Genome size evolution in the beetle genus *Diabrotica*.

Dimpal Lata[1], Brad S Coates[2], Kimberly K O Walden[3], Hugh M Robertson[3] and Nicholas J Miller[1*]

1. Department of Biological Sciences, Illinois Institute of Technology, Chicago, Illinois, USA.

2. USDA-ARS, Corn Insects & Crop Genetics Research Unit, Ames, Iowa, USA.

3. Department of Entomology, University of Illinois at Urbana-Champaign, Urbana, Illinois USA.

*Correspondence: Nicholas J Miller, 298 R A Pritzker Science Center 3101 S Dearborn St, Chicago, Illinois 60616, nmiller11@iit.edu, +1 312 567 3288.

Running head: Genome size evolution in *Diabrotica*.

Abstract

Diabrocite corn rootworms are one of the most economically significant pests of maize in the United States and Europe and an emerging model for insect-plant interactions. Genome sizes of several species in the genus *Diabrotica* were estimated using flow cytometry along with that of *Acalymma vittatum* as an outgroup. Genome sizes ranged between 1.56 - 1.64 gigabase pairs (Gb) and between 2.26- 2.59 Gb, respectively, for the *Diabrotica* subgroups fucata and virgifera; the *Acalymma vittatum* genome size was around 1.65 Gb. This result indicated that a substantial increase in genome size occurred in the ancestor of the virgifera group. Further analysis of fucata group and virgifera group genome sequencing reads indicated that the genome size difference between the *Diabrotica* subgroups could be attributed to a higher content of transposable elements, mostly miniature inverted-transposable elements (MITEs) and gypsy-like long terminal repeat (LTR) retroelements.

Introduction

The family Chrysomelidae is one of the largest families of phytophagous beetles (Order: Coleoptera), with nearly 40,000 species. A large number of species are important agricultural and forestry pests causing negative economic impacts (Reid 1995; Nie *et al.* 2020). A subtribe of Chrysomelidae, Diabroticina includes important agricultural pests from the genera *Acalymma, Cerotoma,* and *Diabrotica* (Toepfer *et al.* 2009). Many species of the genus *Acalymma* are specialists on Cucurbitaceae, with *Acalymma vittatum*, the striped cucumber beetle, being one of the key pests of cucurbits in the northeastern United States (Lewis *et al.* 1990). The bean leaf beetle, *Cerotoma*

51 *trifurcata,* is an important pest of leguminous crops such as peas and soybeans
52 throughout the eastern USA (Koch *et al.* 2004). *Diabrotica*, the most diverse genus
53 (Eben and Espinosa de los Monteros 2013), includes some of the most destructive
54 insect pests impacting US agriculture. *Diabrotica spp.* are divided into three groups:
55 signifera, fucata, and virgifera, with the latter two containing recognized pest species
56 (Krysan 1986). The species in the fucata group are multivoltine and polyphagous, while
57 species in the virgifera group are univoltine and oligophagous (Branson and Krysan
58 1981; Krysan 1982). *Diabrotica undecimpunctata* (southern corn rootworm) within the
59 fucata group is a generalist feeder that feeds on several crops, including cucurbits,
60 peanuts, and maize in the southern USA (Jackson *et al.* 2005). *Diabrotica virgifera*
61 *virgifera* (western corn rootworm), *Diabrotica barberi* (northern corn rootworm), and
62 *Diabrotica virgifera zeae* (Mexican corn rootworm) from the virgifera group are specialist
63 feeders and are pests of maize. *Diabrotica virgifera virgifera* is most abundant in the US
64 Corn Belt but is found throughout much of the United States as well as parts of Canada
65 and Mexico. *Diabrotica virgifera virgifera* and *Diabrotica barberi* are sympatric in the
66 northern part of the US Corn Belt, while *D. v. zeae* is sympatric with *D. v. virgifera* over
67 part of their range in Texas, Arizona, and Mexico (Bragard *et al.* 2019).
68 As the name “corn rootworm” suggests, several *Diabrotica* species cause
69 substantial economic damage to maize agriculture. The western corn rootworm,
70 *Diabrotica virgifera virgifera*, is considered one of the most destructive pests of maize
71 throughout the US Corn Belt in the United States, and it accounts every year for over $1
72 billion in yield losses and pest management costs (Sappington *et al.* 2006). The species
73 has also been introduced into Europe and has become widespread because of a

74 combination of transatlantic introductions and intra-continental movement (Ciosi *et al.*
75 2008, 2011; Miller *et al.* 2010). *Diabrotica barberi* is a serious maize pest but is less
76 widespread than *D. v. virgifera* (Capinera 2008). Maize-specialist corn rootworms have
77 proven to be highly adaptable to a variety of pest management tactics. Resistance has
78 evolved to a variety of synthetic insecticides (Ball and Weekman 1962; Meinke *et al.*
79 1998; Pereira *et al.* 2015; Souza *et al.* 2019, 2020) to most rootworm active transgenic
80 maize varieties (Gassmann *et al.* 2014; Calles-Torrez *et al.* 2019) and to cultural control
81 methods (Krysan *et al.* 1984; Gray *et al.* 2009)**.** Although *D.undecimpunctata* is more
82 widely distributed, it is unable to survive the winter temperatures of the US Corn Belt,
83 and it is considered an occasional pest of maize.

84 The genus *Diabrotica* is emerging as a model for insect-plant interactions in
85 generalist *versus* specialist herbivory. The ancestral state for the genus is thought to be
86 generalist feeding with a host plant range that includes Cucurbitaceae, Fabaceae, and
87 Poaceae, which is retained in the fucata group (Eben and Espinosa de los Monteros
88 2013). Following the split between the fucata and virgifera groups, around 30 million
89 years ago, the virgifera group specialized on Poaceae (Eben and Espinosa de los
90 Monteros 2013). Consequently, pest *Diabrotica* includes both generalist and specialist
91 species that share a common, experimentally-tractable host plant in maize.

92 The benzoxazinoid DIMBOA (2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one) is
93 one of the major secondary metabolites produced by maize plants (Sasai *et al.* 2009).
94 Specialist virgifera group species and generalist fucata group species respond distinctly
95 to this metabolite. *Diabrotica virgifera virgifera* larvae gained significantly more dry
96 weight when fed wild-type plants compared to larvae fed mutant plants, deficient for

97 DIMBOA biosynthesis. However, *D. undecimpunctata* performed equally well when fed
98 on both types of plants (Alouw and Miller 2015). The enhanced performance of
99 specialist *D. v. virgifera* may be related to its ability to use DIMBOA as a signal to locate
100 nutritious parts of roots, while the generalist from the fucata group does not (Robert *et*
101 *al.* 2012). Further, RNA-Seq studies showed transcripts encoding for a CYP9-like
102 cytochrome P450 monooxygenase were expressed in *D. v. virgifera* larvae feeding on
103 wild type plants but not in larvae feeding on benzoxazinoid-deficient mutant plants
104 (Miller and Zhao 2015), suggesting a cytochrome P450 mediated adaptation to
105 benzoxazinoids in *D. v. virgifera*.

106 Given the economic importance of and growing research interest in *Diabrotica*
107 beetles, there has been considerable interest in obtaining sequences of their genomes
108 and understanding genetic mechanisms for adaptations. Most *Diabrotica* genetics and
109 genomics research so far has been concentrated on *D. v. virgifera* (Gray *et al.* 2009;
110 Miller *et al.* 2010). The genome of *D. v. virgifera* is one of the larger genomes among
111 beetles and is estimated to be around 2.58 Gb (Coates *et al.* 2012), whereas the
112 average genome size for Coleoptera is 0.76Gb (Schoville *et al.* 2018; Gregory 2021).
113 Increased sizes of eukaryotic genomes are generally attributed to corresponding
114 increased numbers of repetitive DNA elements (Kidwell 2002), where a large proportion
115 of repeats are composed of different transposable element (TE) sequences (Kojima
116 2019). Eukaryotic transposons are divided into retroelements that propagate by an RNA
117 intermediate (class I) and DNA elements (class II) that mobilize by a “cut-and-paste”
118 mechanism (Finnegan 1989; Wicker *et al.* 2007).

