

RESEARCH

Open Access



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

Léa Frachon^{1*} and Florian P. Schiestl¹

Abstract

Background Insect pollinators shape rapid phenotypic evolution of traits related to floral attractiveness and plant reproductive success. However, the underlying genomic changes remain largely unknown despite their importance in predicting adaptive responses to natural or to artificial selection. Based on a nine-generation experimental evolution study with fast cycling *Brassica rapa* plants adapting to bumblebees, we investigate the genomic evolution associated with the previously observed parallel phenotypic evolution. In this current evolve and resequencing (E&R) study, we conduct a genomic scan of the allele frequency changes along the genome in bumblebee-pollinated and hand-pollinated plants and perform a genomic principal component analysis (PCA).

Results We highlight rapid genomic evolution associated with the observed phenotypic evolution mediated by bumblebees. Controlling for genetic drift, we observe significant changes in allelic frequencies at multiple loci. However, this pattern differs according to the replicate of bumblebee-pollinated plants, suggesting putative non-parallel genomic evolution. Finally, our study underlines an increase in genomic variance implying the putative involvement of multiple loci in short-term pollinator adaptation.

Conclusions Overall, our study enhances our understanding of the complex interactions between pollinator and plants, providing a stepping stone towards unravelling the genetic basis of plant genomic adaptation to biotic factors in the environment.

Keywords Fast cycling *Brassica rapa*, Experimental evolution, Evolve and resequencing study, Non-parallel genomic evolution, Rapid genomic evolution, Bumblebees' selection

*Correspondence:

Léa Frachon
lea.frachon@inrae.fr

¹Department of Systematic and Evolutionary Botany, University of Zürich,
Zürich, Switzerland



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

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 pollinator insects has been detected in different geographical regions and insect families [2–4], the understanding of the adaptive potential of plants to such changes remains in its infancy. Plant adaptation to pollinators typically involves traits associated with flower attractiveness such as (1) flower morphology [5–7], flower colour [8, 9], flower scent [10–12], and (2) traits associated with mating system like herkogamy [13, 14] or selfing [15, 16]. While most studies assessed the result of long-term evolutionary adaptation to pollinators, tracking the adaptive processes across generations remains scarce. Both the 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 [12] 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 these evolved traits have also been shown to carry substantial heritability in this study system [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, understanding the genetic architecture involved in plant response to pollinators is essential to understand their adaptive potential in changing environments [20–22]. 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, 25–27], nectar [28, 29] and volatiles [30–34]. 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 environmental changes, genetic correlation among traits may allow the synchronous response of different phenotypic traits to varying patterns of selection [19, 36, 37]. While essential to predict the adaptive potential of plants to pollinators and enable breeding of crop that are more attractive to pollinators, we are still in the beginning of understanding the genetic basis involved in the adaptive response to pollinators. By combining experiment evolution and next generation sequencing (NGS), evolve and resequencing (E&R) studies have proven their ability to unravel the genomic basis involved in rapid evolution [38–40].

Here, based on previous experimental evolution performed by Gervasi and Schiestl [12] with outcrossing fast-cycling *Brassica rapa* plants, we tracked the genomic changes involved in the adaptive response of plants to bumblebee selection compared to hand pollinated control plants in an E&R study (Fig. 1). Gervasi and Schiestl [12] initiated their experimental evolution

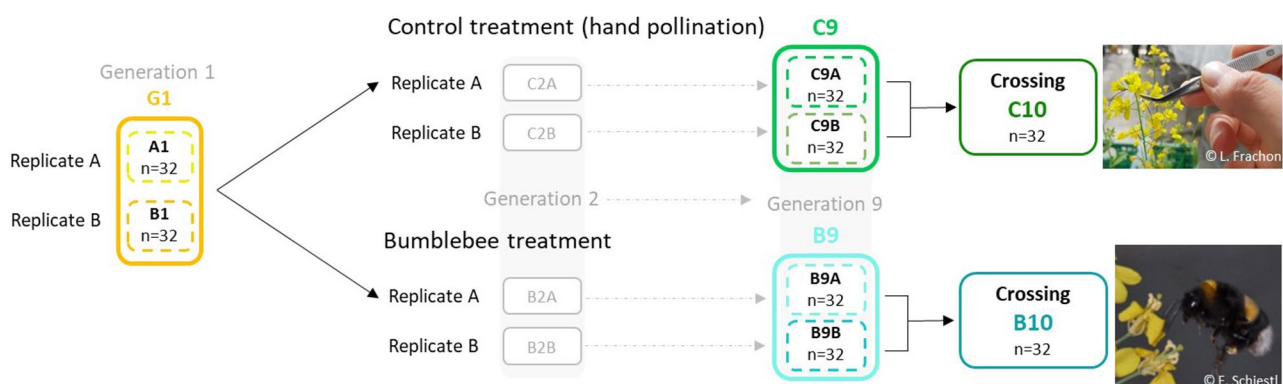


Fig. 1 Experimental evolution design. Partial experimental evolution design from Gervasi and Schiestl [12] from the first generation (G1) to the ninth generation in control (C9) and bumblebee treatment (B9). The tenth generation was obtained by inter-replicate crossing within treatment (C10 and B10). Population of *Brassica rapa* fast cycling used for seedling sequencing are coloured. The abbreviation and the colours assigned to the populations are unchanged along the manuscript. The number of seedlings (n) sequenced per population is indicated. Images: courtesy of F.P. Schiestl and L. Frachon

with 108 full sib seed families in the greenhouse. Over nine generations of selection, they grew plants in three isolated replicates (36 plants per replicate). The selection was mediated by three pollination treatments during the experiment: bumblebee (*Bombus terrestris*) pollination, hoverfly (*Episyrphus balteatus*) pollination, and hand pollination. Overall, they observed parallel phenotypic evolution with bumblebee-pollinated plants evolving taller, more fragrant flowers, and being more attractive to bumblebees. In our E&R study, we re-sequenced plants from two replicates in bumblebee and hand-pollination treatments. We performed a genome scan of allele frequency changes and a genomic principal component analysis to observe putative genomic evolution shaped by bumblebees.

