *G3 (Bethesda)* 6(3):695-708.

Deakin JE. 2018. Chromosome Evolution in Marsupials. *Genes*. *9*: 72.

Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Shamim MS, Machol I, Lander ES, Aiden AP, Aiden EL. 2017. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* 356:92-95. https://doi.org/10.1126/science.aal3327.

Dudchenko O, Shamim MS, Batra S, Durand NC, Musial NT, Mostofa R, Pham M, et al. 2018. The Juicebox Assembly Tools Module Facilitates de Novo Assembly of Mammalian Genomes with Chromosome-Length Scaffolds for under $1000. *bioRxiv*, January, 254797. https://doi.org/10.1101/254797

Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, Aiden EL. 2016. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst.* 3(1):95-8. doi:10.1016/j.cels.2016.07.002.

Eagen KP, Aiden EL, Kornberg RD. 2017. Polycomb-mediated chromatin loops revealed by a subkilobase-resolution chromatin interaction map. *Proc. Natl. Acad. Sci.* 114:8764-8769.

Fallon TR, Lower SE, Chang CH, Bessho-Uehara M, Martin GJ, Bewick AJ, Behringer M, Debat HJ, Wong I, Day JC, et al. 2018. Firefly genomes illuminate parallel origins of bioluminescence in beetles. *eLife*. 2018;7:e36495. pmid:30324905

Ghurye J, Pop M, Koren S, Bickhart D, Chin CS. 2017. Scaffolding of long-read assemblies using long range contact information. *BMC genomics*. 18(1):527.

Ghurye J, Rhie A, Walenz BP, Schmitt A, Selvaraj S, Pop M, Phillippy AM, Koren S. Integrating Hi-C links with assembly graphs for chromosome-scale assembly. 2019. *PLoS Comput Biol.* 15(8):e1007273. doi:10.1371/journal.pcbi.1007273.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 29(7):644-52. doi: 10.1038/nbt.1883.

Goto B, Okazaki K, Niwa O. 2001. Cytoplasmic microtubular system implicated in de novo formation of a Rabl-like orientation of chromosomes in fission yeast. *J Cell Sci.* 114:2427-35.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Philip D, Bowden J, Couger MB, Eccles D, Li B, Macmanes MD, Ott M, Orvis J, Pochet N: *Reference Generation and Analysis with Trinity. Volume 8*; 2014.

Haas BJ. 2008. Analysis of Alternative Splicing in Plants with Bioinformatics Tools. Nuclear pre-mRNA Processing in Plants.

Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. 2008. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* 9(1):R7. doi: 10.1186/gb-2008-9-1-r7.

Hammond PM. 1992. Species inventory. Global Biodiversity. Status of the Earth's Living Resources. A Report Compiled by the World Conservation Monitoring Centre, ed Groombridge B (Chapman and Hall, London), pp 17-39.

Hao Z, Lv D, Ge Y, Shi J, Weijers D, Yu G, Chen J. 2020. RIdeogram: drawing SVG graphics to visualize and map genome-wide data on the idiograms. *PeerJ Comput. Sci*. 6:e251 http://doi.org/10.7717/peerj-cs.251

Harrop T, Le Lec MF, Jauregui R, Taylor SE, Inwood SN, van Stijn T, Henry H, Skelly J, Ganesh S, Ashby RL, Jacobs J, Goldson SL, Dearden PK. 2020. Genetic Diversity in Invasive Populations of Argentine Stem Weevil Associated with Adaptation to Biocontrol. *Insects*, *11*(7), 441. https://doi.org/10.3390/insects11070441

Heaney LR, 1985. Zoogeographic evidence for middle and late pleistocene land bridges to the philippine islands. *Mod Quatern Res SE Asia* 9:127-144.

Hill J, Rastas P, Hornett EA, Neethiraj R, Clark N, Morehouse N, de la Paz Celorio-Mancera M, Cols JC, Dircksen H, Meslin C, et al., 2019. Unprecedented reorganization of holocentric chromosomes provides insights into the enigma of lepidopteran chromosome evolution. *Sci. Adv.* 5:1-13.

Edelman NB, Frandsen PB, Miyagi M, Clavijo B, Davey J, Dikow R, García-Accinelli G, Van Belleghem SM, Patterson N, Neafsey DE, et al. Genomic architecture and introgression shape a butterfly radiation. 2019. *Science*. 366(6465):594-599. doi: 10.1126/science.aaw2090

Herndon N, Shelton J, Gerischer L, Ioannidis P, Ninova M, Dönitz J, Waterhouse RM, Liang C, Damm C, Siemanowski J, et al. 2020. Enhanced genome assembly and a new official gene set for *Tribolium castaneum*. *BMC Genomics*. 21(1):47. doi: 10.1186/s12864-019-6394-6.

Hoff KJ, Lomsadze A, Borodovsky M, Stanke M. 2019. Whole-Genome Annotation with BRAKER. *Methods Mol Biol*. 1962:65-95. doi:10.1007/978-1-4939-9173-0_5

Hsu CF, Tseng HY, Hsiao Y, Ko CC. 2017. First record of the host plant and larvae of *Pachyrhynchus sonani* (Coleoptera: Curculionidae) on Lanyu Island, Taiwan. *Entomological Science*, 20:288-291. doi: 10.1111/ens.12262.

Idziak D, Robaszkiewicz E, Hasterok R. 2015. Spatial distribution of centromeres and telomeres at interphase varies among Brachypodium species. *Journal of Experimental Botany*. *66*(21): 6623–6634. https://doi.org/10.1093/jxb/erv369

Inger RF, 1954. Systematics and zoogeography of Philippine Amphibia. *Fieldiana* 33:182-531.

Jin Q, Trelles-Sticken E, Scherthan H, Loidl J. 1998. Yeast nuclei display prominent centromere clustering that is reduced in nondividing cells and in meiotic prophase. *The Journal of Cell Biology* 141:21-29.

Kingan SB, Urban J, Lambert CC, Baybayan P, Childers AK, Coates B, Scheffler B, Hackett K, Korlach J, Geib SM. 2019. A high-quality genome assembly from a single, field-collected spotted lanternfly (Lycorma delicatula) using the PacBio Sequel II system. *GigaScience*, 8(10): giz122, https://doi.org/10.1093/gigascience/giz122

Kemkemer C, Kohn M, Cooper DN, Froenicke L, Högel J, Hameister H, Kehrer-Sawatzki H. 2009. Gene synteny comparisons between different vertebrates provide new insights into breakage and fusion events during mammalian

karyotype evolution. *BMC evolutionary biology*, *9*, 84. https://doi.org/10.1186/1471-2148-9-84

Kent WJ. 2002. BLAT—The BLAST-Like Alignment Tool. Genome Research 12:656-664. DOI: 10.1101/gr.229202.

Kent WJ. 2012. BLAT—The BLAST-Like Alignment Tool. Version 35. Available at https://users.soe.ucsc.edu/~kent.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* 14**:**R36 (2013). https://doi.org/10.1186/gb-2013-14-4-r36.

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol.* 37(8):907-915. doi: 10.1038/s41587-019-0201-4.

