nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-018-0553-z

Evolution of invasiveness by genetic accommodation

Dan G. Bock[1,5*], Michael B. Kantar[1,2], Celine Caseys[1,3], Remi Matthey-Doret[4] and Loren H. Rieseberg[1]

Invasion success of species introduced to novel environments may be facilitated by adaptive evolution and by phenotypic plasticity. Here we investigate the independent and joint contribution of both mechanisms as drivers of invasiveness in the perennial sunflower *Helianthus tuberosus*. We show that invasive genotypes have multiple origins, and that invasive spread was facilitated by the repeated evolution of extreme values in a single trait, clonality. In line with genetic accommodation theory, we establish that this evolutionary transition occurred by refining a preexisting plastic response of clonality to water availability. Further, we demonstrate that under the non-drought conditions typically experienced by this plant in its introduced range, invasive spread is mediated by hybrid vigour and/or two major additive-effect loci, and that these mechanisms are complementary. Thus, in *H. tuberosus*, evolution of invasiveness was facilitated by phenotypic plasticity, and involved the use of multiple genetic solutions to achieve the same invasiveness trait.

The long-held view that evolutionary change is a slow process and therefore unlikely to contribute to the spread of invasive species has now repeatedly been challenged[1–4]. Increasingly, studies are demonstrating that rapid evolution of invasive populations can have a large impact on their success, either independent of the local environment[5,6] or through adaptation along climatic gradients[7,8]. Despite this progress, few studies have identified the traits that drive invasive spread or the underlying genetic mechanisms (see refs [9–11]). Additionally, phenotypic plasticity may facilitate range expansions, by allowing invasive genotypes to maintain high fitness across environments, or to more efficiently exploit favorable conditions[12–14]. Importantly, adaptive evolution and phenotypic plasticity are not mutually exclusive, and may also interact. For example, if plasticity exposes genetic variation in invasiveness traits, it can jump-start adaptive evolution, by a process known as genetic accommodation[15]. During genetic accommodation, a preexisting plastic response is refined by selection in the introduced range. As a result, the slope and/or elevation of reaction norms are increased for invasive genotypes, which then outperform non-invasive genotypes at one end of the environmental gradient[16–18]. While support for the contribution of adaptive evolution to invasion success has been accumulating, the role of plasticity is still disputed[19]. As well, whether both mechanisms contribute to invasion success via genetic accommodation remains to be established[17].

Here, we investigate these topics in the perennial sunflower *Helianthus tuberosus* (Jerusalem artichoke). Phylogenetic evidence indicates *H. tuberosus* is a polyploid hybrid species, containing a diploid subgenome derived from *H. grosseserratus* and a tetraploid subgenome derived from *H. hirsutus*[20,21]. Native to North America (Supplementary Fig. 1), *H. tuberosus* was introduced to Europe in 1607, and cultivated for its tubers[22]. It remained a minor crop until the 1900s, when a number of breeding programmes were established across the continent[22,23]. During the second half of the twentieth century, as these breeding efforts were abandoned, reports started to accumulate regarding the spread of the species into natural habitats along watercourses[23–25]. These expansions have become increasingly aggressive during the past three decades, such that *H. tuberosus* is now one of the most invasive plants in central-eastern Europe[23–26].

To understand the origin of invasive genotypes, we analyse single-nucleotide polymorphism (SNP) data from a diverse collection of *H. tuberosus* (Supplementary Table 1). We then combine greenhouse and common garden experiments to identify genetically based trait shifts between invasive and non-invasive lineages, and to infer what the important phenotypic determinants of invasion success are. To assess the contribution of phenotypic plasticity to invasive spread, we compare the performance of invasive and non-invasive lineages at putative invasiveness traits across environmental conditions. Lastly, we combine genomic and phenotype data to pinpoint genetic mechanisms of invasiveness.

Results and discussion

Origin of invasive genotypes. A principal component analysis (PCA) based on 27,396 SNPs clustered *H. tuberosus* between its proposed progenitor species (Fig. 1a). This analysis further indicated that the main axis of divergence within *H. tuberosus* is between wild North American samples (hereafter 'native samples') and cultivated samples (Fig. 1a). Wild European (hereafter 'invasive') samples spanned a large fraction of the PC space delineated by native and cultivated samples (Fig. 1a; see also Supplementary Fig. 2), suggesting that invasive genotypes may have a diverse origin. Further supporting this possibility, a maximum-likelihood phylogeny based on the diploid subgenome of *H. tuberosus* did not group invasive samples monophyletically (Supplementary Fig. 3). Because relationships among *H. tuberosus* genotypes may not follow a tree-like pattern, we verified this result using a phylogenetic network. Here as well, invasive *H. tuberosus* grouped in four phylogenetic clusters (origins 1–4; Fig. 1b). These results therefore point towards at least four derivations of invasive genotypes.

[1]Department of Botany, University of British Columbia, Vancouver, British Columbia, Canada. [2]Department of Tropical Plant and Soil Sciences, University of Hawaii, Honolulu, HI, USA. [3]Department of Plant Sciences, University of California Davis, Davis, CA, USA. [4]Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada. [5]Present address: Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA. *e-mail: dan.g.bock@gmail.com

