



HAL
open science

Unraveling genetic load dynamics during biological invasion: insights from two invasive insect species

Eric Lombaert, Aurelie Blin, Barbara Porro, Thomas Guillemaud, Julio S Bernal, Gary Chang, Natalia Kirichenko, Thomas W Sappington, Stefan Toepfer, Emeline Deleury

► To cite this version:

Eric Lombaert, Aurelie Blin, Barbara Porro, Thomas Guillemaud, Julio S Bernal, et al.. Unraveling genetic load dynamics during biological invasion: insights from two invasive insect species. 2024. hal-04694490

HAL Id: hal-04694490

<https://hal.science/hal-04694490v1>

Preprint submitted on 11 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

1 **Unraveling genetic load dynamics during biological invasion: insights** 2 **from two invasive insect species**

3
4 Eric Lombaert¹, Aurélie Blin¹, Barbara Porro^{1,2}, Thomas Guillemaud¹, Julio S. Bernal³, Gary
5 Chang⁴, Natalia Kirichenko^{5,6,7}, Thomas W. Sappington⁸, Stefan Toepfer^{9,10} and Emeline
6 Deleury¹

7
8 ¹ INRAE, Université Côte d'Azur, ISA, Sophia-Antipolis, France

9 ² Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-
10 Saclay, Evry, France

11 ³ Department of Entomology, Texas A&M University, College Station, TX 77843-2475, USA

12 ⁴ Biology Department, Gonzaga University, WA, USA

13 ⁵ Sukachev Institute of Forest, Siberian Branch of the Russian Academy of Sciences, Federal Research
14 Center “Krasnoyarsk Science Center SB RAS”, Akademgorodok 50/28, 660036 Krasnoyarsk, Russia

15 ⁶ Institute of Ecology and Geography, Siberian Federal University, Svobodny pr. 79, 660041
16 Krasnoyarsk, Russia

17 ⁷ All-Russian Plant Quarantine Center, Krasnoyarsk branch, Krasnoyarsk, Zhelyabova Str., 6/6, 660020,
18 Russia

19 ⁸ USDA-Agricultural Research Service, Corn Insects & Crop Genetics Research Unit, Genetics
20 Laboratory, Iowa State University, Ames, Iowa, USA

21 ⁹ CABI Switzerland, c/o Plant Protection and Soil Conservation Directorate, Hodmezovasarhely,
22 Hungary

23 ¹⁰ MARA-CABI Joint Laboratory for Biosafety, Institute of Plant Protection, Chinese Academy of
24 Agricultural Sciences, Beijing, China

25

26 **Corresponding author:**

27 Eric Lombaert – Institut Sophia Agrobiotech – 400, route des chappes – BP 167 – 06903 Sophia-
28 Antipolis Cedex - France

29 E-mail: eric.lombaert@inrae.fr

30 Tel: +33 4 92 38 64 81

31

32 **Keywords:** target enrichment, non-model organism, population genomics, pool-sequencing,
33 *Harmonia axyridis*, *Diabrotica virgifera virgifera*, genetic load

34

35 **Running title:** Evolution of genetic load in two invasive species

36 **Abstract**

37

38 Many invasive species undergo a significant reduction in genetic diversity, i.e. a genetic
39 bottleneck, in the early stages of invasion. However, this reduction does not necessarily
40 prevent them from achieving considerable ecological success and becoming highly efficient
41 colonizers. Here we tested the so-called purge hypothesis, which suggests that demographic
42 bottlenecks may indeed help to purge deleterious mutations, thereby reducing genetic load.
43 We used a transcriptome-based exome capture protocol to identify thousands of SNPs in
44 coding regions of native and invasive populations of two highly successful invasive insect
45 species, the western corn rootworm (Chrysomelidae: *Diabrotica virgifera virgifera*) and the
46 harlequin ladybird (Coccinellidae: *Harmonia axyridis*). We categorized and polarized SNPs to
47 investigate changes in genetic load between invasive populations and their sources. Our
48 results varied between species. In *D. virgifera virgifera*, although there was a general
49 reduction in genetic diversity in invasive populations, including that associated with genetic
50 load, we found no clear evidence for genetic load purging, except marginally for strongly
51 deleterious mutations in one European population. Conversely, in *H. axyridis*, the reduction in
52 genetic diversity was minimal, and we detected signs of genetic load fixation in invasive
53 populations. These findings provide new insights into the evolution of genetic load during
54 invasions, but do not offer a definitive answer to the purge hypothesis. Future research should
55 include larger genomic datasets and a broader range of invasive species to further elucidate
56 these dynamics.

57

58 Introduction

59

60 Biological invasions represent a significant aspect of global change, profoundly impacting
61 biodiversity through the alteration of species distributions worldwide, particularly in recent
62 times due to the significant increase in human-assisted dispersal (Seebens et al. 2017). The
63 key factors determining the success of invasive species remain largely hypothetical. While
64 numerous hypotheses have been proposed to explain invasion success (Sherpa and Després
65 2021; Daly et al. 2023), most fail to simultaneously account for several persistent
66 characteristics of invasions: (i) the rarity of successful invasions resulting from introductions
67 (Williamson and Fitter 1996), (ii) the lag time between initial introduction and invasion (Sakai
68 et al. 2001), (iii) the frequent reduction in genetic diversity due to demographic bottlenecks
69 (Nei et al. 1975) and (iv) the prevalence of multiple invasions originating from an initial
70 invasive population (i.e. the bridgehead effect; Lombaert et al. 2010). These characteristics
71 suggest that the success of invasions may stem from partially stochastic biological processes
72 spanning multiple generations and combining both demographic and genetic mechanisms.

73 An intriguing hypothesis that aligns with these characteristics is that the genetic load
74 is purged in the initial stages of invasions, i.e. deleterious alleles that cause inbreeding
75 depression are eliminated from the introduced population (Estoup et al. 2016). During the
76 demographic bottleneck induced by the introduction of a limited number of individuals into a
77 new environment, genetic drift intensifies and homozygosity increases. This situation often
78 results in a "mutational meltdown" (Lynch et al. 1995; Simberloff 2009), culminating in failure
79 of the introduced population to establish. Conversely, bottlenecks may also purge deleterious
80 alleles, thereby increasing the mean fitness of the introduced individuals. Theoretically, such
81 purges can occur under specific demographic (i.e., reduction of population size of
82 intermediate intensity) and genetic (i.e., strongly deleterious, recessive alleles) conditions,
83 optimizing exposure to natural selection (Crow 1970; Charlesworth et al. 1990; Glémin 2003;
84 Robinson et al. 2023). Such purging processes within introduced populations are particularly
85 interesting for their potential to facilitate successful invasions by sharply decreasing
86 inbreeding depression so that even inbred individuals may exhibit high fitness levels.

87 Purging has been demonstrated empirically by measuring the evolution of various
88 traits in artificially bottlenecked populations (Crnokrak and Barrett 2002; Avila et al. 2010).
89 However, in the context of naturally occurring bottlenecks during biological invasions,
90 measuring life history traits can be challenging, and evidence of purging has been documented
91 in only a few invasive species (Parisod et al. 2005; Mullarkey et al. 2013; Fountain et al. 2014;
92 Marchini et al. 2016). One notable formal test of this hypothesis was conducted on the
93 invasive Asian ladybird *Harmonia axyridis*, where measurement of life history traits revealed
94 that invasive populations did not experience the inbreeding depression suffered by native
95 ones, suggesting that deleterious alleles were purged during the invasion process (Facon et al.
96 2011). Overall, case studies focusing on the evolution of genetic load during biological
97 invasions have either examined a few life history traits (e.g. Facon et al. 2011) or used a single
98 locus approach (e.g. Zayed et al. 2007), making it difficult to generalize the results.

99 Advances in population genomics over the last decade have provided promising
100 avenues for investigating genetic load on large scales. Initially explored in humans (Lohmueller
101 et al. 2008; Henn et al. 2015), these approaches have been expanded into the fields of
102 domestication (e.g. Schubert et al. 2014; Marsden et al. 2016; Makino et al. 2018; Wang et al.
103 2021) and conservation (Xue et al. 2015; Robinson et al. 2016; Grossen et al. 2020; Dussex et
104 al. 2021; Ochoa and Gibbs 2021). These studies typically involve SNP-calling within coding

105 regions, determination of ancestral SNP states, categorization of the severity of fitness losses
106 due to derived alleles, and population comparisons while accounting for genetic drift using
107 synonymous and/or intergenic polymorphisms. A good knowledge of the genome is essential
108 for such studies, which may explain the limited application of these methods to invasive
109 species, which are mostly non-model organisms. However, this limitation is gradually
110 diminishing as progress in genome sequencing and computational techniques make these
111 methods more accessible and cost-effective (Bertorelle et al. 2022).

112 In this study, we investigated whether deleterious mutations were purged in two
113 successful insect invaders, namely the western corn rootworm, *Diabrotica virgifera virgifera*,
114 and the harlequin ladybird, *Harmonia axyridis*, by directly assessing genetic load using
115 genomic data from feral native and invasive populations. Importantly, the purpose of this
116 study did not include testing the instrumental role of purging in invasion success. We used a
117 pool-seq transcriptome-based exome capture protocol previously developed for non-model
118 species (Deleury et al. 2020) to identify SNPs within coding sequences and categorize them as
119 synonymous, moderately deleterious, or strongly deleterious. Our results offer insights into
120 the fate of the genetic load and provide valuable perspectives on the purge hypothesis on a
121 broader scale.

