Evolutionary Biology

Repeatability of adaptation in sunflowers: genomic regions harbouring inversions also drive adaptation in species lacking an inversion

Shaghayegh Soudi, Mojtaba Jahani, Marco Todesco, Gregory L. Owens, Natalia Bercovich, Loren H. Rieseberg, Sam Yeaman

Department of Biological Sciences, University of Calgary, 2500 University Drive NW, Calgary, Canada, T2N 1N4 • Department of Botany, University of British Columbia, 6270 University Blvd, Vancouver, Canada, V6T 1Z4 • Michael Smith Laboratories, University of British Columbia, 6270 University Blvd, Vancouver, Canada, V6T 1Z4 • Irving K. Barber Faculty of Science, University of British Columbia Okanagan, 3187 University Way, Kelowna, Canada, V1V 1V7 • Department of Biology, University of Victoria, 3800 Finnerty Rd, Victoria, Canada, V8P 5C2

Reviewed Preprint

Revised by authors after peer review.

About eLife's process

Reviewed preprint version 2
November 13, 2023 (this version)

Reviewed preprint version 1
July 3, 2023

Sent for peer review
May 3, 2023

Posted to bioRxiv
April 14, 2023

Abstract

Local adaptation commonly involves alleles of large effect, which experience fitness advantages when in positive linkage disequilibrium (LD). Because segregating inversions suppress recombination and facilitate the maintenance of LD between locally adapted loci, they are also commonly found to be associated with adaptive divergence. However, it is unclear what fraction of an adaptive response can be attributed to inversions and alleles of large effect, and whether the loci within an inversion could still drive adaptation in the absence of its recombination-suppressing effect. Here, we use genome-wide association studies to explore patterns of local adaptation in three species of sunflower: *Helianthus annuus*, *H. argophyllus*, and *H. petiolaris*, which each harbour a large number of species-specific inversions. We find evidence of significant genome-wide repeatability in signatures of association to phenotypes and environments, which are particularly enriched within regions of the genome harbouring an inversion in one species. This shows that while inversions may facilitate local adaptation, at least some of the loci can still harbour mutations that make substantial contributions without the benefit of recombination suppression in species lacking a segregating inversion. While a large number of genomic regions show evidence of repeated adaptation, most of the strongest signatures of association still tend to be species-specific, indicating substantial genotypic redundancy for local adaptation in these species.

eLife assessment

This is a **valuable** comparative study of adaptation across multiple species. The results provide a **solid** example of the application of genotype-environment associations to demonstrate that local adaptation is repeatable.

Introduction

The genetic basis of local adaptation is sometimes highly repeatable, with examples of large effect genes driving responses in multiple species, such as *FT* affecting flowering time in numerous plants (Izawa 2007; Auge et al. 2019) or *Mc1r* driving colour polymorphisms in vertebrates (Manceau et al. 2010; Rosenblum et al. 2014). Local adaptation can also be repeated for polygenic traits, with significant patterns of similar association found across many loci for comparisons of conifers (Yeaman et al. 2016), maize and its wild relative teosinte (Tittes et al. 2021; Wang et al. 2021), and *Brassicaceae* (Bohutínská et al. 2021), to name a few. While we have a growing number of examples of repeatability in the basis of adaptation, it is also interesting to know if species use different genes to adapt to the same selection pressure. Genotypic redundancy – the potential for many genotypes to yield a given phenotype – is one critical factor affecting the repeatability of adaptation, as high redundancy would be expected to result in lower repeatability (Yeaman et al. 2018). Another critical factor affecting repeatability is shared standing variation, whether present due to introgression or incomplete lineage sorting. In either case, variants shared among lineages are much more likely to contribute to a repeated response than new mutations (MacPherson and Nuismer 2018; Ralph and Coop 2015). Consistent with this, Bohutínská and colleagues (2021) found that repeatability was negatively related to phylogenetic distance. While we are now accumulating more studies about the repeatability of adaptation, we still have very few examples and much remains unknown about the relative importance of these factors (Yeaman 2022).

Inversions have been implicated in local adaptation in many species (Wellenreuther and Bernatchez 2018), likely due to their effect to suppress recombination among inverted and non-inverted haplotypes, and thereby maintain LD among beneficial combinations of locally adapted alleles (Rieseberg 2001; Noor et al. 2001; Kirkpatrick and Barton 2006). This has been approached by models studying the establishment of inversions that capture combinations of locally adapted alleles present as standing variation (*e.g.*, Kirkpatrick and Barton 2006), as well as models examining the accumulation of locally adapted mutations within inversions (*e.g.*, Schaal et al. 2022). If there is variation in the density of loci that can potentially contribute to local adaptation, inversions would be expected to preferentially establish and be retained in regions harbouring a high density of such loci (and this expectation would hold for both the capture and accumulation models). We would also expect to see stronger signatures of repeated local adaptation in such high density regions. Despite mounting evidence of their importance in adaptation, it is unclear how inversions may covary with repeatability of adaptation among species. A fundamental parameter of importance in these models is the relationship between migration rate and strength of selection on individual alleles, which may not make persistent contributions to local adaptation without the suppressing effects of recombination if selection is too weak (Yeaman and Whitlock 2011; Bürger and Akerman 2011). If most alleles have small effects relative to migration rate and can only contribute to local adaptation via the benefit of the recombination-suppressing effect of an inversion, then we would expect little repeatability at the site of an inversion – other species lacking the inversion would not tend to use that same region for adaptation because selection would be too weak for alleles to persist. On the other hand, if some loci are particularly important for local adaptation and regularly yield mutations of large effect, with these patterns being conserved among species, repeatability within regions harbouring inversions may be substantial. Thus, studying whether adaptation at the same genomic region harbouring an inversion is observed in other species lacking the inversion can give insights about the underlying architecture of adaptation, and the evolution and maintenance of inversions.

Here, we explore the repeatability of local adaptation in three species of sunflowers, *Helianthus annuus*, *Helianthus argophyllus* and *Helianthus petiolaris* (**Figure 1**), which harbour large regions of suppressed recombination ("haploblocks"), most of which are inversions, and are often associated with adaptive traits (Todesco et al. 2020). *Helianthus annuus*, the common sunflower, is the closest wild relative of cultivated sunflower, which was domesticated from it around 4,000 years ago (Blackman et al. 2011). Populations of *H. annuus* are distributed throughout the central and western USA and generally found on mesic soils, but can grow in a variety of disturbed or extreme habitats, such as semi-desertic or frequently flooded areas, as well as salt marshes. *Helianthus petiolaris*, the prairie sunflower, prefers sandier soils, and ecotypes of this species are adapted to sand sheets and sand dunes (Ostevik et al. 2016). Here we include samples from two subspecies: *H. petiolaris ssp. petiolaris*, which is commonly found in the southern Great Plains, and *H. petiolaris ssp. fallax*, which is limited to more arid regions in Colorado, Utah, New Mexico and Arizona (Heiser et al. 1969). *Helianthus petiolaris* and *H. annuus* have broad and overlapping distributions throughout the central and western United States and appear to have adapted to similar changes in temperature, moisture and photoperiod regimes. There is also evidence indicating *H. annuus* and *H. petiolaris* have likely been exchanging genes during much of their history of divergence (Strasburg and Rieseberg 2008), although partially isolated by strong pre- and post-zygotic barriers (Sambatti et al. 2012). The third species, *Helianthus argophyllus*, the silverleaf sunflower, is found exclusively in the southeast coast of Texas and includes both an early flowering ecotype on the coastal barrier islands and a late flowering ecotype inland (Moyers and Rieseberg 2016; Todesco et al. 2020). *H. argophyllus* is thought to have undergone cycles of sympatry and allopatry with *H. annuus* and *H. petiolaris* over time, but currently only overlaps with *H. annuus,* with thus overlap likely being a recent event (Heiser, 1951).

Using broad sampling across the ranges of these species (**Figure 1**) and sequencing data first published by Todesco et al. (2020), we study the genomic basis of local adaptation by conducting genome-wide scans for associations with climatic and soil environmental variables and phenotypes measured in a common garden. An inherent problem in studying the basis of local adaptation is accounting for the covariance between genome-wide population genetic structure and environment. When the selection pressure driving local adaptation tends to parallel a main axis of demographic expansion or isolation by distance, neutral alleles will have spatial patterns resembling those of causal alleles. Methods that do not correct for population structure have large numbers of false positives because allele frequencies at neutral loci tend to correlate with the environment more than expected by chance (Lotterhos and Whitlock 2014). By contrast, methods that use structure correction tend to suffer from false negatives because the causal loci have similar patterns of allele frequency variation as the genomic background, and so their statistical signatures are decreased (DeRaad et al. 2021; Booker et al. 2023b). When a non-corrected approach is applied to multiple species, the false positive problem can be mitigated when testing for loci that contribute to repeated adaptation, as the same gene should not tend to be a false positive in multiple species more often than expected by chance (Yeaman et al. 2016). Here we use a conventional GWAS approach with structure correction to study the basis of phenotypes, which vary both within and among populations, and use an uncorrected approach to test association to environment, which only varies among populations (as all individuals within the same population have the same value of environmental variable). For the environmental association portions of the study, we also explore the effect of structure correction in some analyses, to contrast the false negative vs. false positive problem inherent in each approach. We apply these approaches individually to *H. annuus*, *H. argophyllus*, and the two *H. petiolaris* subspecies, and make pairwise and higher-order comparisons among combinations of these four lineages to study repeatability.

Our overall aim is to characterize the extent of repeatability of local adaptation for different traits and environments, and explore the importance of inversions and other factors that may drive differences in repeatability. This analysis builds upon the work of Todesco et al. (2020), which

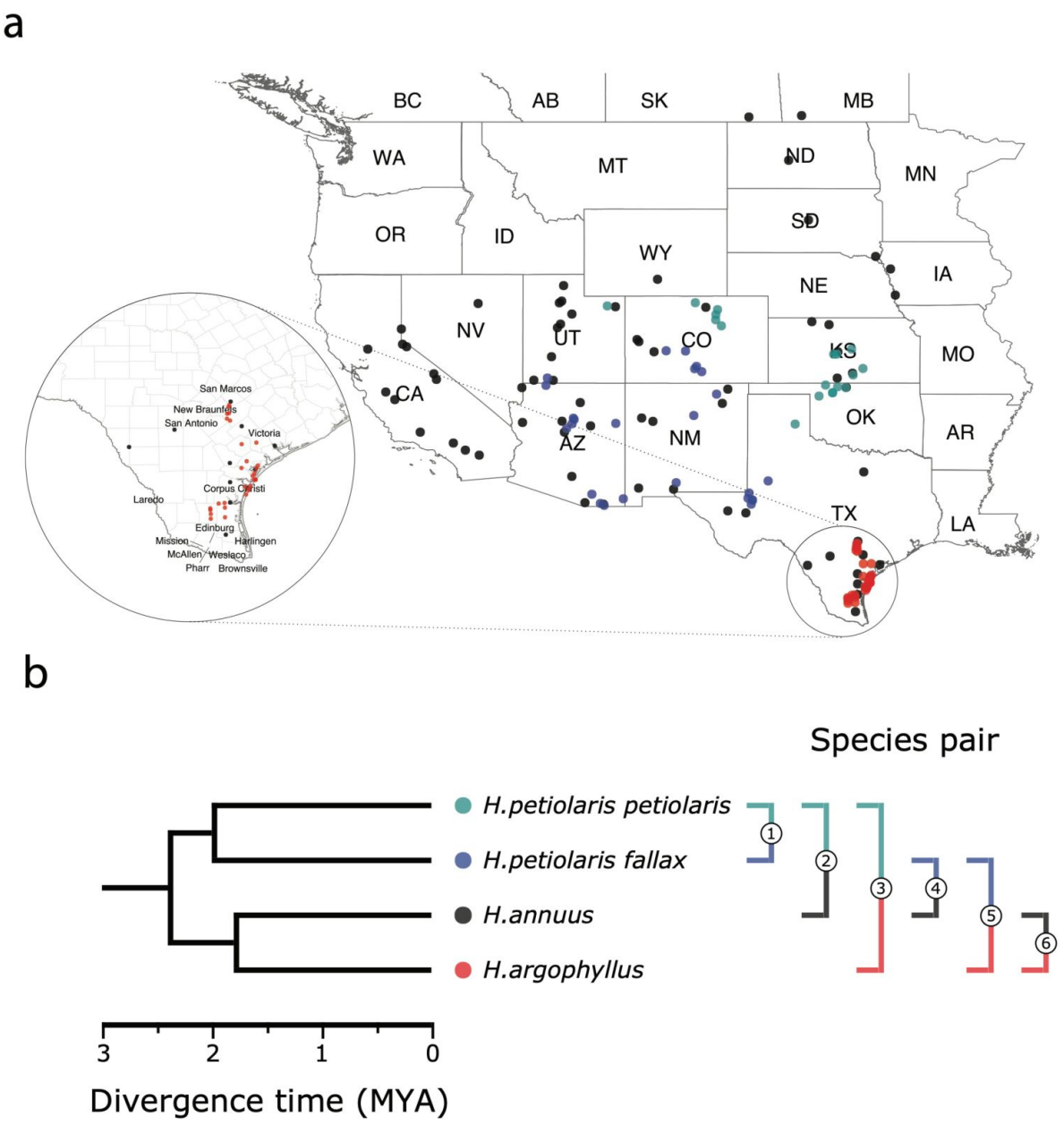

Figure 1.

