Title: Rapid genomic evolution in *Brassica rapa* with bumblebee selection in experimental evolution

Author affiliations

Léa Frachon*, Florian P. Schiestl

Department of Systematic and Evolutionary Botany, University of Zürich, Switzerland.

****Corresponding author***: Léa Frachon

Email: lea.frachon@systbot.uzh.ch, Phone: +41 44 63 48435

Key words: fast cycling *Brassica rapa*, experimental evolution, genomic variance, rapid genomic evolution, bumblebees' selection

Abstract

Insect pollinators shape rapid phenotypic evolution of traits related to floral attractiveness and plant reproductive success. However, the underlying genomic changes and their impact on standing genetic variation remain largely unknown despite their importance in predicting adaptive responses in nature or in crop's artificial selection. Here, based on a previous, nine generation experimental evolution study with fast cycling *Brassica rapa* plants adapting to bumblebees, we document genomic evolution associated to the adaptive process. We performed a genomic scan of the allele frequency changes along the genome and estimated the nucleotide diversity and genomic variance changes. We detected signature of selection associated with rapid changes in allelic frequencies on multiple loci. During experimental evolution, we detected an increase in overall genomic variance, whereas for loci under selection, a reduced variance was apparent in both replicates suggesting a parallel evolution. Our study highlights the polygenic nature of short-term pollinator adaptation and the importance of a such genetic architecture in the maintenance of genomic variance during strong natural selection by biotic factors.

Introduction

Pollinator insects are important selective agents for wild- and crop plant species due to their essential role in the reproduction of most flowering plants [1]. While a decline of insect pollinators has been detected in different geographical regions and insect families [2, 3, 4], the understanding of the adaptive potential of plants to such changes in their biotic environment remains in its infancy. Plant adaptation to pollinators typically involves traits associated to flower attractiveness such as (1) flower morphology [5, 6, 7], flower colour [8, 9], flower scent [10, 11, 12], and (2) traits associated to mating system like herkogamy [13, 14] or selfing [15, 16]. While most of the studies assessed the result of long-term evolutionary adaptation to pollinators, tracking the adaptive processes across generations remains poorly described. Both a resurrection approach in natural populations, growing seeds from different generations together, or experimental evolution studies, applying the same selective pressure for multiple generations, can bridge this gap. For instance, using a resurrection approach, Thomann et al. [17] observed phenological and reproductive trait changes over 18 years in *Adonis annua* plants in response to the loss of wild bees. While this approach benefits from ecological realism in natural populations, it makes it difficult to differentiate the effect of the factor of interest from other factors such as climate, also shaping plant evolution. Gervasi and Schiestl [10] performed experimental evolution with fast-cycling *Brassica rapa* plants evolving with different pollinators and under controlled conditions, to identify the evolutionary response to pollinator-mediated selection. They showed, within nine generations of experimental evolution, rapid plant adaptation to bumblebee pollination in phenotypic traits, such as floral volatiles, UV reflection and plant height. However, while the evolved traits are known to be heritable [18, 19], the genomic changes underlying these rapid plant phenotypic changes are still unknown.

In the current context of pollinator decline and the associated changes in pollinator communities, analyzing the genetic architecture of adaptation to pollinators and its association to standing genetic

variation is essential to understand the adaptative potential of plants in changing environments [20, 21, 22]. Molecular genetic studies have uncovered the molecular and genetic bases of several traits involved in pollination and pollinator attractiveness such as selfing [23], pollination syndromes [24, 8, 25, 26, 27], nectar [28, 29] and volatiles [30, 31, 32, 33, 34]. However, insects use a combination of signals (shape, colour, scent) and rewards for identifying suitable flowers leading to plant adaptation based on multiple traits [35]. For instance, honest signals (signals associated with reward) and pollination syndromes (convergent evolution of specific signal combinations selected by pollinators) are good examples of evolution of multiple traits. In a context of rapid changes, genetic correlation among traits may allow the synchronous response of different phenotypic traits to varying patterns of selection [36, 37, 19]. However, we are still in the infancy of understanding the genetic basis involved in the rapid evolutionary response of plants to pollinator changes. Identification of genomic regions involved in plant adaptation to pollinators is essential to predict the adaptive potential of plants to pollinator changes and enable breeding of more attractive crop plants. In addition, an important aspect in conservation is to understand the mechanisms maintaining genetic diversity within populations. In fact, standing genetic variation is an important resource for rapid response to environmental changes [38, 39], and strong directional selection is considered to lead to the loss of genetic diversity resulting in a loss of adaptive potential in populations [40].

Here, based on previous experimental evolution performed by Gervasi and Schiestl [10] with outcrossing fast-cycling *Brassica rapa* plants, we tracked the genomic changes involved in the adaptative response of plants to bumblebee selection compared to hand pollinated control plants. We dissected the main changes observed in the genetic architecture during selection via bumblebee-pollination compared to hand-pollination. Finally, we documented the changes in genetic diversity observed in the context of strong selection by estimating the average nucleotide diversity and genomic variance before and after experimental evolution.

Results

Genomic changes during bumblebee selection. In our study, we observed allele frequency changes (Δh) over nine generations in both bumblebee and control treatments. For instance, 214 alleles (4.5% of all SNPs) were monomorphic after nine generations in the control treatment, against 344 alleles (7.3% of total SNPs) in the bumblebee treatment. Overall, larger genomic changes were observed in the bumblebee treatment compared to the control treatment (**Figure 1A**). Controlling for random genetic drift, we observed significant changes (*pvalue* < 0.05) for 195 SNPs (4.1% of the 4'713 SNPs) in the control treatment (**Figure 1C**), and for 353 SNPs (7.5% of the 4'713 SNPs) in the bumblebee treatment (**Figure 1D**). The most important changes (pvalue < 0.05 and Δh > 0.5) were observed to be 3.2-times higher under bumblebee selection (76 SNPs) than in the control group (24 SNPs, **Figure 1BCD**). The most significant allele frequency change (pvalue < 0.01) was absolute(Δh) = 0.70 with a mean of 0.41 ± 0.12 SD in the control treatment, while it was absolute(Δh) = 0.80 with a mean of 0.52 ± 0.2 SD in the bumblebee treatment (**Figure 2, Figure S3**).

As expected, the selective process is associated with an increase of linkage disequilibrium in the bumblebee treatment, but also in the control treatment. In fact, while the median linkage disequilibrium decay was slower in the first generation (r^2~0.2), this decay increased during selection in both control (r^2~0.3) and the bumblebee treatment (r^2~0.35) in the ninth and the inter-replicate crossing generation (**Figure 3A**). For instance, in the two genomic regions most under bumblebee selection, we observed an important increase of the LD in the bumblebee treatment, stronger than in the control treatment (**Figure 2B & C**). As expected with the increase of LD, we observed a decrease of the number of LD blocks over nine generations (from 949 LD blocks in first generation, to 818 LD block in the control treatment and 791 in the bumblebee treatment) and an increase of their length (**Figure 3B**, **Table S1**).

Identity of candidate genes underlying genomic evolution to bumblebees. After retrieving the annotated genes around 4.2kb (2.1kb upstream, 2.1kb downstream) of the 76 SNPs with allele frequency changes for the bumblebee treatment (pvalue < 0.05, $\Delta h > 0.5$), we obtained a list of 32 candidate genes (**Table S2**). Briefly, most of these genes are involved in encoding receptor kinases (LRR_3, RIOK2), transporters (ABGG35, ABCG38, SLAH1) and signalling (PTI1, PP2C16, PEX3, PSMA7, LOG1, LOG3). ABC transporters may play a role in floral scent production in *Brassica,* as recently shown for *Petunia* [41], a trait under strong selection in the study of Gervasi and Schiestl [10]. Interestingly, the cytokinins encoded by LONELY GUY (like LOG1 and LOG3 in our results) are known for their importance in reproductive development in *Arabidopsis* [42]. Some candidate genes are also involved in pectin synthesis and pollen tube growth (GALT6, 43). However, in view of the low number of markers used in our study, complementary analyses are needed to validate the implication of these genes or biological processes in plant response to bumblebee selection.

Genetic diversity and overall genomic variance. In order to assess the changes in genetic diversity during experimental evolution, we estimated the nucleotide diversity across the genome for the 256 individuals in a sliding window of 4.2kb (median LD block length in ninth generation after bumblebee selection). We did not observe any shift of the nucleotide diversity over nine generation of selection (**Figure 3B**). The average nucleotide diversity (π) in the first generation (mean $_{\pi} = 15.8.10^{-5}$, $\pm$ 11.10^{-5} SD) is similar to the ninth generations (mean $_{\pi} = 15.7.10^{-5}$, $\pm$ $11.5.10^{-5}$ SD for control treatment, mean $_{\pi} = 15.6.10^{-5}$, $\pm$ $11.2.10^{-5}$ SD for bumblebee treatment). The average nucleotide diversity remained similar after the inter-replicate crossing (mean $_{\pi} = 16.10^{-5}$, $\pm$ $11.8.10^{-5}$ SD for control treatment, mean $_{\pi} = 15.8.10^{-5}$, $\pm$ $11.5.10^{-5}$ SD for bumblebee treatment).

Using all SNPs in a genomic principal component analysis (PCA), we observed a structuring of our samples determined by generations, treatments, and replicates (**Figure 4A** and **figures S4**). Along the two first principal components (PCs), explaining 24.5% of the total genomic variance, all individuals

from the two replicates of generation one were well grouped together, and individuals from the latest generations were clearly separated from the first generation, but they were also separated between treatments (**Figure 4A**). We observed the same pattern in the principal components 3 and 4, explaining 17% of the total genomic variance (cumulative variance of the first four axes is 41%, **Figures S4A**).