122

123 **Methods**

124

125 *Species choice*

126

127 The two species, *Diabrotica virgifera virgifera* (hereafter called DVV) and *Harmonia axyridis*
128 (HA), are good candidates for testing the genetic load purging hypothesis. Both species are
129 highly successful invaders with extensive invasive ranges (Gray et al. 2009; Roy et al. 2016).
130 Additionally, their invasion routes are well-documented (Ciosi et al. 2008; Lombaert et al.
131 2014, 2018) and they exhibit bridgehead effects, meaning that certain invasive populations
132 have played a crucial role in global dissemination (Miller et al. 2005; Lombaert et al. 2010).
133 Additionally, genetic load purging was detected in HA through a laboratory study focusing on
134 life-history traits (Facon et al. 2011).

135

136 *Design of exome capture probes for target enrichment*

137

138 Design of the exome capture probes was performed as described in Deleury et al. (2020). We
139 searched for peptide-coding sequences using FRAMEDP (v1.2.2; Gouzy et al. 2009) on the *de*
140 *novo* transcriptomes described in Coates et al. (2021) and Vogel et al. (2017) for DVV and HA,
141 respectively. BLASTX results (e-value $\leq 1e^{-7}$) of transcripts were used against the insect
142 proteomes of *Tribolium castaneum*, *Anoplophora glabripennis*, *Dendroctonus ponderosae*,
143 *Drosophila melanogaster* and the SwissProt database (v2016-02-17) for training. From the
144 obtained coding sequences (CDS), we eliminated (i) *Wolbachia* and other putative
145 endosymbiont sequences, (ii) CDS with > 1% missing nucleotide bases (Ns) or with more than
146 four consecutive Ns, (iii) CDS with a GC% below 25 or above 75 and (iv) CDS with lengths < 400
147 bp or > 3500 bp. From the remaining 7,132 and 12,739 CDS for DVV and HA respectively, we
148 drew at random c.a. 5.5 Mb for each species.

149

149 Probes based on the selected CDS were designed and manufactured by NimbleGen. In
150 the case of DVV, repetitiveness of the probes was assessed based on the highest 15-mer
151 frequency among the genomes of *Tribolium castaneum* (GCA_000002335.3), *Dendroctonus*

152 *ponderosae* (GCA_000355655.1 and GCA_000346045.2) and DVV (the one available at the
153 time, GCA_003013835.2). Probes with more than five close matches, defined as no more than
154 five single-base insertions, deletions or substitutions using the SSAHA algorithm (Ning et al.
155 2001), in the DVV genome were then discarded. Probes that matched sequences in the
156 mitochondrial genome were also discarded. Random nucleotides were used to replace
157 residual Ns in target sequences. For HA, the method used to ensure probe uniqueness is
158 described in Deleury *et al.* (2020).

159 Overall, the final probe sets corresponded to a total of 4,151 CDS (5,282,603 bases
160 across 12,017 regions of overlapping probes) and 5,717 CDS (5,347,461 bases across 6,400
161 regions of overlapping probes; Deleury et al. 2020) for DVV and HA respectively. This final set
162 of probes was manufactured in the form of biotinylated DNA oligomers. One capture reaction
163 contained 2,100,000 overlapping probes of 50 to 99 bp in length (mean length of 73.86 ± 4.46
164 bp for DVV, and 74.71 ± 4.92 bp for HA).

165

166 *Sample collection*

167

168 For both species, the choice of populations to be sampled was based on previously known
169 invasion routes (Figure S1; Miller et al. 2005; Lombaert et al. 2010, 2018), so that each invasive
170 population could be compared to its source. We sampled adult DVV at 6 sites: two in the
171 native area (Mexico), two in North America (Colorado and Pennsylvania, respectively
172 corresponding to the core and front of the first invasive population; Lombaert et al. 2018),
173 and two in Europe (Hungary, referred to the Central-Southeastern European population, and
174 north-western Italy, which are independently derived from the same source area of eastern
175 North America; Miller et al. 2005). Adult HA were sampled at four sites: two in the native area
176 (Russia (Siberia) and China), and two in North America (Pennsylvania and Washington,
177 corresponding to two independent outbreaks from the native area; Lombaert et al. 2010). Due
178 to mixed results regarding the origin of the east North American population of HA in a previous
179 study using 18 microsatellite markers (Lombaert et al. 2011), we chose to identify its source
180 through an ABC (approximate Bayesian computation; Beaumont et al. 2002) analysis using
181 synonymous SNPs from the current study (see Appendix S1 for details).

182 To determine the ancestral and derived alleles for each SNP (i.e. to polarize the alleles),
183 outgroup species were also sampled. For DVV, we selected the closely related species
184 *Diabrotica adelpha*, as well as the more phylogenetically distant *Ceratomyza trifurcata* (Eben
185 and de los Monteros 2013). In the case of HA, from phylogenetically closest to farthest
186 (Tomaszewska et al. 2021), the sampled outgroup species were *Harmonia yedoensis*,
187 *Harmonia conformis* and *Harmonia quadripunctata*. Complete information about samples is
188 provided in Table 1 and Table S1.

189

Species	Population code name	Status (source)	Sampling site (Lat.; long.)	Sampling date	Haploid sample size
<i>Diabrotica virgifera virgifera</i>					
	D-N-MX1	Native	Canatlán, Mexico (24.554; -104.741)	Oct. 2015	80
	D-N-MX2	Native	Durango, Mexico (23.999; -104.614)	Oct. 2015	80
	D-I-COL	Invasive (D-N-MX1)	Fort Morgan, CO, USA (40.218; -103.869)	Aug. 2015	80
	D-I-PEN	Invasive	Landisville, PA, USA	Aug. 2015	80

	(D-I-COL)	(40.119; -76.432)		
D-I-HUN	Invasive	Kondoros, Hungary	July 2015	80
	(D-I-PEN)	(46.736; 20.816)		
D-I-ITA	Invasive	Cuneo, Italy	July 2015	80
	(D-I-PEN)	(44.463; 7.571)		
<i>Harmonia axyridis</i>				
H-N-CHI	Native	Beijing, China	Oct. 2015	72
		(40.057; 116.540)		
H-N-RUS	Native	Krasnoyarsk, Russia	Sept. 2015	72
		(55.992; 92.757)		
H-I-PEN	Invasive	Landisville, PA, USA	Aug. 2015	72
	(H-N-CHI)	(40.119; -76.432)		
H-I-WAS	Invasive	Spokane, WA, USA	Nov. 2015	72
	(H-N-CHI)	(47.665; -117.403)		
<i>Diabrotica adelpha</i>				
D-OG-ADE	Outgroup DVV	Oaxaca State, Mexico	Jan. 2017	4
		(15.926; 97.151)		
<i>Cerotoma trifurcata</i>				
D-OG-TRI	Outgroup DVV	Weldon, CA, USA	Oct. 2016	4
		(35.660; -118.330)		
<i>Harmonia yedoensis</i>				
H-OG-YED	Outgroup HA	Beijing, China	Feb. 2015	2
		(40.031; 116.265)		
<i>Harmonia conformis</i>				
H-OG-CON	Outgroup HA	Mouans-Sartoux, France	Oct. 2009	2
		(43.620; 6.934)		
<i>Harmonia quadripunctata</i>				
H-OG-QUA	Outgroup HA	Rearing, Valbonne, France	July 2007	2
		(NA)		

190 **Table 1:** Description of *Diabrotica virgifera virgifera* (DVV) and *Harmonia axyridis* (HA) samples,
 191 including outgroup species (see Table S1 for BioSample accessions). Sources of invasive populations
 192 were determined based on known invasion routes (Figure S1; Miller et al. 2005; Lombaert et al. 2010,
 193 2018), except for H-I-PEN for which we retraced invasion routes in this study (see Appendix S1 for
 194 details). Because the two native DVV populations are genetically almost indistinguishable, we selected
 195 D-N-MX1 as the source of D-I-COL based on the smallest pairwise F_{ST} value between native and invasive
 196 populations (see Results section).

197

198 *DNA extraction, exome capture and pool sequencing*

199

200 For each population, 4 legs from each of 40 individuals (DVV) or 2 legs from each of 36
 201 individuals (HA) were pooled for DNA extraction with the Qiagen DNeasy Blood & Tissue kit,
 202 in accordance with the manufacturer's recommendations. For the outgroup species, the same
 203 kit was used to extract DNA from a single individual each for the three *Harmonia* species, and
 204 from a pool of two individuals each for the *Diabrotica* and *Cerotoma* species.

205 Genomic libraries were prepared using NimbleGen SeqCap EZ HyperCap Library v2.0
 206 and NimbleGen SeqCap EZ Library v5.0 for DVV and HA, respectively. In brief, for each of the
 207 population and outgroup samples, DNA (2 µg in 100 µl) was mechanically sheared to an
 208 average size of 200 bp using a Covaris S2 E210 device (6 cycles of 30 seconds each). In the case
 209 of DVV, the fragmented DNA was then divided into three technical replicates for each
 210 population. Subsequently, the fragments were subjected to end-repair, A-tailing and indexing

211 (with one unique index per sample) using the KAPA Library Preparation kit designed for
212 Illumina platforms. Following the ligation of Illumina adapters and indexes, only fragments
213 falling within the size range of 250 to 450 bp were retained. A PCR amplification step consisting
214 of 7 cycles was performed using standard Illumina paired-end primers, and the resulting
215 amplicons were purified using AMPure XP beads (Beckman). The length, quality, and
216 concentration of the prepared DNA fragments were assessed using a BioAnalyzer with Agilent
217 High Sensitivity DNA Assay, along with a Qubit.