Kim S, Liachko I, Brickner DG, Cook K, Noble WS, Brickner JH, Shendure J, Dunham MJ. 2017. The dynamic three-dimensional organization of the diploid yeast genome. *eLife* 2017;6:e23623 DOI: 10.7554/eLife.23623

Kingan SB, Heaton H, Cudini J, Holroyd N, Tracey A, Lambert CC, Baybayan P, Galvin B, Korlach J, Berriman M, Lawniczak, MKN. 2019. A Low DNA Input Protocol for High-quality PacBio De Novo Genome Assemblies from Single Invertebrate Individuals. https://www.pacb.com/proceedings/a-low-dna-input-protocol-for-high-quality-pacbio-de-novo-genome-assemblies-from-single-invertebrate-individuals/

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences, *Bioinformatics*, 34(18):3094-3100. https://doi.org/10.1093/bioinformatics/bty191

Liao Y, Zhang X, Chakraborty M, Emerson JJ. 2020. Topologically associating domains and their role in the evolution of genome structure and function in *Drosophila* bioRxiv 2020.05.13.094516; doi: https://doi.org/10.1101/2020.05.13.094516

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326:289-293.

Liu Q, Guo Y, Zhang Y, Hu W, Li Y, Zhu D, Zhou Z, Wu J, Chen N, Zhou X-N. 2019. A chromosomal-level genome assembly for the insect vector for Chagas disease, *Triatoma rubrofasciata*. *GigaScience*, 8(8): giz089 https://doi.org/10.1093/gigascience/giz089

828 Lu S, Yang J, Dai X, Xie F, He J, Dong Z, Mao J, Liu G, Chang Z, Zhao R, Wan W,
829 Zhang R, Li Y, Wang W, Li X. 2019. Chromosomal-level reference genome of
830 Chinese peacock butterfly (*Papilio bianor*) based on third-generation DNA
831 sequencing and Hi-C analysis. *GigaScience*, 8(11):giz128
832 https://doi.org/10.1093/gigascience/giz128
833
834 Lukyanchikova V, Nuriddinov M, Belokopytova P, Liang J, Reijnders MJMF, Ruzzante
835 L, Waterhouse RM, Tu Z, Sharakhov IV, Fishman V. 2020. *Anopheles*
836 *mosquitoes* revealed new principles of 3D genome organization in insects.
837 bioRxiv 2020.05.26.114017; doi: https://doi.org/10.1101/2020.05.26.114017
838
839 Lukhtanov VA, Dincă V, Friberg M, Šíchová J, Olofsson M, Vila R, Marec F, Wiklund
840 C. 2018. Versatility of multivalent orientation, inverted meiosis, and rescued
841 fitness in holocentric chromosomal hybrids. *Proc Natl Acad Sci. 115(41):E9610-*
842 *E9619*.
843
844 Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, Wicker T, Radchuk V,
845 Dockter C, Hedley PE, Russell J, Bayer M, et al. 2017. A chromosome
846 conformation capture ordered sequence of the barley genome. *Nature*
847 544(7651):427-433. https://doi.org/10.1038/nature22043.
848
849 Marec F, Tothova A, Sahara K, Traut W. 2001. Meiotic pairing of sex chromosome
850 fragments and its relation to atypical transmission of a sex-linked marker in
851 *Ephestia kuehniella* (Insecta: Lepidoptera). Heredity. 87:659–671.
852
853 McKenna D.D., Sequeira A.S., Marvaldi A.E., Farrell B.D. 2009. Temporal lags and
854 overlap in the diversification of weevils and flowering plants. *Proceedings of the*
855 *National Academy of Sciences*. 106:7083-7088.
856
857 McKenna DD, Shin S, Ahrens D, Balke M, Beza-Beza C, Clarke DJ, Donath A, Escalona
858 HE, Friedrich F, Letsch H, et al. 2019. The evolution and genome basis of beetle
859 diversity. *Proceedings of the National Academy of Sciences* 116(49):24729-
860 24737.
861
862 Mathers TC, Wouters RHM, Mugford ST, Swarbreck D, van Oosterhout C, Hogenhout
863 SA. 2020. Chromosome-scale genome assemblies of aphids reveal extensively
864 rearranged autosomes and long-term conservation of the X chromosome.
865 BioRxiv. doi: 10.1101/2020.03.24.006411
866
867 Matthews BJ, Dudchenko O, Kingan SB, Koren S, Antoshechkin I, Crawford JE, Glassford WJ,
868 Herre M, Redmond SN, Rose NH, et al. 2018. Improved reference genome of *Aedes*
869 *aegypti* informs arbovirus vector control. *Nature*. 563(7732):501-507. doi:
870 10.1038/s41586-018-0692-z.
871
872 Mesquita RD, Vionette-Amaral RJ, Lowenberger C, Rivera-Pomar R, Monteiro FA, Minx P,
873 Spieth J, Carvalho AB, Panzera F, Lawson D, et al. 2015. Genome of *Rhodnius prolixus*,

874 an insect vector of Chagas disease, reveals unique adaptations to hematophagy and
875 parasite infection. *Proceedings of the National Academy of Sciences*. 112(48):14936-
876 14941. DOI: 10.1073/pnas.1506226112
877
878 Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, Frandsen PB, Ware J,
879 Flouri T, Beutel RG, et al. 2014. Phylogenomics resolves the timing and pattern
880 of insect evolution. *Science*. 346:763-767.
881
882 Mizuguchi T, Barrowman J, Grewal SIS. 2015. Chromosome domain architecture and
883 dynamic organization of the fission yeast genome. *FEBS Letters*. 589, doi:
884 10.1016/j.febslet.2015.06.008

885 Oberprieler RG. 2010. A reclassification of the weevil subfamily Cyclominae
886 (Coleoptera: Curculionidae). *Zootaxa* 2515:1-35.

887 Obbard DJ, Maclennan J, Kim K-W, Rambaut A, O'Grady PM, Jiggins FM. 2012.
888 Estimating Divergence Dates and Substitution Rates in the *Drosophila*
889 Phylogeny. *Molecular Biology and Evolution*, 29(11): 3459–3473.
890 https://doi.org/10.1093/molbev/mss150
891
892 Oberprieler RG, Marvaldi AE, Anderson RS. 2007. Weevils, Weevils, Weevils
893 Everywhere. Zootaxa 1668: 491-520. In Zhang, Z.-Q. and Shear, W.A. (eds.)
894 (2007) Linnaeus Tercentenary: Progress in Invertebrate Taxonomy. *Zootaxa*
895 1668:1-766.

896 Paradis E, Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and
897 evolutionary analyses in R. *Bioinformatics* 35:526-528.

898 Pouokam M, Cruz B, Burgess S, Segal MR, Vazquez M, Arsuaga J. 2019. The Rabl
899 configuration limits topological entanglement of chromosomes in budding yeast.
900 *Scientific Reports* 9:1-10. https://doi.org/10.1038/s41598-019-42967-4/

901 Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn
902 AL, Machol I, Omer AD, Lander ES, Aiden EL. 2014. A 3D Map of the Human
903 Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell*
904 159:1665-1680.
905
906 Renschler G, Richard G, Valsecchi CIK, Toscano S, Arrigoni L, Ramírez F, Akhtar A.
907 2019. Hi-C guided assemblies reveal conserved regulatory topologies on X and
908 autosomes despite extensive genome shuffling. Genes Dev. 33(21-22):1591-1612.
909 doi: 10.1101/gad.328971.119.
910
911 Rabl C. 1885. Ueber Zelltheilung Morphol Jahrbuch, 10:214-330
912
913 Roach MJ, Schmidt SA, Borneman AR. 2018. Purge Haplotigs: allelic contig
914 reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics*
915 19:460. https://doi.org/10.1186/s12859-018-2485-7

Rukmane A. 2018. An annotated checklist of genus *Pachyrhynchus* (Coleoptera: Curculionidae: Pachyrhynchini). *Acta Biol. Univ. Daugavp*. 18 (1):63-68.