Results

Allele frequency changes during experimental evolution

The study of allele frequency dynamics provides valuable information on the underlying genetic processes behind phenotypic changes observed. To disentangle the effect of genetic drift, which is particularly pronounced in small populations, from natural selection driven by bumblebee pollination, we compared the observed allele frequency changes with those obtained by simulating genetic drift. For our purpose, we defined as significant allele

frequency changes (colored dots in Fig. 2) the observed changes near or outside the ranges of 10'000 drift simulations assessed using $\text{fdr} < 0.05$ (upper and lower grey lines in Fig. 2). We observed allele frequency changes (Δh) between ancestor and descendants in both bumblebee and control treatments (Fig. 2). For instance, we found 32 SNPs with strong ($\Delta h > 0.5$) and significant ($\text{fdr} < 0.05$) changes between the first and the last generation in the replicate A of the bumblebee treatment. In contrast, in the replicate B of the same treatment (bumblebee), we observed only 4 SNPs with significant allele frequency changes between the first and the last generation. In the control (hand-pollinated) treatment, there were 3 SNPs with significant allele frequency changes in the replicate B and zero SNPs with allele frequency changes in the replicate A between the first and the last generation. Subsequently, we compared the number of SNPs exhibiting significant allele frequency changes between bumblebee and control treatments. In replicate A, we found a 42-fold increase in the number of SNPs displaying significant allele frequency changes ($\text{fdr} < 0.05$) with the bumblebee treatment compared to the control treatment, regardless of its strength i.e., considering all the value of Δh (4 SNPs with significant changes in the control treatment, and 171 SNPs in the bumblebee treatment, Fig. 2). However, no difference in the number of SNPs involved

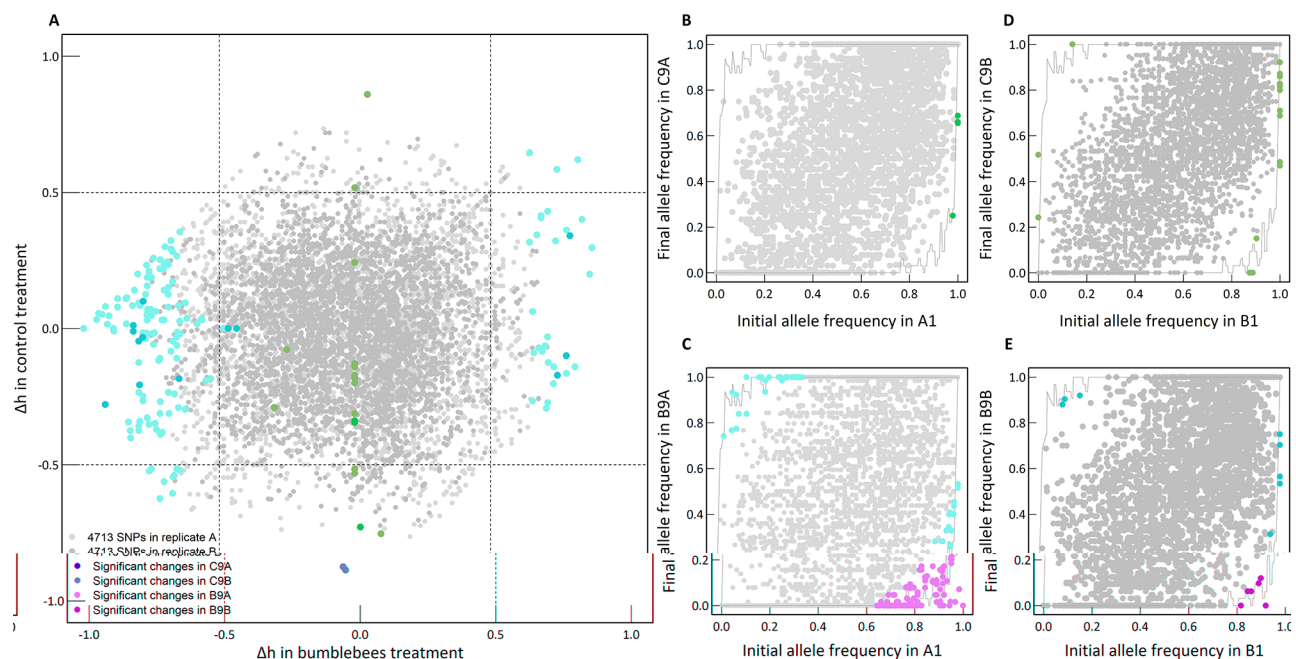


Fig. 2 Allele frequency changes during experimental evolution. **(A)** Comparison of the allele frequency changes (Δh) between the bumblebee treatment (x-axis) and the control treatment (y-axis). The grey dots represent the 4713 SNPs. Replicate A is in light grey, replicate B in darker grey. The significant changes are highlighted in blue and green (see the legend in the figure for details). Comparison of initial (first generation) and final (ninth generation) allele frequencies in the control treatment for both replicate A **(B)** and replicate B **(D)**, and in the bumblebee treatment for both replicate A **(C)** and replicate B **(E)**. 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 coloured dots represent significant changes in our study with $\text{fdr} < 0.05$

in significant allele frequency changes between bumblebee and control treatment were observed in replicate B (21 SNPs with significant changes in the control, and 19 SNPs in the bumblebee treatment considering $\text{fdr} < 0.05$, considering all value of Δh , Fig. 2). Overall, our study identified several SNPs exhibiting potential adaptive changes in response to bumblebees, particularly in the replicate A, that differed from the observed changes in the control treatment.

Changes in genetic linkage structure

During selective processes, changes in beneficial allele frequency are often linked to changes in surrounding alleles. This non-random association among alleles is called linkage disequilibrium (LD) and tends to be inherited together with the beneficial allele. In this study, we observed an increase of linkage disequilibrium both in the bumblebee treatment, and in the control treatment. In both treatments, we observed a slower decay of the median linkage disequilibrium in the ninth generation compared to the first generation (Fig. S3A). Within the bumblebee treatment, we observed a slower decay of LD in replicate A than in replicate B associated with the observed stronger allele frequency changes (Fig. S3A). The decays of LD in the next generation with inter-replicate crossing (generation 10th) were slower in both treatments compared to the ninth generation (Fig. S3A). Moreover, linkage disequilibrium is interconnected with haplotype blocks, which are defined as specific sets of alleles that tend to be inherited together. We observed an increase of the haplotype block length associated with a decrease of their number over nine generations (Fig. S3BC, Table S1, from 786 to 780 haplotype blocks for replicate A and B respectively in the first generation, to 563 and 648 haplotype blocks for replicate A and B respectively in the ninth generation of the control treatment, and 478 and 600 for replicate A and B respectively in the ninth generation of the bumblebee treatment). We observed 743 and 696 haplotype blocks in control and bumblebee treatments respectively in the generation with inter-replicate crossing (Fig. S3BC, Table S1). This result suggested that the sets of alleles inherited together were larger due to selective processes. As the size of the haplotype blocks increased, the total number of haplotype blocks decreased simultaneously, probably due to merging of adjacent blocks. Overall, an increase in LD and a decrease in the number of haplotype blocks indicated changes in the genetic architecture. It is worth noting that the moderate density of genetic markers (18 SNPs per Mb in mean among the 10 chromosomes, Fig. S1) suggested that the estimate of LD and haplotype length may be slightly underestimated, although this had no effect on the trends observed.