Sampling sites and phylogenetic relationship among surveyed species.
(a) Sampling locations of wild sunflower populations studied in the present study, and (b) phylogenetic relationship of the four (sub-)species. Numbered brackets represent the six pairwise comparisons performed in this study.

focused on the evolution of haplobocks within each species, but did not systematically study patterns of repeatability in signatures of association among species. We begin by using the strength of correlations between environment and phenotypes measured in a common garden to identify which environmental variables are likely driving local adaptation within each species, and use an index to quantify the relative similarity in these patterns among pairs of species (see Figure S1 for a schematic overview of methods and research questions). We then conduct Genome-Wide Association tests to identify regions most strongly associated with phenotypes and environments within each species. To assess the similarity in these association statistics among species, we use methods similar to Yeaman et al. (2016) to identify regions of the genome that exhibit greater similarity in signatures of association among pairs of lineages than expected by chance, which are likely driven by repeated adaptation using the same genes (but not necessarily the same alleles). We then explore how the number and total size of these regions of repeated association covary with the index of similarity in environmental importance. We also explore whether these regions of repeated association tend to covary with patterns of shared standing variation and test if they are enriched within previously identified haploblocks, to explore the role of inversions in adaptation. We then identify candidate genes that show particularly strong signatures of repeated association across multiple lineages. Finally, based on the number of signatures of association that are shared vs. non-shared among lineages, we estimate the number of regions genome-wide that could potentially contribute to local adaptation for each variable and phenotype, which is related to the genotypic redundancy. Taken together, our results show that local adaptation in sunflowers tends to involve both strong repeatability at a small number of genes, often associated with inversions, coupled with high redundancy and non-repeated responses across a much larger number of loci.

Results

Local adaptation at the phenotypic level

The four taxa vary considerably in the breadth of environmental variation spanned by their respective ranges, with *H. annuus* spanning the widest niche and *H. argophyllus* spanning the narrowest (**Figure 2A**). Using the correlation between phenotype and environment as a metric of the strength of local adaptation on individual traits, we observe strong correlations for many combinations (Figure S2), suggesting that local adaptation is quite pronounced for some traits. To quantify the similarity among pairs of species in the strength of local adaptation at the phenotypic level (and to later compare this to results at the genomic level), we calculate an index we refer to as the Similarity in Phenotype-Environment Correlation (SIPEC), which is maximized when both species have strong correlations (in either direction) between a phenotype and environmental variable (see Methods). For each environmental variable, we take the maximum value of SIPEC across all phenotypes as a relative measure of the importance of the variable for driving local adaptation in both species (such that when SIPEC is low, the variable is not a strong driver of phenotypic adapatation in at least one of the two species). We find the largest values of SIPEC for temperature variables and the smallest values for soil types, and generally find higher values for comparisons between the *petiolaris* subspecies (Figure S3).

Genome-wide analysis of repeated local adaptation

To search for regions of the genome driving repeated adaptation in multiple taxa, we first identified windows of the genome within each species that showed strong signatures of association to either phenotype (GWAS; with structure correaction) or environment (GEA; without structure correction), referred to as “top candidate” windows. To identify Windows of Repeated Association (WRAs) between pairs of species, we then assessed whether top candidate windows identified in one focal taxon also tended to be enriched for strong signatures of association to the same environment or phenotype in each of the other taxa (using the null-W test, which tends to be

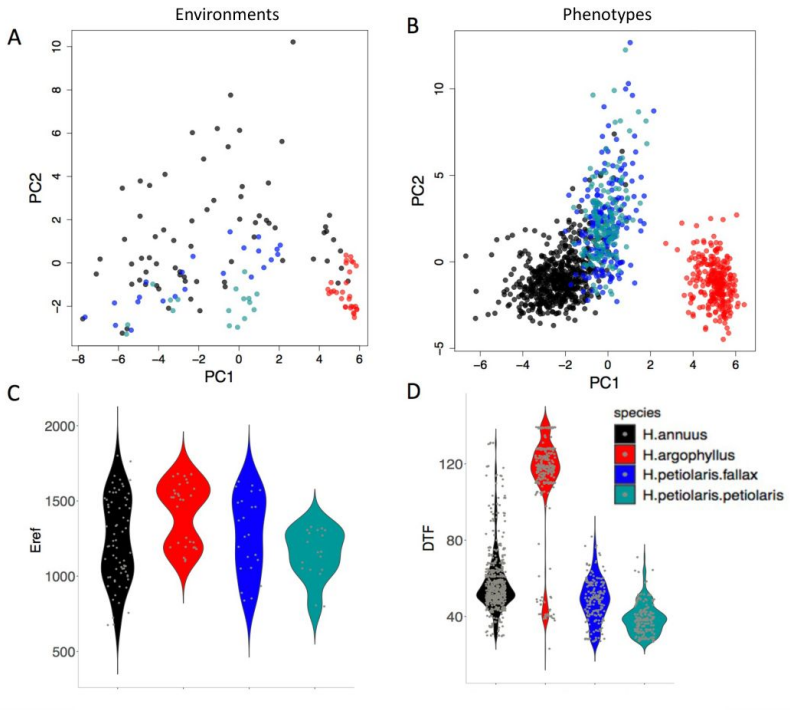

Figure 2.

The range of environmental and phenotypic variation in the studied species. Variation in environment (a) and phenotype (b) for the studied species, along the two largest axes of a Principle component analysis (PCA). Violin plots show two examples of variation in environment (Hargreaves reference evapotranspiration index; Eref) and phenotype (Days to Flower; DTF) within and among the taxa (c,d).

more sensitive than just finding the overlap between the top candidate windows; See Methods). As recombination rate can affect the sensitivity of these kinds of window-based genome scans (Booker et al. 2020), the identification of WRAs was conducted after binning windows by recombination rate, although in many cases we did not observe substantial differences in the null distributions among these bins (*e.g.*, Figure S4). This analysis revealed many windows with signatures of repeated association, with the strongest repeated signature found for Number of Frost-Free Days (NFFD; **Figure 3A**; S4), but also showed that many of the windows with the strongest association in one species did not have strong signatures in other species (**Figure 3B**).

Under the null hypothesis that all regions of the genome evolve neutrally due to drift and independently in each species, ~5% of the top candidate windows identified in one species would be expected to fall into the tail of the null distribution of another species (*i.e.*, be classified as WRAs). While we find significantly more WRAs than the 5% expected by chance, ranging from 6.3% to 44.1% (mean = 14.5%; Figure S5), multiple factors can violate the assumptions of this test and increase the proportion of windows classified as WRAs. Most importantly, similarity in signatures of association can be driven by shared ancestral variation or ongoing introgression, and this must always be considered as an alternative explanation. It is clear that most regions of the genome harbour some shared standing variation between all 6 pairwise comparisons of the taxa, whether due to segregating ancestral variation or introgression, based on an index quantifying the proportion of shared to non-shared SNPs in each window (Figure S6). However, we do not find consistent differences in the amount of shared standing variation within windows when comparing WRAs with the rest of the genome (Figure S6). Thus, while introgression may make subtle contributions to WRAs, a lack of increased shared standing variation in the significant windows suggests that it is not a primary driver of these patterns.

A second source of inflation in the number of WRAs is due to hitchhiking: if a given locus is a true positive driving adaptation, LD with other windows in tight physical linkage will result in spurious correlations causing them to also be classified as WRAs increasing the genome-wide proportion above 5%, and this effect will be particularly exaggerated in haploblocks. To control for this, we binned neighbouring WRAs together that exhibited high LD (>95th percentile) over short spans of the genome (<1 cM) in either species, yielding regions we refer to as Clusters of Repeated Association (CRAs). While this clustering method should provide a partial control for the effect of hitchhiking, we note that the number of CRAs cannot be taken as a reliable estimate of the number of independent targets of natural selection. Instead, we treat the number and size of these CRAs as proxies for the relative genome-wide similarity in association between each pair of species, and conduct downstream analyses to test hypotheses about the factors driving local adaptation.

CRAs varied in number and size across the 6 pairwise-species comparisons, with particularly large CRAs identified at the sites of known haploblocks (Todesco et al. 2020). For CRAs identified for phenotypic associations, their extent ranged from 168 clusters spanning 48,524,903 bp (covering only 1.6% of the genome) in *H.argophyllus-H.pet.petiolaris* to a maximum of 1,667 clusters spanning 426,409,647 bp (14.2 % of the genome) in *H. pet. petiolaris-H. pet. fallax* (Figure S7). CRAs identified for environmental variables tended to be more numerous and cover a larger region of the genome, varying from 1154 clusters comprising 15% of the total genome for *H. argophyllus-H. pet. fallax* up to 2,260 distinct clusters covering 29% of the genome between *H. petiolaris* subspecies (Figure S8). The large extent of the genome covered by CRAs is mainly driven by their occurrence within many different haploblocks, which tend to be unique to each species and cover substantial portions of their genomes (Todesco et al. 2020). Genes within CRAs tend to be enriched for a large number of GO terms spanning many different categories (Figure S9, S10). When we use an approach that corrects for population structure in the genotype-environment association tests (Baypass), we find substantially reduced numbers of CRAs (Figure S11). This reduction is likely due to reduced power to detect true positive causal loci when population structure covaries with environment, as true positives do not “stand out” relative to the genomic background (Booker et al. 2023b). While single-species association tests without correction for

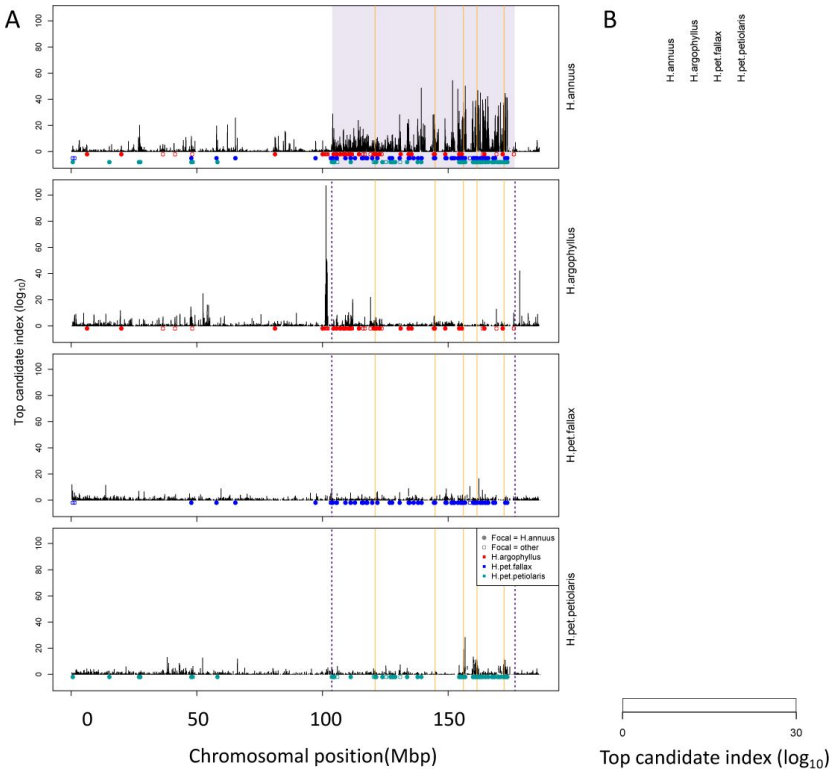

Figure 3.

Signatures of association for Number of Frost-Free Days (NFFD) in the four taxa on chromosome 15 (A) and genome-wide (B). Panel A shows Windows of Repeated Association (WRAs; coloured bars) for comparisons between the focal species, *H. annuus*, and each of the other 3 taxa, with the haploblocks in *H. annuus* shaded in violet and the regions with significant PicMin hits as vertical orange lines. Panel B shows the value of the top candidate index for each of the 1000 windows with the strongest signatures of association in at least one species (~top 2% of genome-wide windows). Rows are ordered using hierarchical clustering to group windows with similar patterns across multiple species, illustrating the extent of overlap/non-overlap in the windows with strongest signatures in each species (*i.e.*, position in the figure does not reflect chromosomal position).

population structure typically yield large numbers of false positives, the null-W test comparing association results across species accounts for the chance that the same region is a false positive in both species (Yeaman et al. 2016).

Repeated association in the genome reflects patterns at the phenotypic level

If natural selection is driving repeated patterns of local adaptation in the genomes of two species, we should see greater similarity for associations with environmental variables that are also strongly associated with phenotypic variation in both species. Consistent with this prediction, we found that environmental variables with a high maximum SIPEC index (*i.e.*, high correlation between environment and some axis of phenotypic variation in both species; Figure S3) also tended to have a larger number of CRAs (**Figure 4**), with higher repeatability at both the phenotype and genome level for temperature-related variables and much lower repeatability for soil-related variables (with similar patterns for mean SIPEC index; Figure S12). Unfortunately, given the non-independence of environmental variables there is pseudo-replication in these data so it is not possible to conduct a formal significance test of these patterns. In all cases, the number of CRAs exhibited a weak negative relationship with the index of standing variation, with environmental variables that had the greatest number of CRAs having the lowest relative amounts of shared standing variation (Figure S13). This suggests that the observed similarity in patterns of association among species is not being strongly driven by incomplete lineage sorting or introgression, as we would expect a positive relationship between the index of shared standing variation and the number of CRAs. We found similar but weaker patterns in the association between SIPEC and size of CRAs (Figure S14). The weakening of these patterns relative to those for number of CRAs may reflect comparatively greater contribution of haploblocks driving patterns with size rather than number (*i.e.*, if a particularly large haploblock is a CRA, then that variable will have a particularly large value relative to its actual importance for local adaptation).