SAMclip. https://github.com/tseemann/samclip. Accessed 1 May 2019

Sanborn AL, Rao SSP, Huang S-C, Durand NC, Huntley MH, Jewett AI, Bochkov ID, Chinnappan D, Cutkosky A, Li J, Geeting KP, Gnirke A, Melnikov A, McKenna D, Stamenova EK, Lander ES, Aiden EL. 2015. Extrusion predicts 3D genome engineering results. *Proc. Natl Acad. Sci. USA*. 112(47):E6456-E6465. DOI: 10.1073/pnas.1518552112

Schneider C, Woehle C, Greve C, D'Haese CA, Wolf M, Janke A, Bálint M, Hüttel B. 2020. Biodiversity genomics of small metazoans: high quality de novo genomes from single specimens of field-collected and ethanol-preserved springtails. bioRxiv 2020.08.10.244541; doi: https://doi.org/10.1101/2020.08.10.244541

Schultze W. 1922. Beitrag zur coleopteren fauna der Philippinen. *Philippine Journal of Science* 21(6): 569-596.

Schultze W. 1923. A monograph of the Pachyrrhynchid group of the Brachyderinae, Curculionidae: Part I. *Philippine Journal of Science* 23: 609-673.

Seppey M, Ioannidis P, Emerson BC, Pitteloud C, Robinson-Rechavi M, Roux J, Escalona HE, McKenna DD, Misof B, Shin S, Zhou X, Waterhouse RM, Alvarez N. 2019. Genomic signatures accompanying the dietary shift to phytophagy in polyphagan beetles. *Genome Biol* 20, 98 https://doi.org/10.1186/s13059-019-1704-5

Sheffer MM, Hoppe A, Krehenwinkel H, Uhl G, Kuss AW, Jensen L, Jensen C, Gillespie RG, Hoff KJ, Prost S. 2020 Chromosome-level reference genome of the European wasp spider *Argiope bruennichi*: a resource for studies on range expansion and evolutionary adaptation *bioRxiv* 2020.05.21.103564; doi: https://doi.org/10.1101/2020.05.21.103564

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs, *Bioinformatics*, 31(19):3210-3212. https://doi.org/10.1093/bioinformatics/btv351

Simison BW, Parham JF, Papenfuss TJ, Lam AW, Henderson JB. 2020. An Annotated Chromosome-Level Reference Genome of the Red-Eared Slider Turtle (*Trachemys scripta elegans*). *Genome Biol Evol.* 12(4):456-462. doi: 10.1093/gbe/evaa063.

Smit AFA, Hubley R, Green P. 2015. *RepeatMasker Open-4.0*. 2015 <http://www.repeatmasker.org>.

Song C, Liu Y, Song A, Dong G, Zhao H, Sun W, Ramakrishnan S, Wang Y, Wang S, Li T, Niu Y, Jiang J, Dong B, Xia Y, Chen S, Hu Z, Chen F, Chen S. 2018. The *Chrysanthemum nankingense* genome provides insights into the evolution and diversification of chrysanthemum flowers and medicinal traits. *Mol Plant*. 11(12):1482-1491. doi:10.1016/j.molp.2018.10.003

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies, *Bioinformatics*, 30:(9)1312-1313. https://doi.org/10.1093/bioinformatics/btu033

Stork NE, McBroom J, Gely C, Hamilton AJ. 2015. New approaches narrow global species estimates for beetles, insects and terrestrial arthropods. *Proc. Natl Acad. Sci. USA* 112, 7519–7523.

Tribolium Genome Sequencing Consortium, Richards S, Gibbs RA, Weinstock GM, Brown SJ, Denell R, Beeman RW, Gibbs R, Beeman RW, Brown SJ, et al. 2008. The genome of the model beetle and pest *Tribolium castaneum*. *Nature.* 452(7190):949-55. doi: 10.1038/nature06784.

Tseng H-Y, Lin C-P, Hsu J-Y, Pike DA, Huang W-S 2014. The Functional Significance of Aposematic Signals: Geographic Variation in the Responses of Widespread Lizard Predators to Colourful Invertebrate Prey. *PLoS ONE* 9(3):e91777. doi:10.1371/journal.pone.0091777

Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res*. 27(5):737-746. doi:10.1101/gr.214270.116

Van Belleghem SM, Vangestel C, De Wolf K, De Corte Z, Möst M, Rastas P, De Meester L, Hendrickx F. 2018. Evolution at two time frames: Polymorphisms from an ancient singular divergence event fuel contemporary parallel evolution. *PLoS Genet.* 14(11):e1007796. doi: 10.1371/journal.pgen.1007796.

Wang W, Guan R, Liu X, Zhang H, Song B, Xu Q, Fan G, Chen W, Wu X, Liu X, Wang J. 2019. Chromosome level comparative analysis of Brassica genomes. *Plant Mol Biol.* 99(3):237-249. doi: 10.1007/s11103-018-0814-x.

Wu TD, Watanabe CK. 2005. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics*. 21(9):1859-1875. https://doi.org/10.1093/bioinformatics/bti310

Yap SA, Gapud VP. 2007. Taxonomic review of the genus *Metapocyrtus* Heller (Coleoptera: Curculionidae: Entiminae: Pachyrrhynchini). *Philip. Entomol.* 21:115-135.