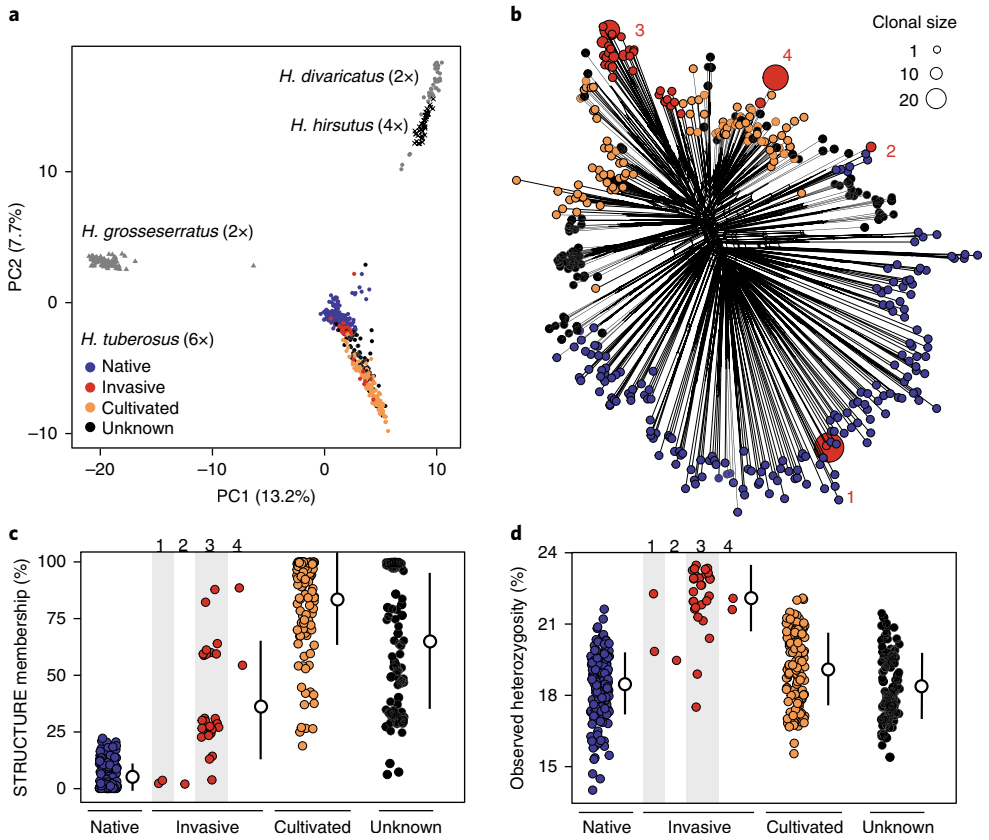

Fig. 1 | Population structure and genetic diversity of *H. tuberosus*. **a**, PCA of 427 *H. tuberosus* non-clonal genotypes and 175 genotypes of the progenitor species. **b**, Phylogenetic network built using the neighbour-net method. For invasive samples only, the size of the circles indicates clonal size, or the number of ramets recovered in invasive populations for each genotype. **c**, STRUCTURE membership coefficients for the $K=2$ model, considering the dominant cluster in cultivated samples. White circles represent the averages (±s.d.) for each category. Numbers above the plot indicate the inferred origins of invasive genotypes. **d**, Genome-wide heterozygosity calculated using 82,957 filtered SNPs. Invasive samples are divided by inferred origin. Kruskal–Wallis tests for differences among groups: $\chi^2_3 = 88.705$, $P < 2.2\times10^{-16}$; pairwise Kruskal–Wallis comparisons using the Nemenyi test, and sequential Bonferroni adjusted P values: $P_{\text{native-cultivated}} = 0.06$, $P_{\text{native-invasive}} = 6.66\times10^{-16}$, $P_{\text{cultivated-invasive}} = 6.65\times10^{-9}$. For all panels, colours indicate membership to the native (blue), invasive (red), cultivated (orange) or unknown (black) groups.

Bayesian STRUCTURE analyses supported the distinction between invasive origins 1–2, represented by three genotypes of native ancestry, and invasive origins 3–4, represented by 32 genotypes with native and cultivated ancestry (Fig. 1c; Supplementary Fig. 4). The assignment of mixed ancestry for most invasive genotypes could be explained by incomplete lineage sorting or hybridization. Two lines of evidence indicate that the latter scenario, that most invasive genotypes have an admixed native by cultivated origin, is more likely. First, most (~85%) invasive genotypes, part of inferred origin 3, are closely related to a number of breeding lines that were obtained from crosses involving native and cultivated genotypes (Supplementary Fig. 3). Second, heterozygosity was elevated in invasive samples (Fig. 1d). This result is consistent throughout much of the genome (Supplementary Fig. 5), and is robust to calculations of heterozygosity based on the complete marker set, or based on markers assigned to the two subgenomes (Supplementary Fig. 6).

Identification and evolution of invasiveness traits. To identify determinants of invasive spread, we compared invasive and non-invasive genotypes grown in a common environment at 20 traits. These traits included allelopathy and clonality, two previously proposed invasiveness traits in this system (Supplementary Table 2; see also Supplementary Discussion). Overall, the first two PC axes of morphological variation recapitulated patterns recovered using SNP markers, and indicated that invasive *H. tuberosus* are intermediate relative to native and cultivated samples (Fig. 2a). Analyses of variance (ANOVAs) confirmed this result. In most cases (~71% of comparisons), including all analyses of allelopathy traits, invasive samples were not differentiated from native and/or cultivated conspecifics (Supplementary Tables 3–5). In interspecific comparisons as well, allelopathy of *H. tuberosus* did not exceed values observed in its non-invasive parental taxa (Supplementary Table 6). These results therefore do not point towards a role of allelopathy in invasiveness for this species.

By contrast, trait analyses supported the possibility that clonality (estimated as number of tubers per plant) is an important invasiveness trait in *H. tuberosus*. In ANOVAs performed without corrections for population genetic structure, invasive *H. tuberosus* was significantly differentiated at seven traits. For three of these, namely tuber number, average tuber weight and tuber shape index, differences were maintained when accounting for genetic structure (Supplementary Tables 4,5). Of these, tuber number was extreme in invasive samples, while average tuber weight and tuber shape index were intermediate. In MCMCglmm analyses also, tuber number, as well as tuber area and tuber perimeter, were identified as significant predictors of invasive status, with and without corrections for genetic structure (Supplementary Table 7). Lastly, the extent of differentiation in tuber number among invasive and non-invasive samples is substantial, and comparable with domestication-related divergence in this species for the total tuber weight, as well as