There is evidence that the *D. v. virgifera* genome contains a high proportion of repetitive elements (Coates *et al.* 2012, 2014). The cadherin gene of *D. v. virgifera* is approximately 13.3 fold larger than the *Tribolium castaneum* ortholog due to much larger introns. The presence of numerous MITE-like elements within the cadherin gene of *D. v. virgifera* indicates that the difference in the gene size is due to the insertion of transposable elements in the *D. v. virgifera* introns (Coates *et al.* 2012). Class I BEL-like long terminal repeat (LTR) retrotransposons have been also found in the *D. v. virgifera* genome (Coates *et al.* 2014). Initially, MITEs were found as key components of plant genomes, as they are frequently associated with genes with high copy numbers indicating a possible role in gene expression and genome evolution (Santiago *et al.* 2002; Oki *et al.* 2008). They are also found in animals, including mosquitoes, *Drosophila*, fish, and humans (Deprá *et al.* 2012). Similar to MITES, LTR retrotransposons were also first discovered in plants. They are usually located largely in intergenic regions and are often the single largest component of plant genomes (Kumar and Bennetzen 1999; Feschotte *et al.* 2002). Previous studies have also revealed that both MITES and LTR retrotransposons can modify gene expression by inserting into promoter regions (Butelli *et al.* 2012; Wang *et al.* 2017; Liu *et al.* 2019). TE integration and excision can introduce novel variation (McClintock 1950; Wendel and Wessler 2000). Transposons can cause mutations by inserting themselves into functional regions and causing change by either modifying or eliminating gene expression (Feschotte 2008; Oliver and Greene 2009). They may also lead to genomic rearrangement (Maumus *et al.* 2015; Mat Razali *et al.* 2019).

141 Although the genome size of *D. v. virgifera* has been reported, no genome size
142 data have been obtained for the other species in the genus *Diabrotica* and related
143 genera. Since the genome size of *D. v. virgifera* is relatively large, we hypothesized that
144 there has been a recent expansion in genome size in the lineage leading to it. We
145 tested this hypothesis by estimating and comparing the genome sizes of *D. v. virgifera*
146 with those of several *Diabrotica* species and an outgroup species, *A. vittatum.* As a high
147 proportion of repetitive elements were found in the cadherin gene of *D. v. virgifera,* we
148 further hypothesized that genome size expansion in the lineage leading to *D.v. virgifera*
149 was due to a general increase in repetitive elements. To test this hypothesis, we looked
150 at the nature and quantity of repetitive elements in the virgifera group and compared it
151 with the fucata group.
152

153 **Materials and Methods**

154 **Sample Collection for Flow Cytometry**

155 Specimens of *D. v. virgifera* and *A. vittatum* were collected from a maize field in Illinois
156 in 2017, while *D. barberi* were collected from Wisconsin by Tracy Schilder, Wisconsin
157 Department of Agriculture. Specimens of *D. v. zeae* were collected from Texas by
158 Thomas Sappington, US Department of Agriculture Agricultural Research Service and
159 *D. balteata* were provided by Blair Siegfried and Heather McAuslane, University of
160 Florida, from a laboratory colony. *Diabrotica undecimpunctata* were obtained from Crop
161 Characteristics (Farmington, Minnesota, USA). Adult male *Periplaneta americana,*
162 which were used as an external reference (Guo *et al.* 2015; He *et al.* 2016) for flow
163 cytometric measurement, were obtained from Carolina Biological Supply (Burlington,

North Carolina, USA). All samples were flash-frozen in liquid nitrogen and preserved at -80°C.

Sample Preparation for Flow Cytometry

Genome size estimates were generated for eight individuals from five species of *Diabrotica* and one species of *Acalymma*. Preparations of nuclei were based on the method of Hare and Johnston (2012). The heads of single individuals were homogenized in 1ml of cold Galbraith buffer (4.26 g of $MgCl_2$, 8.84 g of sodium citrate, 4.2g of MOPS, 1ml of Triton-x, and 1mg of boiled ribonuclease A into 1 liter of ddH_20) placed in a 7ml Kontes Dounce. The homogenate was filtered through a 20 µm nylon mesh in a 1.5ml Eppendorf tube. Nuclei were stained with propidium iodide (PI) at 50 µg/ml in the dark at 4°C for an hour. In addition to the test sample, the brain tissue of *P. americana* was used as a standard (Hanrahan and Johnston 2011). The brain tissue of *P. americana* was dissected out, and the nuclear suspension was prepared and stained as described above.

Flow Cytometric Analysis

Stained nuclei were analyzed using an Attune NxT Flow Cytometer (Thermo Fisher Scientific, Waltham, Massachusetts). The propidium iodide-stained nuclei were excited by exposing them to the 488 nm blue laser. Red fluorescence from the propidium iodide was collected using the YL2 detector channel. The calibration of the flow cytometer was performed using a standard manufacturer's protocol before use. During each sample run, the linearity of the fluorescence measurement was confirmed by checking that the

187 mean channel number of the 4C nuclei (G2 phase) was double that of 2C nuclei (G1
188 phase). At least 1000 nuclear events were collected under each unknown and standard
189 2C peak. The nuclei peak (PI fluorescence histogram) and coefficient of variation (CV)
190 for each peak of interest (sample and standard) were obtained using the gating function
191 in the Attune Software. The coefficient of variation (CV) was less than 5 percent which
192 is considered appropriate for accurate genome size estimates (Dolezel *et al.* 2007;
193 Tomaszewska *et al.* 2021). The known genome size of the external standard (3.34Gb,
194 Harahan and Johnson 2011) and the relative fluorescence obtained from the sample
195 and external standard were then used to estimate the genome size using the following
196 formula:
197 Sample 1C DNA content = [(sample 2C mean peak position) / (standards 2C mean
198 peak position)] * standards 1C.
199 Genome size variations were analyzed using analysis of variance (ANOVA) and
200 Tukey's honest significant difference (HSD) post-hoc analyses using R statistical
201 software (version: 4.10) (R Core Team 2021). Letters were assigned showing
202 significance based on Tukey HSD post-hoc test using R statistical software (version:
203 4.10) (R Core Team 2021).
204
205 **Sampling and Genomic DNA Sequencing**
206 Separate sample collection and preparation were done to obtain the data from Illumina
207 whole-genome shotgun sequencing which were used to analyze the repetitive DNA
208 content of *D. barberi, D.undecimpunctata,* and *D. v. virgifera*. Adult *D. barberi* (*n* = 71)
209 and *D. undecimpunctata* (*n* = 50) were collected from maize fields near Ames, Iowa,

210 and Monmouth, Illinois, respectively. Each sample was pooled by species, flash-frozen
211 in liquid nitrogen and ground in liquid nitrogen, and then DNA extracted from ~3.0 mg of
212 tissue using the Qiagen DNeasy Blood and Tissue Extraction kit, with modifications as
213 described (Coates *et al.* 2014). Two micrograms of extracted DNA was submitted to the
214 Iowa State University DNA Facility (Ames, IA, USA) from which ~500bp insert indexed
215 sequencing libraries were generated using the Illumina TruSeq v2 Library Construction
216 Kit (Illumina, San Diego, CA). Single-end 100-bp reads were generated from *D. barberi*
217 and *D. undecimpunctata* libraries in separate lanes of an Illumina HiSeq2500. Raw
218 reads were submitted to the Biotechnology Information (NCBI) Short Read Archive
219 (SRA) under accessions SRR13363759 and SRR13364002 for *D. barberi* and *D.*
220 *undecimpunctata*, respectively.
221 *Diabrotica virgifera virgifera* adult females of inbred line Ped12, developed by the
222 USDA-ARS North Central Agricultural Research Laboratory were used for the genomic
223 DNA isolation. Briefly, whole beetles were homogenized in an SDS-based cell lysis
224 solution followed by overnight incubation with Proteinase K at 55°C. Cellular debris was
225 pelleted and RNA was digested with RNaseA. The homogenate was mixed with a high-
226 salt solution and incubated overnight at 4°C. The DNA in the supernatant was
227 precipitated overnight with ethanol at -20°C. DNA was quantified on an Invitrogen Qubit.
228 A paired-end short-insert genomic DNA library was prepared at the Roy J. Carver
229 Biotechnology Center at the University of Illinois at Urbana-Champaign using an
230 Illumina TruSeq DNAseq Sample Prep kit. Reads were sequenced to 100bp with the
231 Illumina TruSeq SBS sequencing kit version 3 on an Illumina HiSeq 2000 instrument