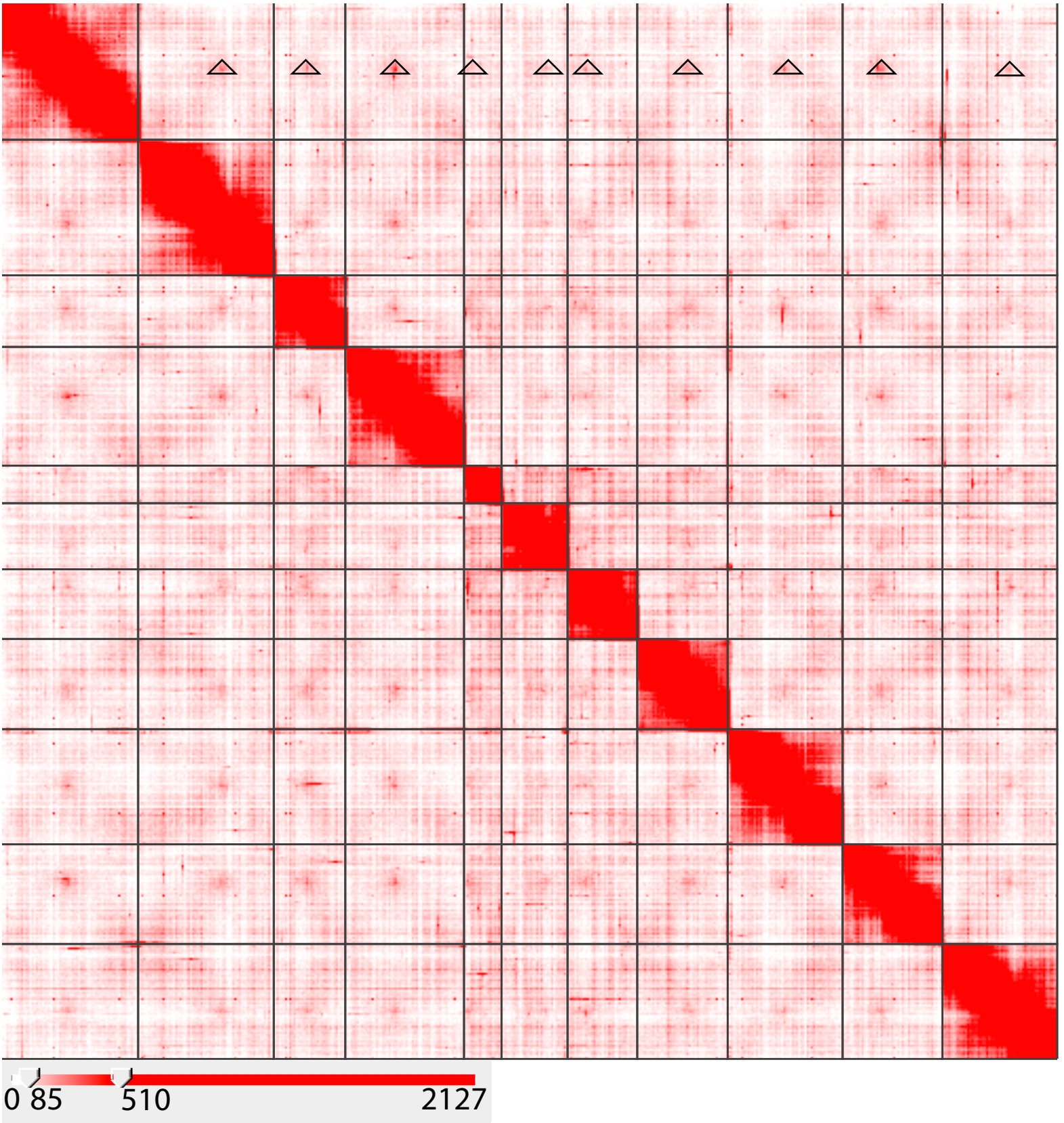
Yoshitake H. 2013. A New Genus and Two New Species of the Tribe Pachyrhynchini (Coleoptera: Curculionidae) from Palawan Island, the Philippines. *Esakia*. 53:1-8.

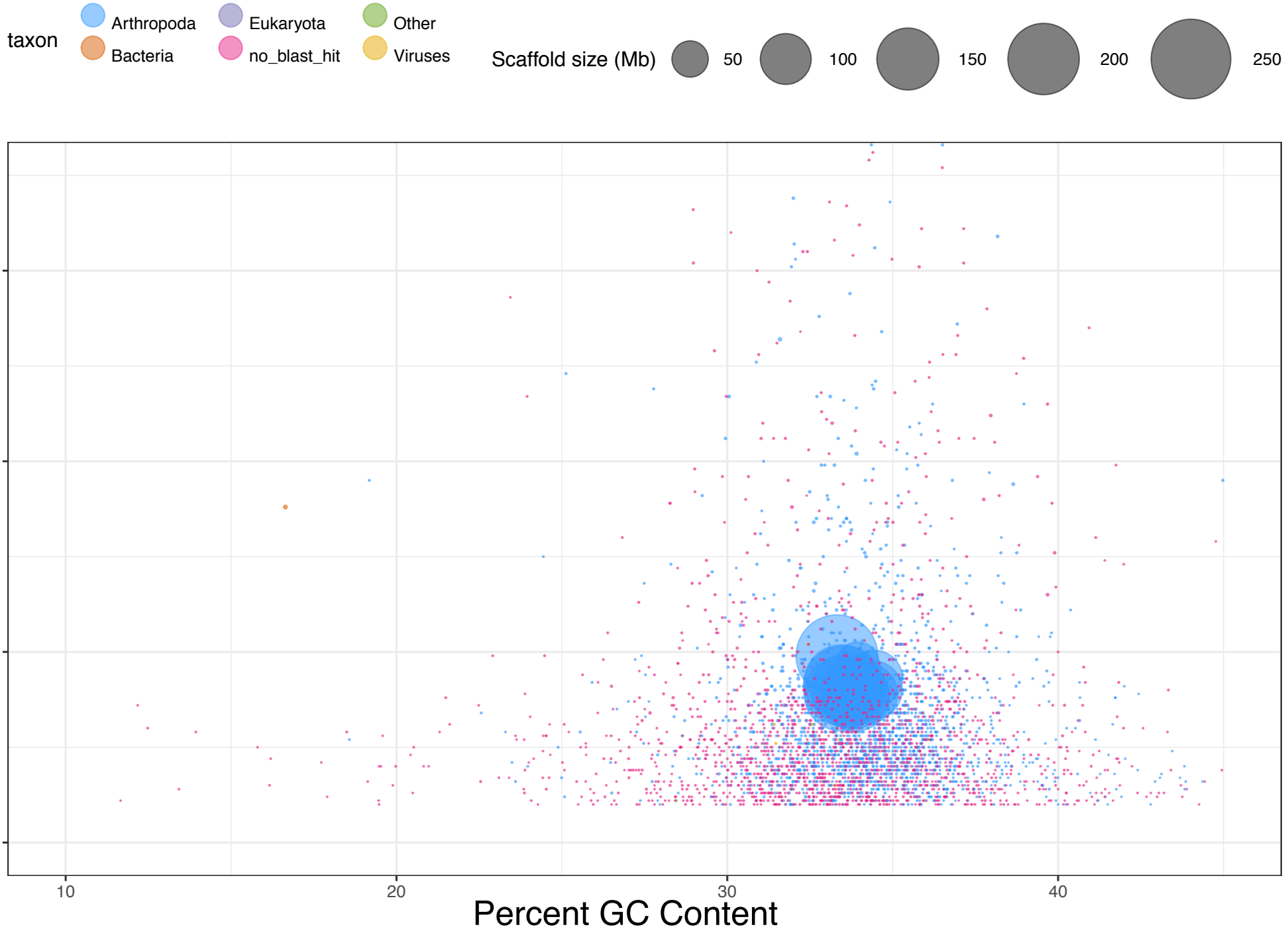
Yoshitake H. 2018. A new genus and two new species of the tribe Pachyrhynchini (Coleoptera, Curculionidae, Entiminae) from the Philippines. *Elytra* 8(1):5-14.