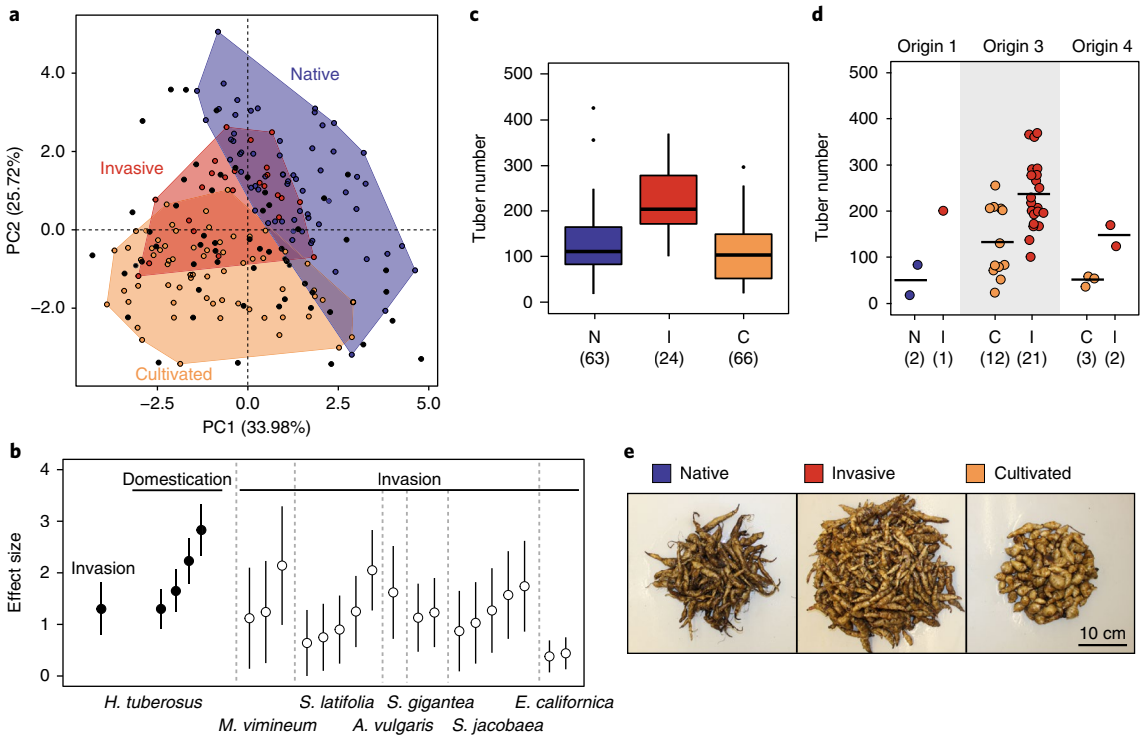

Fig. 2 | Phenotypic basis of invasiveness in *H. tuberosus*. All panels include only tuber-propagated *H. tuberosus* samples. **a**, PCA based on 11 quantitative traits for 160 native (blue), invasive (red) and cultivated (orange) samples, and 60 unknown (black) samples. Polygons enclose native, invasive and cultivated samples. **b**, Hedges' effect size estimates (±95% C.I.) for tuber number in *H. tuberosus* (filled circle, invasion), four candidate domestication traits in *H. tuberosus* (filled circles, domestication) and 18 candidate invasiveness traits in six other species (open circles, invasion; see Supplementary Table 7 for references and trait information). **c**, Box plots of tuber number for native (N), invasive (I) and cultivated (C) *H. tuberosus* indicating the median, the interquartile range and outliers (black circles). **d**, Tuber numbers for invasive samples from three lineages (origins 1, 3 and 4; Fig. 1b; Supplementary Fig. 3) with available phenotype data, and closely related native or cultivated samples with phenotype data. Lines represent the means for each group. For **c** and **d**, sample sizes are given in parentheses. **e**, Tuber yield from single native, invasive and cultivated plants.

divergence at other traits associated with invasiveness in other taxa (Fig. 2b,c; Supplementary Table 8).

Both the extent and the direction of differentiation in tuber number between invasive and non-invasive lineages are as expected if this trait diverged under selection for increased invasiveness in the introduced range (that is, invasive samples are extreme and superior tuber number producers). Similar trends could, however, have been generated by random genetic drift[27]. Two lines of evidence argue against this possibility. First, native and cultivated *H. tuberosus* have been grown extensively in Europe as part of breeding programmes and field trials[22]. As such, it is unlikely that the shift we observe in the distribution of tuber number values in invasive samples is due to biased sampling during a limited number of introduction events. Second, invasive *H. tuberosus* from each of the three origins with available phenotype data showed increased tuber number production relative to their corresponding non-invasive ancestral lineages (Fig. 2d). By contrast, trait divergence for average tuber weight, tuber area, tuber perimeter or tuber shape index was not in the direction predicted by ANOVA or MCMCglmm analyses for all ancestral lineage comparisons (not shown). Note, however, that because of limited sample size, statistical significance was tested for only one comparison (origin 3; $t = 3.68$, d.f. = 22.102, $P = 0.0013$; Fig. 2d). With this caveat, evidence of parallelism in increased clonality among independently derived invasive lineages provides additional, strong support for a role of natural selection, and not post-introduction genetic drift in driving tuber number differentiation.

If spread of invasive genotypes was indeed driven by the production of extreme tuber numbers, we expect invasive populations to be highly clonal. This has previously been observed for a number of important weeds that rely predominantly on vegetative propagation[28–30]. SNP data confirmed this expectation. Namely, 84% (112 of 133) of invasive *H. tuberosus* samples obtained from tubers shared a clonal connection with another sample (Supplementary Table 9). Also, levels of genotypic richness (calculated as per ref. [31]) were low for invasive populations (mean 0.09; range 0–0.5), with 70% of sampling sites containing a single clone. Ramets of invasive genotypes were not restricted to single localities. Three genotypes were particularly widespread, collectively accounting for 55% of all surveyed populations. In some cases, sites occupied by the same clone were separated by over 500 km (for example, clonal series 3, Supplementary Table 9). Finally, two of the invasive clones contain samples we obtained from Europe in 2013, as well as multiple accessions collected in Europe since at least 1975 and archived in public repositories (clonal series 3 and 4; Supplementary Table 9, Supplementary Fig. 7). This confirms that clones of invasive *H. tuberosus* are both widespread and long-lived. Our results therefore consistently indicate that clonality is an important component of invasion success in *H. tuberosus*. Hereafter, we considered clonality as a major invasiveness trait in this system.

Plasticity of invasiveness traits. To investigate whether the contribution of clonality to invasive spread is contingent on environment, and whether genetic accommodation occurred during evolution of invasiveness, we compared invasive and non-invasive genotypes under different levels of water availability (see Methods for a justification of using this environmental variable). We first quantified spread for plants that overwintered in the field, under contrasting levels of soil water content (Supplementary Fig. 8a). Clonal size