232 using Casava 1.8 for basecalling. Raw reads were submitted to the Biotechnology
233 Information (NCBI) Short Read Archive (SRA) under accession SRR6985755.
234 Sequencing generated 90 million single-end reads of 100bp for *D. barberi,* 118 million
235 single-end reads of 100bp for *D.undecimpunctata*, and 116 million 100-bp paired-end
236 reads for *D. v. virgifera*.
237
238 **Annotation and Quantification of Repeat Content from sequencing data.**
239 Raw reads were quality-filtered using fastp software (version 0.20.1) with a minimum 20
240 average Phred score. Reads mapping to mitochondrial genome sequences of
241 *Diabrotica* species available through the NCBI website (KF658070.1, KF669870.1) were
242 identified (minimap2 v2.17) and filtered out as implemented in the SSRG workflow
243 (Pombert, 2021). Repetitive elements in the genomes of *D.undecimpunctata, D. barberi,*
244 and *D. v. virgifera* were assembled and quantified using dnaPipeTE v1.3 (Goubert *et al.*
245 2015) and annotated using the DeepTE tool (Yan *et al.* 2020). To quantify the
246 proportion of TEs, dnaPipeTE uses samples of sequence reads instead of genome
247 assemblies, making this pipeline (dnaPipeTE) applicable for genomes with lower
248 sequencing depth. The pipeline performed assembly of repetitive reads into contigs
249 from low coverage sampling of raw reads using Trinity (Grabherr *et al.* 2011) and
250 annotated them using RepeatMasker (Smit *et al.* 1996-2010) with built-in Repbase
251 libraries (Bao *et al.* 2015, version 2017-01-27). Quantification was done by mapping a
252 random sample of reads onto the assembled repeats. The parameters set as the
253 benchmark for repeat content analysis for genomes greater than 500Mb (Goubert *et al.*
254 2015), including the coverage parameter, were used to run dnaPipeTE. The pipeline

255 was run for all three species using 0.1x coverage. Additionally, 0.1x coverage was
256 chosen based on the high N50 metric and plateauing point of transposable elements i.e,
257 increasing the coverage beyond 0.1x only marginally increased the proportion of
258 transposable elements for all three species. The dnaPipeTE pipeline does not annotate
259 novel repeats that do not match an entry in the included Repbase library. A high
260 proportion of repeats from each of the three beetle species were not annotated by
261 dnaPipeTE. DeepTE, a deep learning method based on convolutional neural networks,
262 was used to classify and annotate the unknown TEs. DeepTE uses eight trained models
263 to classify TEs into superfamilies and orders. All the TE contigs assembled by
264 dnaPipeTE were analyzed using DeepTE, whether or not they had been previously
265 classified by dnaPipeTE. Combining the results of the assembly and quantification by
266 dnaPipeTE with the classification results from DeepTE allowed the abundance of repeat
267 families in the genomes of all three species to be determined.

268 For comparison with the dnaPipeTE *de-novo* assembly of repetitive elements, the
269 percentage of repetitive elements in the *D. v. virgifera* genome assembly (NCBI RefSeq
270 accession GCF_003013835.1) was analyzed with RepeatModeler version 2.0.1 (Flynn
271 *et al.* 2019). Repeatmodeler is a *de-novo* transposable element identification package
272 that uses three repeat finding programs (RECON, RepeatScout, and
273 LtrHarvest/Ltr_retriever) to discover repetitive DNA sequences in the genome. These
274 repetitive DNA sequences were annotated by repeatClassifier based on the similarity to
275 RepBase and Dfam databases. The annotated library produced was used as input to
276 RepeatMasker to detect and mask repeats in the genome. Default parameters were
277 used to run RepeatModeler.

278

279 **Results**

280 **Genome size:**

281 The genome sizes of *D. v. zeae*, *D. v. virgifera,* and *D. barberi* from the virgifera group
282 were estimated to be 2.59Gb ± 0.01, 2.58 Gb ± 0.02, and 2.26Gb ± 0.04, respectively.
283 The genomes of *D. balteata* and *D.undecimpunctata* from the fucata group were
284 estimated at 1.64Gb ± 0.01 and 1.56 Gb ± 0.00. The outgroup species *A. vittatum*
285 genome size was estimated to be 1.65 Gb ± 0.01. An analysis of variance (ANOVA)
286 showed a significant difference ($F_{(5, 42)} = 597$, $p < 0.001$) in the genome sizes of the
287 species under study. A subsequent Tukey HSD test showed that there were no
288 significant differences in genome size between *D. v. virgifera* and *D. v. zeae*, between
289 *D. balteata* and *D.undecimpunctata,* or between *D. balteata* and *A. vittatum*. The
290 estimated genome size for each species with their phylogenetic relationships is shown
291 in Figure 1.

292

293 **Repeat Content analysis:**

294 The repeatomes of *D.v.virgifera, D. barberi,* and *D.undecimpunctata* comprised 72.4%,
295 70.3%, and 52.7% of their genomes, respectively. The repeat content obtained via
296 Repeat Modeler for the draft *D. v. virgifera* genome was 57.4%. As the assembly of the
297 *D.v. virgifera* genome was based on short-reads, TEs were expected to be under-
298 represented because of the difficulty of assembling individual copies. TE-rich large
299 genomes are difficult to assemble and often end up with high levels of fragmentation
300 around repetitive regions leading to underestimation of TE content (Green 2002).

301 To further investigate the classes of repeat families that contributed to the genome size
302 variation between the two groups of *Diabrotica*, the DeepTE annotations were coupled
303 with the dnaPipeTE abundance quantification to estimate the abundance of different
304 repeat elements in the genomes of *D. v. virgifera, D.barberi,* and *D.undecimpunctata*.
305 The TEs that accounted for most of the difference in the genome size of the two groups
306 were annotated as class II DNA Tc1-mariner Miniature-repeat Transposable Elements
307 (MITEs) (ClassII_DNA_TcMar_MITE) and class II DNA hAT MITE
308 (ClassII_DNA_hAT_MITE) of transposable elements and class I long terminal repeat
309 (LTR) Gypsy (ClassI_LTR_Gypsy) transposable elements and is shown in Figure 2. The
310 *D. v. virgifera* genome contained a large amount of ClassII_DNA_TcMar_MITE(0.61
311 Gb), ClassII_DNA_hAT_MITE (0.23 Gb), and ClassI_LTR_Gypsy(0.31 Gb) (Figure 2).
312 Similarly, *D. barberi* also had a high amount of ClassII_DNA_TcMar_MITE(0.50 Gb),
313 ClassII_DNA_hAT_MITE (0.19 Gb), and ClassI_LTR_Gypsy(0.25 Gb) (Figure 2). The
314 *D. undecimpunctata* had a lower amount of ClassII_DNA_TcMar_MITE(0.15 Gb),
315 ClassII_DNA_hAT_MITE (0.06 Gb), and ClassI_LTR_Gypsy(0.19 Gb) (Figure 2).
316 Transposable elements from other repeat families such as nLTRS, helitrons, and others
317 from class I and class II DNA elements were not as prominent as those mentioned
318 above. Single-low copy sequences representing the non-repetitive portion of the
319 genomes of *D. v. virgifera*, *D.barberi*, and *D.undecimpunctata* totaled 0.73 Gb, 0.68 Gb,
320 and 0.75 Gb, respectively (Figure 2).
321
322 **Discussion:**

323 *Diabrotica* is the most diverse genus within the subtribe Diabroticina and includes 354
324 species native to America. Only species from the fucata and virgifera groups occur in
325 the United States (Krysan 1986), while species from the signifera group are endemic to
326 South America. The signifera group species are not of economic importance, and their
327 biology is also mostly unknown (Clark *et al.* 2001) and so the group is understudied.
328 The economically significant pest species in this genus either belong to the virgifera
329 group or to the fucata group, justifying the need to study them comprehensively. The
330 expansion and adaptability of the pest species generates a sense of urgency to study
331 them.

332 Our results demonstrated that the genome size of virgifera group species are
333 approximately 1Gb larger than that of the fucata group and *Acalymma* species. The
334 genome size for *D. v. virgifera* obtained in this study and in a previous study (Coates *et*
335 *al.* 2012) are consistent. Our genome size results, when coupled with the species'
336 phylogenetic relationship, indicated that an expansion in genome size occurred in the
337 common ancestor of the virgifera group leading to *D. barberi, D. v. virgifera,* and *D. v.*
338 *zeae*. There was also a significant difference in the genome sizes of *D. barberi* and *D.*
339 *virgifera subsp.* suggesting a possible further expansion of the genome size in the
340 common ancestor of *D. v. virgifera* and *D. v. zeae*.