For plants of generation nine, the pattern of genomic evolution was different between the selection treatments (control *vs* bumblebee). Interestingly, the individuals from the bumblebee treatment were more dispersed in genomic space represented by principal components PC1, PC2 and PC4 than the individuals from the first generation, highlighting a considerable increase of genomic variance (filled orange and yellow dots/bars (G1) and blue dots/bars (B9 and B10), **Figure 4ACD** and **figure S4ACD**). In the control treatment, the variance of the samples of the genomic space created by the principal components (*i.e.* the genomic variance) was similar between the first and last generations (filled orange and yellow dots/bars (G1) and green dots/bars (C9 and C10), **Figure 4ACD** and **figure S4ACD**). The average genomic variance among individuals from the bumblebee treatment increased 75-fold in PC1 and 191-fold in PC2 (2-fold in PC3 and 61-fold in PC4) over nine generations (**Table S3**). In the ninth generation, the average genomic variance of samples on the PC1 was 3-fold greater and 140-fold for PC2 (0.04-fold for PC3 and 39-fold for PC4) in the bumblebee treatment than in the control treatment (**Table S2**).

In the dataset with the 76 SNPs under strongest bumblebee selection (1.6% of the total SNPs), we observed a different pattern. Using this subset of SNPs, only individuals of the bumblebee treatment were clearly separated from the individuals of the first generation and the control treatment in the genomic space represented by the first two PCs axes (**Figure 4B**). In the bumblebee treatment, while the replicates were separated in the dataset with all SNPs (4'713 SNPs, **Figure 4A**), they were clustered closely together (**Figure 4B**) for the 76 SNPs under selection, indicating shared directional

145 selection and parallel evolution (this pattern appears to be unrelated to a sub-sampling of the dataset,
146 **Figure S5**).

147 In terms of overall genetic variance, in the bumblebee treatment, we observed a lower relative
148 variance among individuals, on the PC1, PC2 and PC4, for the subset of selected SNPs (76 SNPs)
149 compared to the all-SNPs dataset (**Figure 4CD** and **figure S4CD**). In the subset of selected SNPs, we
150 observed no clear patterns among treatment, generation, and replicate in the variance of individuals
151 for PC1, PC3 and PC4 (**Figure 4C**), while for PC2, a decrease of genomic variance was detected (**Figure**
152 **4D**).

153

154 **Discussion**

155 Understanding how and how fast selection affects standing genetic variation within the
156 genome remains an important challenge in conservation as well as in evolutionary genomics. Here, we
157 screened for the genomic consequences of biotic selection in an experimental evolution experiment
158 by sequencing genome-wide SNP markers in *Brassica rapa* plant individuals before and after nine
159 generations of selection by bumblebees, and under random hand-pollination. As shown previously at
160 the level of the phenotype, this primarily outcrossing plant shows rapid adaptation to specific
161 pollinators [10]. We documented signature of directional selection driven by bumblebee pollinators,
162 with allele frequency changes at several loci, as well as parallel genomic evolution. Interestingly, we
163 documented a maintenance of standing genetic variation across the genome, a finding that challenges
164 the assumption of a general loss of genetic variation in evolving small populations.

165 In agreement with the previously demonstrated phenotypic selection and evolution in the fast-
166 cycling *Brassica rapa* experimental system [10] in traits known to be heritable [18, 17], we have shown
167 genomic evolution across nine generations associated with the signature of selection. The here
168 documented changes in allele frequencies, increasing linkage disequilibrium, as well as parallel

169 evolution of genomic regions the most under selection, underline the importance of pollinators in
170 shaping plant rapid genomic evolution. Such rapid adaptive evolution is in line with previous results
171 documented in plants responding during only few generations to environmental changes in response
172 to climate variation [44, 45]. However, while many studies reveal the genomic architecture involved in
173 plant adaptation to climate [46, 47, 48, 49], few studies have investigated the genomic regions involved
174 in plant adaptation to biotic factors (pollinators, plants, microorganisms, herbivores, etc.) despite their
175 obvious importance given their direct interaction with plants.

176 Our study highlighted the potential involvement of multiple loci in rapid adaption to bumblebees,
177 which agrees with studies highlighting a polygenic genetic architecture underlying floral evolution [27,
178 37]. On the short evolutionary timescale applied in our study, the involvement of multiple loci can be
179 explained both by the selective agent itself, selecting for combinations of different phenotypic traits,
180 and/or by the complexity of pathways regulating them i.e., many loci underlying a single trait [50, 51].
181 Studies based on population divergence or reproductive isolation between closely related species have
182 shown the importance of polygenic genetic architecture in floral- and reproductive trait evolution [52,
183 53, 54, 55]. Whereas these studies focus on adaptation to abiotic parameters, other recent studies
184 have also highlighted a complex genetic architecture involved in the evolution of mating system or
185 petal colour during pollinator shifts leading to reproductive barriers between species [56, 57]. In our
186 study, the polygenic architecture underlying adaptation to bumblebees could be explained by the
187 observed combination of phenotypic traits involved in the increase attractiveness to bumblebees
188 shown by Gervasi and Schiestl [10]. Among our 32 candidate genes under selection, interestingly,
189 several are associated to transporters, signalling pathways and potential kinase receptors. The function
190 of these genes could be associated to the production and emission of volatile organic compounds
191 involving complex biosynthetic pathways [58], as volatiles were prominently evolving traits in Gervasi
192 and Schiestl [10]. However, the low-density of markers used in our study, the low number of genome-
193 by-sequencing tags and the absence of phenotypic trait data do not allow us to unravel with certainty

194 the number and the identity of the genes under selection. Moreover, among loci under-selection, for
195 some of them, no annotation could be found, or with unknown functions, highlighted the need to
196 deepen our knowledge on function of genes involved in plant-pollinator interactions. Then, the relative
197 contribution of individual loci to the phenotypic variation that matters for bumblebee attraction is still
198 unknown and deserves more attention in the future.

199 In genomic regions most under selection, a loss of genetic variance was observed, as expected by
200 the selective sweep model where beneficial mutations' frequency increases in a population until
201 fixation [59]. It is well known that selective sweeps lead to loss of diversity in favorable alleles and in
202 their surrounding region due to a hitchhiking effect [60]. However, in our study we also observed a
203 maintenance of nucleotide diversity, and an increase of overall genomic variance despite the strong
204 selection imposed by bumblebees during experimental evolution [10], and the small effective size of
205 the populations. This increase of overall genomic variance was observed among individuals in both
206 replicates (B9A and B9B), as well as in the inter-replicate crossing (B10). This pattern might be
207 explained by a weaker selection acting on multiple standing variants (soft sweep) and by the multiple
208 loci underlying individual phenotypic trait evolutionary changes. An increased number of studies
209 demonstrate the importance of polygenic adaptation [61, 62, 50] related to the infinitesimal model
210 (reviewed in Barton et al. [63]), where local adaptation is driven by small allele frequency changes in
211 multiple loci. This outcome is supported by recent work observing the maintenance of standing genetic
212 variation during long-term artificial selection on chicken weight, mainly explained by selection acting
213 on highly polygenic architecture [64]. Multiple genes underlying phenotypic variation are widely
214 emphasized in plants with the advances of GWAs [61], however their involvement in the maintenance
215 of standing genetic variation is still poorly understood and deserves further studies.

216

217

Conclusion

We revealed important genomic changes on multiple loci during bumblebee selection during only nine generations. We hypothesize that the observed complexity of the genetic architecture allows the maintenance of a high standing genetic variation essential for rapid adaptation to future changes. Our study is a first step in the understanding of the complex genomic mechanisms involved during rapid evolutionary adaptation to biotic factors, and we advocate further analyses to understand (1) the genetic architecture underlying phenotypic variation, (2) pleiotropic effects of quantitative-trait locus in rapid adaptation and (3) the mechanisms behind a maintenance of genetic variance. We also underline the importance of better characterizing the gene functions involved in plant-pollinator interactions. Overall, pollinators constitute complex patterns of selection which deserve more attention for predicting the adaptive responses of wild and crop plant species to their decline.

Material and methods

Plant material and experimental design. *Brassica rapa* (Brassicacea) is an outcrossing plant with genetic self-incompatibility, pollinated by diverse insects such as bumblebees, flies or butterflies [65]. Our study used rapid-cycling *Brassica rapa* plants (Wisconsin Fast Plants) selected for its short life cycle of approximately two months from seed to seed. The plants used in this study were grown from seeds produced by the study of Gervasi and Schiestl [10], performing experimental evolution with bumblebees and control hand pollination. We used one seed per individual from 64 plants (half of replicate A and half of replicate B) of (1) the starting generation (generation 1; here called A1 and B1); (2) the ninth generation selected by bumblebees (bumblebee treatment; here called B9A and B9B); (3) the ninth generation of control hand pollination plants (control treatment; here called C9A and C9B; **Figure S1**). Finally, we performed crossings between replicates A and B within each treatment, (generation 10) yielding 32 individuals from the bumblebee treatment (inter-replicate crossing in

bumblebee treatment; here called B10) and 32 individuals from the control treatment (inter-replicate crossing in control treatment; here called C10). These manual crossing are commonly used for reducing the effect of potential inbreed depression on trait changes. Pollen donors and receivers were randomly assigned in these crossing. Each combination of generation*treatment*replicate is called a population (*e.g.* ninth generation, treatment bumblebees, replicate A called B9A is a population). A total of 256 seeds from these 8 populations (first, ninth and tenth generation) were sown out in a phytotron (first generation in 2017 and ninth generation as well as the inter-replicate crossing in 2019) and the leaf tissue of each plant was collected for DNA extraction and whole genomic sequencing.