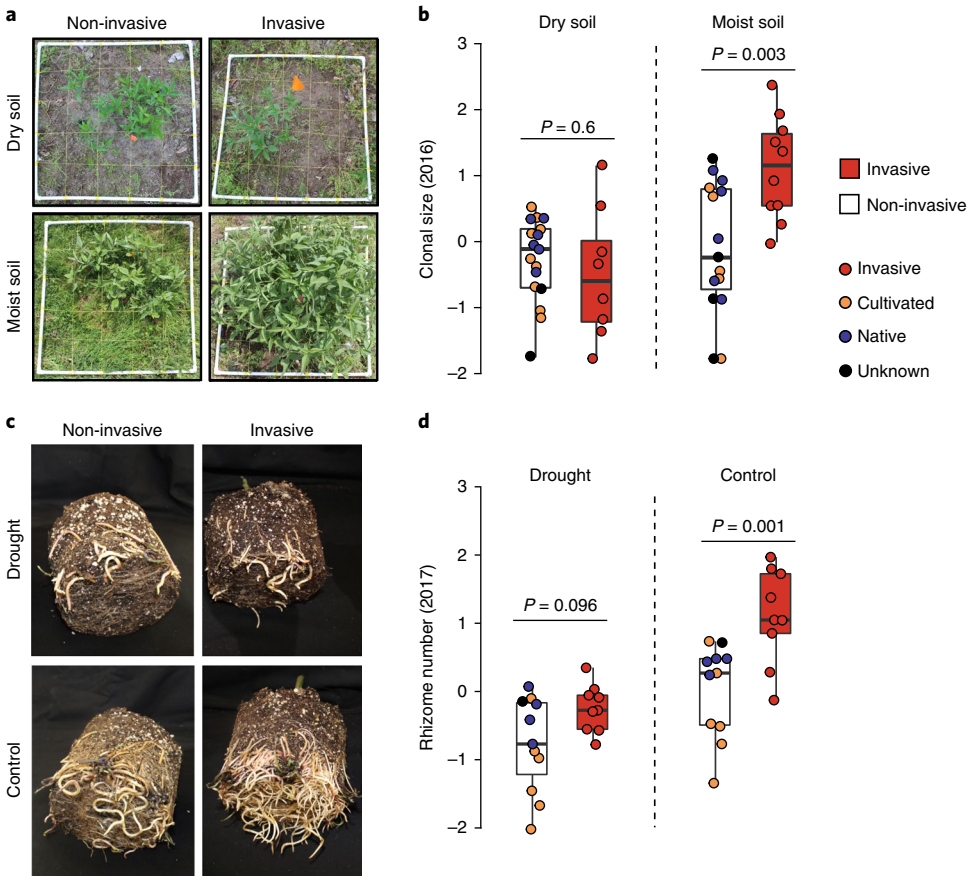

Fig. 3 | Effect of environment on the contribution of clonality to invasiveness. a, Ramets sprouted by non-invasive and invasive genotypes in the field, in dry and moist soil. The white frame in each picture covers a 1 m^2 area, is centred on the location of planting for a parental genotype, and encompasses all ramets produced by the parental genotype. **b**, Box plots of clonal size (rank-transformed) for plants in dry and moist soil showing the median, interquartile ranges and *P* values from one-way ANOVAs performed for each plot. A two-way ANOVA identified significant effects of sample category ($P=0.021$), plot ($P=0.006$) and category-by-plot interaction ($P=0.006$). **c**, Rhizomes produced by non-invasive and invasive genotypes in the greenhouse under drought and control conditions. **d**, Box plots of rhizome number (rank-transformed) for drought and control plants plots, with *P* values from one-way ANOVAs performed for each treatment. A two-way ANOVA identified significant effects of sample category ($P=0.0067$), treatment ($P<0.0001$) and category-by-treatment interaction ($P=0.0007$).

(that is, the number of ramets produced per genotype) was the best predictor of invasiveness, and showed significantly larger values in invasive samples, in the moist soil (ANOVA, $F_{1,23}=10.44$, $P=0.003$; Fig. 3a,b; see Supplementary Fig. 8 for results on a second invasiveness metric). These differences were dependent on environment, however (Fig. 3b; two-way ANOVA interaction term: $F_{1,46}=8.28$; $P=0.006$). Namely, in the dry soil, invasive samples did not maintain an advantage over non-invasive samples (ANOVA, $F_{1,23}=0.28$, $P=0.6$; Fig. 3a,b).

We evaluated these results in the greenhouse using a drought experiment. A plastic response of clonality to water availability was mounted by non-invasive genotypes (ANOVA, $F_{1,73}=58.29$, $P<0.0001$) and by invasive genotypes (ANOVA, $F_{1,62}=96.9$, $P<0.0001$). In agreement with observations made in the field, the slope of reaction norm for the number of clonal propagules produced per plant was significantly larger for invasive samples (two-way ANOVA interaction term: $F_{1,135}=12.13$; $P=0.0007$). As a result, and in line with expectations from genetic accommodation theory, invasive genotypes produced significantly more clonal propagules than non-invasive genotypes under well-watered conditions (ANOVA, $F_{1,18}=13.39$, $P=0.001$), but not under drought conditions (ANOVA, $F_{1,18}=3.08$, $P=0.096$; Fig. 3c,d).

Genetic mechanisms of accommodation. We then asked what genetic mechanisms contributed to accommodation. Theory predicts that these should be visible to selection in the derived (that is, invasive) environment[16–18]. We therefore investigated the genetic control of tuber number production under well-watered conditions. We used the 2015 common garden data, which are representative of clonality estimates for plants grown in the field in the moist soil plot (Supplementary Fig. 8b), or under well-watered conditions in the greenhouse (Supplementary Fig. 9). We considered heterosis (hybrid vigour[32]) a prime candidate for two reasons. First, heterozygosity is significantly elevated in invasive samples (Fig. 1d; Supplementary Figs. 5,6). Second, observational and experimental evolution evidence from other systems suggests that hybrid vigour can enhance invasion success[33–35].

As expected if genome-wide heterozygosity has an important fitness contribution in *H. tuberosus*, we detected significant positive correlations between heterozygosity and trait values for seven phenotypes (Supplementary Fig. 10a). For tuber number, correlations were significant for all samples, as well as for the native and cultivated subsets (Spearman correlations $r_{all}=0.46$, $P=6.3\times10^{-13}$; $r_{native}=0.31$, $P=0.012$, $r_{cultivated}=0.27$, $P=0.028$; Fig. 4a; see also Supplementary Fig. 10). Consistent with theoretical predictions that polyploids experience reduced inbreeding depression[36], the tuber number heterosis signal was more pronounced for markers assigned to the diploid than the tetraploid subgenome (Supplementary Fig. 10b,c).