218 For each capture (five for DVV and two for HA), we used a total of approximately 1 µg
219 of amplified DNA for exome enrichment, combining multiple samples in proportions that
220 allowed for equimolarity between population samples (see Table S2 and Table S3). This
221 enrichment was performed using the capture probes described above, following the
222 guidelines of either the NimbleGen SeqCap EZ HyperCap Library Protocol v2.0 or the SeqCap
223 EZ Library Protocol v5.0. Following each capture, we conducted two parallel PCRs, each
224 comprising 14 cycles, on the elution solution. In the case of HA, the resulting PCR products
225 were combined. Subsequently, all PCR products underwent purification using AMPure XP
226 beads. The length, quality, and concentration of the final DNA fragments were assessed using
227 a BioAnalyzer equipped with the Agilent High Sensitivity DNA Assay and a Qubit fluorometer.

228 For sequencing, we used one lane of an Illumina HiSeq3000 sequencer per species,
229 following the manufacturer's instructions, in paired-end mode for 150 cycles. After
230 sequencing, the data were demultiplexed and exported as FastQ files, and the libraries were
231 processed independently. For DVV, the FastQ files of technical replicates were merged prior
232 to subsequent analysis.

233

234 *Mapping, SNP calling and annotation*

235

236 Sequence quality assessment was conducted using FastQC v0.11.5 (Andrews 2010).
237 Subsequently, adapter sequences were removed, and low-quality base pairs were eliminated
238 using *Trimmomatic* v0.35 (Bolger et al. 2014), with the following parameter settings:
239 ILLUMINACLIP:TruSeq-file.fa:2:30:10 LEADING:25 TRAILING:25 SLIDINGWINDOW:5:30
240 MINLEN:75. We used almost the same parameters for the outgroup species sequences, but
241 with less stringency on quality of reads (SLIDINGWINDOW:5:20).

242 Filtered reads were mapped onto the genome assemblies PGI_DIABVI_V3a (GenBank
243 identifier: GCA_917563875.2) and icHarAxyr1.1 (GenBank identifier: GCA_914767665.1), for
244 DVV and HA respectively, using default options of the *bwa-mem* aligner v0.7.15 (Li 2013). We
245 used the *SAMtools* v1.15.1 software package (Li et al. 2009) and its *fixmate*, *view*, and
246 *markdup* tools to perform the following operations sequentially: (1) removal of unmapped
247 reads and secondary alignments, (2) filtering out read alignments with a mapping quality
248 Phred-score <20 and improperly paired reads, and (3) identification and removal of PCR
249 duplicates. Processing filtered reads for the outgroup species followed the same steps,
250 including mapping to the genome of the corresponding focal species, except that PCR
251 duplicates were not removed.

252 For each species, we conducted variant calling on the resulting *bam* alignment files
253 using Freebayes v1.3.6 software (Garrison and Marth 2012), with a focus on exonic regions
254 (options `-t exons.bed --pooled-continuous --min-alternate-count 1 --min-alternate-fraction`
255 `0.001 --min-alternate-total 1 --min-coverage 80 --use-best-n-alleles 3`). The resulting *vcf* file
256 was subsequently filtered using the *view* tool within *bcftools* v1.13 software (Danecek et al.

257 2021) and an in-house script to retain only true bi-allelic SNPs. Annotation was then
258 performed with the *SnpEff* program v5.0 (Cingolani et al. 2012b).

259

260 *Allele polarization*

261

262 SNP positions were extracted from both *vcf* files using the *bcftools* query tool. Subsequently,
263 for each outgroup species, we identified the nucleotides (or absence of typing) at positions
264 matching those in their respective focal species using the *samtools mpileup* tool on the
265 previously generated *bam* alignment files.

266 To polarize the SNPs, we used *est-sfs* software v2.03 (Keightley and Jackson 2018) with
267 the Kimura 2-parameter model. The input files contained data from one native population of
268 the focal species, together with the corresponding outgroup species, ensuring a consistent
269 phylogenetic topology for the software. Given that *est-sfs* computes the probability of the
270 most frequent allele being ancestral solely for polymorphic SNPs within the native population,
271 we used probabilities from a parallel analysis for monomorphic loci, incorporating an extra
272 haplotype for each allele. The complete process was repeated twice for each focal species,
273 considering the availability of two native populations for each. This resulted in two *est-sfs*
274 probabilities per focal species, both reflecting the likelihood of the most frequent allele being
275 the ancestral one.

276 To consider a SNP as polarized, we applied the following rules. If both probabilities
277 were above 0.5 for the same allele, and at least one of the probabilities exceeded 0.75, we
278 considered the most frequent allele as the ancestral one. If both probabilities were below 0.5
279 for the same allele, with at least one below 0.25, and if there were no more than two distinct
280 nucleotides present in total across the focal and outgroup species combined, we considered
281 the less frequent allele to be the ancestral one. SNPs that did not meet these criteria were
282 considered as not polarized but were retained for computations that did not require
283 polarization.

284

285 *Additional filters and categorization of deleterious mutations*

286

287 After extracting essential information from the *vcf* file using the *SnpSift* program v5.0
288 (Cingolani et al. 2012a), we applied additional filters through an in-house R script (R Core Team
289 2021). First, we selected only biallelic SNPs with coverage greater than 50 reads, falling below
290 the 95th percentile of overall coverage in each pool, and exhibiting a minor allele frequency
291 exceeding 0.01 in at least one population. Second, we retained SNPs that were located in
292 coding regions and had unambiguous annotations. Finally, we excluded SNPs located on the
293 X-chromosome in the specific case of HA, for which this information was available.

294 *SnpEff* annotations were used to categorize the severity of fitness loss due to
295 mutations. Synonymous variants were used as proxies for neutral polymorphism. Non-
296 synonymous mutations were categorized as either “missense” (considered as potentially
297 moderately deleterious, involving amino acid changes) or “LoF” (loss-of-function, considered
298 as potentially highly deleterious, involving gain or loss of stop codons).

299

300 *Genetic diversity and genetic load analyses*

301

302 We computed several descriptive statistics on the whole set of SNPs. At the inter-population
303 level, we assessed genetic differentiation by calculating pairwise F_{ST} values using the R package

304 *poolfstat* v2.1.1 (Gautier et al. 2022). At the intra-population level, we used in-house R scripts
305 to compute the synonymous expected heterozygosity H_{eS} , as a measure of diversity, and the
306 ratio of non-synonymous to synonymous expected heterozygosity H_{eN}/H_{eS} , as a measure of
307 the efficacy of selection. We used a 100-block jackknife resampling approach to estimate
308 means and standard errors.

309 Using polarized SNPs, we reported for each population and each severity category the
310 proportion of positions for which the derived allele is absent (frequency $f = 0$), rare ($f > 0$ and
311 $f \leq 0.1$), common ($f > 0.1$ and $f < 1$), or fixed ($f = 1$). The proportion of rare derived alleles
312 serves as a proxy for the masked load (i.e. the load which may become express in future
313 generations; Bertorelle et al. 2022), whereas common and fixed alleles serve as a proxy for
314 the realized load (i.e. the load which reduces the fitness in the current generation; Bertorelle
315 et al. 2022). Additionally, within each severity category and population, we calculated the
316 mean derived allele frequency and estimated standard errors via a 100-block jackknife
317 resampling approach. Within each species and severity category, mean derived allele
318 frequencies were compared among populations using z-tests with false discovery rate (FDR;
319 Benjamini and Hochberg 1995) correction for multiple comparisons.

320 Finally, polarized SNPs were used to compute the R_{XY} statistic, which allows assessment
321 of relative excess or deficit of derived alleles within specific categories of deleterious
322 mutations in one population compared to another (Appendix S2; Do et al. 2015; Xue et al.
323 2015). This statistic effectively corrects for variation due to demography using neutral
324 mutations (synonymous SNPs in our case) as a reference point. In our analyses, population X
325 is an invasive population, and population Y is its source population (provided in Table 1). An
326 R_{XY} value of 1 indicates an equivalent relative genetic load in both populations, whereas values
327 greater or less than 1 suggest an excess (including fixation) or a deficit (including purge) of the
328 genetic load in the invasive population X , respectively. We assessed the significance of any
329 deviation from 1 using a z-score two-tailed test based on the 100-block jackknife resampling
330 method.

331

332 Results

333

334 *SNP-calling and genomic variation*

335

336 In the target enrichment experiments, we obtained mean numbers of raw read sequences of
337 50,196,834 and 63,360,676 for DVV and HA, respectively (BioProject PRJNA1079689; Table
338 S1). After trimming, 85.0% and 93.8% of the sequences were retained for DVV and HA
339 respectively. For outgroup species, the total number of raw reads was highly variable, with a
340 mean of 2,631,775 of which 97.8% were conserved after trimming (see detailed information
341 for all libraries, including outgroup species, in Table S4). Mapping, calling and filtering
342 identified 66,274 SNPs (of which 62,034 could be polarized) and 169,755 SNPs (of which
343 169,102 could be polarized) within coding sequences for DVV and HA respectively (see Table
344 S5 for details).