DNA extraction and genomic characterization. Because leaf tissue was collected in 2017 for the first generation and 2019 for the last generations, we adapted the collection storage (drying vs freezing). Leaf material from the first generation was dried in vacuum at 40 °C for 20 hours, and leaf material from the ninth and tenth generation was stored in -80°C. A high molecular weight DNA extraction (average DNA concentration of 48 ng/µL, LGC extraction protocol) and library preparation for genotyping-by-sequencing (restriction enzyme MsII, insert size mean range ~215bp) was performed by the LGC Genomics group Berlin. Samples were sequenced with Illumina NextSeq 500 V2 sequencer using 150 paired-end reads; the alignment of our samples was performed with BWA version 0.7.12 against the reference genome sequence of *Brassica rapa* FPsc v1.3, Phytozome release 12 (https://phytozome.jgi.doe.gov/pz/portal.html) by the LGC Genomics group Berlin. The variant discovery and the genotyping were realized using Freebayes v1.0.2-16 with the following parameters by the LGC Genomic Group Berlin: --min-base-quality 10 --min-supporting-allele-qsum 10 --read-mismatch-limit 3 --min-coverage 5 --no-indels --min-alternate-count 4 --exclude-unobserved-genotypes --genotype-qualities --ploidy 2 or 3 --no-mnps --no-complex --mismatch-base-quality-threshold 10. We then performed a quality trimming on chromosomes (we discarded the scaffolds) using vcftools, removing SNPs with missing data in more than 5% of the individuals (function --max-missing 0.95, *i.e.* genotype calls had to be present for at least 243 samples out of 256 for a SNP to be

included in the downstream analysis), and retained only bi-allelic SNPs with a minimum average Phred quality score of 15 (function --minGQ 15) and a maximum mean depth value of 100 (function --max-meanDP 100, distribution in **Figure S3**). Finally, we discarded SNPs with a minor allele frequency (MAF) lower than 0.1 (function --maf 0.1, distribution in **Figure S3**). The final dataset contained 4'713 SNPs in ~ 215Mb genome size.

Allele frequency changes. The allele frequencies of the reference allele for the 4'713 SNPs were estimated within each populations using VCFtools (function --freq). To control for potential genetic drift during the nine generations of evolution, we simulated random final allele frequencies 10'000-fold for different ranges of initial allele frequencies (from 0 to 1 by an interval window of 0.01). The simulations were performed using the R environment package "learnPopGen" (function "drift.selection", 66) over eight transitions between generations (*i.e.* from the first generation to the ninth generation) considering 32 individuals within each population for an effective size (N_e) of 16 (*i.e.* individuals contributing to the next generation, see details of experimental evolution in Gervasi and Schiestl [10]), and considering an equal fitness for each individual. From these simulations, a *P value* was estimated for each SNP using the following equations:

(1) For a decrease of reference allelic frequency *i.e.* ($AF_{initial} - AF_{final}$) > 0, *pvalue* = (number of simulation with $AF_{simulated} \geq AF_{final}$)/10'000

(2) For an increase of reference allelic frequency *i.e.* ($AF_{initial} - AF_{final}$) < 0, *pvalue* = (number of simulation with $AF_{simulated} \leq AF_{final}$)/10'000

(3) For ($AF_{initial} - AF_{final}$) = 0, *pvalue* = 1

With $AF_{simulated}$ = simulated final allele frequency, $AF_{initial}$ = initial allele frequency from the reference allele (first generation), and AF_{final} = observed final allele frequency for the reference allele (ninth generation). Using these parameters, the minimum expected pvalue is ~ 1.10^{-4} (1/10'000), except for final allelic frequency completely out of the simulated range (*i.e.* zero simulated allele

291 frequencies are al higher or lower than the final observed allele frequency) would be associated with
292 pvalue=0.

293 Finally, we estimated the allele frequency changes (Δh) from the reference allele according to
294 the equation (1) for both bumblebee and control treatments:

295 $$\Delta h = \mathrm{AF}_{\mathrm{final}} - \mathrm{AF}_{\mathrm{initial}} \qquad (1)$$

296 Where Δh is the allelic frequency change between the first and the ninth generation, $AF_{initial}$ is the initial
297 observed allele frequency, and AF_{final} is the observed final allele frequency at the ninth generation.

298 **Estimation of linkage disequilibrium (LD) changes across rapid evolution.** During selective process,
299 an increase of the linkage disequilibrium is expected, especially in genomic regions strongly under
300 selection. First, we calculated pairwise linkage disequilibrium (LD) among all set of SNPs using
301 VCFtools (function --geno-r2) for 256 samples within each population. The associated median LD was
302 then estimated and plotted. Second, a pairwise LD among SNPs in the surrounding of SNPs highly
303 under selection in bumblebee treatment were calculated using plink1.9 (function -r2). Finally, we
304 calculated the LD blocks in each population using plink1.9 with the following parameters: --blocks no-
305 pheno-req --maf 0.07 --blocks-max-kb 200.

306 **Candidate genes.** We identified candidate genes associated with 43 SNPs with the highest significant
307 allele frequency changes (pvalue < 0.01 and abs(Δh) > 0.5) for the bumblebee treatment. Because the
308 median linkage disequilibrium in bumblebee treatment is 4.2kb, we retrieved the annotated genes
309 around 4.2kb (2.1 kb upstream and 2.1 kb downstream) for 43 SNPs and extracted the gene description
310 using phytozome.jgi.doe.gov. Because some gene descriptions were missing or partially incorrect, we
311 double checked the description of genes using well documented *Arabidopsis thaliana* (Brassicaceae)
312 databases. The gene sequences were extracted for the different transcripts from phytozome, blasted

on TAIR (www.arabidopsis.org), and the description of gene record as well as the GO biological process available for *A. thaliana* was extracted.

Nucleotide diversity and genome-wide variance. We estimated the genetic diversity using a marker-based index *i.e.* the nucleotide diversity (π) over 4'713 SNPs using a 4.2kb sliding windows in VCFtools (--window-pi). The choice of the window size was made according to the median LD in treatment B9 (median LD = 4.2kb).

The genomic variance among individuals per population (i.e., generations* treatment) was estimated performing a principal component analysis (PCA) on scaled and centered genotype data (pcadapt package in R environment, function pcadapt, 67). In order to unravel the changes in genomic variance over nine generations, we performed the PCA on different sub-datasets:

1. On the total number of SNPs (*i.e.* 4'713 SNPs).
2. On 76 SNPs with higher significant (pvalue<0.05 and Δh>0.5) allele frequency changes *i.e.* the genomic regions the most under bumblebee selection.
3. 1'000 times on 76 randomly choose SNPs controlling sub-sampling effect (details and results in SI).

Acknowledgements

We thank Jörg Vogt for the first-generation DNA extraction, and the LGC Genomic group Berlin for the DNA extraction of other samples, the sequencing and bioinformatic treatment. We thank Cyril Zipfel for discussion regarding candidate genes. We thank Loren Rieseberg and Marius Rösti for their useful comments on the draft. Finally, we thank the University of Zürich for their financial support.

References

1. Van der Niet, T., Peakall, R. & Johnson, S.D. Pollinator-driven ecological speciation in plants: new evidence and future perspectives. *Annals of Botany*. **113**, 199–211 (2014).
2. Biesmeijer, J. C. *et al*. Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. *Science*. **313**, 351-354 (2006).
3. Hallmann, C. *et al*. More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS ONE*. **12**(10), e0185809 (2017).
4. Wagner, D. L., Grames, E. M., Forister, M. L., Berenbaum, M. R. & Stopak, D. Insect decline in the Anthropocene: Death by a thousand cuts. *PNAS*. **118**(2), e2023989118 (2020).
5. Anderson, B., Pauw, A., Cole, W. W. & Barrett S. C. H. Pollination, mating and reproductive fitness in a plant population. *Journal of Evolutionary Biology*. **29**, 1631–1642 (2016).
6. Minnaar, C., Anderson, V., de Jager, M. L., & Karron, J. D. Plant–pollinator interactions along the pathway to paternity. *Annals of Botany*. **123**, 225–245 (2019).
7. Wessinger, C. A. & Hileman, L. C. Accessibility, constraint, and repetition in adaptive floral evolution. *Developmental Biology*. **419**, 175–183 (2016).
8. Bradshaw, H. D. & Schemske, D. W. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature*. **426**, 176-178 (2003).
9. Sobral, M., Veiga, T., Domínguez, P., Guitián, J.A., Guitián, P., & Guitián, J. M. Selective pressures explain differences in flower color among Gentiana lutea populations. *PLoS ONE*. **10**(7), e0132522 (2015).
10. Gervasi, D. D. L. & Schiestl, F. P. Real-time divergent evolution in plants driven by pollinators. *Nature communications*. **8**, 14691 (2017).
11. Schiestl, F. P. On the success of a swindle: pollination by deception in orchids. *Naturwissenschaften*. **92**, 255–264 (2005).
12. Xu, S., Schlüter, P. M., Grossniklaus, U. & Schiestl, F. P. The genetic basis of pollinator adaptation in a sexually deceptive Orchid. *PLoS Genet*. **8**(8), e1002889 (2012).
13. Luo, Y. & Widmer, A. Herkogamy and its effects on mating patterns in *Arabidopsis thaliana. PLoS ONE*. **8**(2), e57902 (2013).
14. Moeller, D. A. & Geber, M. A. Ecological context of the evolution of self-pollination in *Clarkia xantiana*: population size, plant communities, and reproductive assurance. *Evolution*. **59**(4), 786–799 (2005).
15. Bodbyl Roles, S. A. & Kelly, J. K. Rapid evolution caused by pollinator loss in *Mimulus guttatus. Evolution*. **65**(9), 2541–2552 (2011).
16. Ramos, S. E. & Schiestl, F. P. Rapid plant evolution driven by the interaction of pollination and herbivory. *Science*. **364**, 193–196 (2019).
17. Thomann, M., Imbert, E. & Cheptou, P.O. Is rapid evolution of reproductive traits in Adonis annua consistent with pollinator decline? *Acta Oecologica*. **69**, 161-166 (2015).
18. Zu, P., Blanckenhorn, W. U. & Schiestl, F. P. Heritability of floral volatiles and pleiotropic responses to artificial selection in *Brassica rapa*. *New Phytologist*. **209**, 1208–1219 (2016).
19. Zu, P. & Schiestl, F. P. The effects of becoming taller: direct and pleiotropic effects of artificial selection on plant height in Brassica rapa. *The Plant Journal*. **89**, 1009–1019 (2017).
20. Bay, R. A. *et al*. Predicting responses to contemporary environmental change using evolutionary response architectures. *The American naturalist*. **189**(5), 463-473 (2017).
21. Bergelson, J. & Roux, F. Towards identifying genes underlying ecologically relevant traits in Arabidopsis thaliana. *Nature Reviews Genetics.* **11**, 867-879 (2010).
22. Hansen, T. F. The evolution of genetic architecture. *Annual Review of Ecology, Evolution, and Systematics*. **37**, 123–57 (2006).
23. Sicard, A. & Lenhard, M. The selfing syndrome: a model for studying the genetic and evolutionary basis of morphological adaptation in plants. *Annals of Botany*. **107**, 433–1443 (2011).