Overall, the SNPs involved in the evolutive responses were independent between replicates, and between treatments (Fig. 3). For instance, among the 171 SNPs with significant allele frequency changes in replicate A in bumblebee treatment, 164 SNPs were unique to this population, whereas five SNPs were shared with the control treatment in replicate B, and three SNPs were shared with the bumblebee treatment in replicate B (Fig. 3).

Identity of candidate genes underlying genomic evolution in bumblebee treatment

After retrieving the annotated genes around 4 kb (2 kb upstream, 2 kb downstream) for 187 SNPs with significant allele frequency changes in the bumblebee treatment ($\text{fdr} < 0.05$), we obtained a list of 171 candidate genes (Dataset 1). Briefly, we found genes with biological functions that may be involved in the emission of volatiles (ABC transporter G), in the regulation of flowering time (B3), or in shoot architecture (UCH). In addition, some genes are involved in the response to different biotic and abiotic stresses (METACASPASE-2, PUB, Pum, PAT8, BON). Finally, some candidate genes do not have a known biological function. However, due to a putative underestimate of LD and haplotype length (see previous section), we may not have been able to correctly identify all the candidate genes involved in this evolutionary process.

Genome-wide population structure

To assess the relatedness among individuals across generation and treatments, we performed a genomic principal component analysis (PCA) on the full set of SNPs (4713 SNPs). We observed a structuration of our samples determined by populations i.e. generations, treatments, and replicates (Fig. 4). Along the first six principal components (PCs) axis, explaining 56.07% of the total genomic variance (Fig. 4A), all individuals from the two replicates of generation one were grouped together, and individuals from the ninth generations were clearly separated from the first generation (Fig. 4). In this genomic space, individuals from the inter-replicate crossing generation fell between those from the scatterplot formed by replicates A and B of generation nine for both treatments (Fig. 4). Interestingly, in the genomic space created by the first six principal components explaining the most genetic variance, the individuals resulting from selection by bumblebees were always more scattered than those from the control treatment or the first generation. Indeed, the polygon area significantly increased between the first and the last generation in bumblebee treatment in the genomic space created by the first six principal components (Wilcoxon test: $V=3$, $p\text{-value}=9.31\text{e-}09$, Table S2). Conversely, no significant increase was observed for the control treatment in the same genomic space (Wilcoxon

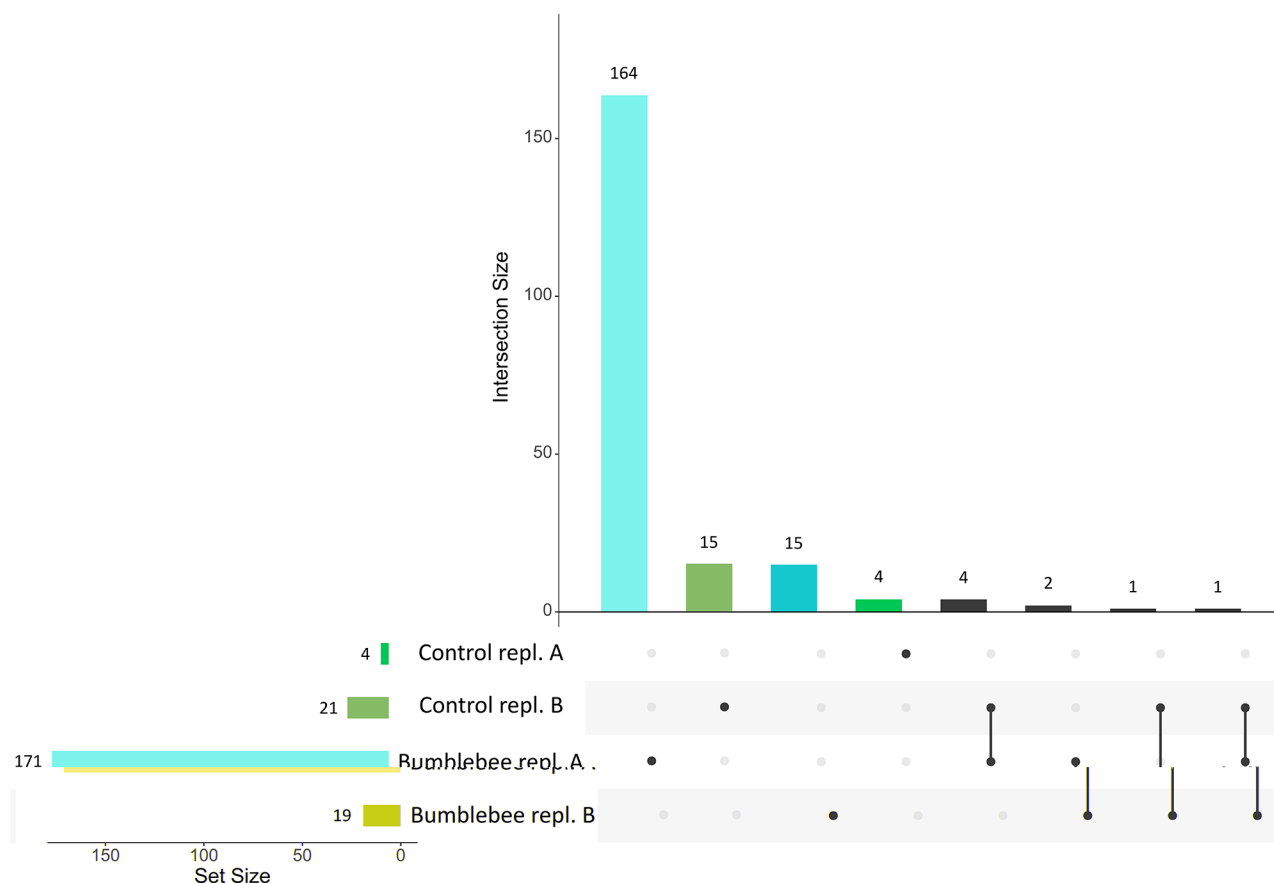


Fig. 3 Intersection of genomic variants under selection between treatment and replication. The UpSet plot illustrates the genomic variants involved in the evolution of *Brassica rapa* in control and bumblebee treatment for both replicate A and replicate B. On the left (Set Size), the number of significant genomic variants involved in evolutionary response of *Brassica rapa* in different populations is indicated. The dots indicate that the genomic variants are involved in the evolutionary response only in a particular population, while the vertical lines indicate that the genomic variants are involved in the evolutionary processes in several populations. The number of genomic variants involved in unique population or in different populations is indicated in the upper part of the figure

test: $V=26$, $p\text{-value}=0.53$, Table S2). Similar patterns were observed for the inter-replicate crossing generation, revealing a significant increase in the bumblebee treatment ($V=0$, $p\text{-value}=1.86\text{e-}09$), and no increase in the control treatment ($V=255$, $p\text{-value}=0.66$), as compared to the first generation.