Overlap of signatures of repeated association with low recombination haploblocks

Inversions can facilitate local adaptation by suppressing recombination between locally adapted alleles within the inverted vs. ancestral haplotypes (Kirkpatrick and Barton 2006). If a given species exhibits a signature of association to environment at a segregating inversion, it is interesting to test whether similar signatures of association are found in other species that lack a segregating inversion at the same region of the genome. This would suggest that particularly strong selection is acting on loci in this region, as signatures of association can still evolve even without the recombination suppressing effect of the inversion. While not all of the low-recombination haploblocks identified by Todesco et al. (2020) have been validated as inversions, for simplicity we treat each haploblock as representative of a segregating inversion. The majority of these haploblocks are present as segregating variation in only one of the three species, but when they occur in *H. petiolaris*, they tend to be found as segregating variation within both subspecies. Thus, if we find significant enrichment of the regions of repeated association (WRAs and CRAs) within haploblocks, this suggests that both the species with a segregating inversion and the species lacking an inversion are using this region of the genome to drive local adaptation. As a first test of this relationship, we assessed two-way contingency tables for whether top candidate windows within a species were also significant windows of repeated association (WRA/non-WRA) and whether they fell within haploblocks (yes/no). This approach controls for the potential enrichment of top candidate windows within haploblocks relative to the rest of the genome and asks if the proportion of WRAs within haploblocks is even higher than non-WRA top candidate windows. For most phenotypes and environments, we found that the proportion of WRAs occurring within haploblocks was much higher than the proportion of non-WRAs occurring within haploblocks (**Figure 5**). Across all species comparisons and environmental/phenotypic

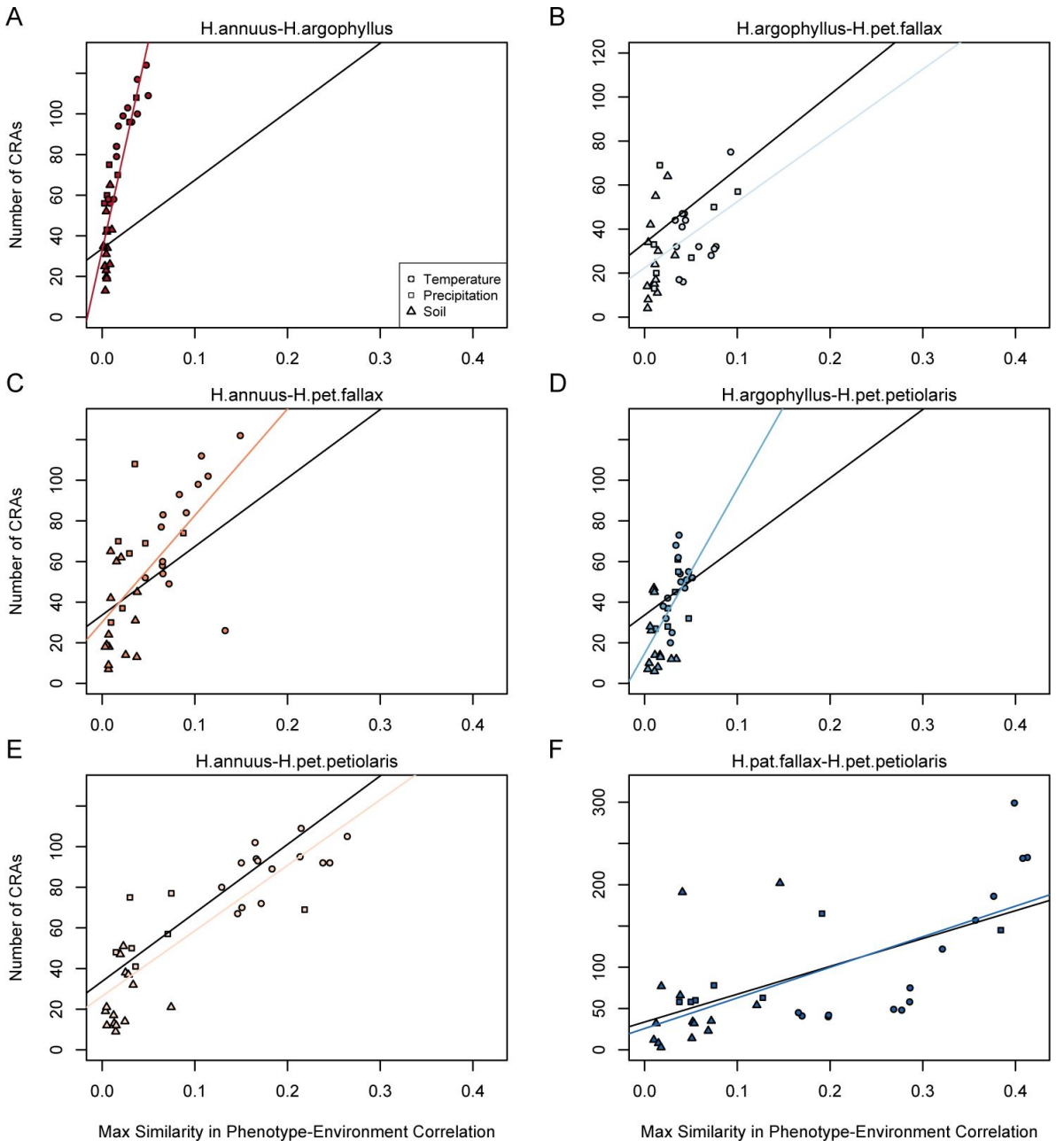

Figure 4.

Relationship between maximum Similarity in Phenotype-Environment Correlation (SIPEC) and number of Clusters of Repeated Association (CRAs). SIPEC was calculated for each phenotypic PCA axis with the maximum taken across the axes that cumulatively explain 95% of the phenotypic variance, for each environment. Each panel includes both a linear model fit to the data within the panel (coloured lines), and a linear model fit to all data simultaneously (black lines) for comparison. Note that because environmental variables are correlated, these points are not independent and therefore represent a source of pseudoreplication, preventing formal statistical tests of this relationship.

variables, a total of 310 contrasts showed significant enrichment for WRAs in haploblocks, compared to only 23 contrasts showing a significantly higher proportion of non-WRAs in haploblocks (note that because of non-independence of the phenotypes and environmental variables, there is pseudo-replication inherent in these estimates). In *H. argophyllus*, while non-WRA top candidates tended to fall within haploblocks less commonly than in the other taxa, WRAs tended to be very strongly enriched within haploblocks (**Figure 5B**). As a follow-up, we also tested whether CRAs were enriched within haploblocks (without controlling for whether top candidate windows also tended to be enriched), also finding strongly significant enrichment of CRAs within haploblocks for many traits and environments (Figure S7C & D). Broad patterns revealed by these analyses are similar, as both show that genomic regions harbouring haploblocks tend to be enriched for signatures of association to environment in species lacking the haploblock. It is noteworthy that few significant signatures of enrichment in haploblocks were found for *H. petiolaris petiolaris* vs. *H. petiolaris fallax* (Figure S7C & D), perhaps because extensive introgression across non-locally adapted regions of the genome obscures true signal.

Multi-species signatures of repeated association

As a complement to the pair-wise analyses described above, we also conducted an analysis using PicMin, which simultaneously considers any number of lineages to identify genes with particularly strong and repeated signatures of association (Booker et al. 2023a). After identifying significant windows clustering together those < 1 Mbp apart, we found a total of 145 regions that were significant for at least one environmental variable or phenotype (at FDR < 0.1 applied within each variable). The largest number of significant regions for environment was found with the Number of Frost-Free Days variable (20 with FDR < 0.1; 44 with FDR < 0.2) and for phenotypes with Total Leaf Number, a measure of developmental timing of flowering (9 with FDR < 0.1; 19 with FDR < 0.2). Considering the 720 individual 5000 bp windows with FDR < 0.2 for at least one variable, 387 (53.8%) are within 500 bp of a genic region, representing a 2.3x enrichment relative to the rate for windows that were not significant hits using PicMin (23.3% of other windows fall within 500 bp of a genic region; Ξ^2 test $p < 10^{-15}$).

Estimating the number of potentially adaptive loci

Even for the environment with the highest repeatability (Number of Frost-Free Days; NFFD), most of the strongest signals of association are found in only one lineage (**Figure 3B**). Unfortunately, while the null-W test provides strong support and controls false positives when identifying repeated association, regions of the genome with strong associations in only a single species may include a large (and unknown) number of false positives. Thus, while it is biologically interesting to know how much adaptation is non-repeated among species, it is difficult to quantify with certainty (Booker et al. 2023a; Booker et al. 2023b). Under the assumption that at least some of the non-repeated signatures of association are true positives, this implies that there is considerable genotypic redundancy, with many different ways for these species to adaptively respond to variation in the same environment. To estimate the number of windows that could potentially contribute to local adaptation for each variable (L_{eff}), we modified the method of Yeaman et al. (2018) to partially account for the effect of linkage among nearby windows (see methods). This method assumes that the windows with signatures of association in each species are a random draw from a larger number of potentially contributing windows (L_{eff}), which can be inferred based on the ratio of shared vs. non-shared windows with strong association signatures (see Methods). We find that estimates of L_{eff} tend to be large, always well over 1000 windows regardless of the trait or environment (Figure S15A). There are numerous sources of error that affect the estimation of L_{eff}, which will tend to be overestimated due to linkage (Figure S15A) or when many signatures of association are false positives (see supplementary results), and will be underestimated when some repeatability is due to shared standing variation (Yeaman et al.

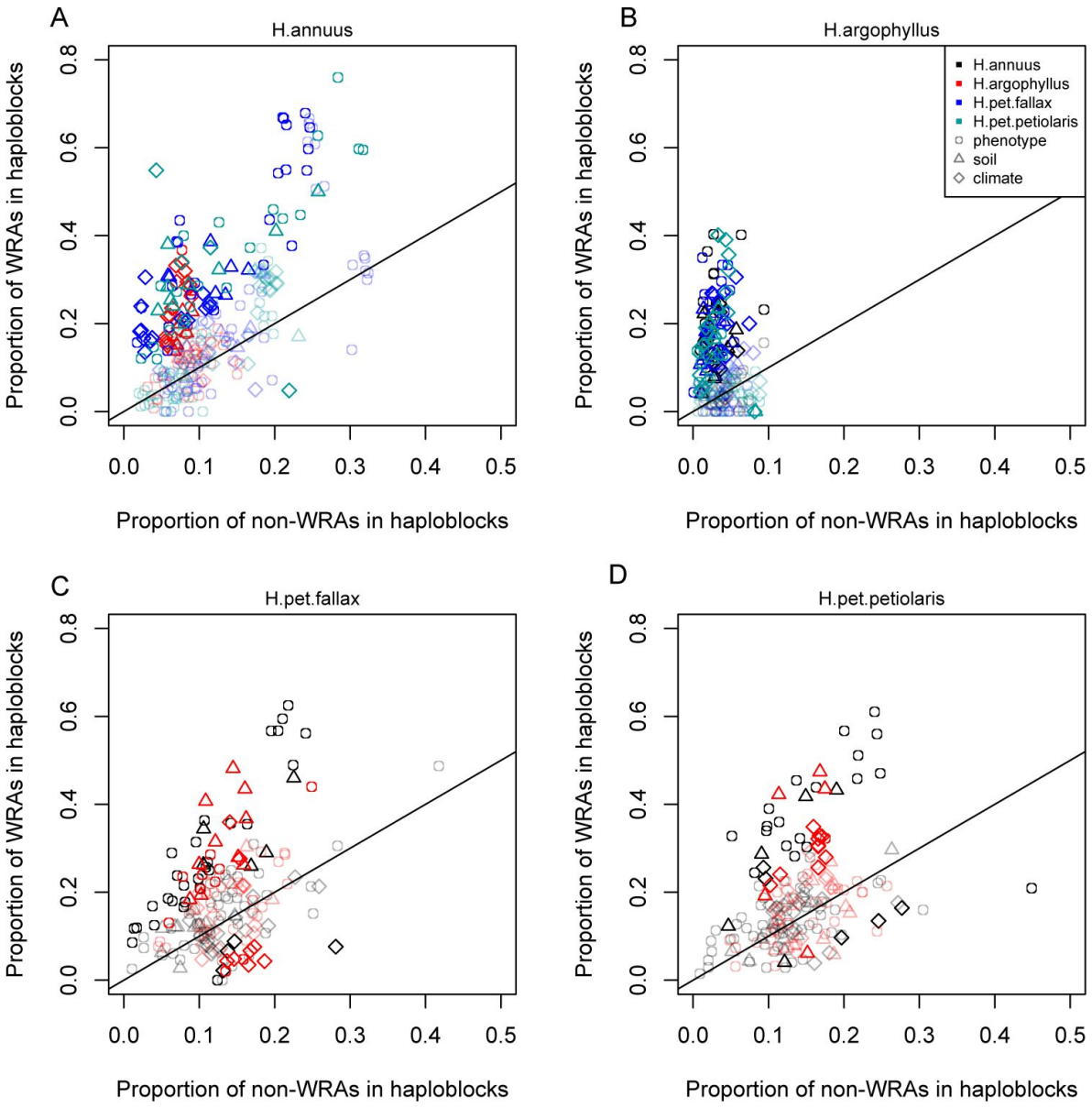

Figure 5.

Enrichment of signatures of repeated association within genomic regions harbouring a haploblock in one of the two compared lineages. Each panel shows the proportion of top candidate windows that fall within haploblocks for windows with significant signatures of repeated association by the null-W test (WRAs) vs. those with non-significant signatures (non-WRAs). Comparisons of *H. petiolaris petiolaris* vs. *H. petiolaris fallax* are omitted as they share segregating haploblocks. Each point corresponds to the results for a single phenotype or environment, with dark shading used for cases where the deviation from random for the contingency table is significant by a permutation test ($p < 0.05$), and lighter shading indicating a non-significant result. Note that because many environmental variables and phenotypes are correlated with each-other, these points are not independent and therefore represent a source of pseudoreplication, preventing formal statistical tests of the overall relationship within each panel.