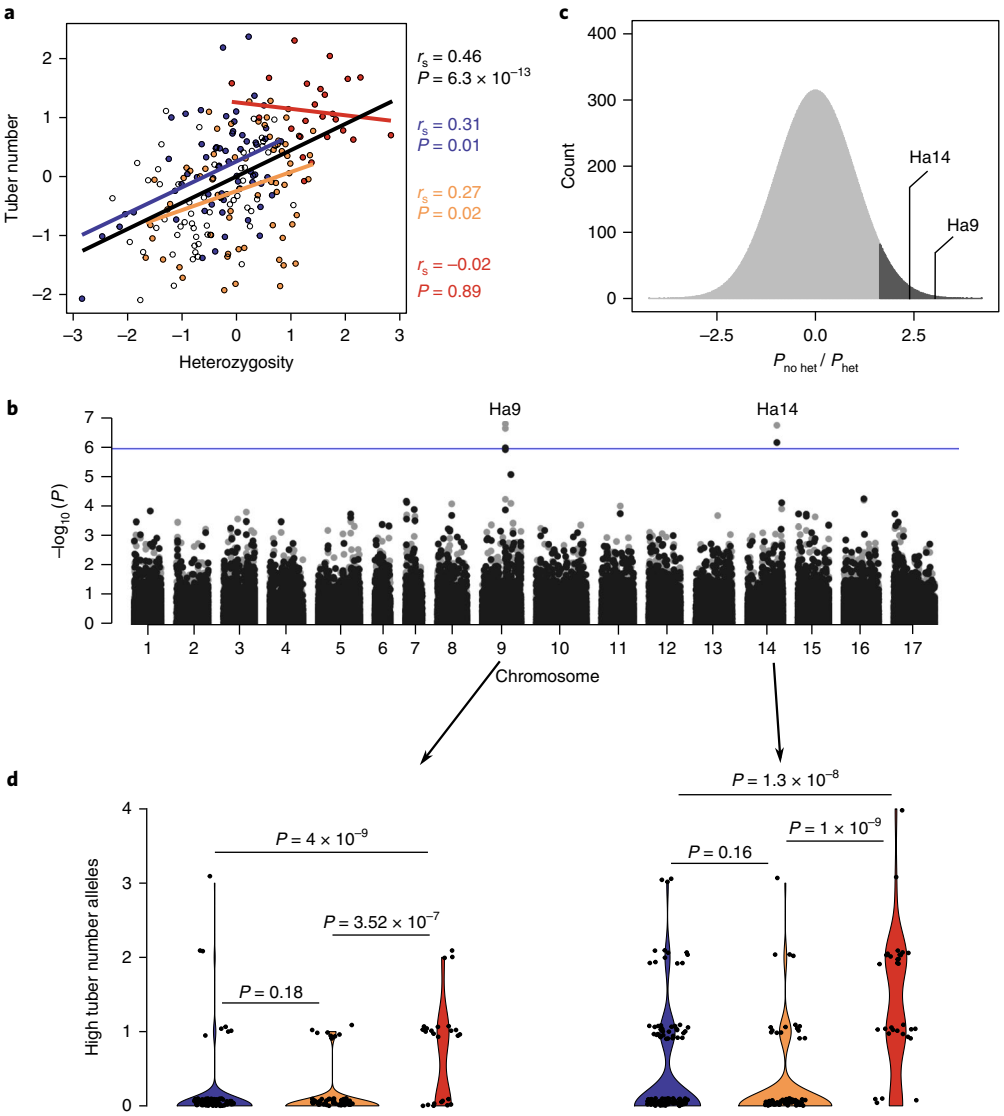

Fig. 4 | Genetic architecture of tuber number production in *H. tuberosus*. **a**, Correlations between rank-transformed values for heterozygosity and tuber number for all samples combined (black), and for the native (blue), cultivated (orange) and invasive (red) samples. Open circles indicate unknown samples. r_s, Spearman's rank correlation coefficient. **b**, Manhattan plot for association analyses using an additive marker-effect model, performed with and without corrections for genome-wide heterozygosity (grey and black, respectively). The blue line indicates the Bonferroni significance threshold. **c**, Histogram of rank-transformed ratios between *P* values obtained for association models without ($P_{no\ het}$) and with (P_{het}) corrections for genome-wide heterozygosity. Values above 0 indicate markers for which the association was improved when correcting for heterozygosity. Dark grey is used to indicate values higher than 95% of ratios. **d**, Violin plots showing the number of high tuber number alleles at the Ha9 and Ha14 QTLs for each sample (black dots) within the native (blue), cultivated (orange) and invasive (red) categories. *P* values were computed using a permutation test with 10^9 replicates, based on the chi-square statistic with fixed margins.

Invasive samples did not follow the same trend, with most individuals maintaining high tuber number production irrespective of heterozygosity (Fig. 4a; Supplementary Fig. 10b,c). This could have resulted if invasive *H. tuberosus* do not display hybrid vigour. We consider this unlikely, however, because invasive samples show trait values that are, on average, higher than those observed in native and cultivated samples at most other heterotic traits (Supplementary Table 5). Increased values at multiple component traits, and in particular at biomass traits, such as those reported here, are a well-known characteristic of heterotic phenotypes[32,37]. Moreover, in *H. tuberosus*, heterosis at these same traits (for example, tuber number, branch number) has been reported following intraspecific and interspecific crosses[22,38]. Alternative explanations for a lack of correlation between heterozygosity and tuber number for invasive *H. tuberosus* include the reduced sample size for this sample category. Indeed, the estimated power to detect an effect similar to the one we observed in non-invasive samples was low for the invasive group (0.3; Supplementary Methods). Also, there may not be an added benefit for tuber number production beyond a given heterozygosity level. Another possibility is that additional genetic factors make an important contribution to tuber number production in *H. tuberosus*. This is suggested by the fact that, after accounting for the contribution of heterozygosity to total phenotypic variance, ~65% of broad-sense heritability for tuber number remains unexplained. If these additional genetic factors were common in invasive samples, they would distort the heterosis signal preponderantly for this group.

To investigate this possibility, we performed genome-wide association (GWA) mapping. For tuber number, two quantitative trait loci (QTLs) were identified (hereafter referred to as Ha9 and Ha14; Fig. 4b), on linkage groups 9 and 14 of the *H. annuus* genome

(Supplementary Tables 10,11). Three lines of evidence indicate the signal at these QTLs is independent of heterosis. First, contrary to the pattern observed for most other markers along the genome, associations at both loci were improved when including genome-wide heterozygosity as an additional covariate (Fig. 4b,c). Second, both QTLs were identified using an additive marker-effect model, and were not supported in association analyses assuming dominant or overdominant effects. Third, the two QTLs do not interact epistatically with one another, or with SNPs that were significantly associated with other traits in our GWA analysis ($P > 0.05$ for all epistasis tests).