Discussion

Understanding pollinator-mediated evolutionary genomics of flowering plants remains an important challenge in biodiversity conservation and crop improvement. Here, we screened for the genomic consequences of biotic selection in an E&R study by sequencing genome-wide SNP markers in *Brassica rapa* plant individuals before and after nine generations of selection by bumblebees, and under random hand-pollination. The previous experimental evolution has shown, at the phenotypic level, that this primarily outcrossing plant rapidly adapts to specific pollinators [12]. For instance, Gervasi and Schiestl [12] observed taller plants with an increase in

total scent emission per flower in bumblebee-pollinated plants. The pattern of phenotypic evolution was similar between replicates from the same pollination treatment. In this E&R study, we documented the putative signature of directional selection driven by bumblebee pollinators with significant allele frequency changes at several loci. Moreover, we have shown different pattern of genomic evolution between replicates, and an increase of genomic differentiation among individuals.

In agreement with the demonstrated phenotypic evolution in the fast-cycling *Brassica rapa* experimental system [12], we have shown genomic evolution across nine generations. The here documented changes in allele frequencies, the increase of linkage disequilibrium and the decrease of number of haplotype blocks, underscore the importance of pollinators in shaping plant rapid genomic evolution. The limited number of individuals (36 per replicate and per treatment) and replicates (two replicates) could impact the evolutionary processes and the identification of the underlying genomic bases. Indeed, small

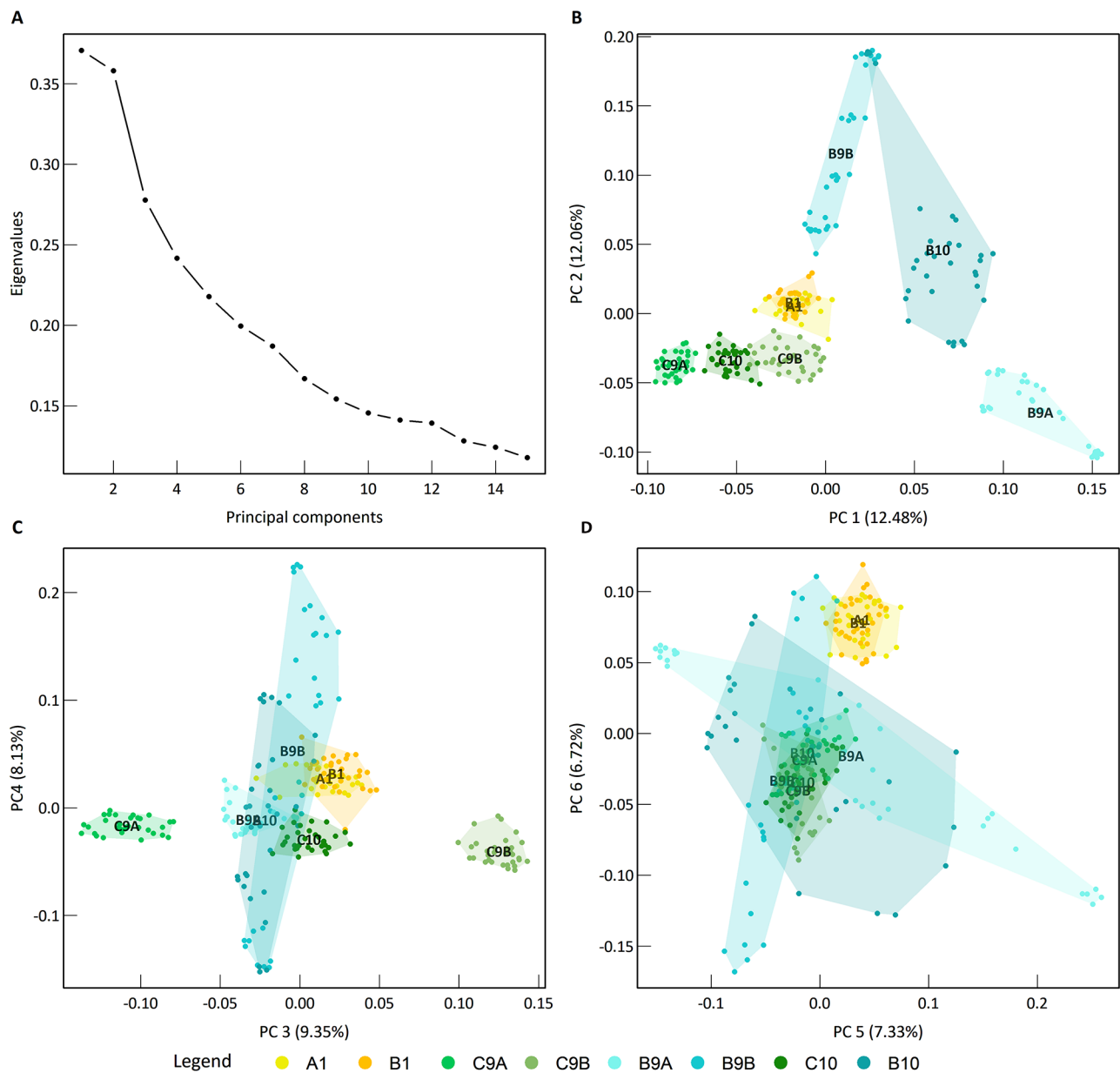


Fig. 4 Genome-wide population structure. **(A)** Eigenvalues of the 15 principal components of the genomic PCA based on 4713 SNPs. **(B–D)** Position of the 256 individuals in the genomic space from the principal component analysis (PCA) performed on their genotypes (GT). The label of the population is shown on their centroid. Polygons linking outer points of the scatter plot are displayed for all populations. The legend colours are indicated in the bottom of the figure

population size can lead to an increase in the effect of genetic drift, and a low number of replicates to a poor estimate of loci under selection [41]. However, at the phenotypic level, the observed differences among the three replicates were smaller than those observed among treatments, indicating that pollinator selection played a larger role than random drift [12]. Moreover, in our study, we accounted for genetic drift bias and identified changes in allele frequency that were greater than those expected from 10'000 genetic drift simulations. While we