24. Bradshaw, H. D., Wilbert, S. M., Otto, K. G. & Schemsket, D. W. Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (*Mimulus*). *Nature*. **376**, 762-765 (1995).
25. Streisfeld, M. A. & Rausher, M. D. Genetic changes contributing to the parallel evolution of red floral pigmentation among Ipomoea species. *New Phytologist*. **183**, 751–763 (2009).
26. Yarahmadov, T., Robinson, S., Hanemian, M., Pulver, V. & Kuhlemeier, C. Identification of transcription factors controlling floral morphology in wild Petunia species with contrasting pollination syndromes. *The Plant Journal*. **104**(2), 289–301 (2020).
27. Yuan, Y. W., Byers, K. J. R. P. & Bradshaw, H. D. J. The genetic control of flower–pollinator specificity. *Current Opinion in Plant Biology*. **16**, 422–428 (2013).
28. Lin, W. *et al*. Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature*. **508**, 546-549 (2015).
29. Solhaug, E. M. *et al*. An integrated transcriptomics and metabolomics analysis of the Cucurbita pepo nectary implicates key modules of primary metabolism involved in nectar synthesis and secretion. *Plant Direct*. **3**, e00120 (2019).
30. Cai, J., Zu, P. & Schiestl F. P. The molecular bases of floral scent evolution under artificial selection: insights from a transcriptome analysis in Brassica rapa. *Scientific Reports*. **6**, 36966 (2016).
31. Dudareva, N., Negre, F., Nagegowda, D. A. & Orlova, I. Plant Volatiles: Recent Advances and Future Perspectives. *Critical Reviews in Plant Sciences*. **25**, 417–440 (2006).
32. Fattorini, R. & Glover, B. J. Molecular Mechanisms of Pollination Biology. *Annual Review of Plant Biology*. **71**, 487-515 (2020).
33. Klahre, U. *et al*. Pollinator choice in Petunia depends on two major genetic loci for floral scent production. *Current Biology*. **21**, 730-739 (2011).
34. Lo, S., Fatokun, C., Boukar, O., Gepts, P., Close, T. J., & Muñoz-Amatriain, M. Identification of QTL for perenniality and floral scent in cowpea (*Vigna unguiculata* [L.] Walp.). *PLoS ONE*. **15**(4), e0229167 (2020).
35. Schiestl, F.P. & Johnson, S. D. Pollinator-mediated evolution of floral signals. *Trends in Ecology & Evolution*. **28**, 307-315 (2013).
36. Hermann, K., Klahre, U., Moser, M., Sheehan, H., Mandel, T. & Kuhlemeier, C. Tight genetic linkage of prezygotic barrier loci creates a multifunctional speciation island in *Petunia*. *Current Biology*. **23**, 873–877 (2013).
37. Smith, S. D. Pleiotropy and the evolution of floral integration. *New Phytologist*. **209**, 80–85 (2015).
38. Barrett, R. D. H. & Schluter, D. Adaptation from standing genetic. *TRENDS in Ecology and Evolution*. 23(1), 38-44 (2008).
39. Chevin, L. M., Lande, R. & Mace, G. M. Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLOS biology*. **8**, e1000357 (2010).
40. Barton, N. H. & Keightley, P.D. Understanding quantitative genetic variation. *Nature Reviews Genetics*. 3, 11-21 (2002).
41. Adebesin, F. *et al.* Emission of volatile organic compounds from petunia flowers is facilitated by an ABC transporter. *Science*. **356**, 1386-1388 (2017).
42. Terceros G. C., Resentini F., Cucinotta M., Manrique S., Colombo L. & Mendes M. A. The importance of cytokinins during reproductive development in Arabidopsis and beyond. *International Journal of Molecular Sciences*. **21**, 8161 (2020).
43. Lund, C. H. *et al*. Pectin synthesis and pollen tube growth in Arabidopsis involves three GAUT1 Zolgi-Anchoring Proteins: GAUT5, GAUT6, and GAUT7. *Frontiers in Plant Science*. **11**, 585774 (2020).
44. Frachon, L. *et al*. Intermediate degrees of synergistic pleiotropy drive adaptive evolution in ecological time. *Nature Ecology & Evolution*. **1**, 1551–1561 (2017).
45. Franks, S. J., Kane, N. C., O'hara, N. B., Tittes, S. & Rest, J. S. Rapid genome-wide evolution in Brassica rapa populations following drought revealed by sequencing of ancestral and descendant gene pools. *Molecular Ecology*. **25**, 3622–3631 (2016).

46. Frachon, L. *et al*. A genomic map of adaptation to local climate in *Arabidopsis thaliana*. *Frontiers in Plant Science*. **9**, 967 (2018).
47. Lasky, J. R. *et al.* Genome-environment associations in sorghum landraces predict adaptive traits. *Science Advances*. **1**, e1400218 (2015).
48. Li, Y., Cheng, R., Spokas, K. A., Palmer, A. A. & Borevitz, J. O. Genetic variation for life history sensitivity to seasonal warming in Arabidopsis thaliana. *Genetics*. **196**, 569–577 (2014).
49. Yoder, J. B., Stanton-Geddes, J., Zhou, P., Briskine, R., Young, N. D. & Tiffin, P. Genomic signature of adaptation to climate in *Medicago truncatula. Genetics*. **196**, 1263–1275 (2014).
50. Csilléry, K., Rodriguez-Verdugo, K., Rellstab, C. & Guillaume, F. Detecting the genomic signal of polygenic adaptation and the role of epistasis in evolution. *Molecular Ecology*. **27**, 606-612 (2018).
51. Hendry A. P. Key questions in the genetics and genomics of eco-evolutionary dynamics. *Heredity*. **111**, 456-466 (2013).
52. Ferris, K. G, Barnett, L. L., Blackman, B. K. & Willis, J. H. The genetic architecture of local adaptation and reproductive isolation in sympatry within the *Mimulus guttatus* species complex. *Molecular Ecology*. **26**, 208–224 (2017).
53. Fishman, L., Kelly, A. J. & Willis, J. H. Minor quantitative trait loci underlie floral traits associated with mating system divergence in mimulus. *Evolution*. **56**(11), 2138–2155 (2002).
54. Hall, M. C., Basten, C. J. & Willis, J. H. Pleiotropic quantitative trait loci contribute to population divergence in traits associated with life-history variation in *Mimulus guttatus*. *Genetics*. **172**, 1829–1844 (2006).
55. Liu, X. & Karrenberg, S. Genetic architecture of traits associated with reproductive barriers in Silene: Coupling, sex chromosomes and variation. *Molecular Biology*. **27**, 3889–3904 (2018).
56. Campitelli, B. E., Kenney, A. M., Hopkins, R., Soule, J., Lovell, J. T. & Juenger, T.E. Genetic mapping reveals an anthocyanin biosynthesis pathway gene potentially influencing evolutionary divergence between two subspecies of scarlet gilia (*Ipomopsis aggregata*). *Molecular Biology and Evolution*. **35**(4), 807–822 (2017).
57. Fishman, L., Beardsley, P. M., Stathos, A., Williams, C. F. & Hill, J. P. The genetic architecture of traits associated with the evolution of self-pollination in Mimulus. *New Phytologist*. **205**, 907–917 (2015).
58. Borghi, M., Fernie, A. R., Schiestl, F. P. & Bouwmeester, H. J. The sexual advantage of looking, smelling, and tasting good: the metabolic network that produces signals for pollinators. *Trends in Plant Science*. **22**(4), 338-350 (2017).
59. Przeworski, M., Coop, G. & Wall, J. D. The signature of positive selection on standing genetic variation. *Evolution*. **59**(11), 2312–2323 (2005).
60. Smith, J.M. & Haigh, J. The hitch-hiking effect of a favourable gene. *Genetics research Cambridge*. **23**, 23-35 (1974).
61. Barghi, N., Hermisson, J. & Schlötterer, C. Polygenic adaptation: a unifying. *Nature Reviews Genetics*. 21, 769-781 (2020).
62. Pritchard, J. K. & Di Rienzo, A. Adaptation – not by sweeps alone. *Nature Reviews Genetics*. **11**(10), 665–667 (2010).
63. Barton, N. H., Etheridge, A.M. & Véber, A. The infinitesimal model: Definition, derivation, and implications. *Theoretical Population Biology*. 118, 50-73 (2017).
64. Sheng, Z., Pettersson, M. E., Honaker, C. F., Siegel, P. B. & Carlborg, O. Standing genetic variation as a major contributor to adaptation in the Virginia chicken lines selection experiment. *Genome Biology*. **16**, 219 (2015).
65. El-Esawi, M. A. Genetic diversity and evolution of Brassica genetic resources: from morphology to novel genomic technologies – a review. *Plant Genetic Resources: Characterization and Utilization*. **15**(5), 388–399 (2017).

66. Revell, L. J. learnPopGen: An R package for population genetic simulation and numerical analysis. *Ecology and Evolution*. **9**, 7896–7902 (2019).
67. Privé, F., Luu, K., Vilhjálmsson, B. J. & Blum, M. G. B. Performing highly efficient genome scans for local adaptation with R package pcadapt version 4. *Molecular Biology and Evolution*. **37**(7), 2153–2154 (2020).

Legends

Figure1. Allele frequency changes during experimental evolution. (**A**) Comparison of the allele frequency changes (Δh) between the bumblebee treatment (*x-axis*) and the control treatment (*y-axis*). The grey dots represent the 4'713 SNPs. (**B**) Ven diagram for the number of SNPs with highest allele frequency changes (absolute(Δh) > 0.5) and under significant selection (*pvalue* < 0.05) in the control treatment (green circle) and the bumblebee treatment (blue circle). Comparison of initial (first generation) and final (ninth generation) allele frequencies in the control **(C)** and the bumblebee treatment **(D)**. The grey dots represent the non-significant changes in allele frequencies between generations. The grey solid lines indicate the maximum (upper line) or minimum (lower line) of final simulated allele frequencies obtained by 10'000 simulations of random genetic drift (over nine generations, N_e=16). The green gradient dots represent significant changes (light green for a *pvalue* < 0.05, medium green *pvalue* < 0.01, and dark green *pvalue* < 0.001) calculated from the 10'000 simulations.