341 Repeat elements of species from fucata and virgifera were studied to understand
342 the basis of genome size expansion in *Diabrotica.* Differences in the total amount of
343 repetitive DNA accounted almost entirely for the differences in genome size between
344 the three species that were studied. The total amounts of low copy number DNA (which
345 includes most genes) were very similar, only differing by a few 10s of Mb. These results

346 strongly support the hypothesis of TE proliferation as the driver of genome expansion in
347 the virgifera group. Our data do not support a role for genome duplication, as this
348 mechanism would also produce substantial differences in the quantity of low copy
349 number DNA.
350 The genomes of the virgifera and fucata group of *Diabrotica* species contained many
351 common TEs, but the abundance of three TE families differed substantially and
352 accounted for approximately 74% of the difference in genome size between the two
353 groups and also 66% of the difference within the virgifera group between *D.barberi* and
354 *D. virgifera subsp*. The MITE-like Tc1-mariner and hAT elements and LTR Gypsy
355 retroelements were more abundant in the virgifera group (45% in *D. virgifera,* 42% in
356 *D.barberi*). Miniature-repeat Transposable Elements (MITEs) are short AT-rich (<0.5kb)
357 derivatives of DNA elements whose internal sequence lacks an open reading frame (Lu
358 *et al.* 2012), contain conserved terminal inverted repeats, are flanked by target site
359 duplications, and are closely associated with euchromatic genes (Kuang *et al.* 2009).
360 Gypsy elements are one of the most abundant classes of the long terminal repeat
361 (LTR)-retrotransposons superfamily with large numbers of copies found in almost all the
362 plants, animals, and fungi tested (Thomas-Bulle *et al.* 2018). Generally, class I
363 transposable elements have been reported to be in high abundance in insect genomes,
364 such as *Tribolium castaneum* (Wang *et al.* 2008), *Drosophila* (Clark *et al.* 2007) and
365 *Bombyx mori* (Osanai-Futahashi *et al.* 2008) in comparison to class II elements.
366 However, in *Diabrotica*, we discovered an abundance of class II elements and, to some
367 extent, class I elements.

368 Genome size varies enormously among eukaryote species (Hidalgo *et al.* 2017),
369 including beetles in the family Chrysomelidae (Petitpierre *et al.* 1993; Hanrahan and
370 Johnston 2011). Some have speculated that variation in genome size *per se* is related
371 to variation in phenotypic traits in insects. Correlations have been reported between
372 genome size and traits, including body size (Ferrari and Rai 1989; Finston *et al.* 1995;
373 Palmer and Petitpierre 1996; Palmer *et al.* 2003), development rate (Carreras *et al.*
374 1991; Gregory *et al.* 2003), and indeed, host plant range (Matsubayashi and Ohshima
375 2015; Calatayud *et al.* 2016; Zhang *et al.* 2019). However, in many of these examples,
376 different studies have produced contradictory results, with both positive and negative
377 correlations for the same trait. In the case of our data, there was no relationship
378 between host plant range and genome size. We found the cucurbit-specialist *Acalymma*
379 had a similar-sized genome to generalists in the fucata group of *Diabrotica*, whereas the
380 other specialist species in the study (virgifera group *Diabrotica*) had substantially larger
381 genomes. The data from our study support the view that correlations between genome
382 size and phenotypic traits are generally coincidental.
383 Our data showed that the increased genome size in the virgifera group of *Diabrotica*
384 species was the result of the proliferation of a few TE families. Comparative genomic
385 studies in insects have revealed that repeat elements can make large contributions to
386 genome size variation. Variation in TE abundance can be seen both within and among
387 species (Lynch 2007). Honeybees, with a genome size of 230Mb, show very few repeat
388 elements, representing a case of TE extinction (Weinstock *et al.* 2006). Similarly, the
389 small genome of *Belgica antarctica,* the Antarctic midge, (99 Mb) is also due to the
390 reduction of repeats in the genome (Kelley *et al.* 2014). There are also cases of TE

391 proliferations, such as in *Locusta migratoria,* that led to its large genome size of 6.5 Gb
392 (Wang *et al.* 2014). Cases of increase in genome size as a consequence of
393 transposable elements have also been reported in wood white (Leptidea) butterflies and
394 North American fireflies (Lampyridae) (Lower *et al.* 2017; Talla *et al.* 2017). Overall, it
395 appears that genome size in insects is fairly plastic and largely driven by the loss and
396 gain of transposable elements.

397 It is likely that the proliferation of TEs responsible for the increase in genome size
398 observed in our study occurred sometime between the divergence of the ancestors of
399 the virgifera and fucata groups 30 million years ago, but before the radiation of the
400 virgifera group species, around 17 million years ago (Eben and Espinosa de los
401 Monteros 2013). Host genomes have several mechanisms to suppress transposable
402 element expression and mobility (Bourque *et al.* 2018), including epigenetic silencing
403 through histone modifications or DNA methylation, targeted mutagenesis, small RNA
404 interference, as well as sequence-specific repressors such as the recently profiled
405 KRAB zinc-finger proteins (Fouché *et al.* 2020; Maupetit-Mehouas and Vaury 2020). At
406 the same time, some transposable elements have evolved regulatory sequences
407 controlling their own copy number to autonomously replicate in the genome (Lohe and
408 Hartl 1996; Saha *et al.* 2015). TE de-repression is triggered by environmental stimuli, in
409 particular stress (Bundo *et al.* 2014; Voronova *et al.* 2014; Fouché *et al.* 2020),
410 impacting transcription levels and increasing transpositional activity (Dubin *et al.* 2018).
411 Additionally, there are other factors influencing TE mobilization, for example,
412 demethylation and the removal of repressive histone marks during epigenetic
413 reprogramming stages (Russell and LaMarre 2018). The factors that led to de-

414 repression and subsequent proliferation of TEs in the ancestor of the virgifera group of
415 species are unknown.

416 Although gross variation in genome size, driven by gain and loss of TEs, does
417 not correlate with phenotypic adaptation, mutations associated with specific TE
418 insertions or excisions can be adaptive. Mostly, TE insertions are presumed to be
419 deleterious or neutral, but some have been shown to be selectively advantageous.
420 There are several studies showing that TE-mediated insertions have led to insecticide
421 resistance. In pink bollworm *Pectinophora gossypiella*, a major pest of cotton (Rostant
422 *et al.* 2012), several independent TE insertions in the *PgCad1* gene conferred
423 resistance to *Bt* Cry1Ac toxin (Fabrick *et al.* 2011; Wang *et al.* 2019). Cases of
424 resistance to *Bacillus thuringiensis* (Bt) toxins have also been reported in *Heliothis*
425 *virescens* which is caused by disruption of a cadherin-superfamily gene by TE insertion
426 (Gahan 2001). TE insertions in xenobiotic metabolism related genes such as those
427 encoding cytochrome P450 monooxygenases and glutathione S-transferases in
428 *Helicoverpa armigera* are the cause of resistance to insecticides (Klai *et al.* 2020).
429 Another example demonstrating that TEs can produce adaptive mutations has been
430 reported in *D. melanogaster* (Rostant *et al.* 2012; Gilbert *et al.* 2021). An increased
431 resistance to dichlorodiphenyltrichloroethane (DDT) in *D. melanogaster* has been
432 reported due to Cyp6g1 upregulation caused by insertion of the Accord transposon in
433 the 5′ regulatory region of the Cyp6g1 gene (Chung *et al.* 2007). Similarly, TE insertion,
434 which truncates the CHKov1 gene in *D. melanogaster*, confers resistance towards
435 organophosphate (Aminetzach 2005).

436 The degree to which mutations caused by the proliferation of TEs in the virgifera
437 group of *Diabrotica* contributed directly to the evolution of the group is unknown.
438 Tackling this question will require a comparative evolutionary genomic analysis of the
439 *Diabrotica* genus. To date, significant genome sequence data are only available for one
440 species, *Diabrotica virgifera virgifera*. However, the USDA-ARS Ag100Pest initiative
441 aims to sequence additional *Diabrotica* genomes. The information on the evolution of
442 size and repeat content of *Diabrotica* genomes presented in this paper will help to
443 inform the optimum strategies for sequencing additional genomes for the genus and,
444 potentially, Diabroticite genomes more generally.
445

446 **Data availability**

447 Nucleic acid sequencing data are available from the NCBI Sequence Read Archive
448 under the accession numbers provided above. All flow cytometry data are available for
449 download from Figshare.
450

451 **Acknowledgments:**

452 We would like to thank Dr. Thomas Sappington, Dr. Blair Siegfried, Dr. Heather
453 McAuslane, and Tracy Schilder for providing insect samples. We thank Dr. J. Spencer
454 Johnston and David Leclerc for their advice and guidance on flow cytometry methods
455 and Dr. Clément Goubert's advice on dnaPipeTE. Additionally, we also thank two
456 anonymous reviewers for their constructive feedback on an earlier version of the
457 manuscript.
458

Funding and Competing Interests:

This work was supported and funded by USDA NIFA award number 2011-67009-30134 and by the USDA Agricultural Research Service (ARS) (CRIS Project 5030-22000-019-00D). This article reports the results of research only, and mention of any commercial product or service does not constitute a recommendation for its use by USDA. The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or United States government determination of policy. USDA is an equal opportunity employer and provider.