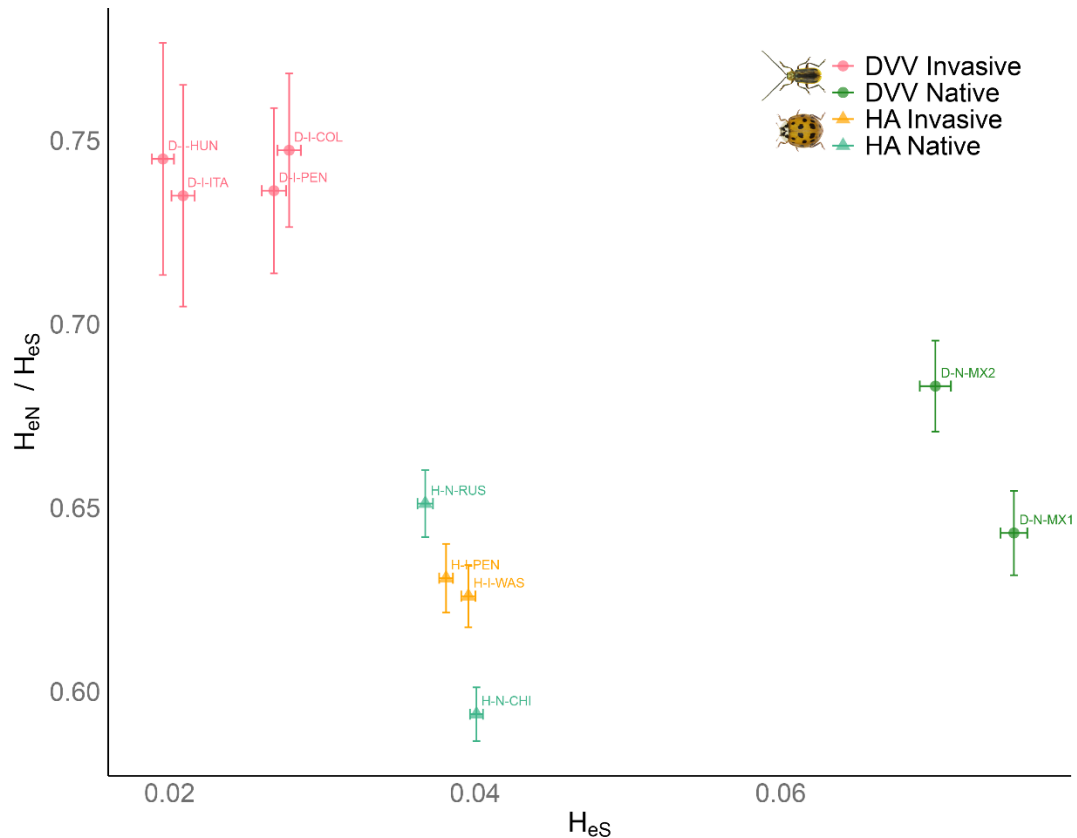
345 In the case of DVV, pairwise F_{ST} computed on the full set of SNPs showed patterns very
346 close to those expected from results of previous analyses of microsatellite datasets (Figure
347 S2). The differentiation between the two native Mexican populations was almost negligible
348 with a mean F_{ST} estimate below 0.001. Given that the F_{ST} estimates between the invasive
349 populations and the native sample D-N-MX1 were consistently lower (mean F_{ST} of 0.121) than
350 those with the sample D-N-MX2 (mean F_{ST} of 0.128), D-N-MX1 was therefore considered the

351 source of the D-I-COL sample (see Table 1). The Pennsylvania and Colorado populations had a
352 low pairwise F_{ST} of 0.012, consistent with an Eastern expansion from Colorado (Figure S1),
353 with plausible continuous gene flow. Conversely, all other F_{ST} estimates were relatively high,
354 particularly between the native and European populations (mean F_{ST} value = 0.152). The
355 highest value was observed between the two European populations at 0.169, confirming
356 previous results indicating that these populations originated from two independent
357 introductions (Miller et al. 2005).

358 In the case of HA, F_{ST} estimates were moderate and fairly homogeneous, ranging from
359 0.020 to 0.050 (Figure S2). The smallest values were found between the eastern native
360 population H-N-CHI and both invasive populations, while the F_{ST} between invasive populations
361 was slightly larger, consistent with previous results indicating independent origins for the two
362 North American outbreaks (Lombaert et al. 2010). The highest values were observed between
363 the western native population H-N-RUS and all other populations.

364 Within each species, invasive populations were characterized by reduced synonymous
365 heterozygosity compared to their native counterparts, reflecting demographic bottlenecks
366 encountered during the invasion process (Figure 1, x axis). One exception was the western
367 native population of HA (H-N-RUS), which stood out for having the lowest diversity within this
368 species. The ratio of non-synonymous to synonymous expected heterozygosity H_{eN}/H_{eS} is
369 typically expected to increase in populations experiencing pronounced drift, as the reduced
370 efficacy of selection allows more non-synonymous mutations, including potentially
371 deleterious ones, to persist. In line with this expectation, we observed a higher ratio in invasive
372 populations (Figure 1, y axis), and a negative correlation between this ratio and synonymous
373 expected heterozygosity within each species (Pearson's $r = -0.95$, $P < 10^{-2}$ for DVV; Pearson's
374 $r = -0.90$, $P < 10^{-1}$ for HA; Figure 1). Overall, the differences between native and invasive
375 populations were more pronounced in DVV, characterized by lower diversities and reduced
376 efficacy of selection in the invasive ranges.

377



378
379

380 **Figure 1:** Ratio of non-synonymous to synonymous expected heterozygosity H_{eN}/H_{eS} vs. the
381 synonymous expected heterozygosity H_{eS} of each *Diabrotica virgifera virgifera* (DVV) and *Harmonia*
382 *axyridis* (HA) population. Means and standard errors (error bars) were determined from 100-block
383 jackknife resampling. See Table 1 for population codes.