2018). Even after controls for linkage and assuming a high false positive rate of 80%, estimates of L_{eff} remain in the hundreds even for the variable with the lowest value of L_{eff} (Eref; Figure S15C, D).

Discussion

Despite its critical importance in shaping the architecture of adaptation, little is known about the extent of genotypic redundancy underlying different traits (Barghi et al. 2020; Láruson et al. 2020; Yeaman 2022). Here, we have shown evidence of significant repeatability in the basis of local adaptation (**Figure 4**, 5), but also an abundance of species-specific, non-repeated signatures (**Figure 3**; S14). In particular, we find that regions of the genome that harbour inversions in one species also tend to be strongly enriched for signatures of association in other species lacking the inversion (**Figure 5**). Taken together, this suggests that local adaptation in these species is highly flexible – different species apparently use quite different sets of loci to adapt to the same environment – yet still involves some component that has minimal redundancy, with inversions playing a particularly important role. Some of the "usual suspects" show up in the set of significantly repeated loci identified by PicMin: in addition to the homologs of the *FT* (*FLOWERING LOCUS T*) gene reported by (Todesco et al. 2020), we also found hits for several other genes involved in circadian regulation and flowering time, including *PRR3* (Para et al. 2007), *TOE1* (Aukerman and Sakai 2003), and *PHYC* (Takano et al. 2005; Chen et al. 2014), which is known to regulate photoperiodic responses in Arabidopsis accessions (Balasubramanian et al. 2006). Other top hits included genes involved in plant development and auxin transport (*PIN3*; (Keuskamp et al. 2010), *ARF4* (Pekker et al. 2005), and *MN*; (Bhatia et al. 2016)), and plant immunity (*CRK13* (Acharya et al. 2007) and *PRR2* (Cheval et al. 2017))

The increased repeatability found in regions of the genome that harbour inversions in only one species is particularly interesting. Inversions are commonly associated with local adaptation (Wellenreuther and Bernatchez 2018), likely because they reduce the rate at which recombination breaks up combinations of co-selected alleles (Kirkpatrick and Barton 2006), which perhaps facilitates contributions by alleles that would be individually too weakly selected to overcome swamping by migration (Bürger and Akerman 2011; Yeaman and Whitlock 2011; Schaal et al. 2022). Evidence here suggests these regions still tend to contribute to adaptation in the species lacking the recombination-suppressing effect of an inversion, consistent with a strong effect of selection relative to migration on at least one locus in the region (rather than adaptation exclusively via many alleles of small effect; Yeaman 2022). As an example, chromosome 15 harbours a large (72 Mbp) haploblock in *H. annuus* that is strongly associated with NFFD (Number of Frost-Free Days), and also shows some signatures of association in the other taxa, with particularly strong signatures of association on either side of the haploblock in *H. argophyllus* (**Figure 3A**). Interestingly, loci of repeated association identified by PicMin within this region include two genes whose homologs are known to regulate responses to cold: *COLD-RESPONSIVE PROTEIN KINASE 1*,*CRPK1* (Liu et al. 2017) and *LATE ELONGATED HYPOCOTYL*, *LHY* (Mizoguchi et al. 2002; Dong et al. 2011). It seems likely that strong selection relative to migration is therefore acting upon several loci in this region, and in many others harbouring inversions.

The observed repeatability associated with inversions further supports the local adaptation model as an explanation for the long-term persistence of segregating inversions (at least in sunflower), rather than mechanisms based on dominance or meiotic drive (Rieseberg 2001). If there is variation across the genome in the density of loci with the potential to be involved in local adaptation, then the establishment and maintenance of inversions would be biased towards regions harbouring a high density such loci under this model. If the genomic basis for local adaptation is conserved amongst species, then these same regions are more likely to have high repeatability. Thus, our observation of genomic regions harbouring inversions also being enriched for WRAs is consistent with this general model for inversion evolution. Unfortunately, our

observations do not provide much insight into whether inversions evolve through the capture (*e.g.*Kirkpatrick and Barton 2006) or accumulation (*e.g.*Schaal et al. 2022) type of model, as either model would be consistent with our results. Most of the the sunflower inversions are >1 My old, and therefore predate any current local adaptation patterns, but likely do not predate the genes underlying local adaptation (which appear to be shared among the species we studied). As for the alleles underlying local adaptation, they may be younger than the inversions, but as our work suggests, these regions are prone to harbouring locally adaptive alleles so it is possible that they also harboured other ancestral locally adaptive alleles. While many studies have demonstrated the importance of inversions for adaptation (Wellenreuther and Bernatchez 2018; Hager et al. 2022), to our knowledge only two other studies have documented the involvement of the same loci making contributions in the absence of the recombination-suppressing effect of the inversion (Lee et al. 2017; Coughlan and Willis 2018). This also highlights how comparative studies of a species lacking an inversion may help identify which genes are driving adaptation in another species with an inversion, as segregating inversions tend to have extensive LD that prevents identification of any potential targets of natural selection within them.

While our results suggest a large number of loci can potentially contribute to adaptation, implying high redundancy (**Figure 3B**; S14), there are several factors that complicate inference. Separating the effects of drift and selection to detect signatures of local adaptation is notoriously difficult, because population structure often covaries with features of the environment that drive adaptation (Lotterhos and Whitlock 2014; Hoban et al. 2016; DeRaad et al. 2021). As found by other analyses (Yeaman et al. 2016; DeRaad et al. 2021), when we use structure correction in our genotype-environment association tests, we find many fewer signatures of repeated association (Figure S7 vs. S11), likely due to the reduced power of Baypass to detect true positives when the environment covaries with population structure (Booker et al. 2023b). Here, we have side-stepped the issues involved in correction for population structure by instead relying on comparisons among species to identify loci with associations more extreme than expected by chance in multiple species, using the null-W and PicMin tests. This assumes that most loci in the genome are not involved in local adaptation, so that the relatively small proportion that are driving adaptation can therefore be picked out due to their tendency to fall into the tail of the distribution of association statistics in at least one other species (which would only occur at the rate of random chance under a purely neutral model). While this approach should be relatively robust for identifying loci with repeated patterns of adaptation, there is no way to formally estimate significance of associations found in only a single species, many of which may be false positives due to covariation of population structure and environment. However, even if we assume that 80% of the observed non-repeated loci are false positives, we still find that estimates of the effective number of loci contributing to adaptation (L_{eff}) are in the hundreds (Figure S15C). As we are unable to detect loci of small effect due to the limited power, but these loci still likely make important contributions to heritability, our estimates of L_{eff} will also be biased downwards by excluding these potentially important drivers. Finally, it is also difficult to exclude the contribution of introgression or incomplete lineage sorting to the observed signatures of repeated association. If a locus tended to be highly introgressed between two species in a restricted region of their range, it is possible it could also covary with environment and therefore result in a signature of repeated association that would be mistakenly interpreted as adaptation. While it is difficult to preclude this from our analysis, regions of the genome with signatures of repeated association do not tend to have higher levels of shared standing variation than background regions (Figure S6), suggesting this is not a broad explanation for our observations. If anything, shared standing variation would be expected to increase repeatability of adaptation (MacPherson and Nuismer 2017), adding further weight to the inference that there is high genotypic redundancy in these species.

In general, we find that temperature is the strongest driver of repeated adaptation at both the phenotypic and genomic levels. We quantified local adaptation at the phenotypic level using correlations between traits measured in common gardens and the home environment of the

population they were sampled from. Across all phenotypes, these correlations tended to be strongest and most similar among species for temperature variables, particularly in comparisons among *H. annuus* and *H. petiolaris* subspecies (Figure S3). We see similar patterns of repeatability reflected in the genome, where temperature variables also tend to have the greatest repeatability (**Figure 4**). The similarity in these phenotypic and genomic signatures is consistent with an effect of strong selection, as other artefactual or drift-based explanations for repeatability would not be expected to reflect patterns found at the phenotypic level. It should be noted that the reduced importance of soil variables in the SIPEC index might be partly driven by the fact that all traits were measured on above-ground features, due to the difficulty of getting non-disruptive phenotypic measurements for roots.

Taken together, these results suggest that some fraction of the genome contributes to adaptation with low redundancy and high repeatability (which tends to be enriched within genomic regions where there exist inversions in at least one species), while the remainder of the adaptive response is driven by loci with high redundancy and species-specific contributions. Theoretical models of adaptive evolution necessarily involve simplified representations of genetics: population genetic approaches explore cases where strong selection deterministically drives a change in allele frequency at one or a few loci (usually without epistasis), whereas quantitative genetic approaches make the infinitesimal assumption that phenotypic change can be realized through small frequency changes at many loci (Barton et al. 2017; Barghi et al. 2020; Yeaman 2022). Such models implicitly make quite different assumptions about redundancy: if a population genetic model of directional selection includes no epistasis then it is implicitly assuming there is no redundancy, as every locus can increase fitness and no locus can affect the potential of another locus to increase fitness. On the other hand, quantitative genetics models implicitly assume complete redundancy, as the loci driving trait variation have interchangeable effects. Given that differences in redundancy result in quite different evolutionary dynamics (Höllinger et al 2019; Yeaman 2022), it is important to re-evaluate the behaviour of theoretical models in light of what we can learn about redundancy from empirical data. Our results suggest that adaptation is a complex process that does not map cleanly onto the assumptions of either approach. At least some component of trait variation experiences sufficiently strong selection and has low enough redundancy to drive the significant signatures of repeatability above the random null that we observe here. On the other hand, most of the strongest signatures of association are specific to a single species, implying high redundancy for most of the genome, even if there are still some detectable signatures of repeatability. Such observations may be consistent with classifying loci into “core” and “peripheral” sets similar to the omnigenic model, but would involve classification based on redundancy rather than gene expression, as advocated by Boyle et al. (2017). Theoretical models adopting a two-class representation of the genetic basis of trait variation may provide more realism than models assuming a simpler distribution, and perhaps yield new insights about qualitatively different dynamical behaviour. Given the difficulty of rigourously quantifying the contribution of alleles of small effect and loci contributing to local adaptation in only one species, it remains an open question what proportion of adaptive response is governed by loci with dynamical behaviour that follow the infinitesimal assumption. This could perhaps be quantified explicitly by decomposing heritabilities for traits driving local adaptation, as has been done for standing variation in humans (Visscher et al. 2017), but this would require a considerable increase in sample size as methods such as GCTA (Yang et al. 2011) do not yield accurate estimates at the sample sizes used here. Future work could perhaps resolve this question using multi-generational crossing designs to break up the LD that tends to accompany local adaptation, allowing a more accurate parsing of the effect sizes of genomic regions and the contribution by repeated vs. redundant components of adaptive trait variation.

Materials and Methods

Data collection, common garden, and phenotyping

Seed samples were collected from 151 wild sunflower populations covering most of their native distributions during the summer of 2015 (*H. annuus*: 61 populations for GWAS and 71 populations for GEA, *H. petiolaris fallax*: 23 populations, *H. petiolaris petiolaris:* 18 populations and *H. argophyllus*: 30 populations, **Figure 1** and Supplementary Table 1). Seeds from ten additional populations of *H. annuus* for the GEA analysis had been previously collected in the summer of 2011. Sample seeds were obtained from randomly chosen mothers, and were first germinated in a greenhouse for two weeks, later moved to an open-sided greenhouse for acclimation. Phenotypic data were collected throughout the growing season, as detailed in Supplementary Table 1. Extensive records of developmental and morphological traits throughout the growth of the plants including leaves, stem, and seeds which were collected and digitally imaged to extract relevant phenotypic data.

Similarity in Phenotype-Environment Correlation (SIPEC)

Locally adapted traits tend to exhibit strong correlations between environment and common-garden phenotype. To estimate which environmental variables are driving local adaptation for the same phenotype in pairs of species, we calculate an index of the similarity in phenotype-environment correlation, SIPEC = $(|r_1| + |r_2|)|r_1||r_2|$ where r_1 and r_2 are the Pearson correlations between the environment and phenotype in the first and second species, respectively. This SIPEC index is maximized when the correlation between a phenotype and an environmental variable is large in both species regardless of the direction, so it does not differentiate phenotypically convergent vs. divergent patterns of local adaptation (*e.g.*, increasing temperature causing an increase in flowering time in one species and a decrease in the other), and provides a means of estimating the relative importance of an environment driving local adaptation in both species in a similar way across all measured phenotypes. To account for non-independence among traits, for each pair of species we fit a PCA to all measured phenotypes and use the Principal Components that collectively explain 95% of the variance, and calculate SIPEC on the correlations of each of these axes with environment. For comparisons including *H. argophyllus*, 95% of the variance was typically explained by 8-10 PC axes (out of 28 or 29 phenotypes), whereas for comparisons among other taxa this included 21 or 22 PC axes (out of 65 or 66 phenotypes). We then report the maximum and mean value of SIPEC for each environmental variable across these phenotypic PCA axes.