References:

Alouw, J. C., and N. J. Miller, 2015 Effects of benzoxazinoids on specialist and generalist *Diabrotica* species. J. Appl. Entomol. 139: 424–431.

Aminetzach, Y. T., 2005 Pesticide resistance via transposition-mediated adaptive gene truncation in *Drosophila*. Science 309: 764–767.

Ball, H. J., and G. T. Weekman, 1962 Insecticide resistance in the adult western corn rootworm in Nebraska. J. Econ. Entomol. 55: 439–441.

Bao, W., K. K. Kojima, and O. Kohany, 2015 Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob. DNA 6: 11.

Bourque, G., K. H. Burns, M. Gehring, V. Gorbunova, A. Seluanov *et al.*, 2018 Ten things you should know about transposable elements. Genome Biol. 19: 199.

Bragard, C., K. Dehnen-Schmutz, F. Di Serio, P. Gonthier, M. Jacques *et al.*, 2019 Pest categorisation of *Diabrotica virgifera zeae* . EFSA J. 17:.

Branson, T. F., and J. L. Krysan, 1981 Feeding and oviposition behavior and life cycle

482 strategies of *Diabrotica* : an evolutionary view with implications for pest
483 management. Environ. Entomol. 10: 826–831.
484 Bundo, M., M. Toyoshima, Y. Okada, W. Akamatsu, J. Ueda *et al.*, 2014 Increased l1
485 retrotransposition in the neuronal genome in schizophrenia. Neuron 81: 306–313.
486 Butelli, E., C. Licciardello, Y. Zhang, J. Liu, S. Mackay *et al.*, 2012 Retrotransposons
487 control fruit-specific, cold-dependent accumulation of anthocyanins in blood
488 oranges. Plant Cell 24: 1242–1255.
489 Calatayud, P.-A., C. Petit, N. Burlet, S. Dupas, N. Glaser *et al.*, 2016 Is genome size of
490 Lepidoptera linked to host plant range? Entomol. Exp. Appl. 159: 354–361.
491 Calles-Torrez, V., J. J. Knodel, M. A. Boetel, B. W. French, B. W. Fuller *et al.*, 2019
492 Field-evolved resistance of northern and western corn rootworm (Coleoptera:
493 Chrysomelidae) populations to corn hybrids expressing single and pyramided
494 Cry3Bb1 and Cry34/35Ab1 Bt proteins in North Dakota. J. Econ. Entomol. 112:
495 1875–1886.
496 Capinera, J. L., 2008 Northern corn rootworm, *Diabrotica barberi* Smith & Lawrence
497 (Coleoptera: Chrysomelidae), pp. 2616–2619 in *Encyclopedia of Entomology*,
498 edited by J. L. Capinera. Springer Netherlands, Dordrecht.
499 Carreras, I., A.-F. A., C. Juan, and E. Petitpierre, 1991 Tasa de desarrollo en *Tribolium*
500 *brevicornis* (Coleoptera: Tenebrionidae) y su relación con el tamaño del genoma.
501 Chung, H., M. R. Bogwitz, C. McCart, A. Andrianopoulos, R. H. Ffrench-Constant *et al.*,
502 2007 *Cis*-regulatory elements in the *Accord* retrotransposon result in tissue-
503 specific expression of the *Drosophila melanogaster* insecticide resistance gene
504 *Cyp6g1*. Genetics 175: 1071–1077.

505 Ciosi, M., N. J. Miller, K. S. Kim, R. Giordano, A. Estoup *et al.*, 2008 Invasion of Europe
506 by the western corn rootworm, *Diabrotica virgifera virgifera*: multiple transatlantic
507 introductions with various reductions of genetic diversity. Mol. Ecol. 17: 3614--
508 3627.
509 Ciosi, M., N. J. Miller, S. Toepfer, A. Estoup, and T. Guillemaud, 2011 Stratified
510 dispersal and increasing genetic variation during the invasion of Central Europe
511 by the western corn rootworm, *Diabrotica virgifera virgifera*. Evol. Appl. 4: 54–70.
512 Clark, A. G., M. B. Eisen, D. R. Smith, C. M. Bergman, B. Oliver *et al.*, 2007 Evolution of
513 genes and genomes on the *Drosophila* phylogeny. Nature 450: 203–218.
514 Clark, T. L., L. J. Meinke, and J. E. Foster, 2001 Molecular phylogeny of *Diabrotica*
515 beetles (Coleoptera: Chrysomelidae) inferred from analysis of combined
516 mitochondrial and nuclear DNA sequences. Insect Mol. Biol. 10: 303–314.
517 Coates, B. S., A. P. Alves, H. Wang, K. K. O. Walden, B. W. French *et al.*, 2012
518 Distribution of genes and repetitive elements in the *Diabrotica virgifera virgifera*
519 genome estimated using BAC sequencing. J. Biomed. Biotechnol. 2012: 1–9.
520 Coates, B. S., L. M. Fraser, B. W. French, and T. W. Sappington, 2014 Proliferation and
521 copy number variation of BEL-like long terminal repeat retrotransposons within
522 the *Diabrotica virgifera virgifera* genome. Gene 534: 362–370.
523 Deprá, M., A. Ludwig, V. L. Valente, and E. L. Loreto, 2012 Mar, a MITE family of hAT
524 transposons in *Drosophila*. Mob. DNA 3: 13.
525 Dolezel, J., J. Greilhuber, and J. Suda, 2007 *Flow Cytometry with Plant Cells: Analysis*
526 *of Genes, Chromosomes and Genomes*.
527 Dubin, M. J., O. Mittelsten Scheid, and C. Becker, 2018 Transposons: a blessing curse.

528 Curr. Opin. Plant Biol. 42: 23–29.
529 Eben, A., and A. Espinosa de los Monteros, 2013 Tempo and mode of evolutionary
530 radiation in Diabroticina beetles (genera *Acalymma*, *Cerotoma*, and *Diabrotica*).
531 ZooKeys 207–321.
532 Fabrick, J. A., L. G. Mathew, B. E. Tabashnik, and X. Li, 2011 Insertion of an intact CR1
533 retrotransposon in a cadherin gene linked with Bt resistance in the pink bollworm,
534 *Pectinophora gossypiella*. Insect Mol. Biol. 20: 651–665.
535 Ferrari, J. A., and K. S. Rai, 1989 Phenotypic correlates of genome size variation in
536 *Aedes albopictus*. Evol. Int. J. Org. Evol. 43: 895–899.
537 Feschotte, C., 2008 Transposable elements and the evolution of regulatory networks.
538 Nat. Rev. Genet. 9: 397–405.
539 Feschotte, C., N. Jiang, and S. R. Wessler, 2002 Plant transposable elements: where
540 genetics meets genomics. Nat. Rev. Genet. 3: 329–341.
541 Finnegan, D. J., 1989 Eukaryotic transposable elements and genome evolution. Trends
542 Genet. TIG 5: 103–107.
543 Finston, T. L., P. D. N. Hebert, and R. B. Foottit, 1995 Genome size variation in aphids.
544 Insect Biochem. Mol. Biol. 25: 189–196.
545 Flynn, J. M., R. Hubley, C. Goubert, J. Rosen, A. G. Clark *et al.*, 2019 RepeatModeler2:
546 automated genomic discovery of transposable element families: Genomics
547 preprint.
548 Fouché, S., T. Badet, U. Oggenfuss, C. Plissonneau, C. S. Francisco *et al.*, 2020
549 Stress-Driven Transposable Element De-repression Dynamics and Virulence
550 Evolution in a Fungal Pathogen. Mol. Biol. Evol. 37: 221–239.