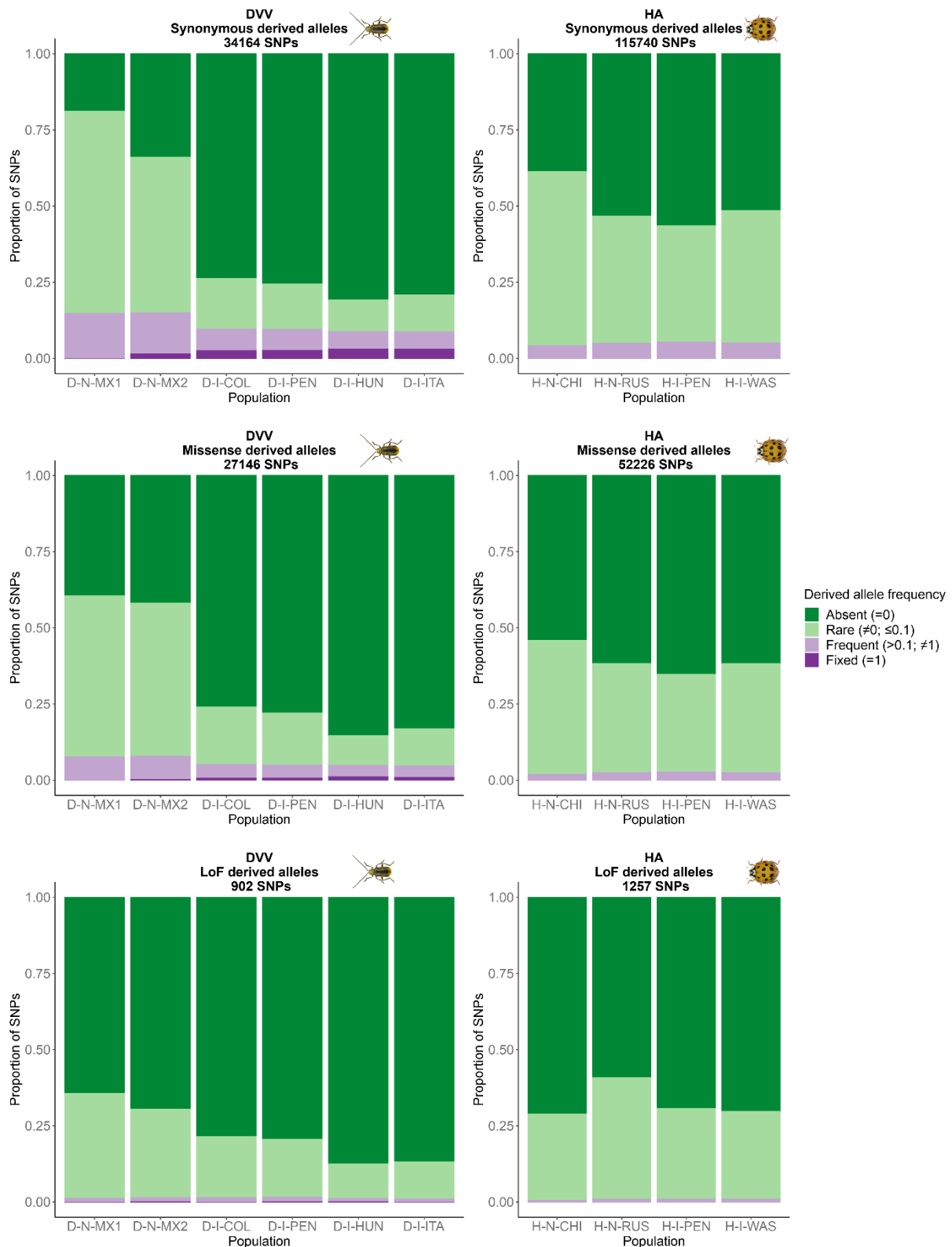
384

385 Genetic load on polarized SNPs

386

387 In all populations studied for each species, derived alleles were mostly rare (with frequencies
388 below 0.1) or absent (Figure 2). The highest derived allele frequencies consistently showed a
389 higher prevalence in synonymous positions compared to non-synonymous positions. In DVV,
390 we observed a substantial loss of non-synonymous derived alleles in invasive populations
391 compared to native ones. Although this was characterized by a reduction in both masked load
392 (proportion of non-synonymous SNPs with rare derived alleles ranging from 0.496 to 0.523 in
393 native populations and from 0.145 to 0.188 in invasive populations) and realized load
394 (proportion of non-synonymous SNPs with frequent or fixed derived alleles ranging from 0.075
395 to 0.077 in native populations and from 0.046 to 0.051 in invasive populations), it also came
396 with slightly higher proportions of fixed derived alleles. This trend was less pronounced for
397 mutations inferred to be strongly deleterious (LoF) than for those inferred to be moderately
398 deleterious (missense). For HA, the disparities between native and invasive populations were
399 considerably less pronounced (Figure 2). Compared to their native Chinese source, HA invasive
400 populations displayed a slight reduction in masked load (proportion of non-synonymous SNPs
401 with rare derived alleles equal to 0.455 in the Chinese population and from 0.347 to 0.381 in
402 invasive populations) and a slight increase in realized load (proportion of non-synonymous
403 SNPs with frequent or fixed derived alleles equal to 0.019 in the Chinese population and from
404 0.024 to 0.027 in invasive populations). Again, this trend was less pronounced for strongly
405 deleterious mutations (Figure 2).

406



407

408

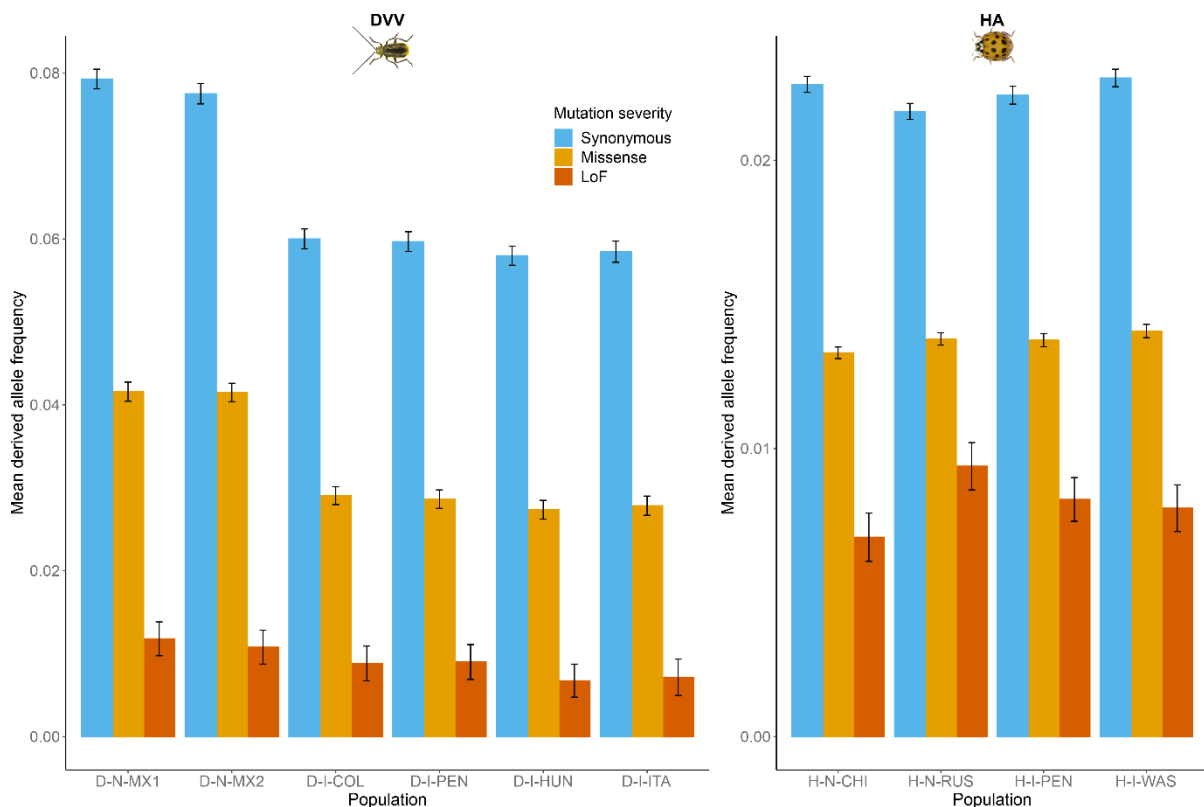
409 **Figure 2:** Proportion of derived alleles found in each population of *Diabrotica virgifera virgifera* (DVV)

410 and *Harmonia axyridis* (HA) within categories of putative severity, classified by their frequency. See

411 Table 1 for population codes.

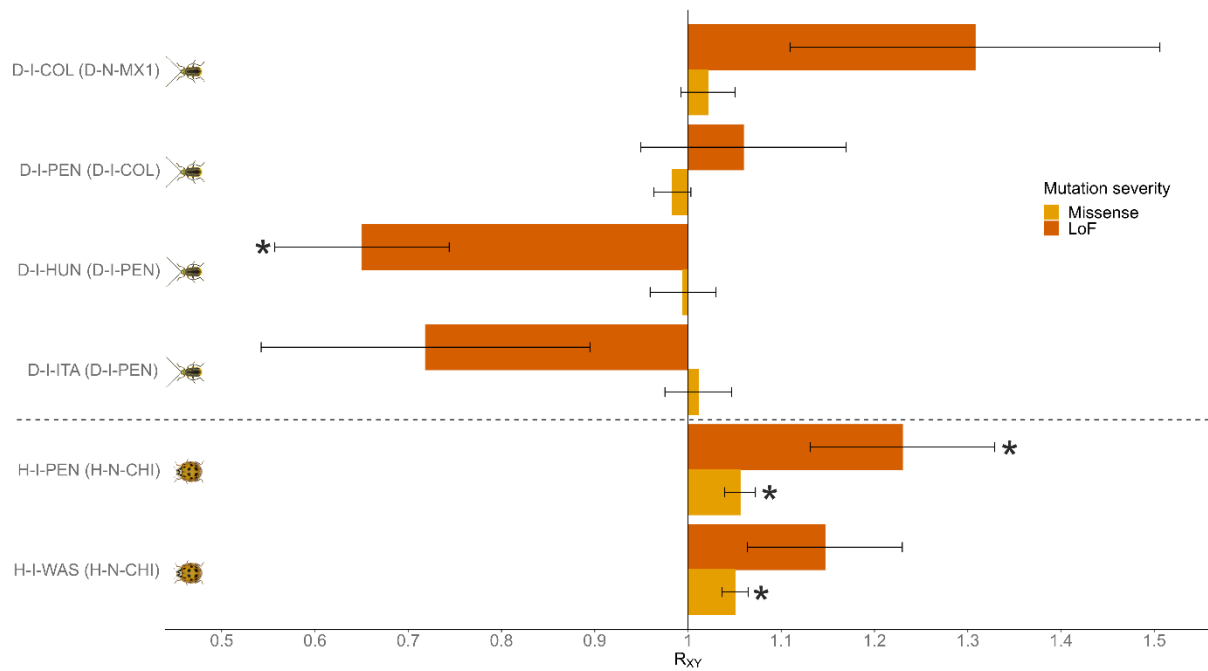
412

413 As expected, the mean derived allele frequencies within populations decline with
414 increasing putative severity in both species (Figure 3). For DVV, synonymous and missense
415 mean derived allele frequencies were consistently significantly lower in the invasive
416 populations than in the native populations ($P < 10^{-4}$ for all comparisons after FDR correction),
417 whereas no significant differences were observed between native populations or between
418 invasive populations ($P > 0.3$ for all comparisons after FDR correction). Within each of the
419 three severity categories, the mean derived allele frequencies were virtually identical in the
420 four HA populations ($P > 0.02$ for all comparisons after FDR correction). No differences in LoF
421 derived allele frequencies were significant ($P > 0.2$ for all comparisons after FDR correction),
422 for either species. Overall, derived allele frequencies were lower in HA than in DVV.
423
424



425
426
427 **Figure 3:** Mean derived allele frequencies in all *Diabrotica virgifera virgifera* (DVV) and *Harmonia*
428 *axyridis* (HA) populations at all three severity categories. Means and standard errors (error bars) were
429 determined from 100-block jackknife resampling. See Table 1 for population codes.
430

431 Finally, the R_{XY} ratio revealed no significant differences in relative frequencies of
432 missense derived alleles between invasive populations and their respective sources in DVV
433 (Figure 4). For putatively highly deleterious loss-of-function alleles, R_{XY} values sometimes
434 deviate sharply from 1, but the number of SNPs is small (Table S5), and most differences are
435 not statistically significant. The only exception is found in the population from Hungary (D-I-
436 HUN), which shows reduced relative frequencies within the putatively highly deleterious
437 category (LoF) of alleles (Figure 4). Regarding HA, slightly higher loads were significant in three
438 out of four comparisons across severity categories between the two invasive populations (H-
439 I-PEN and H-I-WAS) and the native Chinese population (Figure 4).
440



441

442

443 **Figure 4:** Relative mutation load estimated as the R_{XY} ratio of derived alleles for moderate and strong
444 severity categories in populations of *Diabrotica virgifera virgifera* (D) and *Harmonia axyridis* (H). A ratio
445 below or above 1 indicates, respectively, a relative frequency deficit or excess of derived alleles of
446 indicated severity category in the invasive population (listed first) compared to its population source
447 (in parentheses). Means and standard errors (error bars) were determined from 100-block jackknife
448 resampling. A star indicates that the corresponding R_{XY} value is significantly different from 1 (z-score
449 two-tailed test). For population codes, refer to Table 1.