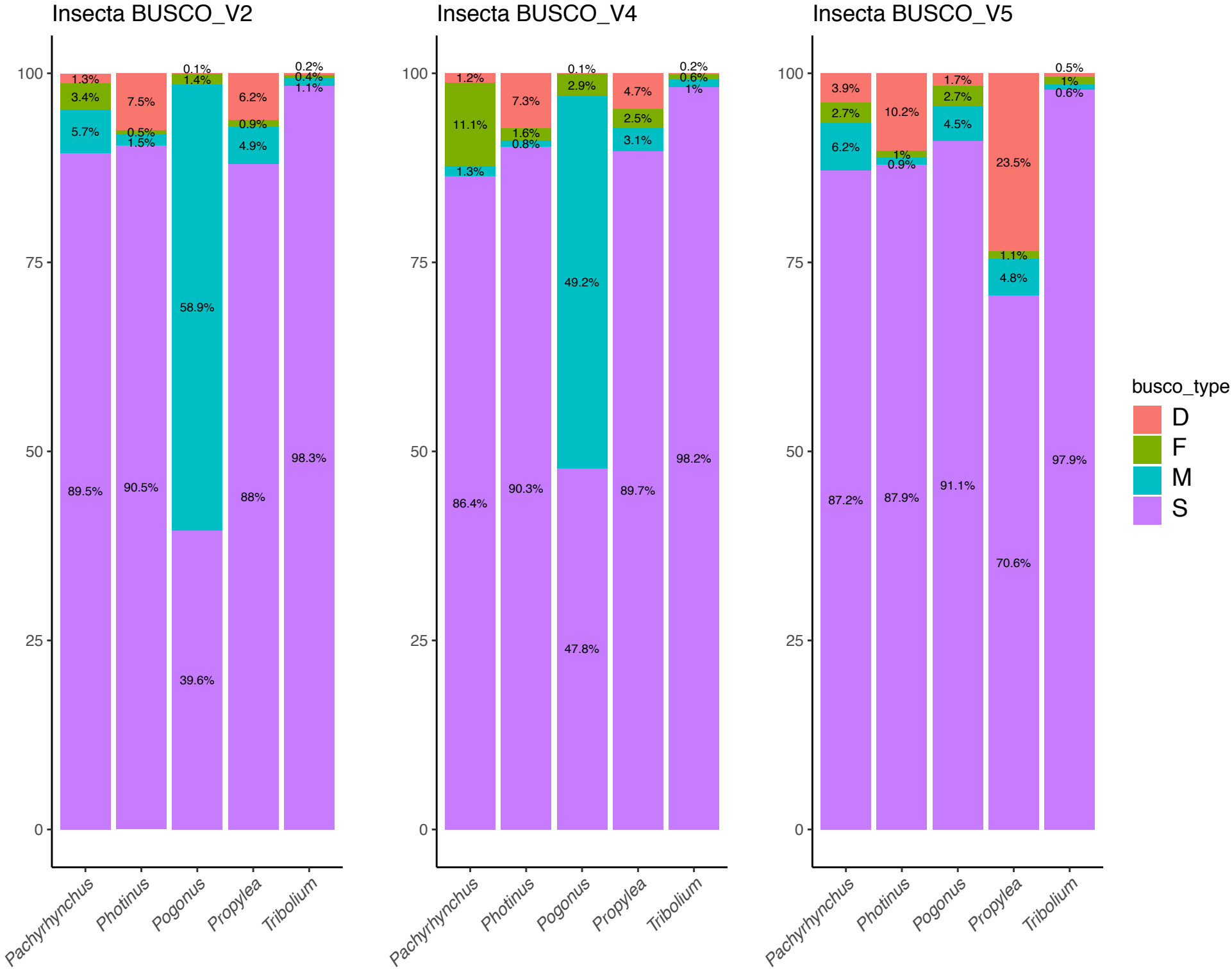
Zhang L, Li S, Luo J, Du P, Wu L, Li Y, Zhu X, Wang L, Zhang S, Cui J. 2020. Chromosome-level genome assembly of the predator *Propylea japonica* to understand its tolerance to insecticides and high temperatures. *Mol Ecol Resour*. 20: 292– 307. https://doi.org/10.1111/1755-0998.13100

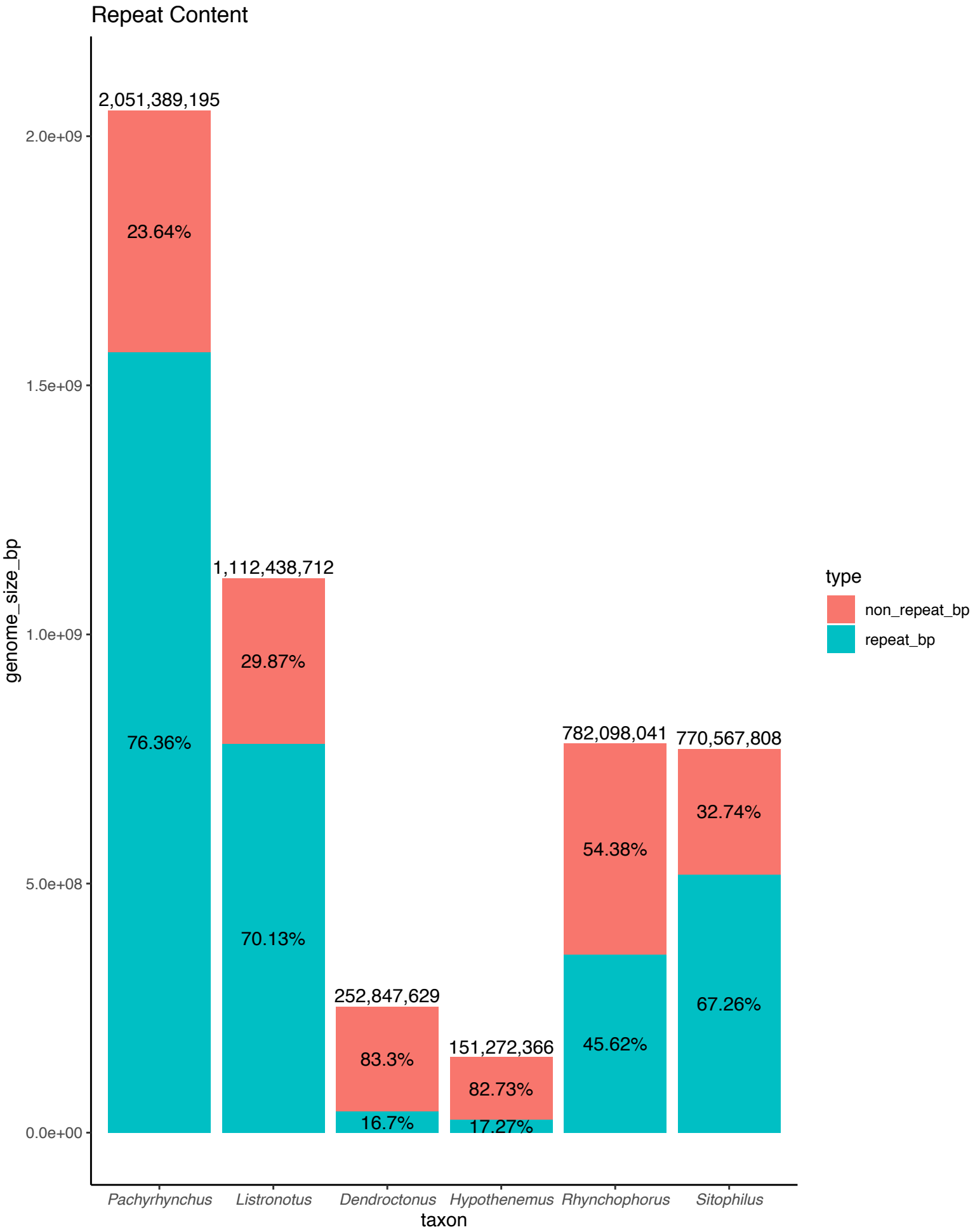
Zhou Y, Liang Y, Yan Q, Zhang L, Chen D, Ruan L, Kong Y, Shi H, Chen M, Chen J. 2020. The draft genome of horseshoe crab *Tachypleus tridentatus* reveals its evolutionary scenario and well developed innate immunity. *BMC Genomics*, 21(1):137-15. http://doi.org/10.1186/s12864-020-6488-1