The Ha9 and Ha14 QTLs account for a large proportion of tuber number variance in our population. The per cent total variance explained (PVE) in models considering each QTL independently was 15% for Ha9 and 17% for Ha14. Similarly, PVE was substantial in models containing terms for both QTLs, either with or without an additional term for genome-wide heterozygosity (35% or 24%, respectively). Finally, as expected if Ha9 and Ha14 have an important contribution to genetic accommodation during the evolution of invasiveness, the frequency of superior tuber number alleles at both loci is significantly elevated in invasive samples (Fig. 4d).

Relative contribution of hybrid vigour and additive QTLs. We next asked whether hybrid vigour and additive QTLs differ in their relative contribution to the evolution of invasiveness in the four invasive lineages (that is, origins 1–4). To answer this, we compared observed heterozygosity and number of superior Ha9/Ha14 alleles for the 32 invasive genotypes. For each invasive genotype, we considered values at either of the two metrics to be extreme if they were higher than the values we observed in the majority (90%) of non-invasive samples (Fig. 5a).

There were 28 ramets assigned to origin 1 invasive *H. tuberosus*, distributed across five European populations. Most of these (97%; 27/28 ramets) belong to a single widespread genotype. This genotype is not extreme in heterozygosity, but contains the Ha14 allele associated with increased tuber number production (Fig. 5a). Origin 4 invasive *H. tuberosus* shows the reverse pattern. There were 29 ramets assigned to this lineage, distributed across five populations. These ramets were partitioned among two genotypes that showed elevated heterozygosity, but lacked superior Ha9 or Ha14 alleles (Fig. 5a). Invasive *H. tuberosus* assigned to origin 2 was restricted to one population and a single genotype (5 ramets; Fig. 5a). This genotype is not extreme in either of the two metrics. Finally, origin 3 invasive *H. tuberosus*, the most common invasive lineage (10 populations; 71 ramets), shows both elevated heterozygosity and number of superior Ha9/Ha14 alleles (Fig. 5a). These results therefore suggest that evolution of invasiveness was driven by additive QTLs for origin 1, by heterosis for origin 4, and by both genetic mechanisms for origin 3 invasive *H. tuberosus*.

The higher apparent success of origin 3 invasive genotypes could have resulted if spread for this lineage is achieved by a combination of clonal propagation and seed dispersal. That sexual reproduction is more frequent for origin 3 is supported by the fact that a larger number of genotypes were recovered for this lineage (30 genotypes) than for origins 1, 2 or 4 (1–2 genotypes). All origin 3 genotypes are closely related, and group monophyletically (Supplementary Fig. 3), indicating they likely originated post-introduction. Furthermore, 50% of European sites occupied by origin 3 *H. tuberosus* contained more than one genotype, whereas this proportion dropped to 20% for origin 1, and to 0% for origins 2 and 4. These observations are in line with previous studies that report contrasting levels of seed production among invasive populations in Europe, and differentiate between 'generative' and 'vegetative' invasive *H. tuberosus*[25].

A second, non-mutually exclusive possibility is that the joint contribution of additive QTLs and heterosis translates to increased tuber number production and invasiveness for origin 3 genotypes under non-drought conditions. To investigate this, we fit three nested linear models that explain tuber number for 220 genetically distinct tuber-propagated accessions as follows. The 'qtl-only' model included one term for genotypes at each of the Ha9 and Ha14 QTLs ($F_{6,186} = 11.43$; $P = 7.22 \times 10^{-11}$; Akaike information criterion (AIC) = 461.4). The 'qtl and hybrid vigour' model additionally included a term for genome-wide heterozygosity ($F_{7,185} = 15.98$; $P = 2.36 \times 10^{-16}$; AIC = 432.69). Finally, the 'interaction' model included all terms from the preceding models, as well as interaction terms between genotypes at each of the two QTLs and heterozygosity ($F_{12,180} = 10.07$; $P = 5.51 \times 10^{-15}$; AIC = 434.84). AIC scores and likelihood ratio tests confirmed that the 'qtl and hybrid vigour' model outperforms the 'qtl-only' model ($P = 1.61 \times 10^{-8}$), and that the 'interaction' model does not provide a significantly better fit ($P = 0.18$). These results thus support the conclusion that additive QTLs and hybrid vigour are complementary, with maximal tuber numbers being achieved through the combined action of both mechanisms (Fig. 5b)

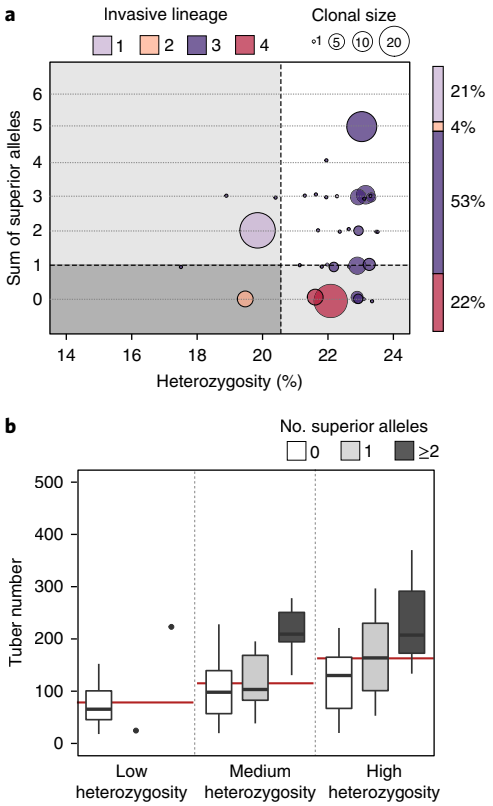

Fig. 5 | Joint contribution of additive QTLs and heterozygosity to invasiveness. a, Distribution of invasive genotypes as a function of heterozygosity and sum of high tuber number alleles at Ha9/Ha14. Circle size indicates clonal size, or the number of ramets observed for each genotype across invasive populations. Areas of grey shading indicate the first 90 percentiles of the distribution of non-invasive samples for both metrics. The stacked bar represents the partitioning of all invasive ramets ($N = 133$) among lineages 1–4. **b**, Boxplots of tuber number, considering all samples propagated from tubers. For visualization only, samples are partitioned in low (14.01–7.15%), medium (17.18–20.34%) and high (20.36–23.52%) heterozygosity bins. Within each bin, samples are further grouped based on the number of high tuber number alleles at Ha9/Ha14. Red lines indicate the average tuber number per heterozygosity bin.