Tests of SNP association with environment (GEA) and phenotype (GWAS)

A total of 39 environmental variables (21 climatic variables, 3 geographic variables, 15 soil variables Supplementary Table 1) were used for the genotype-environment association analysis (GEA). We refer to the climatic, soil and geographic variables collectively as the “environmental variables”, for simplicity. Soil data were collected by taking three to five soil samples collected at each population from across the area in which seeds were collected and submitted to Midwest Laboratories Inc. (Omaha, NE, USA) for analysis. Climate data for each population were collected over a 30-year period (1961-1990) from geographic coordinates of the locations where the samples were collected, using the software package Climate NA (Wang et al. 2012). We used the package BAYPASS version 2.1 (Gautier 2015), which provides a re-implementation of the Bayesian hierarchical model, and explicitly accounts for the covariance structure among the population allele frequencies that originate from the shared history of the populations through the estimation of the population covariance matrix. This renders the identification of SNPs subjected to selection less sensitive to the confounding effect of population demography (Günther and Coop 2013). Population structure was estimated by choosing a random and unlinked set of 10,000 SNPs and running BAYPASS under the core model (*i.e.*, no covariates). Then the Bayes factors (BF) were

calculated running BAYPASS under the STD covariate model to evaluate association of SNPs with environmental variables (*i.e.*, adjusting for population structure). For each SNP, the Bayes factor (denoted BF_{is} as in Gautier 2015) was presented in decibian units (db) via the transformation 10 $\log_{10}$ (BF). BF_{is} relies on the importance sampling algorithm proposed by Coop et al., 2010 and uses MCMC samples obtained under the core model. To produce a narrower set of outlier loci, we followed the popular Jeffreys' rule (Jeffreys 1961) that identified outlier loci with BF ≥ 10. As genome scan methods that correct for population structure can remove some potential signals of local adaption when there is covariation between the demographic history of the species and the environmental variables or phenotypic traits of interest, we also calculated Spearman's rank correlation (ρ, uncorrected GEA) between population allele frequencies for each SNP and each environment variable.

GWAS analysis was performed on 86, 30 and 69 developmental and morphologic traits in *H. annuus, H. argophyllus* and *H. petiolaris,* respectively (Supplementary table 1). 29 variables were measured in all three focal species, and 39 were measured only in *H. annuus* and in *H. petiolaris* subspecies (Supplementary table 1). Seed and flower traits could not be collected for *H. argophyllus*, since most plants of this species flowered very late in our common garden, and failed to form fully developed inflorescences and set seeds before temperatures became too low for their survival. Population structure was controlled for in GWAS by including the first three principal components as covariates, as well as an identity-by-state (IBS) kinship matrix calculated by EMMAX (Kang et al. 2010). We ran each trait GWAS using EMMAX (v07Mar2010), as well as the EMMAX module in EasyGWAS (Grimm et al. 2017). For every SNP/peak above the Bonferroni significance threshold, candidate genes were selected within a 100 Kb interval centered in the SNP with the lowest p-value, or within the boundaries of the GWAS peak, whichever was larger. All variants used for association were initially filtered for VQSR 90% tranche, and then further filtered to only include bi-allelic SNPs genotyped in ≥ 90% of samples and with a minor allele frequency ≥ 3%.

Identification of top-candidate windows

We calculated the bottom 0.05 quantiles for the *p*-values from association tests, Spearman's rank correlation (uncorrected GEA) and GWAS (corrected and uncorrected), yielding two 5% cutoffs. For each environmental and phenotypic variable, we identified all outlier SNPs as those that fell below the respective 5% cutoff. For BayPass, we considered SNPs with Bayes factor ≥ 10 as outlier SNPs. For each 5 kb window that we defined across the whole genome, we counted the number of outlier SNPs (a) and the total number of SNPs (n). To identify top-candidate windows for each variable, we compared the number of outlier SNPs per each 5 kb window to the 0.9999 quantile of the binomial expectation where the expected frequency of outlier SNPs per window is calculated as: $\rho = \sum a_1/n_1$(summation over all 5 kb windows), calculating ρ independently for each environmental and phenotype variable and excluding windows with 0 outliers from the calculation of ρ (as per Yeaman et al. 2016). We also calculated a top candidate index using the same approach to categorize outliers, obtaining a *p*-value for a binomial test for the number of SNPs per window given an expected proportion of outliers (ρ; this *p*-value is not exact due to non-independence of SNPs so we refer to this as an index).

Identifying outlier SNPs detected by genome scans from genome-wide distribution without accounting for local recombination rate variation can promote false positive signals in recombination cold spots (*i.e.*, low recombination regions) and be overly conservative in recombination hot spots (*i.e.*, high recombination regions) (Booker et al. 2020). Therefore, to account for local recombination rate variation, all genomic windows were binned by their estimated recombination rates into 5 equally sized bins (bin1: 0-20% quantile, bin2: 20%-40% quantile, bin3: 40%-60% quantile, bin4: 60%-80% quantile, bin5: 80%-100% quantile). For each recombination bin, we estimated expected frequency of SNPs per window (ρ) and calculated cut-off separately. Windows falling above the threshold were identified as top candidate windows.

Genome-wide survey of repeatability (null-W test)

To explore repeatable genomic signatures of adaptation for each of the six pairwise contrasts among the 4 taxa (**Figure 1**), we used the method developed by Yeaman et al. (2016) with some modifications. A common approach is to identify candidates for adaptation independently in each species and then examine the overlap between these lists, however this approach is quite stringent and may miss many interesting signals. The null-W test is more sensitive, as it takes the list of top candidates from one species and tests whether they tend to show more extreme signatures of association than expected by chance. The null-W test is especially favorable when LD increased divergence of SNPs in tight linkage with causal SNPs but did not raise the test values enough for a window to be classified as an outlier according to the binomial test. For each top candidate window that we identified for each focal species in a pair, we refer to the same window in the other species as "top candidate ortholog". The null distribution for each focal species and variable was constructed by randomly sampling 10,000 background SNPs from non-top candidate ortholog windows. For each non-top-candidate window, we then estimated the test statistic (*W*) for the Wilcoxon-signed rank test vs. the 10,000 background SNPs. This resulted in a null distribution representing the differences between the 10,000 background SNPs and the non-top-candidate ortholog windows. These were then standardized to *Z*-scores using the method in Whitlock and Schluter (2020):

$$Z = \frac{2W - n1n2}{\sqrt{n1n2(n1+n2+1)/3}} \qquad \text{(Equation 2)}$$

where n1 and n2 are the sample sizes being compared. In order to control for heterogeneity in recombination rate and its possible effects on the null distribution, we estimated null distribution for each recombination bin separately (5 in total, see above). We then compared the p-values and bayes factors (BFs) for each focal top-candidate window to the 10,000 background SNPs, calculating the W statistics and converting into a Z-score. Empirical P-values were then calculated by comparing the Z-score for each top-candidate window to the null distribution. When individual windows had values of W that exceeded the bounds of the null distribution, their empirical *p*-value was set to the reciprocal of the number of genes in the null distribution. For each species pair, we refer to the windows identified as significant by this test as "Windows of Repeated Association" or WRAs.

Linkage disequilibrium and detection of clusters of repeated association

Linkage disequilibrium among adjacent genomic windows can result in statistical non-independence and similar GEA/GWAS signatures across many windows. To identify the most significant WRAs and group neighbouring windows with similar signatures of repeatability into a single "Cluster of Repeated Association" (CRA), we used the following approach in each pairwise contrast, and for each environmental variable and phenotype: for each variable, empirical *p*-values for all WRAs were converted to *q*-values to adjust for False Discovery, then beginning with the first significant WRA along the chromosome (*i.e.*, with $q < 0.05$), we compared it to the next closest significant WRA by calculating the squared Pearson correlation coefficient (r^2) on the allele frequencies across all pairs of SNPs, and compared this estimate to a null distribution. To construct the null distribution, we generated a distribution of LD measurements between 10,000 randomly chosen windows with the same physical distance as between two significant WRAs (excluding all WRAs from the null distribution). If the r^2 between the two neighboring significant WRAs was greater than the 95th percentile of the null distribution, we clustered these two windows together.

This process was repeated successively with the next closest neighboring significant WRA, walking out along the chromosome until one of two stopping criteria was reached: 1) the LD between the last two windows did not exceed the 95 percent of the tail distribution or 2) the distance between the initial window and the current window next to it was larger than 1 cM, based on the linkage

map from Todesco et al. (2020). When the first round of clustering stopped due to either of these two criteria, all the clustered windows were removed from the dataset and the process started with the second smallest empirical p-value. By doing this way, each significant WRA will only appear in a single CRA.

Estimating an index of shared standing variation

As genotype calling was conducted separately for each species (due to computational concerns), we estimated the amount of shared standing variation based on counts of shared vs. non-shared SNPs. If two species are evolving independently, the number of shared SNPs should follow a hypergeometric distribution, so we used an approach similar to the C-scores (Yeaman et al. 2018) to calculate the difference between the observed number of shared SNPs and the expectation, scaled by the standard deviation of the hypergeometric. Because of noise and a relatively small number of SNPs per window, we applied this approach on a sliding window basis, including the 5 flanking windows upstream and downstream of each focal window the calculation of its index of shared standing variation.

Correspondence between regions of repeated association and chromosomal rearrangements

To assess the extent of overlap between regions of the genome with repeated signatures of association and previously identified low-recombination haploblocks, we used two approaches. Firstly, for each pairwise species contrast and variable, we constructed a contingency table for the number of top candidates that were significant WRAs vs. not-significant WRAs (by the null-W test), and that did vs. did not fall within a haploblock. We calculated the Pearson's χ^2 statistic on this table, and then permuted the location of haploblocks throughout the genome to construct a null distribution of χ^2 statistics, and calculated the p-value as the proportion of the null that exceeded the observed χ^2 statistic (to account for non-independence of nearby WRAs), which is presented in **Figure 5**. Secondly, we compared observations of the length and number of overlapping regions to expectations based on a randomization approach. For each pair of species, we kept the position of each CRA constant and randomized the position of haploblocks 10,000 times to build a null distribution. By keeping the position of the CRAs constant, we maintained the architecture of adaptation independent from chromosomal rearrangements. We assessed significance by testing whether the observed overlaps between CRAs and haploblocks were more extreme than the 95th percentile of the tail of null distribution, which is presented in Supplementary Figure S7. Details about identifying chromosomal rearrangements can be found in Todesco et al. (2020).

Identifying repeated signatures of association across all taxa

As a complement to the pairwise analysis, we used PicMin (Booker et al. 2022a) to identify windows of the genome with strong signatures of association in multiple (sub-) species. For each environmental variable, association signatures for each window are ranked genome-wide and PicMin identifies windows with extreme ranks in multiple species. We ran the analysis once with each of the two 3-way comparisons involving one *petiolaris* subspecies (*i.e.*, *H. annuus, H. argophyllus,* and either *H. pet. petiolaris* or *H. pet. fallax*) to control for repeatability arising from similar patterns in the two *petiolaris* subspecies (due to high introgression/shared standing variation).

Evaluating genotypic redundancy using C-scores

Genome-wide quantification of the repeatability of association statistics provides insights into the amount of genotypic redundancy underlying a trait or environmental adaptation, and can be assessed using the C-score approach (Yeaman et al. 2018). Briefly, for a given trait or environment, the set of association test scores within each species is classified into "associated" or "non-associated" using the binomial top candidate approach. For a given pair of species (i = 1, 2),

the observed number of associated windows in each species (a_1, a_2) can be compared to the number of windows that are associated in both species (a_b) and the total number of windows being analysed (a_x). Under a null hypothesis where all windows in the genome have equal potential to be involved in adaptation (*i.e.*, associated with the trait or environment), the expected number of windows associated in both species will be described by a hypergeometric distribution, with the expectation ($a_b = a_1 a_2 / a_x$. The difference between the observed and expected amount of overlap in association scores can be quantified as a C-score by scaling the difference by the standard deviation of the hypergeometric (*i.e.*, a C-score of 2 means that the observed amount of overlap is two standard deviations above the expectation under randomness).

We assess the C-scores for each phenotype and environment trait by classifying the top 0.5% of all 5 kb windows within each species as "associated" (a_i; based on the binomial top candidate index), and begin by calculating the C-score obtained when a_x is set to the union of a_i across all four focal species/sub-species (*i.e.*, only those windows associated in at least one species are included in a_x, a number much lower than the total number of windows in the genome). This yields a negative C-score, as a random draw from such a small number of windows tends to yield many more overlapping associations by chance than the observation. When the C-score = 0, it means that the observed overlap between the pair of taxa being considered is consistent with a random draw of their respective "associated" windows from an overall pool of a_x windows. Thus, by adding rows to the matrix with "non-associated" scores for all 4 species/sub-species until finding the matrix that yields C-score = 0, we can estimate the effective number of loci that contributed to variation for the trait or environment being considered (L_{eff}), under the assumption that all such $a_x = L_{eff}$ windows had an equal chance of contributing to the associations.

To run the C-score analysis on LD-clustered windows, we ran the following algorithm for each trait/environmental variable: for each species, we identified all CRAs that also had a top candidate index in top 0.5%. In most cases, a large cluster with associated windows in one species corresponds to either no clusters or a small cluster in another species. To harmonize cluster boundaries across species, we bin any overlapping clusters together to their maximum extent, and take the average top candidate index for each cluster in each species as either: (a) the mean across all windows that were "associated" in that species or (b) the average across all windows associated in any species, if no windows were associated in that species. This yields a matrix that can be submitted to the hypergeometric C-score analysis.

Data availability

All scripts used for analysis and figures are deposited in a Dryad archive at doi:10.5061/dryad.wpzgmsbtd. Raw sequence data are deposited in the Sequence Read Archive (SRA) under BioProject accessions PRJNA532579, PRJNA398560, and PRJNA564337, as described in Todesco et al. (2020).

Acknowledgements

The authors would like to thank the Yeaman and Peichel labs for comments. Computational support was provided by the Digital Resources Alliance of Canada, and funding was provided by Genome Canada and Genome BC (LSARP2014-223SUN), the International Consortium for Sunflower Genomic Resources, Alberta Innovates, and NSERC Discovery. GLO was funded by a Banting Postdoctoral Fellowship.