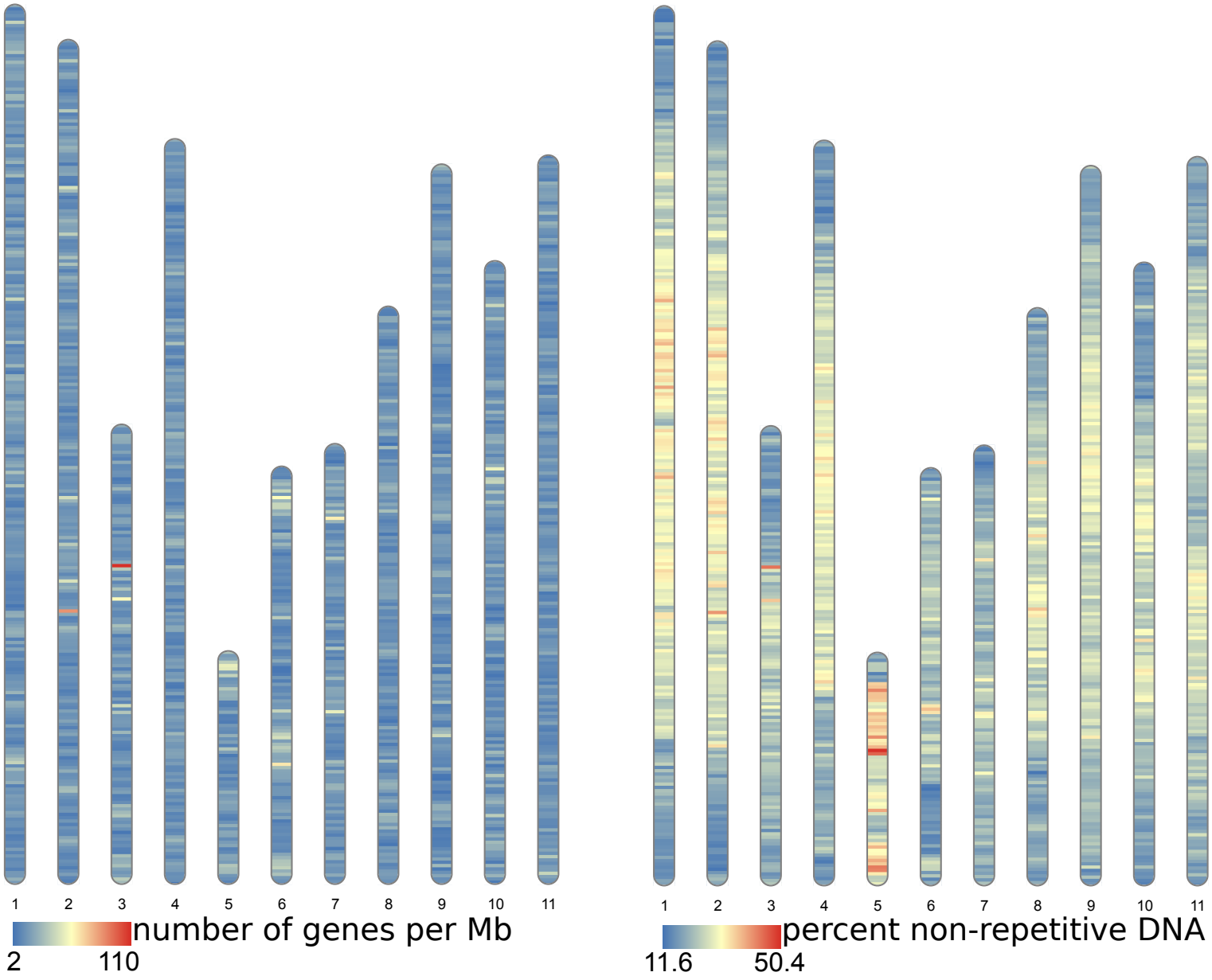
Zimmerman EC. 1994. Australian weevils (Coleoptera: Curculionidae). Volume 1. Orthoceri: Anthribidae to Attelabidae: the primitive weevils. East Melbourne: CSIRO. pp. 741.