Figure 2. Genomic scan of allele frequency changes in bumblebee treatment. The Manhattan plot shows the absolute genomic changes occurring over nine generations (absolute(Δh), *y-axis*) along the genome (*x-axis)* for bumblebee treatment (**A**). The chromosome numbers are indicated below the plots (from I to X), and the different shades of blue encode changes on different chromosomes. The coloured dots are the SNPs under selection (green dots for pvalue < 0.05, and red dots for pvalue < 0.001). We highlighted with the arrows two genomic regions showing important changes near the SNPs 5_136590 and 7_18056205. A zoom of 1Mb in the surrounding regions are plotted in (**B**) for the SNP in the chromosome V, and in (**C**) for the chromosome VII. The meaning of the dots is the same than the Manhattan plot. We added on the plot the median LD (r^2) in these genomic regions using the coloured lines (legend in the bottom left corner of the plot).

Figure 3. Linkage disequilibrium and nucleotide diversity. (A) Distribution of the median pairwise linkage disequilibrium (r^2) for each population by distance between two SNPs (kb). The colour of the population is indicated in the plot (**B**) Number of LD blocks per population (more details Table S1). (**C**) Density plot of the nucleotide diversity (π) measuring in sliding windows of 4.2kb for each population (see legend in 3A for the line colour).

Figure 4. Genomic variance among populations. (A-B) Position of the 256 individuals in the genomic space from the principal component analysis (PCA) performed on their genotypes (GT). The PCAs were performed **(A)** on the total set of SNPs (4'713 SNPs), and **(B)** on the 76 SNPs the most under selection in the bumblebee treatment (pvalue < 0.05 and absolute(Δh) > 0.5). The label of the population is shown on their centroid. The relative variance of the PC1 **(C)** and PC2 **(D)** are represented with the bar plots, where the filled bars are for the PCA performed on the 4'713 SNPs, and the dashed ones on the 76 SNPs. The legend colours are indicated in the bottom right of the figure.

Figures

Figure 1

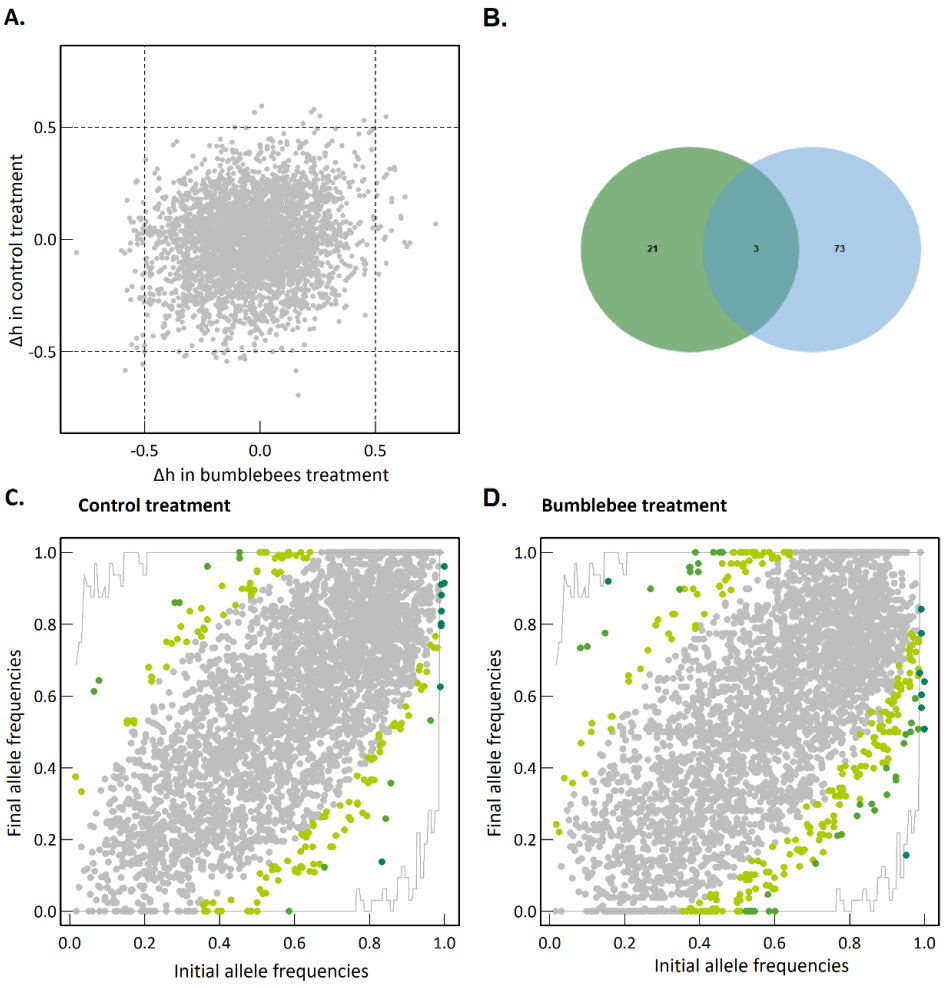


Figure 2

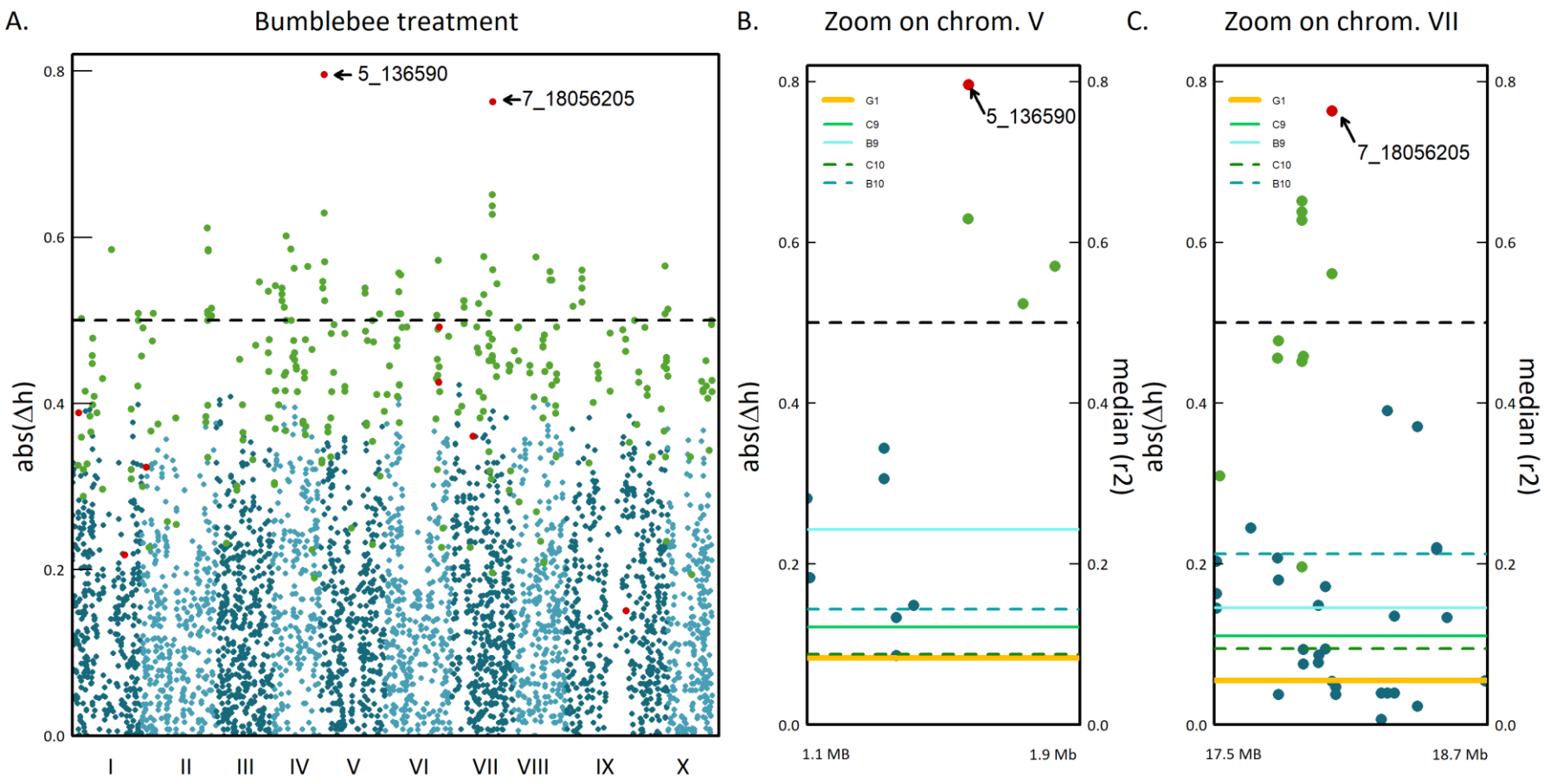

A.
Bumblebee treatment
5_136590
7_18056205
abs(Δh)
0.0
0.2
0.4
0.6
0.8
I
II
III
IV
V
VI
VII
VIII
IX
X
B.
Zoom on chrom. V
G1
C9
B9
C10
B10
5_136590
abs(Δh)
median (r2)
1.1 MB
1.9 Mb
C.
Zoom on chrom. VII
G1
C9
B9
C10
B10
7_18056205
abs(Δh)
median (r2)
17.5 MB
18.7 Mb