References

Acharya BR, Raina S, Maqbool SB, Jagadeeswaran G, Mosher SL, Appel HM, Schultz JC, Klessig DF, Raina R (2007) **Overexpression of CRK13, an Arabidopsis cysteine-rich receptor-like kinase, results in enhanced resistance to Pseudomonas syringae: Overexpression of CRK13 induces pathogen defense** *The Plant Journal* **50**:488–499

Auge GA, Penfield S, Donohue K (2019) **Pleiotropy in developmental regulation by flowering-pathway genes: is it an evolutionary constraint?** *New Phytol* **224**:55–70

Aukerman MJ, Sakai H (2003) **Regulation of Flowering Time and Floral Organ Identity by a MicroRNA and Its APETALA2 -Like Target Genes** *Plant Cell* **15**:2730–2741

Balasubramanian S, Sureshkumar S, Agrawal M, Michael TP, Wessinger C, Maloof JN, Clark R, Warthmann N, Chory J, Weigel D (2006) **The PHYTOCHROME C photoreceptor gene mediates natural variation in flowering and growth responses of Arabidopsis thaliana** *Nat Genet* **38**:711–715

Barghi N, Hermisson J, Schlötterer C (2020) **Polygenic adaptation: a unifying framework to understand positive selection** *Nat Rev Genet* **21**:769–781

Barton NH, Etheridge AM, Véber A (2017) **The infinitesimal model: Definition, derivation, and implications** *Theoretical Population Biology* **118**:50–73

Bhatia N, Bozorg B, Larsson A, Ohno C, Jönsson H, Heisler MG (2016) **Auxin Acts through MONOPTEROS to Regulate Plant Cell Polarity and Pattern Phyllotaxis** *Current Biology* **26**:3202–3208

Blackman BK, Scascitelli M, Kane NC, Luton HH, Rasmussen DA, Bye RA, Lentz DL, Rieseberg LH (2011) **Sunflower domestication alleles support single domestication center in eastern North America** *Proc. Natl. Acad. Sci. U.S.A* **108**:14360–14365

Bohutínská M, Vlček J, Yair S, Laenen B, Konečná V, Fracassetti M, Slotte T, Kolář F (2021) **Genomic basis of parallel adaptation varies with divergence in Arabidopsis and its relatives** *Proc. Natl. Acad. Sci. U.S.A* **118**

Booker TR, Yeaman S, Whitlock MC (2020) **Variation in recombination rate affects detection of outliers in genome scans under neutrality** *Mol Ecol* **29**:4274–4279

Booker TR, Yeaman S, Whitlock MC (2023) **Using genome scans to identify genes used repeatedly for adaptation** *Evolution* **77**:801–811

Booker TR, Yeaman S, Whitlock MC (2023) **The WZA: A window-based method for characterizing genotype-environment association** *Mol Ecol Res (in press)*

Bürger R, Akerman A (2011) **The effects of linkage and gene flow on local adaptation: A two-locus continent–island model** *Theoretical Population Biology* **80**:272–288

Chen A, Li C, Hu W, Lau MY, Lin H, Rockwell NC, Martin SS, Jernstedt JA, Lagarias JC, Dubcovsky J (2014) **PHYTOCHROME C plays a major role in the acceleration of wheat flowering under long-day photoperiod** *Proc. Natl. Acad. Sci. U.S.A* **111**:10037–10044

Cheval C *et al.* (2017) **PRR2, a pseudo-response regulator, promotes salicylic acid and camalexin accumulation during plant immunity** *Sci Rep* **7**

Coughlan JM, Willis HJ (2018) **Dissecting the role of a large chromosomal inversion in life history divergence throughout the Mimulus guttatus species complex** *Mol Ecol* **28**:1343–1357

DeRaad DA, Cobos ME, Alkishe A, Ashraf U, Ahadji-Dabla KM, Nuñez-Penichet C, Peterson AT (2021) **Genome-environment association methods comparison supports omnigenic adaptation to ecological niche in malaria vector mosquitoes** *Molecular Ecology* **30**:6468–6485

Dong MA, Farré EM, Thomashow MF (2011) **CIRCADIAN CLOCK-ASSOCIATED 1 and LATE ELONGATED HYPOCOTYL regulate expression of the C-REPEAT BINDING FACTOR (CBF) pathway in Arabidopsis** *Proc. Natl. Acad. Sci. U.S.A* **108**:7241–7246

Gautier M (2015) **Genome-Wide Scan for Adaptive Divergence and Association with Population-Specific Covariates** *Genetics* **201**:1555–1579

Grimm DG *et al.* (2017) **easyGWAS: A Cloud-Based Platform for Comparing the Results of Genome-Wide Association Studies** *Plant Cell* **29**:5–19

Günther T, Coop G (2013) **Robust Identification of Local Adaptation from Allele Frequencies** *Genetics* **195**:205–220

Hager ER, Harringmeyer OS, Wooldridge TB, Theingi S, Gable JT, McFadden S, Neugeboren B, Turner KM, Jensen JD, Hoekstra HE (2022) **A chromosomal inversion contributes to divergence in multiple traits between deer mouse ecotypes** *Science* **377**:399–405

Heiser C, Smith D, Clevenger S, Martin W (1969) **The North American sunflowers (Helianthus)** *Memoirs of the Torrey Botanical Club* **22**:1–218

Heiser CB (1951) **Hybridization in the Annual Sunflowers: Helianthus annuus x H. argophyllus** *The American Naturalist* **85**:65–72

Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, Poss ML, Reed LK, Storfer A, Whitlock MC (2016) **Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions** *The American Naturalist* **188**:379–397

Höllinger I, Pennings PS, Hermisson J (2019) **Polygenic adaptation: from sweeps to subtle frequency shifts** *PLoS Genet* **15**

Izawa T (2007) **Adaptation of flowering-time by natural and artificial selection in Arabidopsis and rice** *Journal of Experimental Botany* **58**:3091–3097

Jeffreys H (1961) **Theory of probability**

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S, Freimer NB, Sabatti C, Eskin E (2010) **Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S, Freimer NB, Sabatti C, Eskin E. 2010. Variance component model to account for sample structure in genome-wide association studies. Nat Genet 42:348–354.** *Variance component model to account for sample structure in genome-wide association studies. Nat Genet* **42**

Keuskamp DH, Pollmann S, Voesenek LACJ, Peeters AJM, Pierik R (2010) **Auxin transport through PIN-FORMED 3 (PIN3) controls shade avoidance and fitness during competition** *Proc. Natl. Acad. Sci. U.S.A* **107**:22740–22744

Kirkpatrick M, Barton N (2006) **Chromosome Inversions, Local Adaptation and Speciation** *Genetics* **173**:419–434

Láruson ÁJ, Yeaman S, Lotterhos KE (2020) **The Importance of Genetic Redundancy in Evolution** *Trends in Ecology & Evolution* **35**:809–822

Lee C-R *et al.* (2017) **Young inversion with multiple linked QTLs under selection in a hybrid zone** *Nat Ecol Evol* **1**

Liu Z, Jia Y, Ding Y, Shi Y, Li Z, Guo Y, Gong Z, Yang S (2017) **Plasma Membrane CRPK1-Mediated Phosphorylation of 14-3-3 Proteins Induces Their Nuclear Import to Fine-Tune CBF Signaling during Cold Response** *Molecular Cell* **66**:117–128

Lotterhos KE, Whitlock MC (2014) **Evaluation of demographic history and neutral parameterization on the performance of F ST outlier tests** *Mol Ecol* **23**:2178–2192

MacPherson A, Nuismer SL (2017) **The probability of parallel genetic evolution from standing genetic variation** *J. Evol. Biol* **30**:326–337

Manceau M, Domingues VS, Linnen CR, Rosenblum EB, Hoekstra HE (2010) **Convergence in pigmentation at multiple levels: mutations, genes and function** *Phil. Trans. R. Soc. B* **365**:2439–2450

Mizoguchi T, Wheatley K, Hanzawa Y, Wright L, Mizoguchi M, Song HR, Carré IA, Coupland G (2002) **LHY and CCA1 are partially redundant genes required to maintain circadian rhythms in Arabidopsis** *Dev Cell* **2**:629–641

Moyers BT, Rieseberg LH (2016) **Remarkable life history polymorphism may be evolving under divergent selection in the silverleaf sunflower** *Mol Ecol* **25**:3817–3830

Ostevik KL, Andrew RL, Otto SP, Rieseberg LH (2016) **Multiple reproductive barriers separate recently diverged sunflower ecotypes: reproductive barriers between sunflower ecotypes** *Evolution* **70**:2322–2335

Para A, Farré EM, Imaizumi T, Pruneda-Paz JL, Harmon FG, Kay SA (2007) **PRR3 Is a vascular regulator of TOC1 stability in the Arabidopsis circadian clock** *Plant Cell* **19**:3462–3473

Pekker I, Alvarez JP, Eshed Y (2005) **Auxin response factors mediate Arabidopsis organ asymmetry via modulation of KANADI activity** *Plant Cell* **17**:2899–2910

Rosenblum EB, Parent CE, Brandt EE (2014) **The Molecular Basis of Phenotypic Convergence** *Annu. Rev. Ecol. Evol. Syst* **45**:203–226

Sambatti JBM, Strasburg JL, Ortiz-Barrientos D, Baack EJ, Rieseberg LH (2012) **Reconciling extremely strong barriers with high levels of gene exchange in annual sunflowers: gene flow across a strong reproductive barrier** *Evolution* **66**:1459–1473

Schaal SM, Haller BC, Lotterhos KE (2022) **Inversion invasions: when the genetic basis of local adaptation is concentrated within inversions in the face of gene flow** *Phil. Trans. R. Soc. B* **377**

Strasburg JL, Rieseberg LH (2008) **Molecular demographic history of the annual sunflowers helianthus annuus and h. Petiolaris -large effective population sizes and rates of long-term gene flow** *Evolution* **62**:1936–1950

Takano M *et al.* (2005) **Distinct and Cooperative Functions of Phytochromes A, B, and C in the Control of Deetiolation and Flowering in Rice** *The Plant Cell* **17**:3311–3325

Tittes S, Lorant A, McGinty S, Doebley JF, Holland JB, Sánchez-González J de J, Seetharam A, Tenaillon M, Ross-Ibarra J. (2021) **Tittes S, Lorant A, McGinty S, Doebley JF, Holland JB, Sánchez-González J de J, Seetharam A, Tenaillon M, Ross-Ibarra J. 2021. Not so local: the population genetics of convergent adaptation in maize and teosinte. Evolutionary Biology Available from: http://biorxiv.org/lookup/doi/10.1101/2021.09.09.459637** *Not so local: the population genetics of convergent adaptation in maize and teosinte. Evolutionary Biology*

Todesco M *et al.* (2020) **Massive haplotypes underlie ecotypic differentiation in sunflowers** *Nature* **584**:602–607

Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, Yang J (2017) **10 Years of GWAS Discovery: Biology** *Function, and Translation. The American Journal of Human Genetics* **101**:5–22

Wang L, Josephs EB, Lee KM, Roberts LM, Rellán-Álvarez R, Ross-Ibarra J, Hufford MB, Purugganan M (2021) **Molecular Parallelism Underlies Convergent Highland Adaptation of Maize Landraces** *Molecular Biology and Evolution* **38**

Wang T, Hamann A, Spittlehouse DL, Murdock TQ (2012) **ClimateWNA—High-Resolution Spatial Climate Data for Western North America** *Journal Of Applied Meteorology And Climatology* **51**

Wellenreuther M, Bernatchez L (2018) **Eco-Evolutionary Genomics of Chromosomal Inversions** *Trends in Ecology & Evolution* **33**:427–440

Whitlock M, Schluter D (2020) **Whitlock M, Schluter D. 2020. The analysis of biological data. 3rd ed. Macmillan** *The analysis of biological data. 3rd ed. Macmillan*

Yang J, Lee SH, Goddard ME, Visscher PM (2011) **GCTA: A Tool for Genome-wide Complex Trait Analysis** *The American Journal of Human Genetics* **88**

Yeaman S (2022) **Evolution of polygenic traits under global vs local adaptation** *Genetics* **220**

Yeaman S, Gerstein AC, Hodgins KA, Whitlock MC. (2018) **Yeaman S, Gerstein AC, Hodgins KA, Whitlock MC. 2018. Quantifying how constraints limit the diversity of viable routes to adaptation.Ralph PL, editor. PLoS Genet 14:e1007717.** *Quantifying how constraints limit the diversity of viable routes to adaptation.Ralph PL, editor. PLoS Genet* **14**

Yeaman S *et al.* (2016) **Convergent local adaptation to climate in distantly related conifers** *Science* **353**:1431–1433

Yeaman S, Whitlock MC (2011) **The genetic architecture of adaptation under migration-selection balance: the genetic architecture of local adaptation** *Evolution* **65**:1897–1911

Article and author information

Shaghayegh Soudi
Department of Biological Sciences, University of Calgary, 2500 University Drive NW, Calgary, Canada, T2N 1N4

Mojtaba Jahani
Department of Biological Sciences, University of Calgary, 2500 University Drive NW, Calgary, Canada, T2N 1N4, Department of Botany, University of British Columbia, 6270 University Blvd, Vancouver, Canada, V6T 1Z4

Marco Todesco
Department of Botany, University of British Columbia, 6270 University Blvd, Vancouver, Canada, V6T 1Z4, Michael Smith Laboratories, University of British Columbia, 6270 University Blvd, Vancouver, Canada, V6T 1Z4, Irving K. Barber Faculty of Science, University of British Columbia Okanagan, 3187 University Way, Kelowna, Canada, V1V 1V7

Gregory L. Owens
Department of Biology, University of Victoria, 3800 Finnerty Rd, Victoria, Canada, V8P 5C2

Natalia Bercovich
Department of Botany, University of British Columbia, 6270 University Blvd, Vancouver, Canada, V6T 1Z4

Loren H. Rieseberg
Department of Botany, University of British Columbia, 6270 University Blvd, Vancouver, Canada, V6T 1Z4

Sam Yeaman
Department of Biological Sciences, University of Calgary, 2500 University Drive NW, Calgary, Canada, T2N 1N4
For correspondence: samuel.yeaman@ucalgary.ca

Copyright

Editors

Reviewing Editor
Graham Coop
University of California, Davis, United States of America

Senior Editor
Christian Landry
Université Laval, Canada

Reviewer #1 (Public Review):

Soudi, Jahani et al. provide a valuable comparative study of local adaptation in four species of sunflowers, and investigate the repeatability of observed genomic signals of adaptation and their link to haploblocks, known to be numerous and important in this system. The study builds on previous work in sunflowers that have investigated haploblocks in those species and on methodologies developed to look at repeated signals of local adaptations. The authors provide solid evidence of both genotype-environment associations (GEA) and genome-wide association study (GWAS), as well as phenotypic correlations with the environment, to show that part of the local adaptation signal is repeatable and significantly co-occur in regions harboring haploblocks. Results also show that part of the signal is species specific and points to high genetic redundancy. This work will be of interest to evolutionary biologists in general and population geneticists in particular, and constitutes a good example of comparative local adaptation. Importantly, this study helps in advancing our understanding of the genetic architecture implicated in the adaptation process.