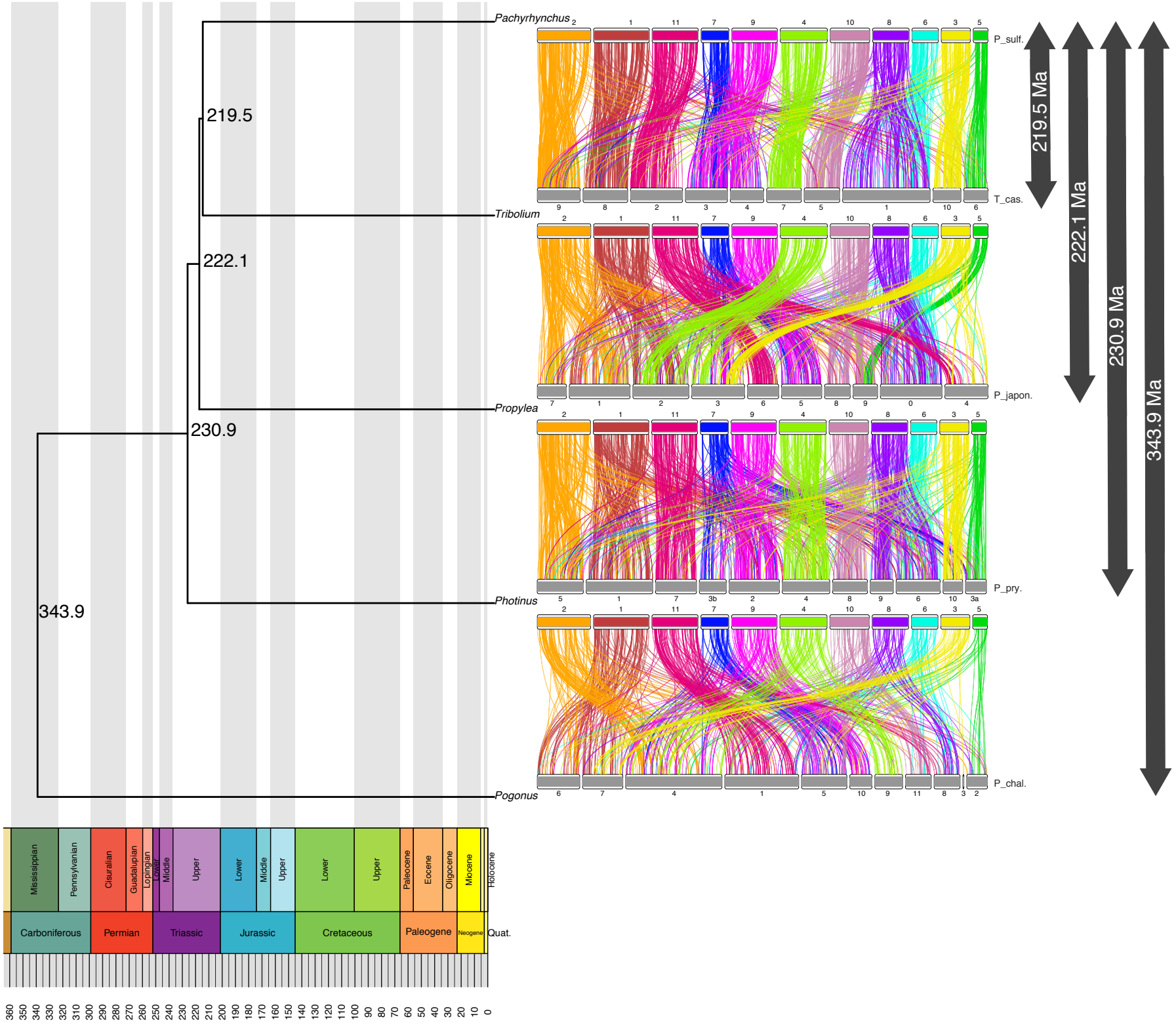

0 85
510
2127



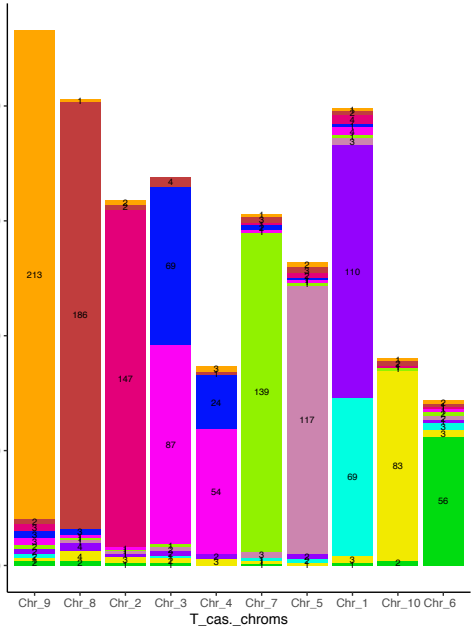


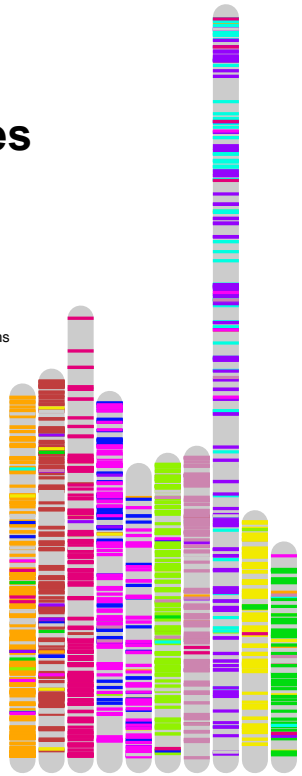




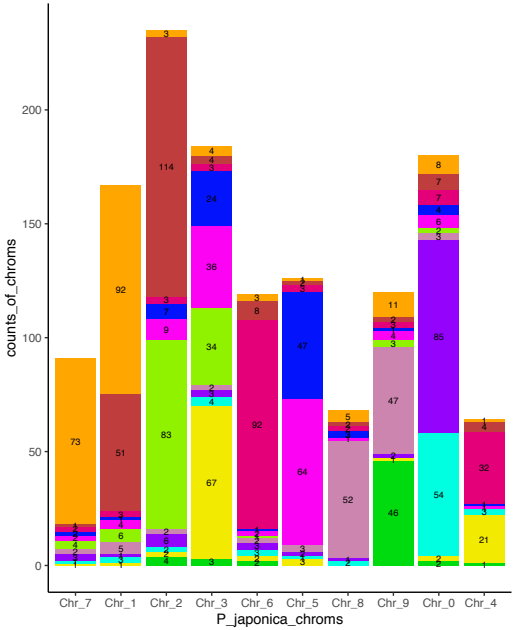


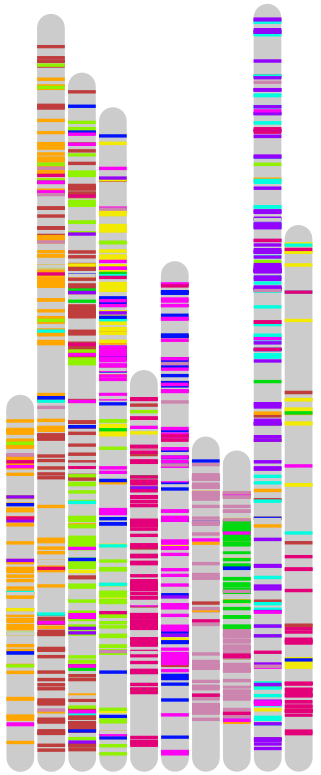
P. sulph. to Tribolium matches



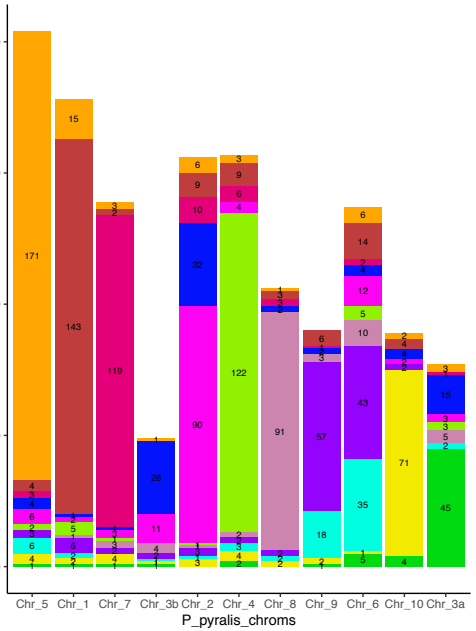


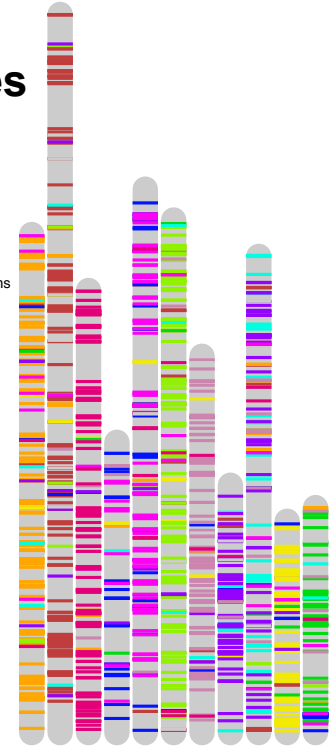
P. sulph. to Propylea matches



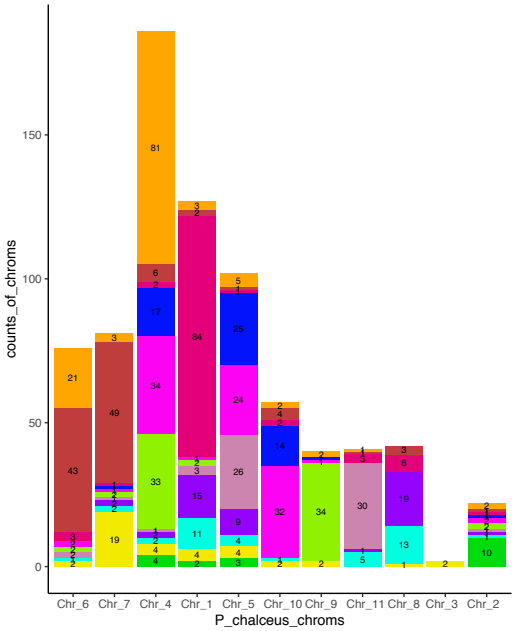


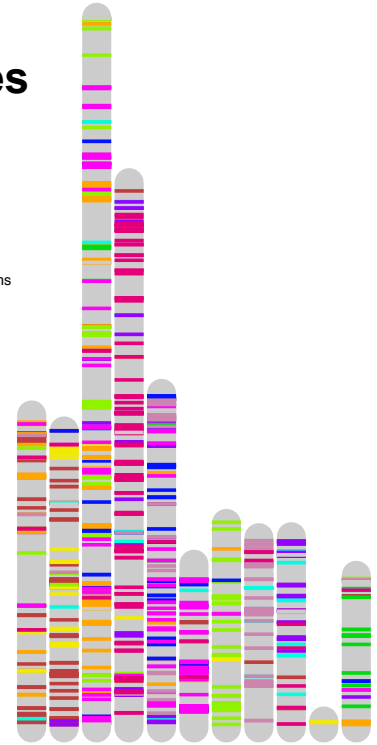
P. sulph. to Photinus matches

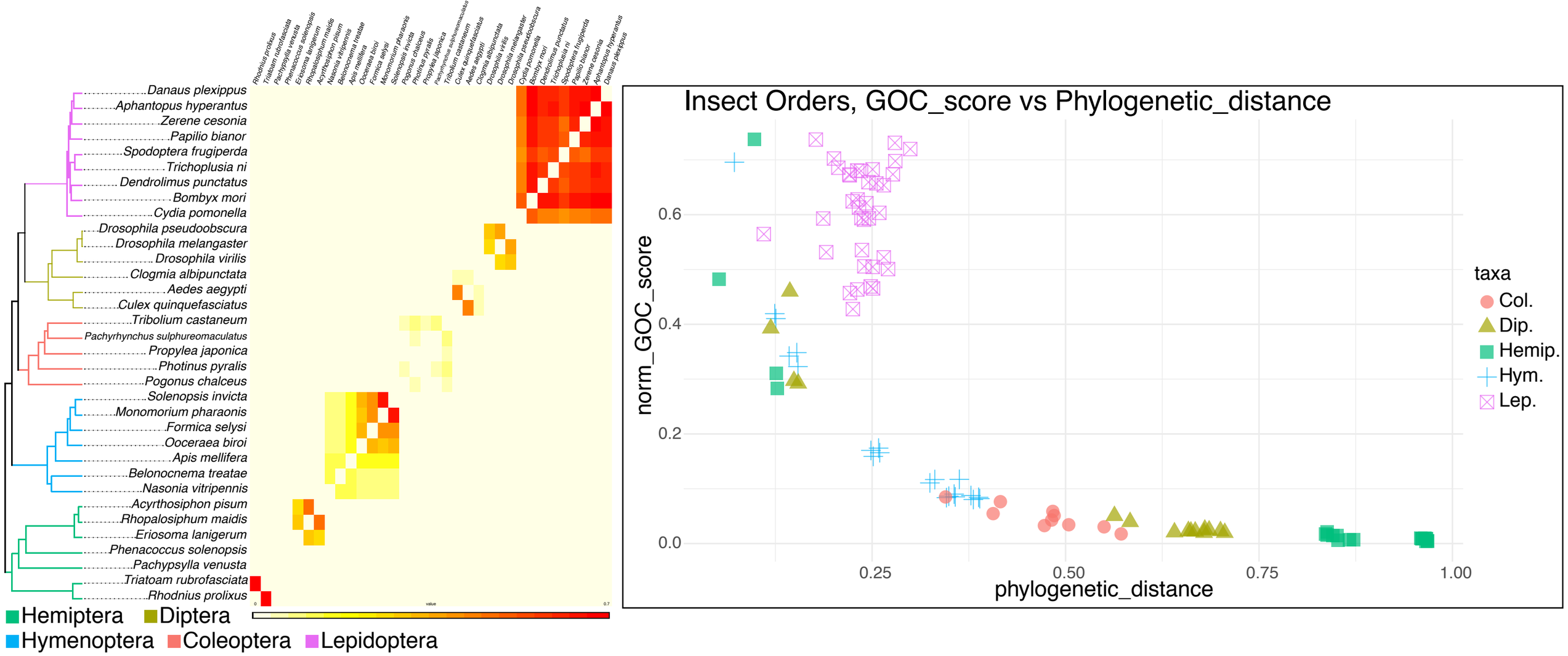




P. sulph. to Pogonus matches






Rhodnius prolixus
Triatoam rubrofasciata
Pachypsylla venusta
Phenacoccus solenopsis
Eriosoma lanigerum
Rhopalosiphum maidis
Acyrthosiphon pisum
Nasonia vitripennis
Belonocnema treatae
Apis mellifera
Ooceraea biroi
Formica selysi
Monomorium pharaonis
Solenopsis invicta
Pogonus chalceus
Photinus pyralis
Propylea japonica
Pachyrhynchus sulphureomaculatus
Tribolium castaneum
Culex quinquefasciatus
Aedes aegypti
Clogmia albipunctata
Drosophila virilis
Drosophila melangaster