551 Gahan, L. J., 2001 Identification of a gene associated with Bt resistance in *Heliothis*
552 *virescens*. Science 293: 857–860.
553 Gassmann, A. J., J. L. Petzold-Maxwell, E. H. Clifton, M. W. Dunbar, A. M. Hoffmann *et*
554 *al.*, 2014 Field-evolved resistance by western corn rootworm to multiple *Bacillus*
555 *thuringiensis* toxins in transgenic maize. Proc. Natl. Acad. Sci. 111: 5141–5146.
556 Gilbert, C., J. Peccoud, and R. Cordaux, 2021 Transposable elements and the evolution
557 of insects. Annu. Rev. Entomol. 66: 355–372.
558 Goubert, C., L. Modolo, C. Vieira, C. ValienteMoro, P. Mavingui *et al.*, 2015 De novo
559 assembly and annotation of the Asian tiger mosquito (*Aedes albopictus*)
560 repeatome with dnaPipeTE from raw genomic reads and comparative analysis
561 with the yellow fever mosquito (*Aedes aegypti*). Genome Biol. Evol. 7: 1192–
562 1205.
563 Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson *et al.*, 2011 Full-
564 length transcriptome assembly from RNA-seq data without a reference genome.
565 Nat. Biotechnol. 29: 644–652.
566 Gray, M. E., T. W. Sappington, N. J. Miller, J. Moeser, and M. O. Bohn, 2009 Adaptation
567 and invasiveness of western corn rootworm: intensifying research on a
568 worsening pest. Annu. Rev. Entomol. 54: 303–321.
569 Green, P., 2002 Whole-genome disassembly. Proc. Natl. Acad. Sci. U. S. A. 99: 4143–
570 4144.
571 Gregory, T. R., 2021 Animal genome size database.
572 Gregory, T. R., O. Nedvěd, and S. Adamowicz, 2003 C-value estimates for 31 species
573 of ladybird beetles (Coleoptera: Coccinellidae). Hereditas 139: 121–7.

574 Guo, L. T., S. L. Wang, Q. J. Wu, X. G. Zhou, W. Xie *et al.*, 2015 Flow cytometry and K-
575 mer analysis estimates of the genome sizes of *Bemisia tabaci* B and Q
576 (Hemiptera: Aleyrodidae). Front. Physiol. 6:.
577 Hanrahan, S. J., and J. S. Johnston, 2011 New genome size estimates of 134 species
578 of arthropods. Chromosome Res. 19: 809–823.
579 Hare, E. E., and J. S. Johnston, 2012 Genome size determination using flow cytometry
580 of propidium iodide-stained nuclei, pp. 3–12 in *Molecular Methods for*
581 *Evolutionary Genetics*, edited by V. Orgogozo and M. V. Rockman. Methods in
582 Molecular Biology, Humana Press, Totowa, NJ.
583 He, K., K. Lin, G. Wang, and F. Li, 2016 Genome sizes of nine insect species
584 determined by flow cytometry and k-mer analysis. Front. Physiol. 7:.
585 Hidalgo, O., J. Pellicer, M. Christenhusz, H. Schneider, A. Leitch *et al.*, 2017 Is there an
586 upper limit to genome size? Trends Plant Sci. 22:.
587 Jackson, D. M., K. A. Sorensen, C. E. Sorenson, and R. N. Story, 2005 Monitoring
588 cucumber beetles in sweetpotato and cucurbits with kairomone-baited traps. J.
589 Econ. Entomol. 98: 12.
590 Kelley, J. L., J. T. Peyton, A.-S. Fiston-Lavier, N. M. Teets, M.-C. Yee *et al.*, 2014
591 Compact genome of the Antarctic midge is likely an adaptation to an extreme
592 environment. Nat. Commun. 5: 4611.
593 Kidwell, M. G., 2002 Transposable elements and the evolution of genome size in
594 eukaryotes. Genetica 115: 49–63.
595 Klai, K., B. Chénais, M. Zidi, S. Djebbi, A. Caruso *et al.*, 2020 Screening of *Helicoverpa*
596 *armigera* mobilome revealed transposable element insertions in insecticide

resistance genes. Insects 11: 879.

Koch, R. L., E. C. Burkness, and W. D. Hutchison, 2004 Confirmation of bean leaf beetle, *Cerotoma trifurcata*, feeding on cucurbits. J. Insect Sci. 4: 5.

Kojima, K. K., 2019 Structural and sequence diversity of eukaryotic transposable elements. Genes Genet. Syst. 94: 233–252.

Krysan, J. L., 1982 Diapause in the nearctic species of the *virgifera* group of *Diabrotica*: evidence for tropical origin and temperate adaptations. Ann. Entomol. Soc. Am. 75: 136–142.

Krysan, J. L., 1986 Introduction: biology, distribution, and identification of pest *Diabrotica*, pp. 1–23 in *Methods for the Study of Pest Diabrotica*, edited by J. L. Krysan and T. A. Miller. Springer Series in Experimental Entomology, Springer, New York, NY.

Krysan, J. L., J. J. Jackson, and A. C. Lew, 1984 Field termination of egg diapause in *Diabrotica* with new evidence of extended diapause in *D. barberi* (Coleoptera: Chrysomelidae). Environ. Entomol. 13: 1237–1240.

Kuang, H., C. Padmanabhan, F. Li, A. Kamei, P. B. Bhaskar *et al.*, 2009 Identification of miniature inverted-repeat transposable elements (MITEs) and biogenesis of their siRNAs in the Solanaceae: new functional implications for MITEs. Genome Res. 19: 42–56.

Kumar, A., and J. L. Bennetzen, 1999 Plant retrotransposons. Annu. Rev. Genet. 33: 479–532.

Lewis, P. A., R. L. Lampman, and R. L. Metcalf, 1990 Kairomonal attractants for *Acalymma vittatum* (Coleoptera: Chrysomelidae). Environ. Entomol. 19: 8–14.

620 Liu, Y., M. Tahir ul Qamar, J.-W. Feng, Y. Ding, S. Wang *et al.*, 2019 Comparative
621 analysis of miniature inverted–repeat transposable elements (MITEs) and long
622 terminal repeat (LTR) retrotransposons in six *Citrus* species. BMC Plant Biol. 19:
623 140.
624 Lohe, A. R., and D. L. Hartl, 1996 Autoregulation of mariner transposase activity by
625 overproduction and dominant-negative complementation. Mol. Biol. Evol. 13:
626 549–555.
627 Lower, S. S., J. S. Johnston, K. F. Stanger-Hall, C. E. Hjelmen, S. J. Hanrahan *et al.*,
628 2017 Genome size in North American fireflies: substantial variation likely driven
629 by neutral processes. Genome Biol. Evol. 9: 1499–1512.
630 Lu, C., J. Chen, Y. Zhang, Q. Hu, W. Su *et al.*, 2012 Miniature inverted-repeat
631 transposable elements (MITEs) have been accumulated through amplification
632 bursts and play important roles in gene expression and species diversity in *Oryza*
633 *sativa*. Mol. Biol. Evol. 29: 1005–1017.
634 Lynch, M., 2007 The origins of genome architecture. Sinauer 7.
635 Mat Razali, N., B. H. Cheah, and K. Nadarajah, 2019 Transposable elements adaptive
636 role in genome plasticity, pathogenicity and evolution in fungal phytopathogens.
637 Int. J. Mol. Sci. 20: 3597.
638 Matsubayashi, K. W., and I. Ohshima, 2015 Genome size increase in the phytophagous
639 ladybird beetle *Henosepilachna vigintioctomaculata* species complex
640 (Coleoptera: Coccinellidae): genome size increase in ladybird beetle. Entomol.
641 Sci. 18: 134–137.
642 Maumus, F., A.-S. Fiston-Lavier, and H. Quesneville, 2015 Impact of transposable

643 elements on insect genomes and biology. Curr. Opin. Insect Sci. 7: 30–36.
644 Maupetit-Mehouas, S., and C. Vaury, 2020 Transposon Reactivation in the Germline
645 May Be Useful for Both Transposons and Their Host Genomes. Cells 9: 1172.
646 McClintock, B., 1950 The origin and behavior of mutable loci in maize. Proc. Natl. Acad.
647 Sci. U. S. A. 36: 344–355.
648 Meinke, L. J., B. D. Siegfried, R. J. Wright, and L. D. Chandler, 1998 Adult susceptibility
649 of Nebraska western corn rootworm (Coleoptera: Chrysomelidae) populations to
650 selected insecticides. J. Econ. Entomol. 91: 594–600.
651 Miller, N. J., S. Richards, and T. W. Sappington, 2010 The prospects for sequencing the
652 western corn rootworm genome. J. Appl. Entomol. 134: 420–428.
653 Miller, N. J., and Z. Zhao, 2015 Transcriptional responses of *Diabrotica virgifera*
654 *virgifera* larvae to benzoxazinoids. J. Appl. Entomol. 139: 416–423.
655 Nie, R., C. Andújar, C. Gómez-Rodríguez, M. Bai, H.-J. Xue *et al.*, 2020 The phylogeny
656 of leaf beetles (Chrysomelidae) inferred from mitochondrial genomes. Syst.
657 Entomol. 45: 188–204.
658 Oki, N., K. Yano, Y. Okumoto, T. Tsukiyama, M. Teraishi *et al.*, 2008 A genome-wide
659 view of miniature inverted-repeat transposable elements (MITEs) in rice, *Oryza*
660 *sativa* ssp.*japonica*. Genes Genet. Syst. 83: 321–329.
661 Oliver, K. R., and W. K. Greene, 2009 Transposable elements: powerful facilitators of
662 evolution. BioEssays News Rev. Mol. Cell. Dev. Biol. 31: 703–714.
663 Osanai-Futahashi, M., Y. Suetsugu, K. Mita, and H. Fujiwara, 2008 Genome-wide
664 screening and characterization of transposable elements and their distribution
665 analysis in the silkworm, *Bombyx mori*. Insect Biochem. Mol. Biol. 38: 1046–

1057.