450

451 Discussion

452

453 Exploring the evolution of genetic load using genomic data has long been restricted to model
454 species. Here, we successfully identified putatively deleterious derived alleles within coding
455 regions in two non-model invasive insect species using a pool-seq transcriptome-based exome
456 capture protocol (Deleury et al. 2020). By comparing native and invasive populations, we were
457 thereby able to test the hypothesis that the genetic load may have been purged during the
458 invasion process, which may enhance invasion success of certain populations (Sherpa and
459 Després 2021; Daly et al. 2023). Given that bottlenecks amplify genetic drift and weaken
460 selection effectiveness (Crow 1970; Glémin et al. 2003), the likelihood of purging is expected
461 to increase for recessive, strongly deleterious mutations, while moderately deleterious
462 mutations may still experience some degree of drift, leading to higher frequencies (Agrawal
463 and Whitlock 2011). Our study exploring the fate of the genetic load in DVV and HA yielded
464 mixed results, underscoring the nuanced interplay of evolutionary forces in invasive species.

465

466 Evolution of the genetic load in *Diabrotica virgifera virgifera*

467

468 DVV exhibits marked genetic differences between native and invasive populations. Consistent
469 with previous studies (Ciosi et al. 2008; Lombaert et al. 2018), invasive populations, including
470 the long-established population in Colorado, display sharp declines in genetic diversity.
471 Compared to their native counterparts, all invasive populations generally exhibit lower

472 frequencies of derived alleles, regardless of their potential fitness impacts, thus confirming
473 the substantial role of genetic drift in the invasion process of the species. While this overall
474 reduction in genetic load within invasive populations is apparent, it is worth emphasizing the
475 presence of a slightly larger fixed load.

476 In North America, despite a strong initial loss of genetic diversity, no significant relative
477 deficit or excess of deleterious mutations was detected in invasive populations compared to
478 their source. This pattern includes the absence of any quantifiable evolution of genetic load
479 during DVV's geographic expansion eastward to Pennsylvania, differing from expectations
480 associated with the concept of "expansion load". Under this concept, rare deleterious alleles
481 may "surf" along the expansion front and ultimately reach high frequencies (Klopfstein et al.
482 2006; Travis et al. 2007; Peischl and Excoffier 2015).

483 Interestingly, the oldest European outbreak in Central-Southeastern Europe (here a
484 sample from Hungary) exhibited a significant deficit of "loss-of-function" deleterious
485 mutations compared to its source population, while "missense" deleterious mutations did not
486 differ significantly. This suggests that the demographic and selective constraints in this
487 population were effective at purging highly deleterious mutations but perhaps not moderately
488 deleterious ones. Such differences in evolutionary trajectories according to mutation severity
489 are theoretically expected (Whitlock 2002; Caballero et al. 2017; Dussex et al. 2023) and have
490 been observed in other species experiencing significant bottlenecks (Xue et al. 2015; Khan et
491 al. 2021; Ochoa and Gibbs 2021; Wang et al. 2023).

492 Finally, the invasive Italian population, despite originating from an independent
493 introduction (Miller et al. 2005), showed patterns similar to that of the Central- Southeastern
494 European population. However, it did not exhibit significant signals of purging or fixation of
495 alleles that contribute to genetic load, regardless of the mutation severity considered.

496

497 *Evolution of the genetic load in Harmonia axyridis*

498

499 In HA, the overall loss of genetic diversity in both invasive populations is relatively low, in
500 agreement with previous findings (Lombaert et al. 2010, 2011). While the putative masked
501 genetic load has decreased slightly, the putative realized load, although consistently low in all
502 populations, showed no sign of decreasing and even exhibits a slight increase. This observation
503 aligns with a trend towards a relative excess of deleterious mutations in invasive populations
504 compared to their native source, regardless of the putative fitness impact. This tendency to
505 fix alleles contributing to genetic load, although subtle, is significant across all populations and
506 severity categories, except for the strongly deleterious mutations in Western North America
507 (sampled in Washington).

508 These findings diverge from a previous study based on life history traits, which
509 demonstrated that invasive populations of HA experienced minimal inbreeding depression
510 while maintaining fitness comparable to outbred native populations (Facon et al. 2011). That
511 study has greatly contributed to the hypothesis of purging genetic load as a key mechanism
512 driving successful biological invasions (Estoup et al. 2016; Sherpa and Després 2021; Daly et
513 al. 2023). Given the markedly different methodologies used in the study by Facon et al. (2011)
514 and our study, it is challenging to ascertain why our results do not align. It is conceivable that
515 a few key genes could play decisive roles in invasion success, an aspect difficult to quantify
516 using our approach, particularly in the context of an overall low genetic load, as previously
517 described in this species (Tayeh et al. 2013). This issue is further complicated by the imperfect
518 nature of our categorization, where missense mutations likely cover a broad spectrum of

519 severity levels, potentially obscuring the influence of these key genes. Finally, our exome
520 capture protocol only targets a fraction of the exome, approximately one-third, and a broader
521 approach may yield qualitatively different results.

522

523 *Conclusions and perspectives*

524

525 We tested the hypothesis of the purge of the genetic load during the invasion of two insect
526 species, a crop pest and a predator. At first glance, our results offer a nuanced perspective. In
527 the case of DVV, we were unable to detect any significant evolution of the genetic load, except
528 for the purge of strongly deleterious mutations in the invasive Central- Southeastern European
529 population. For HA, we observed subtle evidence suggesting a tendency toward fixation,
530 rather than toward purge, of the genetic load. Although these results are not entirely
531 contradictory, they may stem from differences in initial genetic composition, varying
532 demographic history, and divergent ecological niches.

533 While population genomics approaches to assess the evolution of genetic load have
534 become increasingly popular in conservation and domestication biology (Moyers et al. 2018;
535 Bertorelle et al. 2022), our study is one of the first in the context of biological invasions.
536 However, it does not yet provide a comprehensive and definitive answer to the question
537 addressed. To broaden our results and enhance our statistical power, particularly for strongly
538 deleterious mutations, it would be essential to consider the entire exome rather than only a
539 fraction. Additionally, categorizing mutation severity by quantifying evolutionary constraints
540 within a phylogenetic framework could enhance our understanding of genetic load evolution,
541 including in non-coding regions (Davydov et al. 2010). Furthermore, extending this research
542 to a wider range of species would enhance our ability to identify potential patterns in the
543 evolution of the genetic load during biological invasions. Finally, to test not only the
544 occurrence of purging in invasions but also its role in the success of invasions, we must
545 compare the load dynamics in introductions that fail with those that succeed – a comparison
546 that is rarely feasible (Zenni and Nuñez 2013), except in controlled invasions such as in classical
547 biological control (Fauvergue et al. 2012).

548

549 **Acknowledgments**

550

551 We thank our colleagues Pamela Bruno, Charlyne Martignier-Jaccard, Pascal Maignet and
552 Wang Su for *Diabrotica adelpha*, *Harmonia quadripunctata* and *Harmonia yedoensis* samples.
553 We also thank Emmanuelle Murciano-Germain for administrative assistance. Sequencing was
554 performed at the GENOTOUL GeT platform (<https://get.genotoul.fr/en/>).

555

556 **Funding**

557

558 This work was funded by grants from ANR project GENLOADICS, and INRAe SPE department.
559 Sampling in Russia was supported by the basic project of the Sukachev Institute of Forest SB
560 RAS (no. FWES-2024-0029).

561

562 **Conflict of interest disclosure**

563

564 The authors of this preprint declare that they have no financial conflict of interest with the
565 content of this article. TG and EL are both recommenders at PCI Evolutionary Biology and PCI
566 Ecology. TG is co-founder of Peer Community In.

567

568 **Data availability**

569

570 Target capture pool sequencing files and all sequencing data have been deposited in Sequence
571 Read Archive, National Center for Biotechnology Information, under project PRJNA1079689
572 (see also Table S1): <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1079689>. All scripts used
573 in this study have been deposited at Data INRAE: <https://doi.org/10.57745/ESQFDB>

574

575 **Author contributions**

576

577 EL and ED designed the study. JB, GC, NK, TS and ST managed the collection of samples. AB
578 and ED prepared the samples and the libraries. ED and BP developed the bioinformatics
579 pipelines. EL and ED analyzed the data. EL and TG wrote the paper. All authors have revised
580 and approved the final manuscript.

581

582 **References**

583

584 Agrawal AF, Whitlock MC (2011) Inferences about the distribution of dominance drawn from
585 yeast gene knockout data. *Genetics* 187:553–566.