Figure 3

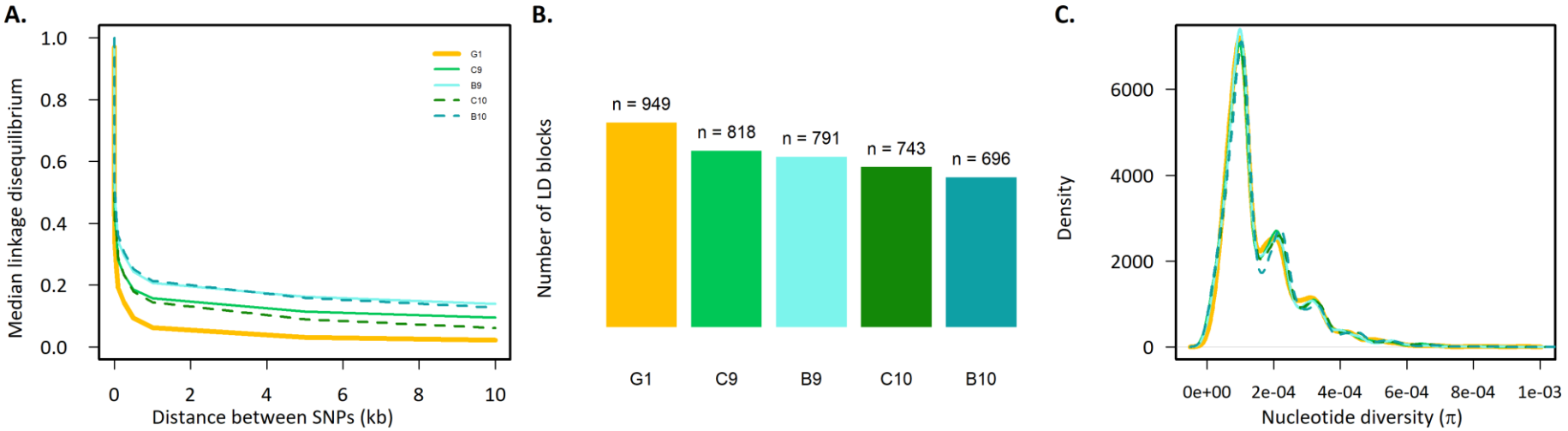
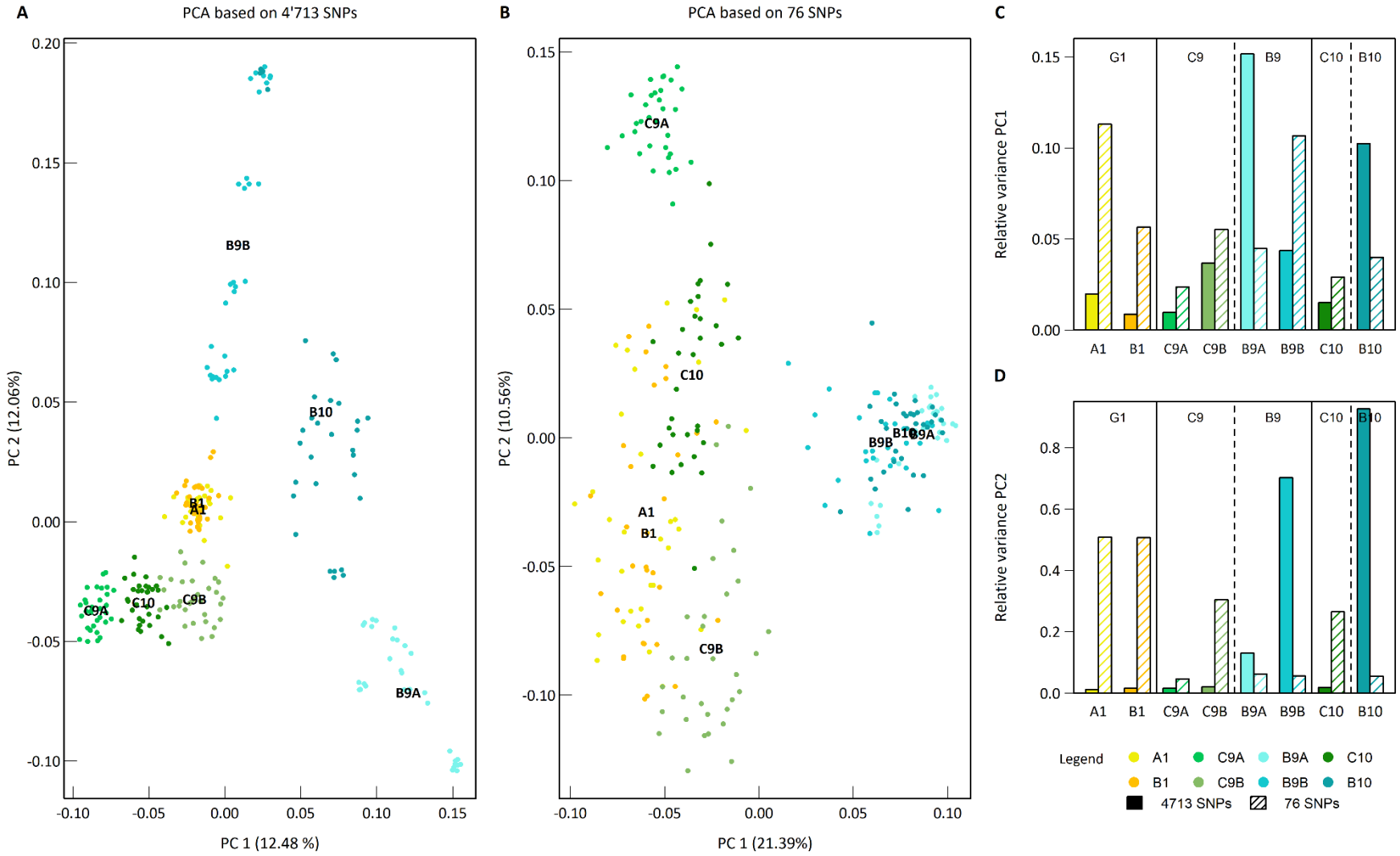

A.
Median linkage disequilibrium
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
Distance between SNPs (kb)
G1
C9
B9
C10
B10
B.
Number of LD blocks
n = 949
n = 818
n = 791
n = 743
n = 696
G1
C9
B9
C10
B10
C.
Density
6000
4000
2000
0
0e+00
2e-04
4e-04
6e-04
8e-04
1e-03
Nucleotide diversity (π)

Figure 4


A
PCA based on 4'713 SNPs
PC 2 (12.06%)
PC 1 (12.48 %)
B9B
B10
B1
A1
C9A
C10
C9B
B9A
B
PCA based on 76 SNPs
PC 2 (10.56%)
PC 1 (21.39%)
C9A
C10
A1
B1
C9B
B9B
B10
B9A
C
Relative variance PC1
G1
C9
B9
C10
B10
A1
B1
C9A
C9B
B9A
B9B
C10
B10
D
Relative variance PC2
Legend
A1
B1
C9A
C9B
B9A
B9B
C10
B10
4713 SNPs
76 SNPs

Supplementary information

Title: Rapid genomic evolution in *Brassica rapa* with bumblebee selection in experimental evolution

Author affiliations

Léa Frachon*, Florian P. Schiestl

Department of Systematic and Evolutionary Botany, University of Zürich, Switzerland.

****Corresponding author***: Léa Frachon

Email: lea.frachon@systbot.uzh.ch, Phone: +41 44 63 48435

Key words: fast cycling *Brassica rapa*, experimental evolution, genomic variance, rapid genomic evolution, bumblebees' selection

Abstract

Insect pollinators shape rapid phenotypic evolution of traits related to floral attractiveness and plant reproductive success. However, the underlying genomic changes and their impact on standing genetic variation remain largely unknown despite their importance in predicting adaptive responses in nature or in crop's artificial selection. Here, based on a previous, nine generation experimental evolution study with fast cycling *Brassica rapa* plants adapting to bumblebees, we document genomic evolution associated to the adaptive process. We performed a genomic scan of the allele frequency changes along the genome and estimated the nucleotide diversity and genomic variance changes. We detected signature of selection associated with rapid changes in allelic frequencies on multiple loci. During experimental evolution, we detected an increase in overall genomic variance, whereas for loci under selection, a reduced variance was apparent in both replicates suggesting a parallel evolution. Our study highlights the polygenic nature of short-term pollinator adaptation and the importance of a such genetic architecture in the maintenance of genomic variance during strong natural selection by biotic factors.

Control of subsampling effect in genomic PCA

In order to control for a potential artefact due to a subsampling, we performed the PCA analysis 1'000-times with 76 SNPs randomly chosen within the full dataset of 4'713 SNPs. The relative variance measured among those 1'000 PCAs was homogenous for the first generation and in the control treatment (ninth and tenth generations, **Figure S4**). However, the measured relative variance for the bumblebee treatment (ninth and tenth generations) among the 1'000 PCA was highly variable on the fourth PCs compared to the first generation or control treatment (**Figure S4**). Moreover, the genomic variance observed based on the 76 SNPs under bumblebee selection (**Figure 5**) was mostly located (not all significantly) in the lower tail of the PC values distribution obtained randomly, confirming that the observed decrease in genomic variance during bumblebee selection is probably not due to an artifact of sub-sampling (**Figure S4**).

Figure S1. Design of the experimental evolution experiment.

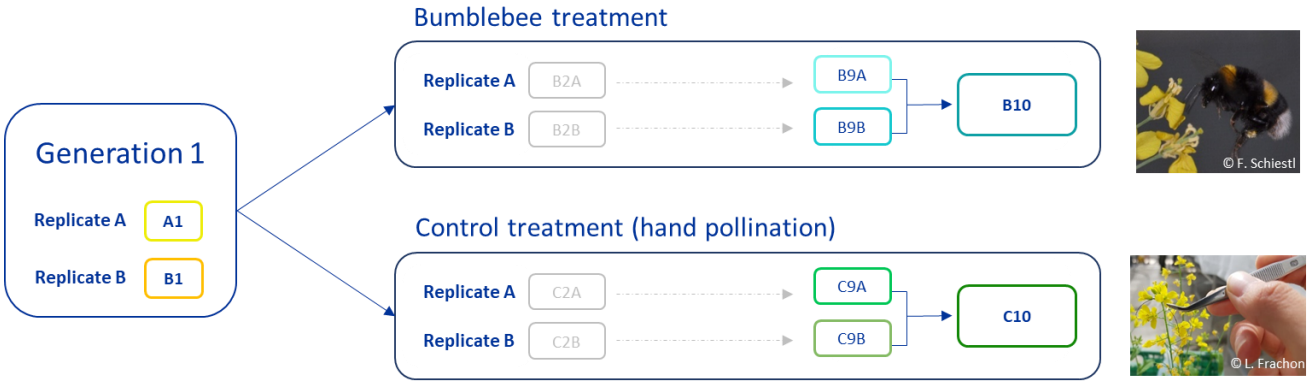


Figure S2. Upper part: histogram of MAF before (5637 SNPs) and after (4713 SNPs) cutting for maf<0.1. Below part: histogram of average read depth (DP) over 256 samples for each 4713 SNPs.