cannot completely exclude some effects of genetic drift, this alone cannot explain the patterns observed. Moreover, some significant candidate genes identified in our E&R study have biological functions in line with the phenotypic evolution of traits observed by Gervasi and Schiestl [12]. For instance, while the previous experimental evolution study highlighted an increase in scent emission and plant height driven by bumblebee pollination, we identified candidate genes who, or at least their family, are known to be involved in emission of volatile organic

compounds, like ABC transporter G family (ABCG35, ABCG38, ABCG1, ABCB19, [42, 43]), in shoot architecture like Ubiquitin C-terminal hydrolase (UCH [44]), or in growth and development like AUXIN-RESPONSIVE PROTEIN-RELATED, or TREHALOSE-PHOSPHATE PHOSPHATASE E-RELATED [45]. The U-box protein is also an interesting candidate gene due to its involvement in different developmental and physiological processes such as self-incompatibility, defence, or abiotic stress response [46]. Moreover, we found significant allele frequency changes associated with flowering time like a B3 transcription regulator [47] or UCH [48]. While Gervasi & Schiestl [12] did not observe any difference in flowering time between bumblebee and control treatments in the last generation, flowering time is an important phenological trait involved in plant-pollinator matching [49, 50]. It is therefore possible that pollinators induce flower phenological shift, which was not detected in the phenotypic data, but picked up in the genomic changes as changes influencing flowering time. Overall, our study highlighted the potential involvement of multiple loci in rapid adaptation to bumblebees, which agrees with studies highlighting such a genetic architecture underlying floral evolution [26, 37, 51–53]. However, further analysis using higher quality of sequencing and the characterization of associated phenotypic and phenological traits (allowing a genome-wide association approach) and functional validation of the genes are required to draw more solid conclusions about the genetic basis of evolutionary changes induced by bumblebee selection.

In addition, we found that the extent of genomic changes observed during the evolutionary processes was different between replicates suggesting non-parallel genomic evolution [54]. While Gervasi and Schiestl [12] have shown convergent (or parallel) phenotypic evolution for the bumblebee treatment, our results indicated that only one replicate exhibited several loci with significant changes in allele frequency (i.e., outside the range of drift simulations assessed by $\text{fdr} < 0.05$). Such a convergent phenotypic evolution associated with non-parallel genomic evolution has been observed in artificial selection of crops [55, 56] and in natural populations. For instance, a recent study highlighted different genomic regions underpinning the evolutionary convergence of herbicide resistance in blackgrass among different populations [57]. Likewise, in natural populations of *Senecio latus*, similar phenotypic variation was reported to be regulated by variation in genomic space across populations in dune-headland coastal habitats [58]. However, the low number of replicates in our study and the significant changes in allele frequencies observed in only one replicate do not provide a clear distinction between potential non-parallel evolution and a “two-speed” parallel genomic evolution (i.e., the second replicate will

follow the same evolutionary path in the upcoming generations). Longer experimental evolution, with more replicates and individuals are needed to clearly understand the evolutionary pattern mediated by biotic factors such as bumblebees.

Finally, we observed an increase of genomic variance, within the genomic space created by the main six differentiation axes, during experimental evolution mediated by bumblebees [12]. This increase in overall genomic variance was observed among individuals in both bumblebee-pollinated replicates (B9A and B9B), as well as in the inter-replicate crossing generation (B10). This pattern might be explained by the polygenic model with a weaker selection acting on multiple standing variants (soft sweep) and by multiple loci underlying individual phenotypic trait evolutionary changes. An increased number of studies demonstrate the importance of polygenic adaptation [56, 60–62] related to the infinitesimal model (reviewed in Barton et al. [63]), where local adaptation is driven by small allele frequency changes in multiple loci. Interestingly, highly polygenic architecture involved in phenotypic evolution could contribute to the maintenance of standing genetic variation, as recently demonstrated in long-term artificial selection on chicken weight [64]. However, deepened analyses are needed using newly developed models to validate the involvement of polygenic architecture in rapid phenotypic evolution [65–67]. Multiple genes underlying phenotypic variation are widely emphasized in plants with the advances of GWAs [59, 62], however the interplay of evolutionary forces on these genes is still poorly understood and deserves further studies.

Conclusion

We revealed important genomic changes on multiple loci associated with non-parallel phenotypic evolution resulting from bumblebee selection in only nine generations. Our study is a first step into the understanding of the complex genomic mechanisms involved in rapid evolutionary adaptation to biotic factors, and we advocate further analyses to understand (1) the genetic architecture underlying phenotypic and phenological 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 pollinator decline.

Methods

Plant material and experimental design

Brassica rapa (Brassicaceae) is an outcrossing plant with genetic self-incompatibility, pollinated by diverse insects such as bumblebees, flies or butterflies [68]. Our study used rapid-cycling *Brassica rapa* plants (Wisconsin Fast Plants™, purchased from Carolina Biological Supply Company, Burlington, USA), selected for its short life cycle of approximately two months from seed to seed in our greenhouse conditions. We used seeds produced by the study of Gervasi and Schiestl [12], performing experimental evolution with bumblebees and control hand pollination. Briefly, starting from 108 full sib seed families, the pollination was carried out over nine generations either by bumblebees (*Bombus terrestris*), hoverflies (*Episyrphus balteatus*) or by random hand cross-pollination (control treatment). This experiment was conducted with 3 isolated replicates (one replicate includes 36 plants) for each treatment. A representative subset of seeds from all pollinated flowers was used for the next generation; the contribution of seeds to the next generation being calculated as 36 divided by the sum of seeds per replicate over all individual seeds. In this study, we focused on bumblebee and hand-pollination treatment, using two most distinct replicates i.e., with the most different phenotypic responses (replicate A and B). We used the offspring of the starting populations and the ninth generation for both treatments. One seed per individual plant were used. In total, we used 32 plants from the starting generation (generation 1) in replicate A (called A1), and 32 plants from replicate B (called B1), 32 plants from the ninth generation selected by bumblebees (bumblebee treatment) in replicate A (called B9A) and 32 plants in replicate B (called B9B) and; 32 plants from the ninth generation of control hand pollination plants (control treatment) in replicate A (called C9A) and replicate B (C9B; Fig. 1). 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). This manual crossing is commonly used for reducing the effect of potential inbreed depression on trait changes. Pollen donors and receivers were randomly assigned. 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. The

study conforms to the institutional, national, and international regulations.