Palmer, M., and E. Petitpierre, 1996 Relationship of genome size to body size in *Phylan semicostatus* (Coleoptera: tenebrionidae). Ann. Entomol. Soc. Am. 89: 221–225.

Palmer, M., E. Petitpierre, and J. Pons, 2003 Test of the correlation between body size and DNA content in *Pimelia* (Coleoptera: Tenebrionidae) from the Canary islands. Eur. J. Entomol. 100: 123–129.

Pereira, A. E., H. Wang, S. N. Zukoff, L. J. Meinke, B. W. French *et al.*, 2015 Evidence of field-evolved resistance to bifenthrin in western corn rootworm (*Diabrotica virgifera virgifera* LeConte) populations in western Nebraska and Kansas. PLOS ONE 10: e0142299.

Petitpierre, E., C. Segarra, and C. Juan, 1993 Genome Size and Chromosomal Evolution in Leaf Beetles (Coleoptera, Chrysomelidae). Hereditas 119: 1–6.

Pombert, J.-F., 2021 *PombertLab/SSRG*. Pombert Lab. https://github.com/PombertLab/SSRG (Original work published 2016)

R Core Team, 2021 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reid, C. A. M., 1995 A cladistic analysis of subfamilial relationships I in the Chrysomelidae sensu lato (Chrysomeloidea). Pakaluk J Slipinski SA Eds Biol. Phylogeny Classif. Coleopt. Pap. Celebr. 80th Birthd. Roy CrowsonMuzeum Inst. Zool. PAN Warszawa 2: 559–631.

Robert, C. A., N. Veyrat, G. Glauser, G. Marti, G. R. Doyen *et al.*, 2012 A specialist root herbivore exploits defensive metabolites to locate nutritious tissues: a root herbivore exploits plant defences. Ecol. Lett. 15: 55–64.

689 Rostant, W. G., N. Wedell, and D. J. Hosken, 2012 Transposable elements and
690 insecticide resistance, pp. 169–201 in *Advances in Genetics*, Elsevier.
691 Russell, S. J., and J. LaMarre, 2018 Transposons and the PIWI pathway: genome
692 defense in gametes and embryos. Reproduction 156: R111–R124.
693 Saha, A., J. A. Mitchell, Y. Nishida, J. E. Hildreth, J. A. Ariberre *et al.*, 2015 A trans-
694 Dominant Form of Gag Restricts Ty1 Retrotransposition and Mediates Copy
695 Number Control. J. Virol. 89: 3922–3938.
696 Santiago, N., C. Herráiz, J. R. Goñi, X. Messeguer, and J. M. Casacuberta, 2002
697 Genome-wide analysis of the *Emigrant* family of MITEs of *Arabidopsis thaliana*.
698 Mol. Biol. Evol. 19: 2285–2293.
699 Sappington, T. W., B. D. Siegfried, and T. Guillemaud, 2006 Coordinated *Diabrotica*
700 genetics research: accelerating progress on an urgent insect pest problem. Am.
701 Entomol. 52: 90–97.
702 Sasai, H., M. Ishida, K. Murakami, N. Tadokoro, A. Ishihara *et al.*, 2009 Species-
703 Specific Glucosylation of DIMBOA in Larvae of the Rice Armyworm. Biosci.
704 Biotechnol. Biochem. 73: 1333–1338.
705 Schoville, S. D., Y. H. Chen, M. N. Andersson, J. B. Benoit, A. Bhandari *et al.*, 2018 A
706 model species for agricultural pest genomics: the genome of the Colorado potato
707 beetle, *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae). Sci. Rep. 8:
708 1931.
709 Smit, A. F. A., R. Hubley, and P. Green, 1996 Repeat-Masker open-3.0.
710 http://www.repeatmasker.org.
711 Souza, D., J. A. Peterson, R. J. Wright, and L. J. Meinke, 2020 Field efficacy of soil

712 insecticides on pyrethroid-resistant western corn rootworms (*Diabrotica virgifera*
713 *virgifera* LeConte). Pest Manag. Sci. 76: 827–833.
714 Souza, D., B. C. Vieira, B. K. Fritz, W. C. Hoffmann, J. A. Peterson *et al.*, 2019 Western
715 corn rootworm pyrethroid resistance confirmed by aerial application simulations
716 of commercial insecticides. Sci. Rep. 9: 6713.
717 Talla, V., A. Suh, F. Kalsoom, V. Dincă, R. Vila *et al.*, 2017 Rapid increase in genome
718 size as a consequence of transposable element hyperactivity in wood-white
719 (*Leptidea*) butterflies. Genome Biol. Evol. 9: 2491–2505.
720 Thomas-Bulle, C., M. Piednoël, T. Donnart, J. Filée, D. Jollivet *et al.*, 2018 Mollusc
721 genomes reveal variability in patterns of LTR-retrotransposons dynamics. BMC
722 Genomics 19: 821.
723 Toepfer, S., T. Haye, M. Erlandson, M. Goettel, J. G. Lundgren *et al.*, 2009 A review of
724 the natural enemies of beetles in the subtribe Diabroticina (Coleoptera:
725 Chrysomelidae): implications for sustainable pest management. Biocontrol Sci.
726 Technol. 19: 1–65.
727 Tomaszewska, P., T. K. Pellny, L. M. Hernández, R. A. C. Mitchell, V. Castiblanco *et al.*,
728 2021 Flow cytometry-based determination of ploidy from dried leaf specimens in
729 genomically complex collections of the tropical forage grass *Urochloa* s. l. Genes
730 12: 957
731 Voronova, A., V. Belevich, A. Jansons, and D. Rungis, 2014 Stress-induced
732 transcriptional activation of retrotransposon-like sequences in the Scots pine
733 (*Pinus sylvestris L.*) genome. Tree Genet. Genomes 10: 937–951.
734 Wang, X., X. Fang, P. Yang, X. Jiang, F. Jiang *et al.*, 2014 The locust genome provides

735 insight into swarm formation and long-distance flight. Nat. Commun. 5: 2957.
736 Wang, S., M. D. Lorenzen, R. W. Beeman, and S. J. Brown, 2008 Analysis of repetitive
737 DNA distribution patterns in the *Tribolium castaneum* genome. Genome Biol. 9:
738 R61.
739 Wang, L., J. Wang, Y. Ma, P. Wan, K. Liu *et al.*, 2019 Transposon insertion causes
740 cadherin mis-splicing and confers resistance to Bt cotton in pink bollworm from
741 China. Sci. Rep. 9: 7479.
742 Wang, X., Y. Xu, S. Zhang, L. Cao, Y. Huang *et al.*, 2017 Genomic analyses of
743 primitive, wild and cultivated citrus provide insights into asexual reproduction.
744 Nat. Genet. 49: 765–772.
745 Weinstock, G. M., G. E. Robinson, R. A. Gibbs, G. M. Weinstock, G. M. Weinstock *et*
746 *al.*, 2006 Insights into social insects from the genome of the honeybee *Apis*
747 *mellifera*. Nature 443: 931–949.
748 Wendel, J. F., and S. R. Wessler, 2000 Retrotransposon-mediated genome evolution on
749 a local ecological scale. Proc. Natl. Acad. Sci. 97: 6250–6252.
750 Wicker, T., F. Sabot, A. Hua-Van, J. L. Bennetzen, P. Capy *et al.*, 2007 A unified
751 classification system for eukaryotic transposable elements. Nat. Rev. Genet. 8:
752 973–982.
753 Yan, H., A. Bombarely, and S. Li, 2020 DeepTE: a computational method for de novo
754 classification of transposons with convolutional neural network. Bioinformatics
755 36: 4269–4275.
756 Zhang, S., S. Gu, X. Ni, and X. Li, 2019 Genome size reversely correlates with host
757 plant range in *Helicoverpa* species. Front. Physiol. 10: 29.