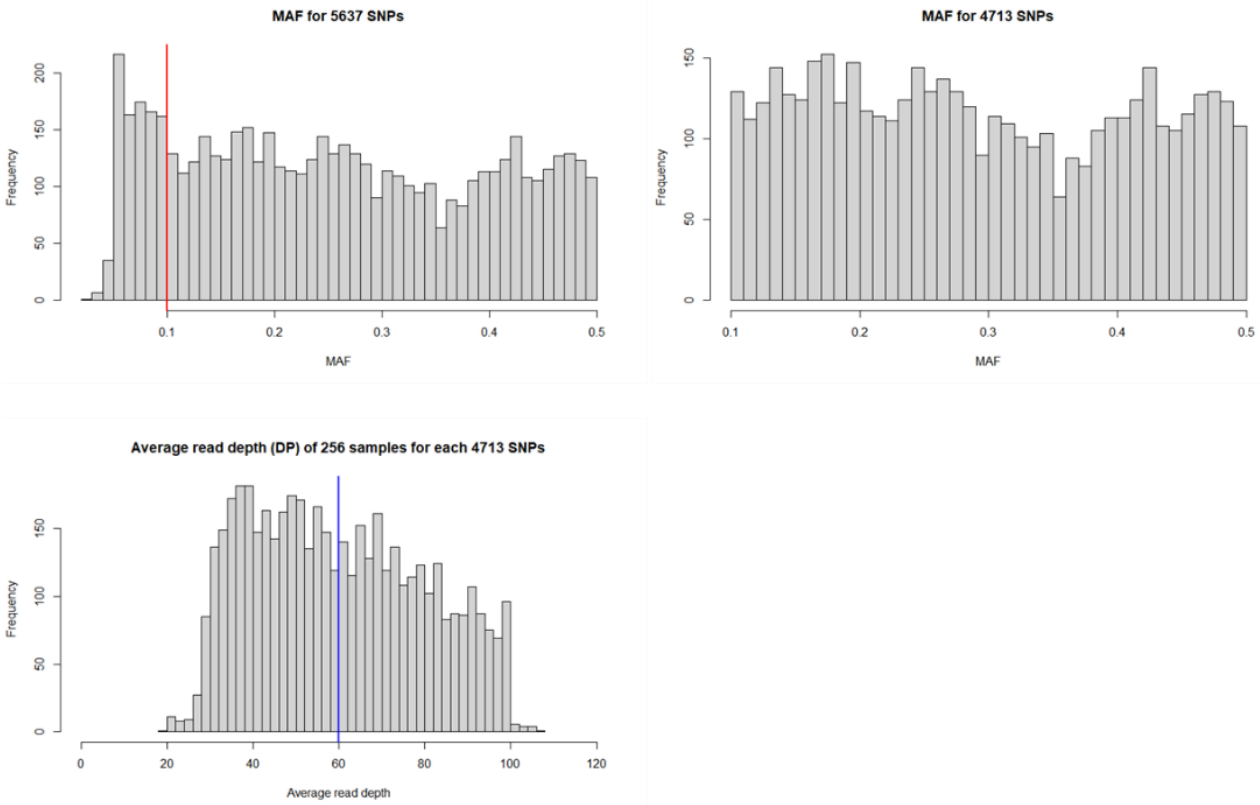


Figure S3. Plot Nucleotide replicates details

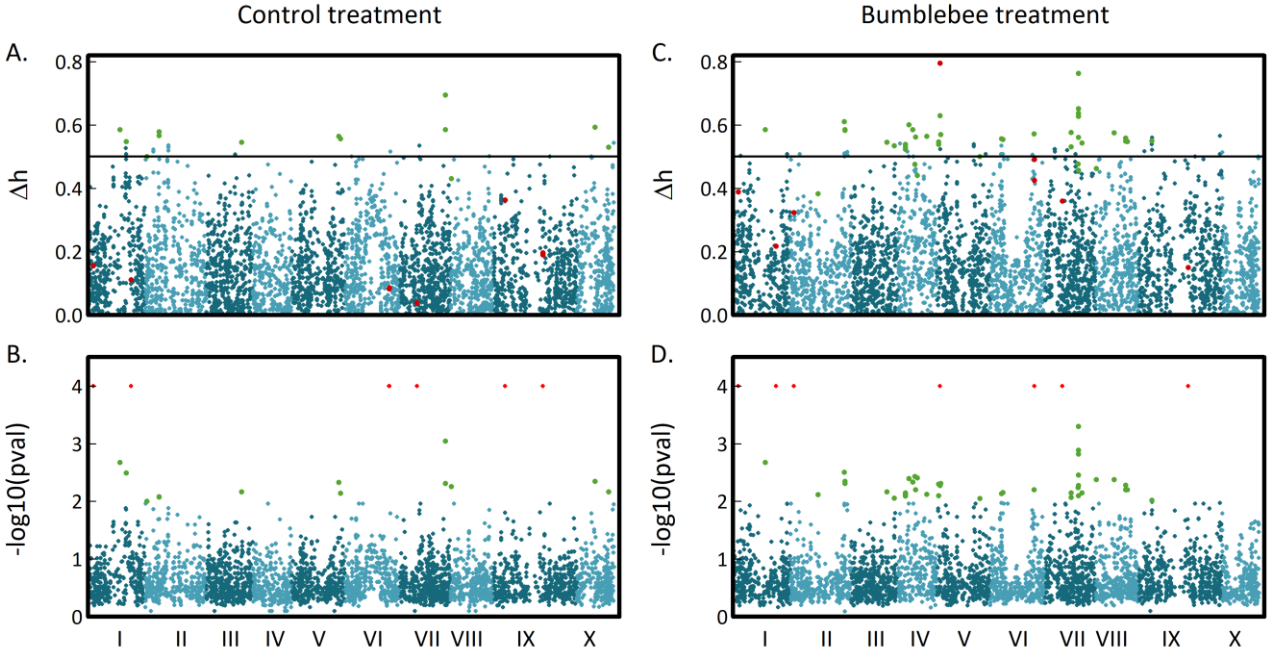


Figure S4. Plot of the PC3 & PC4 performed on 4'713 SNPs and 145 SNPs

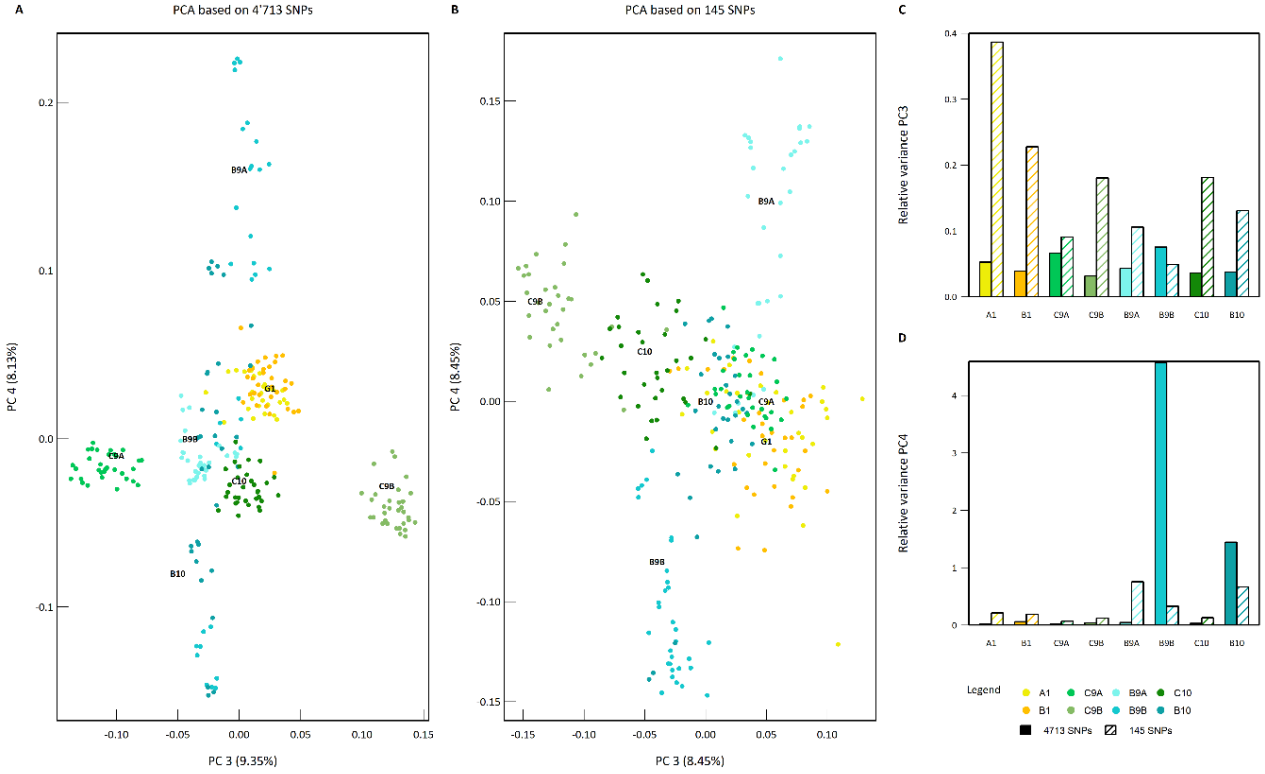


Figure S5. Jitter plots of the variance of the 256 samples on the genomic space from 1'000 PCA performed on 76 random SNPs. The red dots are the variance of these 256 samples on the genomic space from PCA performed on 76 SNPs un bumblebee selection. The significances above the jitter plots indicate whether the variance of the PCs from the 76 SNPs is significantly different from the distribution of the PCs from the 1'000 PCAs.