Conclusions

Our study provides important insights into pathways and mechanisms of biological invasions. First, our results demonstrate that

most invasive genotypes contain wild and cultivated ancestry, and indicate that a substantial fraction of invasive *H. tuberosus* originate from admixed breeding lines. The *H. tuberosus* system therefore extends a growing number of examples of invasive species that escaped from cultivation[39], and reinforces crop development as a high-propagule vector of biological invasions. Aside from a demographic contribution to invasion success, breeding may also introduce genetic variation that is advantageous for invasive spread. Our results illustrate this point by providing evidence that hybrid vigour, commonly exploited in plant breeding, contributes to invasive spread. Following the development of breeding lines, further artificial selection is unlikely to have facilitated invasiveness. This is because the primary target of *H. tuberosus* improvement is the production of large tubers[22], a phenotype that was not identified as an outlier in invasive lineages in our analyses.

Our results further established that evolution of extreme clonality in invasive *H. tuberosus* occurred by genetic accommodation. Specifically, an ancestral plastic response of clonality to water availability likely facilitated the persistence of plants introduced in riparian habitats. Selection on genotypes that produce superior numbers of asexual propagules under sufficient water availability then led to the evolution of invasiveness. These results expand current evidence from natural populations for a role of phenotypic plasticity in invasion success and, more broadly, in adaptive evolution[16,18]. Finally, our study indicates that evolution of invasiveness was achieved via two genetic mechanisms. This result parallels findings from recent studies, which highlight that independent evolution of invasiveness between species and between populations often has a diverse genetic basis[40–42]. Collectively, this confirms worries formulated during early surveys of molecular variation, which indicated invasives are unlikely to be genetically constrained[43]. The success of control strategies for invasive species will therefore depend on a thorough understanding of the range of environmental conditions and genetic solutions that are exploited during the evolution of invasiveness.

Methods

Identification of the origin of invasive genotypes. We performed PCAs using all species combined, as well as only for *H. tuberosus*. We used standard filtering (Supplementary Methods), with the exception of the 1% minor allele frequency filter, which was applied across taxa. We used the R package 'adegenet' (v. 2.0.1[44]), which allows for multiple ploidies in the same analysis. The PCA including only *H. tuberosus* samples was performed using the filtered hexaploid call set, as well as the marker subsets that we assigned to the diploid and the tetraploid subgenomes (Supplementary Methods). To further understand relationships among *H. tuberosus* genotypes, we used a phylogenetic network and Bayesian clustering analyses. For the phylogenetic network, we converted SNP calls from VCF to Nexus format using the same approach as the one used for the phylogenetic tree reconstruction (Supplementary Methods). The network was built in Splitstree4[45] using the neighbour-net method, and was based on uncorrected *P* distances. The Bayesian-model-based clustering was implemented in STRUCTURE v. 2.3.4[46]. To optimize STRUCTURE run times, we used 5,000 SNPs, which were randomly selected for the filtered hexaploid call set using the 'vcfrandomsample' utility of vcflib (https://github.com/ekg/vcflib). We used the admixture model with correlated allele frequencies, with 10^5 burn-in replicates, followed by 10^5 replicates. We performed 20 runs for each value of *K* ranging from 1 to 10. To select the most likely value of *K* according to previosuly published reports[4647], we used Structure Harvester[48]. Results were plotted using the 'pophelper' R package (v. 1.2.1[49]).

Trait data collection and analyses. We scored 6 metrics of allelopathy using a greenhouse experiment, and 14 additional traits including clonality in a common garden (Supplementary Table 2; Supplementary Methods). Across the greenhouse and common garden experiments, we included 305 non-clonal *H. tuberosus* genotypes. Of these, 220 (65 wild, 66 cultivated, 29 invasive, 60 unknown) were propagated from tubers at the UBC horticulture greenhouse for two generations. An additional 85 accessions (70 wild, 15 unknown) were propagated from seeds.

To identify traits with significantly extreme values in invasive *H. tuberosus*, we compared tuber-propagated wild, cultivated and invasive samples (see Supplementary Methods for details on sample classification), using R (v. 3.1.2[50]). Aside from the PCA, which does not use a priori sample categorization, we excluded material labelled as 'unknown'. This is because 'unknown' samples are likely to include a mix of material from the other three categories. We also removed the self-incompatibility trait, because no seeds were recovered for any of the samples.

To improve assumptions of normality, we rank-transformed the data using the 'GenABEL' package[51]. We assessed phenotype differentiation between samples using a PCA computed with the 'prcomp' function. For this analysis, we collapsed replicate measurements available for each genotype at the common garden traits to a single value. To do this, we tested whether there was a significant block effect by fitting linear mixed effect models with maximum likelihood. We used the 'lme' function, treating genotype as a fixed effect and block as random effect. We compared the fit of these models to the fit of linear models without a block term. For traits without a block effect, we used the raw means of replicates. For traits for which block was significant, we used the least-squares means, calculated using the 'lsmeans' package[52]. We then only selected 11 of the 19 traits with data for 90% or more of samples, and samples with data for at least 7 of the 11 traits. Lastly, for each trait, we replaced any remaining missing values with the average observed across all samples.

For allelopathy traits, we ran two types of fixed ANOVAs. The first contained data from *H. tuberosus* and its progenitor taxa (*H. grosseserratus*, *H. divaricatus* and *H. hirsutus*), and had species as predictor variable. The second contained only *H. tuberosus* samples, and had sample category (wild, invasive or cultivated) as the predictor. For the common garden traits, also containing data only for *H. tuberosus*, we used mixed ANOVAs implemented with the function 'lme' in the 'nlme' package[53]. We treated sample category as a fixed effect, and genotype nested within block as random effects. Aside from the ANOVA based on species (above), we ran all models with and without correction for population genetic structure. The versions corrected for population structure included as covariates the scores for the first two PC axes calculated using all filtered markers (82,957 SNPs).

We additionally implemented generalized linear mixed models using Markov chain Monte Carlo in the 'MCMCglmm' package[54]. We used sample category as the response variable, and concomitantly considered all traits as predictor variables. Prior to the analysis, we removed missing data using the same approach described for the PCA above. Of all phenotypes, we selected six through a stepwise withdrawal of traits that have a variance inflation factor greater than 3. MCMCglmm ran for 5×10^8 iterations. We fit equivalent models with correction for population structure, by including the genetic PC1 and PC2 scores as additional fixed effects. Significance levels for all analyses were adjusted using sequential Bonferroni[55].