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 h, 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 *M*sII, insert size mean range ~215 bp) 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 the 10 chromosomes i.e. by discarded unassigned scaffolds from the chromosomes using vcftools [69], by 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 by retaining only bi-allelic SNPs with a minimum average Phred quality score of 15 (function --minGQ 15, Fig. S2DE). We removed SNPs with a mean depth value among individuals less or equal at 100 (function --max-meanDP 100, Fig. S2BC), and discarded SNPs with a minor allele frequency (MAF) lower than 0.1 (function --maf 0.1, Fig. S2A). The final dataset contained 4'713 SNPs in ~283 Mb genome size (<https://phytozome.jgi.doe.gov/pz/portal.html>).

Allele frequency changes and genetic drift simulation

The allele frequencies of the reference allele for the 4'713 SNPs were estimated within each 8 populations using vcftools (function --freq, [69]). 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”, [70]) 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 and considering an equal fitness for each individual. Selection was based on the relative production of each individual within the population [12]. To be conservative and because the effective size varies between generations and treatments, we have chosen a very low value for the number of individuals contributing to the next generation. For each SNP along the genome, we assigned their initial allele frequency value (AF_{initial}) to the range estimated by the final allele frequency simulations (10'000 values). Finally, the observed final allele frequency (AF_{final}) was compared to 10'000 simulated allele frequency values ($AF_{\text{simulated}}$) to estimate a p-value for each SNP using the following equations:

- (1) For a decrease of reference allelic frequency
i.e. $(AF_{\text{initial}} - AF_{\text{final}}) > 0$, $pvalue = (\text{number of simulations with } AF_{\text{simulated}} \geq AF_{\text{final}}) / 10'000$.
- (2) For an increase of reference allelic frequency
i.e. $(AF_{\text{initial}} - AF_{\text{final}}) < 0$, $pvalue = (\text{number of simulations with } AF_{\text{simulated}} \leq AF_{\text{final}}) / 10'000$.
- (3) For $(AF_{\text{initial}} - AF_{\text{final}}) = 0$, $pvalue = 1$.

With $AF_{\text{simulated}}$ = simulated final allele frequency, AF_{initial} = observed initial allele frequency in the first generation (reference allele), and AF_{final} = observed final allele frequency in the ninth generation (reference allele). The p-values were controlled for the False Discovery Rate (fdr) of 5% using Benjamini-Hochberg method implemented in R environment [71].

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

$$\Delta h = AF_{\text{final}} - AF_{\text{initial}} \quad (1)$$

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

Linkage disequilibrium (LD) and haplotype block structure evolution

During the selective process, an increase of the linkage disequilibrium i.e., the non-random association of alleles between loci, is expected in genomic regions strongly under selection. First, we calculated pairwise linkage disequilibrium (LD) among all set of SNPs measured by the correlation coefficient between SNP pair r^2 within each chromosomes using VCFtools (function `-geno-r2`) for each 8 populations containing 32 individuals. The

associated median LD was then estimated and plotted. Finally, to understand whether changes in median LD are due to random allelic associations along the genome, or aggregated in genomic regions under selection, we calculated the haplotype blocks in each population using plink1.9 with the following parameters: `--blocks no-pheno-req -maf 0.07 -blocks-max-kb 200`. This method estimates the length of these blocks with “strong LD” considering the allelic association D' metrics between 0.7 and 0.98 according to Gabriel et al. [72].

Candidate genes

We identified 171 candidate genes associated with 187 SNPs with significant allele frequency changes ($fdr < 0.05$) during bumblebee selection. Because the median linkage disequilibrium in bumblebee treatment is ~ 4 kb, we retrieved the annotated genes around 4 kb (2 kb upstream and 2 kb downstream) for 187 SNPs and extracted the gene description using phytozome.jgi.doe.gov.

Genome-wide variance changes

To obtain a broad picture of genetic variation among individuals within each population, we calculated genomic variance. While commonly used nucleotide diversity, such as Π , specifically quantifies the average number of nucleotide differences per site between two DNA sequences, genomic variance refers to the total genetic variation considering a genome-wide set of SNPs. The genomic variance among individuals within each population was estimated using principal component analysis (PCA) on scaled and centered genotype data (pcadapt package in R environment, function `pcadapt`, [73]). In order to unravel the changes in genomic variance over nine generations, we performed a PCA analysis on the total number of SNPs (4'713 SNPs). To compare genomic variance among populations, we estimated the area of the polygon formed by scatter plots based on the 15 pairs of first six axis within each population using R package tidyverse and splanx [74, 75]. We used the Wilcoxon signed-rank test to compare polygon areas between the first and last generations within the bumblebee treatment, as well as between the first generation and the control treatment. This analysis was conducted using 15 pairs of first six principal components (PC1-PC2, PC1-PC3, PC1-PC4, PC1-PC5, PC1-PC6, PC2-PC3, PC2-PC4, PC2-PC5, etc.).

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12862-023-02194-y>.

Supplementary Material 1: Dataset1. List of 171 candidate genes with significant allele frequency changes during bumblebee selection

Supplementary Material 2: Table S1. Variation of haplotype blocks

Supplementary Material 3: Table S2. Polygon area formed by linking the pairwise combinations of the six first pairwise principal components in genomic PCA

Supplementary Material 4: Figure S1. Density of genetic markers for each 10 chromosomes. The x-axis represents the 10 *Brassica rapa* chromosomes. The y-axis represents the density of genetic markers. **Figure S2.** Distribution of filtered genomic data for our final dataset of 4713 SNPs. (a) Distribution of the minor allele frequency, (b) distribution of average read depth (DP) per SNPs, (c) distribution of the average read depth (DP) per individual, (d) distribution of the average genotype quality (GQ), (e) different thresholds of the minimum average GQ as a function of the number of SNPs in the final dataset. The red line indicates the chosen value in our study ($-\text{minGQ} = 15$). **Figure S3.** Linkage disequilibrium and haplotype block structure. (A) Distribution of the median pairwise linkage disequilibrium (r^2) for each population by distance between two SNPs (kb). (B) Number of haplotype block calculated within each population (C) Average length (kb) of haplotype blocks per population (more details **Table S1**)

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 a previous version of the draft. Finally, we thank the University of Zürich for their financial support.

Author contributions

L.F. and F.S. designed the research. L.F. performed the DNA extraction, statistical analyses, and wrote the manuscript. L.F. and F.S. reviewed and edited the manuscript.

Funding

University of Zürich provided funding.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request. The DNA sequences of all samples will be available at National Library for Biotechnology Information (NCBI) database (BioProject PRJNA931117) after acceptance of the manuscript. Reviewers can access to the data using the following link: <https://dataview.ncbi.nlm.nih.gov/object/PRJNA931117?reviewer=hocdpu4je8n22ueldu7potnjqn>.