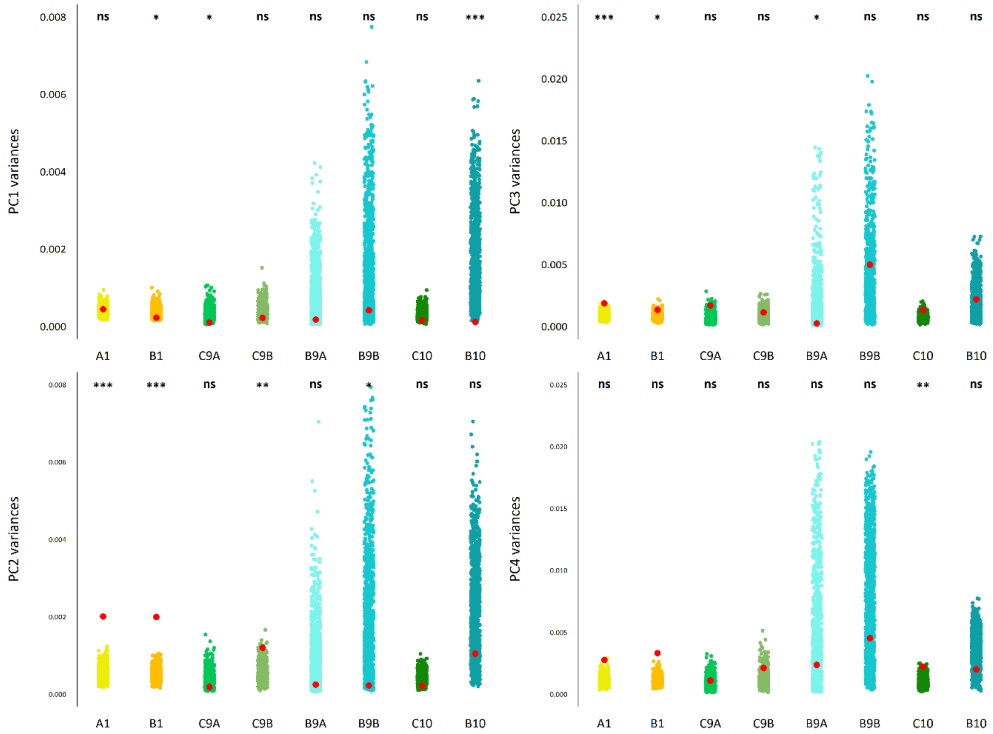


Table S1. Linkage disequilibrium blocks. For each population, the number of LD blocks, the mean (+- sd) of the number of SNPs per LD block, and the length (in kb) of LD blocks.

	Nb. LD blocks	Mean Nb. SNPs / LD block	Median Nb. SNPs / LD block	Mean LD block length (kb)	Median LD block length (kb)
G1	949	3.16	3	13.48	0.13
C9	818	3.85	3	36.04	0.25
B9	791	4.16	4	53.55	4.23
C10	743	3.53	3	24.86	0.15
B10	696	3.87	3	38.94	0.17

Table S2. List of 32 candidate genes.

Gene ID (*B. rapa*)	protein name	Description (Phytozome)	Gene ID (*A. thaliana*)	protein name (*A. thaliana*)	GO term biological Process (*A. thaliana*)
Brara.B03446	LRR_3	Leucine Rich Repeat (LRR_3)	AT5G51630	NA	signal transduction
Brara.B03515	GRF1	GRF1 (growth regulator factor) -INTERACTING FACTOR 1	AT5G28640.3	GRF1, AN3	cell division, leaf development, regulation of gene expression
Brara.B03516	NA	NA	AT5G28610	LOW protein	biological_process
Brara.C03679	PEX3	peroxin-3	AT3G18160	PEX3-1	peroxisome organization, protein import into peroxisome membrane
Brara.C03680	NA	EamA-like transporter family	AT3G18200.2	UMAMIT4	NA
Brara.C04359	PSMA7	20S proteasome subunit alpha 4 (PSMA7)	AT3G51260	PAD1	proteasomal ubiquitin-independent protein catabolic process, proteasome-mediated ubiquitin-dependent protein catabolic process, ubiquitin-dependent protein catabolic process
Brara.C04360	RIOK2	RIO kinase 2 (RIOK2)	AT3G51270	NA	maturation of SSU-rRNA, protein phosphorylation
Brara.D00527	NA	NA	AT3G53490	NA	biological_process
Brara.D00530	LOG3	CYTOKININ RIBOSIDE 5'-MONOPHOSPHATE PHOSPHORIBOHYDROLASE LOG3-RELATED	AT3G53450	LOG4	cytokinin biosynthetic process
Brara.D00531	RPL12	large subunit ribosomal protein L12e (RP-L12e, RPL12)	AT3G53430	NA	translation
Brara.D00706	Zein-binding	Zein-binding	AT4G13630	MYOB13	biological_process
Brara.D00707	NA	ETHYLENE-RESPONSIVE TRANSCRIPTION FACTOR CRF5-RELATED	AT4G13620.1	NA	regulation of transcription, DNA-templated
Brara.D00848	PRR_2	PRR repeat family	AT5G37570	NA	RNA modification
Brara.D00997	GRX	Glutaredoxin (GRX) family	AT3G28850	NA	NA
Brara.D01757	NA	NAD(P)-BINDING ROSSMANN-FOLD SUPERFAMILY PROTEIN-RELATED	AT2G29320.3	NAD(P)-binding R	NA
Brara.E00146	LEA-HRGP	LATE EMBRYOGENESIS ABUNDANT HYDROXYPROLINE-RICH GLYCOPROTEIN	AT2G46300	NA	biological_process
Brara.E00147	DUF1218	Protein of unknown function (DUF1218)	AT1G05291.1	DUF1218	biological_process
Brara.E00344	POP4	ribonuclease P protein subunit POP4 (POP4, RPP29)	AT2G43190	POP4	rRNA processing
Brara.E00345	PTI1	PTI1-LIKE TYROSINE-PROTEIN KINASE 1-RELATED	AT2G43230	CARK6	response to abscisic acid
Brara.F01068	ABCG38	ABC TRANSPORTER G FAMILY MEMBER 38 - Monosaccharide-transporting ATPase	AT1G15520	ABCG40	abscisic acid transport, abscisic acid-activated signaling pathway, cellular response to water deprivation, defense response to oomycetes, import across plasma membrane, import into cell, intercellular transport, lead ion transport, negative regulation of post-embryonic development, response to abscisic acid, response to cold, response to ethylene, response to heat, response to jasmonic acid, response to ozone, response to salicylic acid, response to water deprivation, stomatal closure, terpenoid transport, transmembrane transport
Brara.F01236	PP2C16	PROTEIN PHOSPHATASE 2C 16-RELATED	AT1G17550	ATHAB2	NA
Brara.F01237	NA	large subunit ribosomal protein L14 (RP-L14, MRPL14, rplN)	AT1G17560	HLL	embryo sac development, integument development, negative regulation of cell death, plant ovule development, response to brassinosteroid, translation
Brara.G01361	LOG1	CYTOKININ RIBOSIDE 5'-MONOPHOSPHATE PHOSPHORIBOHYDROLASE LOG1	AT2G28305	LOG1	cytokinin biosynthetic process
Brara.G01976	GATL6	GALACTURONOSYLTRANSFERASE-LIKE 6-RELATED	AT3G62660	GATL7	pectin biosynthetic process
Brara.G01977	SEC61A	protein transport protein SEC61 subunit alpha	AT1G29310.2	SEC61A	SRP-dependent cotranslational protein targeting to membrane, translocation, posttranslational protein targeting to membrane, translocation
Brara.G02000	ABCG35	ABC TRANSPORTER G FAMILY MEMBER 35-RELATED	AT1G59870	ABCG36	response to heat
Brara.G02343	LRR_3	Leucine Rich Repeat (LRR_3)	AT1G72840	NA	signal transduction
Brara.H01604	PAT8	PROTEIN S-ACYLTRANSFERASE 8	AT4G24630	NA	peptidyl-L-cysteine S-palmitoylation, protein targeting to membrane
Brara.H01605	NA	NA	AT4G24590.2	RING	biological_process
Brara.H01737	NA	NA	AT4G37820	NA	biological_process
Brara.H01738	COX6A	cytochrome c oxidase subunit 6a (COX6A)	AT4G37830.1	COX	mitochondrial electron transport, cytochrome c to oxygen
Brara.I01179	SLAH1	S-TYPE ANION CHANNEL SLAH1-RELATED	AT1G62280	SLAC1	cellular ion homeostasis, chloride transport, positive regulation of anion channel activity, response to salt stress, response to water deprivation

Figure S3. Genomic variance among populations. Variance of the 256 samples on the genomic space from PCA performed on the 4'713 SNPs (in e-5).

	A1	B1	G1	C9A	C9B	C9	B9A	B9B	B9	C10	B10
PC1	7.7	3.4	5.5	3.8	14.5	115.9	59.4	17.1	357.2	5.9	40.1
PC2	4.4	6.4	5.5	6.3	7.9	7.5	51.4	275.4	1050.3	7.3	363.4
PC3	20.6	15.4	19.1	26	12.3	1406.8	17	29.5	42.2	14.2	14.6
PC4	9.2	23.2	16.4	5.6	16.2	25.3	16.6	1795.1	995.3	10.7	565.4

Table S4. Genomic variance among populations. Variance of the 256 samples on the genomic space from PCA performed on the 76 SNPs (in e-5).

	A1	B1	G1	C9A	C9B	C9	B9A	B9B	B9	C10	B10
PC1	44.2	22.1	32.7	9.3	21.6	36.6	17.5	41.7	41.4	11.4	15.6
PC2	200.4	199.5	198.5	18.3	119.9	1127.6	24.4	22.2	23.2	104.5	21.6
PC3	185.3	132.3	157	168.8	112.2	240.6	22.1	498.2	299.4	217.4	132.1
PC4	276.7	330.2	299.2	109.7	210.2	227	235.6	450.9	690.7	198.6	221.2