586 <https://doi.org/10.1534/genetics.110.124560>

587 Andrews S (2010) FastQC: a quality control tool for high throughput sequence data.

588 <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>

589 Avila V, Amador C, Garcia-Dorado A (2010) The purge of genetic load through restricted

590 panmixia in a *Drosophila* experiment. *J Evol Biol* 23:1937–1946.

591 <https://doi.org/10.1111/j.1420-9101.2010.02058.x>

592 Beaumont MA, Zhang WY, Balding DJ (2002) Approximate Bayesian computation in
593 population genetics. *Genetics* 162:2025–2035

594 Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - a practical and powerful
595 approach to multiple testing. *J R Stat Soc Ser B-Methodological* 57:289–300

596 Bertorelle G, Raffini F, Bosse M, et al (2022) Genetic load: genomic estimates and

597 applications in non-model animals. *Nat Rev Genet* 23:492–503.

598 <https://doi.org/10.1038/s41576-022-00448-x>

599 Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence
600 data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>

601 Caballero A, Bravo I, Wang J (2017) Inbreeding load and purging: implications for the short-

602 term survival and the conservation management of small populations. *Heredity (Edinb)*

603 1–9. <https://doi.org/10.1038/hdy.2016.80>

604 Charlesworth D, Morgan MT, Charlesworth B (1990) Inbreeding Depression, Genetic Load,
605 and the Evolution of Outcrossing Rates in a Multilocus System With No Linkage.

606 *Evolution (N Y)* 44:1469–1489. <https://doi.org/10.1111/j.1558-5646.1990.tb03839.x>

607 Cingolani P, Patel VM, Coon M, et al (2012a) Using *Drosophila melanogaster* as a model for

608 genotoxic chemical mutational studies with a new program, SnpSift. *Front Genet* 3:1–9.

609 <https://doi.org/10.3389/fgene.2012.00035>

- 610 Cingolani P, Platts A, Wang LL, et al (2012b) A program for annotating and predicting the
611 effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)* 6:80–92.
612 <https://doi.org/10.4161/fly.19695>
- 613 Ciosi M, Miller NJ, Kim KS, et al (2008) Invasion of Europe by the western corn rootworm,
614 *Diabrotica virgifera virgifera*: multiple transatlantic introductions with various
615 reductions of genetic diversity. *Mol Ecol* 17:3614–3627. <https://doi.org/10.1111/j.1365-294X.2008.03866.x>
- 617 Coates BS, Deleury E, Gassmann AJ, et al (2021) Up-regulation of apoptotic- and cell survival-
618 related gene pathways following exposures of western corn rootworm to *B.*
619 *thuringiensis* crystalline pesticidal proteins in transgenic maize roots. *BMC Genomics*
620 22:1–27. <https://doi.org/10.1186/s12864-021-07932-4>
- 621 Crnokrak P, Barrett SCH (2002) Purging the genetic load: A review of the experimental
622 evidence. *Evolution (N Y)* 56:2347–2358. [https://doi.org/10.1554/0014-3820\(2002\)056\[2347:pptgla\]2.0.co;2](https://doi.org/10.1554/0014-3820(2002)056[2347:pptgla]2.0.co;2)
- 624 Crow JF (1970) Genetic loads and the cost of natural selection. In: Krickeberg K, Lewontin RC,
625 Neyman J, Schreiber M (eds) *Mathematical topics in population genetics*. Springer
626 Berlin Heidelberg, Berlin, pp 128–177
- 627 Daly EZ, Chabrerie O, Massol F, et al (2023) A synthesis of biological invasion hypotheses
628 associated with the introduction–naturalisation–invasion continuum. *Oikos* e09645:1–
629 29. <https://doi.org/10.1111/oik.09645>
- 630 Danecek P, Bonfield JK, Liddle J, et al (2021) Twelve years of SAMtools and BCFtools.
631 *Gigascience* 10:1–4. <https://doi.org/10.1093/gigascience/giab008>
- 632 Davydov E V., Goode DL, Sirota M, et al (2010) Identifying a high fraction of the human
633 genome to be under selective constraint using GERP++. *PLoS Comput Biol* 6:1–13.
634 <https://doi.org/10.1371/journal.pcbi.1001025>
- 635 Deleury E, Guillemaud T, Blin A, Lombaert E (2020) An evaluation of pool-sequencing
636 transcriptome-based exon capture for population genomics in non-model species.
637 *bioRxiv* 10.1101/58:v7 peer-reviewed and recommended by PCI Genomics.
638 <https://doi.org/10.1101/583534>
- 639 Do R, Balick D, Li H, et al (2015) No evidence that selection has been less effective at
640 removing deleterious mutations in Europeans than in Africans. *Nat Genet* 47:126–131.
641 <https://doi.org/10.1038/ng.3186>
- 642 Dussex N, Morales HE, Grossen C, et al (2023) Purging and accumulation of genetic load in
643 conservation. *Trends Ecol Evol* 38:961–969. <https://doi.org/10.1016/j.tree.2023.05.008>
- 644 Dussex N, van der Valk T, Morales HE, et al (2021) Population genomics of the critically
645 endangered kākāpō. *Cell Genomics* 100002:1–10.
646 <https://doi.org/10.1016/j.xgen.2021.100002>
- 647 Eben A, de los Monteros AE (2013) Tempo and mode of evolutionary radiation in
648 Diabroticina beetles (genera *Acalymma*, *Cerotoma*, and *Diabrotica*). *Zookeys* 332:207–
649 231. <https://doi.org/10.3897/zookeys.332.5220>
- 650 Estoup A, Ravigné V, Hufbauer R, et al (2016) Is There A Genetic Paradox of Biological
651 Invasion? *Annu Rev Ecol Evol Syst* 47:51–72. <https://doi.org/10.1146/annurev-ecolsys-121415>
- 653 Facon B, Hufbauer RA, Tayeh A, et al (2011) Inbreeding depression is purged in the invasive
654 insect *Harmonia axyridis*. *Curr Biol* 21:424–427.
655 <https://doi.org/10.1016/j.cub.2011.01.068>
- 656 Fauvergue X, Vercken E, Malausa T, Hufbauer RA (2012) The biology of small, introduced

- 657 populations, with special reference to biological control. *Evol Appl* 5:424–443.
658 <https://doi.org/10.1111/j.1752-4571.2012.00272.x>
- 659 Fountain T, Duvaux L, Horsburgh G, et al (2014) Human-facilitated metapopulation dynamics
660 in an emerging pest species, *Cimex lectularius*. *Mol Ecol* 23:1071–1084.
661 <https://doi.org/10.1111/mec.12673>
- 662 Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing.
663 arXiv 1207.3907:1–9. <https://doi.org/10.48550/arXiv.1207.3907>
- 664 Gautier M, Vitalis R, Flori L, Estoup A (2022) f-Statistics estimation and admixture graph
665 construction with Pool-Seq or allele count data using the R package poolstat. *Mol Ecol*
666 *Resour* 22:1394–1416. <https://doi.org/10.1111/1755-0998.13557>
- 667 Glémin S (2003) How are deleterious mutations purged? Drift versus nonrandom mating.
668 *Evolution* (N Y) 57:2678–2687. <https://doi.org/10.1111/j.0014-3820.2003.tb01512.x>
- 669 Glémin S, Ronfort J, Bataillon T (2003) Patterns of Inbreeding Depression and Architecture of
670 the Load in Subdivided Populations. *Genetics* 165:2193–2212.
671 <https://doi.org/10.1093/genetics/165.4.2193>
- 672 Gouzy J, Carrere S, Schiex T (2009) FrameDP: Sensitive peptide detection on noisy matured
673 sequences. *Bioinformatics* 25:670–671. <https://doi.org/10.1093/bioinformatics/btp024>
- 674 Gray ME, Sappington TW, Miller NJ, et al (2009) Adaptation and Invasiveness of Western
675 Corn Rootworm: Intensifying Research on a Worsening Pest. *Annu Rev Entomol* 54:303–
676 321
- 677 Grossen C, Guillaume F, Keller LF, Croll D (2020) Purging of highly deleterious mutations
678 through severe bottlenecks in Alpine ibex. *Nat Commun* 11:1001.
679 <https://doi.org/10.1038/s41467-020-14803-1>
- 680 Henn BM, Botigué LR, Bustamante CD, et al (2015) Estimating the mutation load in human
681 genomes. *Nat Rev Genet* 16:333–343. <https://doi.org/10.1038/nrg3931>
- 682 Keightley PD, Jackson BC (2018) Inferring the Probability of the Derived vs. the Ancestral
683 Allelic State at a Polymorphic Site. *Genetics* 209:897–906. [https://doi.org/10.1007/978-](https://doi.org/10.1007/978-1-59745-494-0_7)
684 [1-59745-494-0_7](https://doi.org/10.1007/978-1-59745-494-0_7)
- 685 Khan A, Patel K, Shukla H, et al (2021) Genomic evidence for inbreeding depression and
686 purging of deleterious genetic variation in Indian tigers. *Proc Natl Acad Sci*
687 118:e2023018118. <https://doi.org/10.1073/pnas.2023018118>
- 688 Klopstein S, Currat M, Excoffier L (2006) The fate of mutations surfing on the wave of a
689 range expansion. *Mol Biol Evol* 23:482–490. <https://doi.org/10.1093/molbev/msj057>
- 690 Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
691 arXiv 1303.3997:1–3. <https://doi.org/10.48550/arXiv.1303.3997>
- 692 Li H, Handsaker B, Wysoker A, et al (2009) The Sequence Alignment/Map format and
693 SAMtools. *Bioinformatics* 25:2078–2079.
694 <https://doi.org/10.1093/bioinformatics/btp352>
- 695 Lohmueller KE, Indap AR, Schmidt S, et al (2008) Proportionally more deleterious genetic
696 variation in European than in African populations. *Nature* 451:994–997.
697 <https://doi.org/10.1038/nature06611>
- 698 Lombaert E, Ciosi M, Miller NJ, et al (2018) Colonization history of the western corn
699 rootworm (*Diabrotica virgifera virgifera*) in North America: insights from random forest
700 ABC using microsatellite data. *Biol Invasions* 20:665–677.
701 <https://doi.org/https://doi.org/10.1007/s10530-017-1566-2>
- 702 Lombaert E, Guillemaud T, Cornuet JM, et al (2010) Bridgehead effect in the worldwide
703 invasion of the biocontrol harlequin ladybird. *PLoS One* 5:e9743. <https://doi.org/e9743>