Declarations

Ethics approval and consent to participate

The study conforms to the institutional, national, and international regulations.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 4 October 2022 / Accepted: 20 December 2023

Published online: 09 January 2024

References

- Van der Niet T, Peakall R, Johnson SD. Pollinator-driven ecological speciation in plants: new evidence and future perspectives. *Ann Bot.* 2014;113:199–211.
- Biesmeijer JC, et al. Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. *Science.* 2006;313:351–4.
- Hallmann C, et al. More than 75% decline over 27 years in total flying insect biomass in protected areas. *PLoS ONE.* 2017;12:e0185809.
- Wagner DL, Grames EM, Forister ML, Berenbaum MR, Stopak D. Insect decline in the Anthropocene: death by a thousand cuts. *Proc Natl Acad Sci U S A.* 2020;118:e2023989118.
- Anderson B, Pauw A, Cole WW, Barrett SCH. Pollination, mating and reproductive fitness in a plant population. *J Evol Biol.* 2016;29:1631–42.
- Wessinger CA, Hileman LC. Accessibility, constraint, and repetition in adaptive floral evolution. *Dev Biol.* 2016;419:175–83.
- Minnaar C, Anderson V, de Jager ML, Karron JD. Plant–pollinator interactions along the pathway to paternity. *Ann Bot.* 2019;123:225–45.
- Bradshaw HD, Schemske DW. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature.* 2003;426:176–8.
- Sobral M, Veiga T, Domínguez P, Guitián JA, Guitián P, Guitián JM. Selective pressures explain differences in flower color among *Gentiana lutea* populations. *PLoS ONE.* 2015;10:e0132522.
- Schiestl FP. On the success of a swindle: pollination by deception in orchids. *Sci Nat.* 2005;92:255–64.
- Xu S, Schlüter PM, Grossniklaus U, Schiestl FP. The genetic basis of pollinator adaptation in a sexually deceptive orchid. *PLoS Genet.* 2012;8:e1002889.
- Gervasi DDL, Schiestl FP. Real-time divergent evolution in plants driven by pollinators. *Nat comm.* 2017;8:14691.
- Moeller DA, Geber MA. Ecological context of the evolution of self-pollination in *Clarkia xantiana*: population size, plant communities, and reproductive assurance. *Evolution.* 2005;59:786–99.
- Luo Y, Widmer A. Herkogamy and its effects on mating patterns in *Arabidopsis thaliana*. *PLoS ONE.* 2013;8:e57902.
- Bodbyl Roles SA, Kelly JK. Rapid evolution caused by pollinator loss in *Mimulus guttatus* Evolution. 2011;65:2541–52.
- Ramos SE, Schiestl FP. Rapid plant evolution driven by the interaction of pollination and herbivory. *Science.* 2019;364:193–6.
- Thomann M, Imbert E, Cheptou PO. Is rapid evolution of reproductive traits in *Adonis annua* consistent with pollinator decline? *Acta Oecol.* 2015;69:161–6.
- Zu P, Blanckenhorn WU, Schiestl FP. Heritability of floral volatiles and pleiotropic responses to artificial selection in *Brassica rapa*. *New Phytol.* 2016;209:1208–19.
- Zu P, Schiestl FP. The effects of becoming taller: direct and pleiotropic effects of artificial selection on plant height in *Brassica rapa*. *Plant J.* 2017;89:1009–19.
- Hansen TF. The evolution of genetic architecture. *Annu Rev Ecol Evol Syst.* 2006;37:123–57.
- Bergelson J, Roux F. Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. *Nat Rev Genet.* 2010;11:867–79.
- Bay RA, et al. Predicting responses to contemporary environmental change using evolutionary response architectures. *Am Nat.* 2017;189:463–73.
- Sicard A, Lenhard M. The selfing syndrome: a model for studying the genetic and evolutionary basis of morphological adaptation in plants. *Ann Bot.* 2011;107:433–1443.
- Bradshaw HD, Wilbert SM, Otto KG, Schemske DW. Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (*Mimulus*). *Nature.* 1995;376:762–5.
- Streisfeld MA, Rausher MD. Genetic changes contributing to the parallel evolution of red floral pigmentation among *Ipomoea* species. *New Phytol.* 2009;183:751–63.
- Yuan YW, Byers KJR, Bradshaw HDJ. The genetic control of flower–pollinator specificity. *Curr Opin Plant Biol.* 2013;16:422–8.
- Yarhamadov T, Robinson S, Hanemian M, Pulver V, Kuhlmeier C. Identification of transcription factors controlling floral morphology in wild *Petunia* species with contrasting pollination syndromes. *Plant J.* 2020;104:289–301.
- Lin W, et al. Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature.* 2015;508:546–9.
- Solhaug EM, 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.* 2019;3:e00120.
- Dudareva N, Negre F, Nagegowda DA, Orlova I. Plant volatiles: recent advances and future perspectives. *Crit Rev Plant Sci.* 2006;25:417–40.
- Klahre U, et al. Pollinator choice in *Petunia* depends on two major genetic loci for floral scent production. *Curr Biol.* 2011;21:730–9.
- Cai J, Zu P, Schiestl FP. The molecular bases of floral scent evolution under artificial selection: insights from a transcriptome analysis in *Brassica rapa*. *Sci Rep.* 2016;6:36966.
- Fattorini R, Glover BJ. Molecular mechanisms of Pollination Biology. *Ann Rev Plant Biol.* 2020;71:487–515.
- Lo S, Fatokun C, Boukar O, Gepts P, Close TJ, Muñoz-Amatriain M. Identification of QTL for perenniality and floral scent in cowpea (*Vigna unguiculata* [L.] Walp). *PLoS ONE.* 2020;15:e0229167.