758

759 **Figures:**

760 Figure 1: Genome size evolution within the genus *Diabrotica*. Phylogeny of *Diabrotica*

761 and outgroup *Acalymma vittatum* is based on (Eben and Espinosa de los

762 Monteros 2013). Letters a - d indicate groups with no significant difference in

763 mean 1C-value (Tukey HSD; α = 0.05).

764

765 Figure 2: Predicted transposable elements (ln Gb) in the genome of three species of

766 *Diabrotica*. The boxed transposable elements are the top three highest

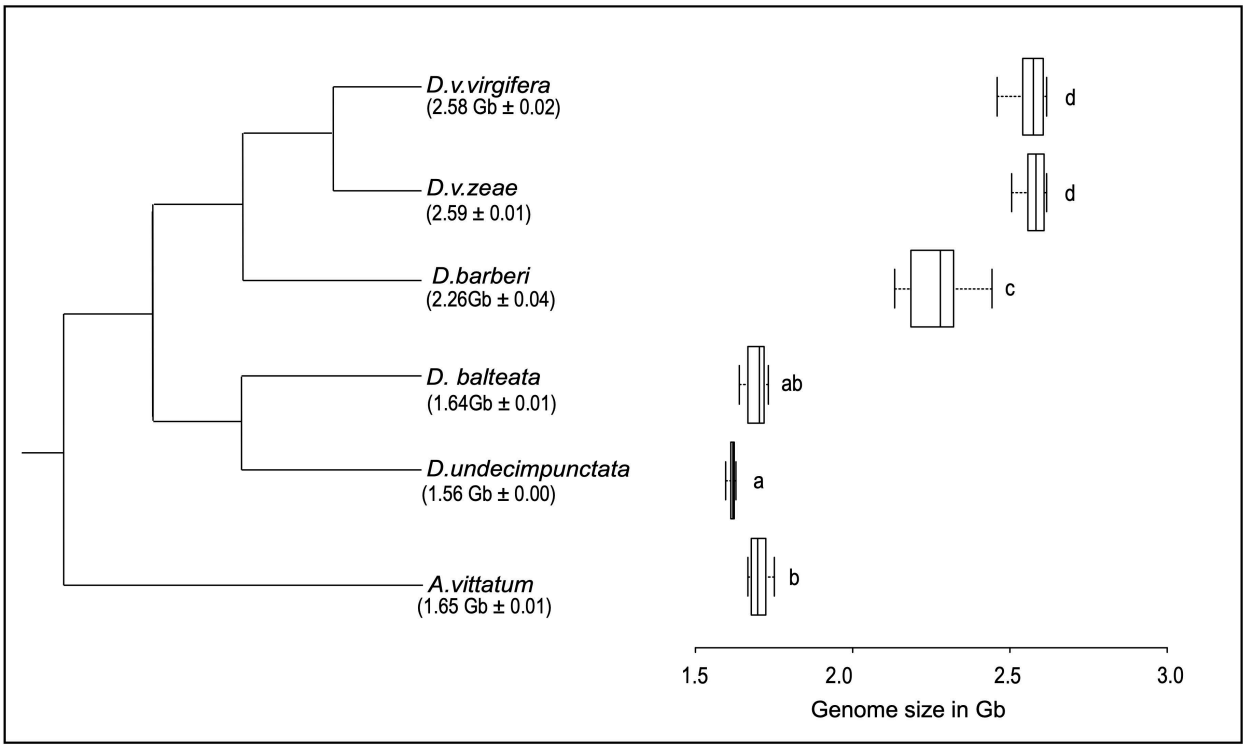
767 contributors to the genome size variation between the groups of fucata and

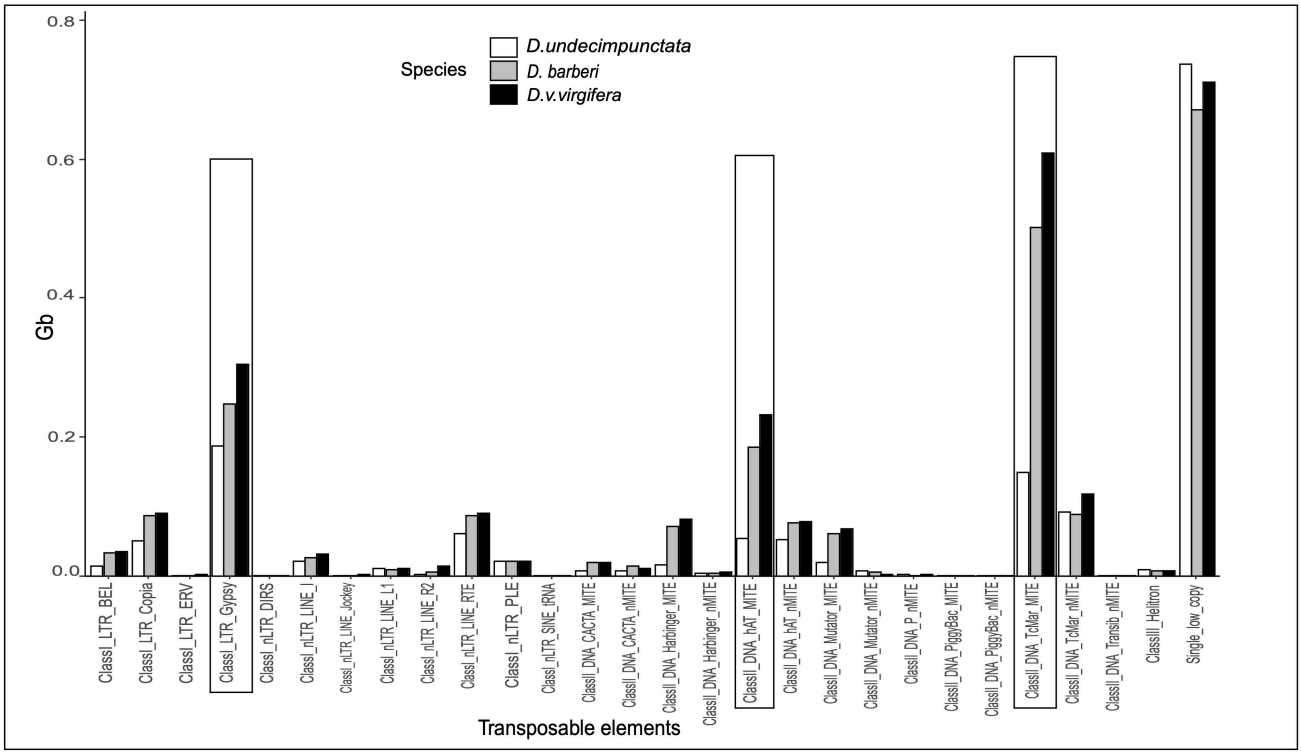
768 virgifera.

769

770

771


D.v.virgifera
(2.58 Gb ± 0.02)
D.v.zeae
(2.59 ± 0.01)
D.barberi
(2.26Gb ± 0.04)
D. balteata
(1.64Gb ± 0.01)
D.undecimpunctata
(1.56 Gb ± 0.00)
A.vittatum
(1.65 Gb ± 0.01)
d
d
c
ab
a
b
1.5
2.0
2.5
3.0
Genome size in Gb


Species
D.undecimpunctata
D. barberi
D.v.virgifera
Gb
0.0
0.2
0.4
0.6
0.8
ClassI_LTR_BEL
ClassI_LTR_Copia
ClassI_LTR_ERV
ClassI_LTR_Gypsy
ClassI_nLTR_DIRS
ClassI_nLTR_LINE_I
ClassI_nLTR_LINE_Jockey
ClassI_nLTR_LINE_L1
ClassI_nLTR_LINE_R2
ClassI_nLTR_LINE_RTE
ClassI_nLTR_PLE
ClassI_nLTR_SINE_tRNA
ClassII_DNA_CACTA_MITE
ClassII_DNA_CACTA_nMITE
ClassII_DNA_Harbinger_MITE
ClassII_DNA_Harbinger_nMITE
ClassII_DNA_hAT_MITE
ClassII_DNA_hAT_nMITE
ClassII_DNA_Mutator_MITE
ClassII_DNA_Mutator_nMITE
ClassII_DNA_P_nMITE
ClassII_DNA_PiggyBac_MITE
ClassII_DNA_PiggyBac_nMITE
ClassII_DNA_ToMar_MITE
ClassII_DNA_ToMar_nMITE
ClassII_DNA_Transib_nMITE
ClassIII_Helitron
Single_low_copy
Transposable elements