- 704 10.1371/journal.pone.0009743
- 705 Lombaert E, Guillemaud T, Lundgren J, et al (2014) Complementarity of statistical treatments
706 to reconstruct worldwide routes of invasion: the case of the Asian ladybird *Harmonia*
707 *axyridis*. *Mol Ecol* 23:5979–5997. <https://doi.org/10.1111/mec.12989>
- 708 Lombaert E, Guillemaud T, Thomas CE, et al (2011) Inferring the origin of populations
709 introduced from a genetically structured native range by approximate Bayesian
710 computation: case study of the invasive ladybird *Harmonia axyridis*. *Mol Ecol* 20:4654–
711 4670. <https://doi.org/10.1111/j.1365-294X.2011.05322.x>
- 712 Lynch M, Conery IJ, Burger R (1995) Mutation Accumulation and the Extinction of Small
713 Populations. *Am Nat* 146:489–518. <https://doi.org/10.1086/285812>
- 714 Makino T, Rubin C-J, Carneiro M, et al (2018) Elevated proportions of deleterious genetic
715 variation in domestic animals and plants. *Genome Biol Evol* 10:276–290.
716 <https://doi.org/10.1093/gbe/evy004/4794728>
- 717 Marchini GL, Sherlock NC, Ramakrishnan AP, et al (2016) Rapid purging of genetic load in a
718 metapopulation and consequences for range expansion in an invasive plant. *Biol*
719 *Invasions* 18:183–196. <https://doi.org/10.1007/s10530-015-1001-5>
- 720 Marsden CD, Ortega-Del Vecchyo D, O’Brien DP, et al (2016) Bottlenecks and selective
721 sweeps during domestication have increased deleterious genetic variation in dogs. *Proc*
722 *Natl Acad Sci* 113:152–157. <https://doi.org/10.1073/pnas.1512501113>
- 723 Miller N, Estoup A, Toepfer S, et al (2005) Multiple transatlantic introductions of the western
724 corn rootworm. *Science* (80-) 310:992. <https://doi.org/10.1126/science.1115871>
- 725 Moyers BT, Morrell PL, McKay JK (2018) Genetic costs of domestication and improvement. *J*
726 *Hered* 109:103–116. <https://doi.org/10.1093/jhered/esx069>
- 727 Mullarkey A a., Byers DL, Anderson RC (2013) Inbreeding depression and partitioning of
728 genetic load in the invasive biennial *Alliaria petiolata* (Brassicaceae). *Am J Bot* 100:509–
729 518. <https://doi.org/10.3732/ajb.1200403>
- 730 Nei M, Maruyama T, Chakraborty R (1975) Bottleneck Effect and Genetic Variability in
731 Populations. *Evolution* (N Y) 29:1–10. <https://doi.org/10.2307/2407137>
- 732 Ning Z, Cox AJ, Mullikin JC (2001) SSAHA: a fast search method for large DNA databases. -
733 Abstract - UK PubMed Central. 1725–1729. <https://doi.org/10.1101/gr.194201.1>
- 734 Ochoa A, Gibbs HL (2021) Genomic signatures of inbreeding and mutation load in a
735 threatened rattlesnake. *Mol Ecol* 30:5454–5469. <https://doi.org/10.1111/mec.16147>
- 736 Parisod C, Trippi C, Galland N (2005) Genetic variability and founder effect in the pitcher
737 plant *Sarracenia purpurea* (Sarraceniaceae) in populations introduced into Switzerland:
738 From inbreeding to invasion. *Ann Bot* 95:277–286. <https://doi.org/10.1093/aob/mci023>
- 739 Peischl S, Excoffier L (2015) Expansion load: Recessive mutations and the role of standing
740 genetic variation. *Mol Ecol* 24:2084–2094. <https://doi.org/10.1111/mec.13154>
- 741 R Core Team (2021) R: A language and environment for statistical computing
- 742 Robinson J, Kyriazis CC, Yuan SC, Lohmueller KE (2023) Deleterious Variation in Natural
743 Populations and Implications for Conservation Genetics. *Annu Rev Anim Biosci* 11:93–
744 114. <https://doi.org/10.1146/annurev-animal-080522-093311>
- 745 Robinson JA, Ortega-Del Vecchyo D, Fan Z, et al (2016) Genomic Flatlining in the Endangered
746 Island Fox. *Curr Biol* 26:1183–1189. <https://doi.org/10.1016/j.cub.2016.02.062>
- 747 Roy HE, Brown PMJ, Adriaens T, et al (2016) The harlequin ladybird, *Harmonia axyridis*:
748 global perspectives on invasion history and ecology. *Biol Invasions* 18:997–1044.
749 <https://doi.org/10.1007/s10530-016-1077-6>
- 750 Sakai AK, Allendorf FW, Holt JS, et al (2001) The population biology of invasive species. *Annu*

- 751 Rev Ecol Syst 32:305–332. <https://doi.org/10.1146/annurev.ecolsys.32.081501.114037>
- 752 Schubert M, Jónsson H, Chang D, et al (2014) Prehistoric genomes reveal the genetic
753 foundation and cost of horse domestication. Proc Natl Acad Sci U S A 111:E5661-9.
754 <https://doi.org/10.1073/pnas.1416991111>
- 755 Seebens H, Blackburn TM, Dyer E, et al (2017) No saturation in the accumulation of alien
756 species worldwide. Nat Commun 8:14435. <https://doi.org/10.1038/ncomms14435>
- 757 Sherpa S, Després L (2021) The evolutionary dynamics of biological invasions: A multi-
758 approach perspective. Evol Appl 1–23. <https://doi.org/10.1111/eva.13215>
- 759 Simberloff D (2009) The Role of Propagule Pressure in Biological Invasions. Annu Rev Ecol
760 Evol Syst 40:81–102
- 761 Tayeh A, Estoup A, Hufbauer RA, et al (2013) Investigating the genetic load of an emblematic
762 invasive species: the case of the invasive harlequin ladybird *Harmonia axyridis*. Ecol Evol
763 3:864–871. <https://doi.org/10.1002/ece3.490>
- 764 Tomaszewska W, Escalona HE, Hartley D, et al (2021) Phylogeny of true ladybird beetles
765 (Coccinellidae: Coccinellini) reveals pervasive convergent evolution and a rapid
766 Cenozoic radiation. Syst Entomol 46:611–631. <https://doi.org/10.1111/syen.12479>
- 767 Travis JMJ, Munkemuller T, Burton OJ, et al (2007) Deleterious mutations can surf to high
768 densities on the wave front of an expanding population. Mol Biol Evol 24:2334–2343.
769 <https://doi.org/10.1093/molbev/msm167>
- 770 Vogel H, Schmidtberg H, Vilcinskis A (2017) Comparative transcriptomics in three ladybird
771 species supports a role for immunity in invasion biology. Dev Comp Immunol 67:452–
772 456. <https://doi.org/10.1016/j.dci.2016.09.015>
- 773 Wang MS, Zhang JJ, Guo X, et al (2021) Large-scale genomic analysis reveals the genetic cost
774 of chicken domestication. BMC Biol 19:1–16. <https://doi.org/10.1186/s12915-021-01052-x>
- 775
- 776 Wang X, Peischl S, Heckel G (2023) Demographic history and genomic consequences of
777 10,000 generations of isolation in a wild mammal. Curr Biol 33:2051–2062.e4.
778 <https://doi.org/10.1016/j.cub.2023.04.042>
- 779 Whitlock MC (2002) Selection, load and inbreeding depression in a large metapopulation.
780 Genetics 160:1191–1202. <https://doi.org/10.1093/genetics/160.3.1191>
- 781 Williamson M, Fitter A (1996) The varying success of invaders. Ecology 77:1661–1666
- 782 Xue Y, Prado-martinez J, Sudmant PH, et al (2015) Mountain gorilla genomes reveal the
783 impact of long-term population decline and inbreeding. Science (80-) 348:242–245.
784 <https://doi.org/10.1126/science.aaa4484>
- 785 Zayed A, Constantin S a., Packer L (2007) Successful biological invasion despite a severe
786 genetic load. PLoS One 2:. <https://doi.org/10.1371/journal.pone.0000868>
- 787 Zenni RD, Nuñez MA (2013) The elephant in the room: The role of failed invasions in
788 understanding invasion biology. Oikos 122:801–815. <https://doi.org/10.1111/j.1600-0706.2012.00254.x>
- 789
- 790
- 791