Strenghts: The authors take great care in acknowledging and investigating the multiple biases inherent to the used methods (GEA and GWAS) and use conservative and well thought statistical approaches to draw their conclusions. Additionally, I appreciated the nuanced discussion and can only agree with the authors that the adaptation process is complex and does not fully fit the classic simplified genetics models of either few large effect genes or only infinitesimal quantitative traits. I find the added Summary figure of this revised version (S1) extremely helpful in better understanding the different analysis steps and how they relate to the different questions.

Weaknesses: After those revisions, I did not find any major weakness and am satisfied with the authors responses.

Reviewer #2 (Public Review):

In this study the authors sought to understand the extent of similarity among species in intraspecific adaptation to environmental heterogeneity at the phenotypic and genetic levels. A particular focus was to evaluate if regions that were associated with adaptation within putative inversions in one species were also candidates for adaptation in another species that lacked those inversions. This study is timely for the field of evolutionary genomics, due to recent interest surrounding how inversions arise and become established in adaptation.

Major strengths-

Their study system was well suited to addressing the aims, given that the different species of sunflower all had GWAS data on the same phenotypes from common garden experiments as well as landscape genomic data, and orthologous SNPs could be identified. Organizing a dataset of this magnitude is no small feat. The authors integrate many state-of-the-art statistical methods that they have developed in previous research into a framework for correlating genomic Windows of Repeated Association (WRA, also amalgamated into Clusters of Repeated Association based on LD among windows) with Similarity In Phenotype-Environment Correlation (SIPEC). The WRA/CRA methods are very useful and the authors do an excellent job at outlining the rationale for these methods.

Weaknesses-

The authors did an excellent job responding to the first set of reviews and overall I found the manuscript more streamlined and easier to read. The main weakness in the manuscript is that correlations among environmental variables were not controlled for in their results, and is a source of potential pseudoreplication. The authors are clear about the results that are affected by pseudoreplication.

The manuscript shows how to integrate many recent methods to study the repeatability of adaptation, and the methods and data are likely to be used in similar studies.

Author Response

The following is the authors' response to the original reviews.

> ***Reviewer #1 (Public Review):***
>
> *Soudi, Jahani et al. provide a valuable comparative study of local adaptation in four species of sunflowers and investigate the repeatability of observed genomic signals of adaptation and their link to haploblocks, known to be numerous and important in this system. The study builds on previous work in sunflowers that have investigated haploblocks in those species and on methodologies developed to look at repeated signals of local adaptations. The authors provide solid evidence of both genotype-environment associations (GEA) and genome-wide association study (GWAS), as well as phenotypic correlations with the environment, to show that part of the local adaptation signal is repeatable and significantly co-occur in regions harboring haploblocks. Results also show that part of the signal is species specific and points to high genetic redundancy. The authors rightfully point out the complexities of the adaptation process and that the truth must lie somewhere between two extreme models of evolutionary genetics, i.e. a population genetics view of large effect loci and a quantitative genetics model. The authors take great care in acknowledging and investigating the multiple biases inherent to the used methods (GEA and GWAS) and use a conservative approach to draw their conclusions. The multiplicity of analyses and their interdependence make them slightly hard to understand and the manuscript would benefit from more careful explanations of concepts and logical links throughout. This work will be of interest to evolutionary biologists and population geneticists in particular, and constitutes an additional applied example to the comparative local adaptation literature.*
>
> *Some thoughts on the last paragraph of the discussion (L481-497): I think it would be fine to have some more thoughts here on the processes that could contribute to the presence/absence of inversions, maybe in an "Ideas and Speculation" subsection. To me, your results point to the fact that though inversions are often presented as important for local adaptation, they seem to be highly contingent on the context of adaptation in each species. First, repeatability results are only at the window/gene level in your results, the specific mutations are not under scrutiny. Is it possible that inversions are only necessary when sets of small effect mutations are used, opposite to a large effect mutation in other species? Additionally, in a model with epistasis, fitness effects of mutations are dependent on the genomic background and it is possible that inversions were necessary in only certain contexts, even for the same mutations, i.e. some adaptive path contingency. Finally, do you have specific demographic history knowledge in this system that maps to the observations of the presence of inversions or not? For example, have the species "using" inversions been subject to more gene flow compared to others?*

Thank you for the great suggestions and helpful comments. Regarding the question of demography, each of the species actually harbours quite a large number of haploblocks (13 in H. annuus spanning 326Mb, 6 in H. argophyllus spanning 114 Mb, and 18 in H. petiolaris spanning 467 Mb; see Todesco et al. 2020 for more details) so there does not seem to be any clear association with demography. We agree about the complexities that might underly the evolution of inversions that you outline above, and have refined some of the text where we discuss their evolution in the Discussion.

> ***Reviewer #2 (Public Review):***
>
> *In this study the authors sought to understand the extent of similarity among species in intraspecific adaptation to environmental heterogeneity at the phenotypic and genetic levels. A particular focus was to evaluate if regions that were associated with adaptation within putative inversions in one species were also candidates for adaptation in another species that lacked those inversions. This study is timely for the field of evolutionary genomics, due to recent interest surrounding how inversions arise and become established in adaptation.*
>
> *Major strengths*
>
> *Their study system was well suited to addressing the aims, given that the different species of sunflower all had GWAS data on the same phenotypes from common garden experiments as well as landscape genomic data, and orthologous SNPs could be identified. Organizing a dataset of this magnitude is no small feat. The authors integrate many state-of-the-art statistical methods that they have developed in previous research into a framework for correlating genomic Windows of Repeated Association (WRA, also amalgamated into Clusters of Repeated Association based on LD among windows) with Similarity In Phenotype-Environment Correlation (SIPEC). The WRA/CRA methods are very useful and the authors do an excellent job at outlining the rationale for these methods.*

Thank you!

> *Major weaknesses*
>
> *The study results rely heavily on the SIPEC measure, but I found the values reported difficult to interpret biologically. For example, in Figure 4 there is a range of SIPEC from 0 to 0.03 for most species pairs, with some pairs only as high as ~0.01. This does not appear to be a high degree of similarity in phenotype-environment correlation. For example, given the equation on line 517 for a single phenotype, if one species has a phenotype-environment correlation of 1.0 and the other has a correlation of 0.02, I would postulate that these two species do not have similar evolutionary responses, but the equation would give a value of (1+0.02)10.02/1 = 0.02 which is pretty typical "higher" value in Figure 4. I also question the logic behind using absolute values of the correlations for the SIPEC, because if a trait increases with an environment in one species but decreases with the environment in another species, I would not predict that the genetic basis of adaptation would be similar (as a side note, I would not question the logic behind using absolute correlations for associations with alleles, due to the arbitrary nature of signing alleles). I might be missing something here, so I look forward to reading the author's responses on these thoughts.*

The reviewer makes a very good point about the range of SIPEC, and we have changed our analysis to reflect this, now reporting the maximum value of SIPEC for each environment (across the axes of the PCA on phenotypes that cumulatively explain 95% of the variance), in Figure 4 and Supplementary Figures S2 and S13. For consistency among manuscript versions and to illustrate the effect of this change, we retain the mean SIPEC value in one figure in the supplementary materials (S12), which shows the small effect of this change on the qualitative patterns. Figure 4 now shows that the maximum SIPEC value is regularly quite strong, which should address the reviewer's concern that this is not being driven by anomalous and small values. We appreciate this point and think this change now more closely reflects how we are trying to estimate the biological feature of interest – that some axis of phenotypic space is strongly (or not) responding to selection from the environmental variable.

With respect to the logic behind using absolute value, we still feel this is justified for traits, because if a trait evolves to be bigger or smaller, it may still use the same genes. For example, flowering time may change to be later or earlier, which would result in opposite correlations with a given environment, but might use the same gene (e.g. FT) for this. As such, we think keeping absolute value is more representative as otherwise species with strong but opposite patterns of adaptation would look like they were very different. We have added a statement on line 584 in the methods section to further clarify the reason for this choice.

> *An additional potential problem with the analysis is that from the way the analysis is presented, it appears that the 33 environmental variables were essentially treated as independent data points (e.g. in Figure 4, Figure 5). It's not appropriate to treat the environmental variables independently because many of them are highly correlated. For example in Figure 4, many of the high similarity/CRA values tend to be categorized as temperature variables, which are likely to be highly correlated with each other. This seems like a type of pseudo replication and is a major weakness of the framework.*

This is a good point and we fully agree. It is for this reason that we didn't present any p-values or statistical tests of the overall patterns that are shown in these figures (i.e. the linear relationship between SIPEC and number of CRAs in figure 4 and the tendency for most points to fall above the 1:1 line in figure 5). But to make sure this is even more clear, we have added statements to the captions of these figures to remind readers that points are non-independent. We still feel that in the absence of a formal test, the overall patterns are strongly consistent with this interpretation. A smaller number of non-pseudo-replicated points in Figure 4 would still likely show linear patterns. Similarly, there are almost no significant points falling below the 1:1 line in Figure 5, and it seems unlikely that pseudoreplication would generate this pattern.

> *Below I highlight the main claims from the study and evaluate how well the results support the conclusions.*
>
> *"We find evidence of significant genome-wide repeatability in signatures of association to phenotypes and environments" (abstract)*
> *Given the questions above about SIPEC, I did not find this conclusion well supported with the way the data are presented in the manuscript.*

We have changed the reporting of the SIPEC metric so that it more clearly reflects whichever axis of phenotypic space is most strongly correlated with environment in both species (using max instead of mean). This shows similar qualitative patterns but illustrates that this happens across much higher values of SIPEC, showing that it is in fact driven by high correlations in each species (or non-similar correlations resulting in low values of SIPEC). While we agree about the pseudo-replication problem preventing formal statistical test of this hypothesis, the visual pattern is striking and seems unlikely to be an artefact, so we think this does still support this conclusion.

> *"We find evidence of significant genome-wide repeatability in signatures of association to phenotypes and environments, which are particularly enriched within regions of the genome harbouring an inversion in one species. " (Abstract) And "increased repeatability found in regions of the genome that harbour inversions" (Discussion)*
> *These claims are supported by the data shown in Figure 4, which shows that haploblocks are enriched for WRAs. I want to clarify a point about the wording here, as my understanding of the analysis is that the authors test if* haploblocks *are enriched with* WRAs*, not whether* WRAs *are enriched for* haploblocks*. The wording of the abstract is claiming the latter, but I think what they tested was the former. Let me know if I'm missing something here.*

We are actually not interested in whether WRAs are enriched for haploblocks; we want to know if WRAs tend to occur more commonly within haploblocks than outside of them. We have tried to clarify that this is our aim in various places in the manuscript. Our analysis for Figure 5 is the one supporting these claims, and it uses the Chi-square test statistic to assess the number of WRAs and non-WRAs that fall within vs. outside of inversions, and a permutation test to assess the significance of this observation, for each environmental variable and phenotype. We don't think that this test has any direction to it – it's simply testing if there is non-random association between the levels of the two factors. Thus, we think the wording we have used is consistent with the test result and our aims. Perhaps the confusion arose from the two methods that we present in the Methods (one is used for Figure 5, the other for Figure S6C & D), so we have added clarifications there.

> *Notwithstanding the concerns about highly correlated environments potentially inflating some of the patterns in the manuscript, to my knowledge this is the first attempt in the literature to try this kind of comparison, and the results does generally suggest that inversions are more likely capturing, rather than accumulating adaptive variation. However, I don't think the authors can claim that repeated signatures are enriched with haploblock regions, and the authors should take care to refrain from stating the relative importance of different regions of the genome to adaptation without an analysis.*

Actually, we don't have a strong feeling about whether inversions are capturing vs. accumulating adaptive variation, as these results could be consistent with either. As described above, we do not understand why we can't claim that repeated signatures are enriched within haploblocks. We thought the reviewer is perhaps referring to the fact that the points are pseudo-replicated in the figures due to environment? We note that a very large number of points are significantly different from random in terms of the distribution of WRAs within vs. outside of haploblocks (light- vs. dark-shaded symbols), and that almost all of them fall above the 1:1 line. While there may be pseudo-replication preventing a test of the bigger multi-environment/multi-species hypothesis across all phenotypes and environments, there is almost a complete lack of significant results in the other direction. This seems like quite strong evidence about enrichment of WRAs within haploblocks, across many environments/species contrasts. We have added some text to the description of patterns in figure 5 to try to clarify this.