Drosophila pseudoobscura
Cydia pomonella
Bombyx mori
Dendrolimus punctatus
Trichoplusia ni
Spodoptera frugiperda
Papilio bianor
Zerene cesonia
Aphantopus hyperantus
Danaus plexippus
Danaus plexippus
Aphantopus hyperantus
Zerene cesonia
Papilio bianor
Spodoptera frugiperda
Trichoplusia ni
Dendrolimus punctatus
Bombyx mori
Cydia pomonella
Drosophila pseudoobscura
Drosophila melangaster
Drosophila virilis
Clogmia albipunctata
Aedes aegypti
Culex quinquefasciatus
Tribolium castaneum
Pachyrhynchus sulphureomaculatus
Propylea japonica
Photinus pyralis
Pogonus chalceus
Solenopsis invicta
Monomorium pharaonis
Formica selysi
Ooceraea biroi
Apis mellifera
Belonocnema treatae
Nasonia vitripennis
Acyrthosiphon pisum
Rhopalosiphum maidis
Eriosoma lanigerum
Phenacoccus solenopsis
Pachypsylla venusta
Triatoam rubrofasciata
Rhodnius prolixus
0
value
0.7
Hemiptera
Diptera
Hymenoptera
Coleoptera
Lepidoptera
Insect Orders, GOC_score vs Phylogenetic_distance
norm_GOC_score
0.0
0.2
0.4
0.6
0.25
0.50
0.75
1.00
phylogenetic_distance
taxa
Col.
Dip.
Hemip.
Hym.
Lep.