35. Schiestl FP, Johnson SD. Pollinator-mediated evolution of floral signals. *Trends Ecol Evol.* 2013;28:307–15.
36. Hermann K, Klahre U, Moser M, Sheehan H, Mandel T, Kuhlmeier C. Tight genetic linkage of prezygotic barrier loci creates a multifunctional speciation island in *Petunia*. *Curr Biol.* 2013;23:873–7.
37. Smith SD. Pleiotropy and the evolution of floral integration. *New Phytol.* 2015;209:80–5.
38. Turner TL, Stewart AD, Fields AT, Rice WR, Tarone AM. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLoS Genet.* 2011;7:e1001336.
39. Schlötterer C, Kofler R, Versace E, Tobler R, Franssen SU. Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. *Hered.* 2015;114:431–40.
40. Long A, Liti G, Luptak A, Tenaillon O. Elucidating the molecular architecture of adaptation via evolve and resequence experiments. *Nat Rev Genet.* 2015;16:567–82.
41. Kofler R, Schlötterer C. A guide for the design of evolve and resequencing studies. *Mol Biol Evol.* 2013;31:474–83.
42. Adebisin F, et al. Emission of volatile organic compounds from petunia flowers is facilitated by an ABC transporter. *Science.* 2017;356:1386–8.
43. Hao R, Yang R, Zhang Z, Zhang Y, Chang J, Qiu C. Identification and specific expression patterns in flower organs of ABCG genes related to floral scent from *Prunus mume*. *Sci Hortic.* 2021;288:110218.
44. Yang P, Smalle J, Le7Je S, Yan N, Emborg TJ, Vierstra RD. Ubiquitin C-terminal hydrolases 1 and 2 affect shoot architecture in *Arabidopsis*. *Plant J.* 2007;51:441–57.
45. Figueroa CM, Lunn JE. A tale of two sugars: Trehalose 6-Phosphate and sucrose. *Plant Physiol.* 2016;172:7–27.
46. Wang C, Duan W, Ramuli Maquina Riquicho A, Jing Z, Liu T, Hou X, Li Y. Genome-wide survey and expression analysis of the *PUB* family in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Mol Genet Genom.* 2015;290:2241–60.
47. Swaminathan K, Peterson K, Jack T. The plant B3 superfamily. *Trends Plant Sci.* 2008;13:647–55.
48. Hayama R, Yang P, Valverde F, Mizoguchi T, Furutani-Hayama I, Vierstra RD, Coupland G. Ubiquitin carboxyl-terminal hydrolases are required for period maintenance of the circadian clock at high temperature in *Arabidopsis*. *Sci Rep.* 2021;9:17030.
49. Forrest JRK. Plant – pollinator interactions and phenological change: what can we learn about climate impacts from experiments and observations? *Oikos.* 2005;124:4–13.
50. Hegland SJ, Nielsen A, Lázaro A, Bjerknes AL, Totland Ø. How does climate warming affect plant-pollinator interactions? *Ecol. Lett.* 2009;12:184–95.
51. Fishman L, Kelly AJ, Willis JH. Minor quantitative trait loci underlie floral traits associated with mating system divergence in *mimulus*. *Evolution.* 2002;56:2138–55.
52. Fishman L, Beardsley PM, Stathos A, Williams CF, Hill JP. The genetic architecture of traits associated with the evolution of self-pollination in *Mimulus*. *New Phytol.* 2015;205:907–17.
53. Campitelli BE, Kenney AM, Hopkins R, Soule J, Lovell JT, Juenger TE. Genetic mapping reveals an anthocyanin biosynthesis pathway gene potentially influencing evolutionary divergence between two subspecies of scarlet gilia (*Ipomopsis aggregata*). *Mol Biol Evol.* 2017;35:807–22.
54. Bolnick DI, Barrett RDH, Oke KB, Rennison DJ, Stuart YE. (Non)parallel evolution. *Ann Rev Ecol Evol Syst.* 2018;19:303–30.
55. Schmutz J, et al. A reference genome for common bean and genome-wide analysis of dual domestications. *Nat Genet.* 2014;46:707–13.
56. Takuno S, Ralph P, Swarts K, Elshire RJ, Glaubitz JC, Buckler ES, Hufford MB, Ross-Ibarra J. Independent molecular basis of convergent highland adaptation in maize. *Genetics.* 2015;200:1297–312.
57. Cai L, Comont D, MacGregor D, Lowe C, Beffa R, Neve P, Saski C. The blackgrass genome reveals patterns of non-parallel evolution of polygenic herbicide resistance. *New Phytol.* 2023;237:1891–907.
58. James ME, Wilkinson MJ, Bernal DM, Liu H, North HL, Engelstädter J, et al. Phenotypic and genotypic parallel evolution in parapatric ecotypes of *Senecio*. *Evolution.* 2021;10:1111–evo14387.
59. Laurie CC, Chasalow SD, LeDeaux JR, McCarroll R, Bush D, Hauge B, Lai C, Clark D, Rocheford TR, Dudley JW. The genetic architecture of response to long-term artificial selection: for oil concentration in the maize kernel. *Genetics.* 2016;168:2141–55.
60. Pritchard JK, Di Rienzo A. Adaptation – not by sweeps alone. *Nat Rev Genet.* 2010;11:665–7.
61. Csilléry K, Rodríguez-Verdugo K, Rellstab C, Guillaume F. Detecting the genomic signal of polygenic adaptation and the role of epistasis in evolution. *Mol Ecol.* 2018;27:606–12.
62. Barghi N, Hermisson J, Schlötterer C. Polygenic adaptation: a unifying. *Nat Rev Genet.* 2020;21:769–81.
63. Barton NH, Etheridge AM, Véber A. The infinitesimal model: definition, derivation, and implications. *Theor Popul Biol.* 2017;118:50–73.
64. Sheng Z, Pettersson ME, Honaker CF, Siegel PB, Carlborg O. Standing genetic variation as a major contributor to adaptation in the Virginia chicken lines selection experiment. *Genome Biol.* 2015;16:219.
65. Berg JJ, Coop G. A population genetic signal of polygenic adaptation. *PloS Genet.* 2014;10:e1004412.
66. Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. *Cell.* 2017;169:1177–86.
67. Wieters B, Steige KA, He F, Koch EM, Ramos-Onsins SE, Gu H, et al. Polygenic adaptation of rosette growth in *Arabidopsis thaliana*. *PloS Genet.* 2021;17:e1008748.
68. El-Esawi MA. Genetic diversity and evolution of Brassica genetic resources: from morphology to novel genomic technologies – a review. *Plant Genet Resour.* 2017;15:388–99.
69. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo, et al. The variant call format and VCFtools. *J Bioinform.* 2011;27:2156–8.
70. Revell LJ, learnPopGen. An R package for population genetic simulation and numerical analysis. *Ecol Evol.* 2019;9:7896–7902.
71. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B.* 1995;57:289–300.
72. Gabriel SB, et al. The structure of haplotype blocks in the human genome. *Science.* 2002;296:2225–9.
73. Privé F, Luu K, Vilhjálmsson BJ, Blum MGB. Performing highly efficient genome scans for local adaptation with R package pcadapt version 4. *Mol Biol Evol.* 2020;37:2153–4.
74. Wickham H, et al. Welcome to the tidyverse. *J Open Source Softw.* 2019;4:1686.
75. Rowlingson B, Diggle P. splancs: Spatial and Space-Time Point Pattern Analysis. *R package version 2.01-43.* 2022; <https://CRAN.R-project.org/package=splancs>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.