> *"While a large number of genomic regions show evidence of repeated adaptation, most of the strongest signatures of association still tend to be species-specific, indicating substantial genotypic redundancy for local adaptation in these species." (Abstract) Figure 3B certainly makes it look like there is very little similarity among species in the genetic basis of adaptation, which leaves the question as to how important the repeated signatures really are for adaptation if there are very few of them. (Is 3B for the whole genome or only that region?). This result seems to be at odds with the large number of CRAs and the claims about the importance of haploblock regions to adaptation, which extend from my previous point.*

Figure 3B is for the whole genome, we have added text to the figure caption to clarify this. We think that both interpretations are possible: that most of the regions of the genome that are driving adaptation are non-repeated, but that a small but significant proportion of regions driving adaptation are repeated above what would be expected at random. Thus, it seems that there is high redundancy, coupled with adaptation via some genes that seem particularly functionally important and non-redundant, and therefore repeated. We added clarifying text on lines 541-548.

> *"we have shown evidence of significant repeatability in the basis of local adaptation (Figure 4, 5), but also an abundance of species-specific, non-repeated signatures (Figure 3)"*
> *While the claim is a solid one, I am left wondering how much of these genomes show repeated vs. non-repeated signatures, how much of these genomes have haploblocks, and how much overlap there really is. Finding a way to intuitively represent these unknowns would greatly strengthen the manuscript.*

We agree, and really struggled to find the best way to communicate both the repeated patterns and the large amount of non-repeated signatures. Unfortunately, we have more confidence in the validity of repeated patterns because for the non-repeated patterns, a strong signature of association to environment in only one species could just be the product of structureenvironment correlation, as we didn't control for population structure. Thus, trying to quantify the proportion of non-repeated signatures is difficult to do with any accuracy and we preferred to avoid putting too much emphasis on the simple calculation of the proportion of top candidate windows that were also WRAs.

> *Overall, I think the main claims from the study, the statistical framework, and the results could be revised to better support each other.*
>
> *Although the current version of the manuscript has some potential shortcomings with regards to the statistical approaches, and the impact of this paper in its present form could be stifled because the biology tended to get lost in the statistics, these shortcomings may be addressed by the authors.*
>
> *With some revisions, the framework and data could have a high impact and be of high utility to the community.*

Thank you for your very helpful comments and suggestions on our paper, we really appreciate it.

> *Recommendations for the authors: please note that you control which revisions to undertake from the public reviews and recommendations for the authors*
>
> *Editor's comments:*
>
> *The reviewers make a series of reasonable suggestions that I echo. I found the paper quite hard to follow, and got fairly lost in the various layers of analyses done. Partially, this represents the complexity of empirical genomic data, which rarely deliver simple stories of convergence at a few genes. However, the properties of the various statistics used to detail local adaptation and convergence are not particularly clear and the figures presented were not intuitive representations of the data. This leaves the reader with an incomplete view of how much weight to put in the various lines of evidence marshaled. I would suggest simplifying the presentation of the results considerably. I add a few additional comments below.*

Great suggestion, we've added a schematic overview of the methods and main research questions to Figure S1 in the supplementary materials.

> *A figure would help showing some of the signals of SNPs with putative signals of convergent environmental correlations across species, e.g. frequencies plotted against climate variables. This would help readers get a sense of how strong these signals were. These could be accompanied by the statistics calculated for these SNPs, that would allow the reader to start to get some intuitive sense of what the numbers mean.*

Great suggestion, we have added a schematic overview of the methods to Figure S1 that shows some of the values and illustrates how the methods work using visual examples from our data.

> *In general, the introduction and some of the discussion of the inversion results feel oddly framed:*
> *Abstract line 36: "This shows that while inversions may facilitate local adaptation, at least some of the loci involved can still make substantial contributions without the benefit of recombination suppression."*

We have changed “some of the loci involved can still make substantial contributions without the benefit of recombination suppression” here to “some of the loci involved can still harbour mutations that make substantial contributions without the benefit of recombination suppression in species lacking a segregating inversion” as it hopefully clarifies that we’re not talking about individual alleles that are present in both species.

> *Models of the role of local adaptation in the establishment of inversions (Kirkpatrick & Barton) assume that there are multiple locally adapted alleles already present. It is the load created by these alleles being constantly maintained in the face of migration and subsequent recombination that allow an inversion to be selected for because it keeps together locally adapted alleles. Thus these models predict that there could well be standing local adaptation at these loci in the absence of the inversion in other species, and that these locally adapted alleles while not fixed may be at high frequency. (After establishment, inversions housing locally adapted alleles, can shield more weakly, locally beneficial alleles from migration allow other alleles to build up.) Empirically it's interesting to find signals of local adaptation in other species that don't contain putative inversions. But the logic of the different predictions is not particularly clear from the introduction, and only becomes somewhat clearer in the discussion.*

Thank you for pointing out this murkiness, we have re-written portions of both the Introduction and Discussion to clarify this aspect.

From the introduction: Inversions have been implicated in local adaptation in many species (Wellenreuther and Bernatchez 2018), likely due to their effect to suppress recombination among inverted and noninverted haplotypes, and thereby maintain LD among beneficial combinations of locally adapted alleles (Rieseberg 2001; Noor et al. 2001; Kirkpatrick and Barton 2006). This has been approached by models studying the establishment of inversions that capture combinations of locally adapted alleles present as standing variation (e.g., Kirkpatrick and Barton 2006), as well as models examining the accumulation of locally adapted mutations within inversions (e.g., Schaal et al. 2022). If there is variation in the density of loci that can potentially contribute to local adaptation, inversions would be expected to preferentially establish and be retained in regions harbouring a high density of such loci (and this expectation would hold for both the capture and accumulation models). We would also expect to see stronger signatures of repeated local adaptation in such high density regions. Despite mounting evidence of their importance in adaptation, it is unclear how inversions may covary with repeatability of adaptation among species. A fundamental parameter of importance in these models is the relationship between migration rate and strength of selection on individual alleles, which may not make persistent contributions to local adaptation without the suppressing effects of recombination if selection is too weak (Yeaman and Whitlock 2011; Bürger and Akerman 2011). If most alleles have small effects relative to migration rate and can only contribute to local adaptation via the benefit of the recombination-suppressing effect of an inversion, then we would expect little repeatability at the site of an inversion – other species lacking the inversion would not tend to use that same region for adaptation because selection would be too weak for alleles to persist. On the other

hand, if some loci are particularly important for local adaptation and regularly yield mutations of large effect, with these patterns being conserved among species, repeatability within regions harbouring inversions may be substantial. Thus, studying whether adaptation at the same genomic region harbouring an inversion is observed in other species lacking the inversion can give insights about the underlying architecture of adaptation, and the evolution and maintenance of inversions.

From the Discussion: The observed repeatability associated with inversions further supports the local adaptation model as an explanation for the long-term persistence of segregating inversions (at least in sunflowers, rather than mechanisms based on dominance or meiotic drive (Rieseberg 2001). If there is variation across the genome in the density of loci with the potential to be involved in local adaptation, then the establishment and maintenance of inversions would be biased towards regions harbouring a high density such loci under this model. If the genomic basis for local adaptation is conserved amongst species, then these same regions are more likely to have high repeatability. Thus, our observation of genomic regions harbouring inversions also being enriched for WRAs is consistent with this general model for inversion evolution. Unfortunately, our observations do not provide much insight into whether inversions evolve through the capture (e.g. Kirkpatrick and Barton 2006) or accumulation (e.g. Schaal et al. 2022) type of model, as either model would be consistent with our results. Most of the sunflower inversions are >1 My old, and therefore predate any current local adaptation patterns, but likely do not predate the genes underlying local adaptation (which appear to be shared among the species we studied). As for the alleles underlying local adaptation, they may be younger than the inversions, but as our work suggests, these regions are prone to harbouring locally adaptive alleles so it is possible that they also harboured other ancestral locally adaptive alleles.

> *As a minor comment, there's a fair number of places where a more nuanced view of the field is needed, e.g.:*
> *"Models in evolutionary genetics tend to focus on extremes: population genetic approaches explore cases where strong selection deterministically drives a change in allele frequency" --This seems like a strange strawman. Population genetic models span a huge parameter range. The empirical approaches of looking for sweeps by detecting genome-wide statistical outliers is predicated on strong selection, but there are numerous papers that have looked for signals of weak selection genome-wide.*

Good point, we have changed our wording here.

> ***Reviewer #1 (Recommendations For The Authors):***
>
> *Comments*
>
> *My main comment on the manuscript is that the different levels and diversity of analyses are slightly hard to follow on the first, and even second, read. As there are several layers of correlations and comparisons, as well as some independent analyses, I wonder if it might be helpful to have a summary schematic figure of how all analyses fit together.*

Great idea, we have added Figure S1 that summarizes the main flow of the methods and research questions.

> - *L169-171: Would it be more accurate to say that SIPEC is maximized when both species have strong correlations for an environmental variable across the same phenotypes? But maybe I misunderstood the index.*

Good point, we have now simplified SIPEC, reporting the max instead of the mean, which we think better reflects when similar patterns are happening in both species for some

phenotype.

> - *L191: Given the discussion in the introduction and elsewhere about the correction for population structure, which version is used here? Same for Figure 3.*

We have added clarification there.

> - *L348: One [environmental] variable?*

Added

> - *L353: Maybe add a percentage indication for 387 so that it is comparable to the following 23.3%.*

Good point, added

-> L388 and paragraph: You mention "significant repeatability" but it is hard from the results at this point to have a broad idea of the amount of signal that is repeatable. Would it be possible to add here some quantitative measure of the proportion of signal repeatable or not, even if approximated?

I wish we could, but I think the precision implied by such an approximation would involve a huge amount of uncertainty and likely inaccuracy. Because it is so hard to conclusively identify how many loci are significant but non-repeated, we really don't have a good handle on the denominator here. We are pretty confident that the repeated loci are strongly enriched for true positives, but the non-repeated loci are also almost certainly strongly enriched for false positives. While we really want to be able to quantify this explicitly, we don't think it's possible given our data.

> *-L415-418: "If there is variation [...] involved in local adaptation", I do not follow this argument, could you rephrase?*

Changed

> *-L447-450: As you say in the supplementary methods, your analyses exclude 3/4 of the genome. Do you think this choice has a large impact on the number of outliers observed here as the genome-wide baseline would change?*

This is a very good question, but one that is quite complex and without a clear answer – we chose not to delve into it in the paper to keep the discussion streamlined. My (SY) feeling is that it is unlikely that regions harbouring transposable elements would contribute much to adaptation, but I think we really don't know if that is true. Even excluding ¾ of the genome harbouring TEs, ¼ of the genome still constitutes a huge amount of sequence and a very large number of genes and it seems plausible that most genes and genic regions would not contribute to adaptation for a given trait, so I don't think this would change the results too much in a qualitative way – but would almost certainly change the number of windows that are significant, etc.

> - *L455-457: "As we are unable [...] potentially important drivers" Could you provide the logical link here between loci of small effect and them being important drivers. I presume you mean that the large effect loci found here only account for a small proportion of the heritability?*

Yes that's what we meant here, so we've added some clarification.

> - *L482: "enriched within inversions" should that be 'in genomic regions where there exist inversions in at least one species'? Thanks for catching that, yes. Changed.*
>
> - *Methods/SIPEC L512: Compared to the Results section it is unclear here what is referred to as an "environment" Is it a variable or a set of environment variables?*

This is done per environmental variable.

> *I find the presence of the PCA for environment variables in Figure 2 misleading as my first interpretation was that PCs for environment were also used.*

Good point, we have clarified this on line 190-193.

> *Maybe one potential addition to the formula would be to add an environment variable j notation such that it reads "$SIPEC_j = \sum_i (|r_{ij,1}| + ...) ...$ where ... between environment variable j". I had initial difficulties to understand how this SIPEC was computed relating to environmental variables and this might help.*

Given the other changes we made to SIPEC, we felt it was simpler to just present it as a single calculation on a given combination of phenotype and environment for a pair of species, and then discuss taking the mean and maximum of this later.

> *Finally, PCA axes explaining 95% of the variance are used, I would find it interesting to see how many PCs are used in comparison to the number of traits being measured.*

We have added the following sentence to the methods describing this:

"For comparisons including H. argophyllus, 95% of the variance was typically explained by 8-10 PC axes (out of 28 or 29 phenotypes), whereas for comparisons among other taxa this included 21 or 22 PC axes (out of 65 or 66 phenotypes."

> *Typos*
>
> *L52: --*

Changed

> *L254: portions [of] their*

Changed

> *L399: additional closing parenthesis*

Changed

> *L458: signatures [of] repeated association*

Changed

> *L554: performed [on]*

Changed

> *L578: 5 kp/kb windows*

Changed

> *L601: casual/causal SNPs*

Changed

> *L615: widow/window*

Changed

> *L732: Banding/Banting Postdoctoral Fellowship*

Changed

> *L1002 & L960: [Supplementary] Figure*

Changed

> *Supplementary: Some figure titles are in bold and others are not.*

Changed

> ***Reviewer #2 (Recommendations For The Authors):***
>
> *Overall I found the writing to be very clear and easy to follow. Despite my comments, it was clear that a lot of thought went into how to conduct the tests and visualize the results. I recommend ending the Discussion on a positive note, rather than an impossible test.*

Thanks for the positive suggestion, we have done this.

> *In Figure 5, is the temperature variable missing in the legend and in the plot?*

No, for this plot we just combined the temperature/precipitation variables into one variable called “climate”.