The role of tuber number in invasive spread under drought. We chose water availability as an environmental variable because previous studies in a range of invasive taxa including the congener *H. annuus* have shown that increased growth of introduced genotypes is often gained at the expense of increased water needs[56–58]. In its native range, *H. tuberosus* is common in woodland ecosystems that experience intermediate levels of precipitation, as compared to those recorded for the genus as a whole[59]. By contrast, invasive populations occur almost exclusively in riparian habitats[24,25,60,61], where water is not limiting. In line with these observations, a previous study[62] estimated that the niche of invasive *H. tuberosus* is significantly narrower than that of native *H. tuberosus*. Moreover, water availability was identified as a major driver of differentiation between native- and invasive-range niches in this species[62]. This indicates water availability may be important for invasive spread in *H. tuberosus*.

Field experiment. We investigated the effect of drought on the contribution of tuber number to invasiveness using a second field experiment conducted at the UBC Totem Field research station. The experiment lasted from early summer of 2015 to the spring of 2016. We used 50 plants (12 native, 14 cultivated, 6 unknown and 18 invasive) sprouted at the UBC horticulture greenhouse from tubers of 2–3 cm in length or from tuber fragments of the same size. On 14 June 2015, we transplanted 1.5-month-old plants (30–40 cm tall) to the field, and flagged the location of each plant. For each genotype, we randomly assigned a position in one of two experimental plots (25 genotypes assigned per plot). Within each plot, plants were spaced on 5 m centres. We did not cover the plots with weed barrier, nor did we remove weeds during the growing season. We watered the plants daily for one week after the transplant, to facilitate establishment. During the growing season, water was provided to plants from both plots only when excessive wilting was observed, to prevent mortality. The plants matured under natural conditions, and produced tubers that were left to overwinter in the field. In the spring of the following year (28–29 April 2016), we recorded the total number of ramets produced by each genotype as well as the maximal distance between ramets and the planting location of the parent genotype.

We rank-transformed the trait data using the 'GenABEL' R package. We then used Pearson correlation to evaluate the relationship between the two invasiveness metrics and rank-transformed values for tuber number. Tuber number estimates were obtained from plants grown during the same season, but under the benign and well-watered conditions of the common garden. These were available for all accessions, except one non-invasive sample in the moist soil plot. Also, for each of the two invasiveness metrics, we used mixed ANOVAs, with sample category, plot and their interaction as fixed effects, and genotype as a random effect. Compared to trait analyses described above, because of the limited samples available, the category term grouped genotypes as either invasive or non-invasive.

Greenhouse drought experiment. In the field experiment described above, different genotypes were used for each plot, and plastic responses are potentially confounded by genetic variation. To verify these results, we performed a greenhouse drought experiment for which clonal replicates were grown with or without water stress. We included 20 genotypes (9 invasive, 6 cultivated, 4 wild and 1 unknown) sprouted from tubers at the UBC horticulture greenhouse. On 17 May 2017, for each genotype, we planted 7 sprouts (for 3 accessions) or 8 sprouts (for 17 accessions; $N = 157$ plants total) of similar size (4–6 leaves stage) in single pots (15 cm diameter, 18 cm height) with potting soil mix. We distributed half of the sprouts for each genotype to a control group, and half to a drought group. Control and drought plants were then assigned a random position on a greenhouse bench, with pots spaced 25 cm apart.

Both control and drought plants were watered as needed using plain water for the first week of the experiment. After 24 May, the control plants were watered, also using plain water, with drip irrigation three times per week during the first three weeks of the experiment, and every day for the remaining five weeks. When wilting was observed for plants assigned to the drought group (usually every three days), we applied, to all drought and control plants, 200 ml of fertilizer water. On 15 July, we harvested the plants and counted the number of rhizomes.

Data were rank-transformed using the 'GenABEL' package. We tested for correlations between number of rhizomes produced in the greenhouse under control and drought conditions and tuber number estimates from common garden plants. Also, we modelled rhizome number production using mixed ANOVAs in the 'nlme' package, with sample category, treatment and their interaction as fixed effects, and genotype as a random effect. Similarly to the field experiment described above, the category term grouped genotypes as either invasive or non-invasive.

Relationship between heterozygosity and trait values. We used Spearman correlation analyses to test the relationship between genome-wide heterozygosity estimated using the 6×, 4× and 2× call sets, and trait values. For tuber number, to account for the possibility that population structure may be driving the observed patterns, we repeated these analyses for each of the three sample categories separately. For invasive samples, we additionally investigated the power available to detect an effect of the same magnitude to the one observed in non-invasive (native, cultivated and unknown) samples. We used the function 'pwr.f2.test' from the 'pwr' R package[63]. We set the critical α level to 0.05. Effect size was estimated as the R^2 from a linear model between rank-transformed heterozygosity and tuber number values. The model considered all non-invasive samples, and was built using the linear model-fitting R function 'lm'.

Association mapping. We used the R package 'GWASpoly'[64], which extends the Q (or P) + K mixed linear model for association analyses to allow the use of polyploid data. We included 305 *H. tuberosus* accessions (native, cultivated, invasive or unknown) for which genotype and phenotype data was available. Of the 305 samples, 220 were propagated from tubers and 85 were propagated from seeds. For all association models, we included mode of propagation (that is, tubers or seeds) as an additional covariate. The genotypes were in hexaploid format, and were based on 43,276 SNPs. These markers were obtained using the same filtering criteria as for the rest of the data (Supplementary Methods), with the exception of the call rate threshold, which we increased from 60% to 90%. We ran mixed linear models that account for kinship (K), kinship and population structure as estimated using the top two principal components ($K+P$), or kinship and population structure as estimated using membership coefficients for the two-populations STRUCTURE model ($K+Q$). The kinship matrix, which is included as a random effect in the associations, was calculated in GWASpoly as the realized relationship matrix, computed after imputing missing marker data with the population mean. We used additive, dominant, diplo-general (heterozygotes have the same effect) and diplo-additive (heterozygotes have the same effect, constrained to be intermediate relative to the two homozygous effects) marker-effect models available in GWASpoly. For the seven heterotic traits, we performed association analyses with and without genome-wide heterozygosity included as an additional covariate. When the same associations were identified by different marker-effect models, we recorded the associations with the lowest P value. To establish the significance level, we used a conservative 5% Bonferroni threshold. Association models were evaluated using quantile–quantile plots. Additionally, to verify that population stratification is correctly accounted for, we calculated inflation factors (λ values) in 'GenABEL'.

For tuber number, to test for QTLs of overdominant effect, we repeated the association analyses described above using a nonstandard SNP encoding scheme, for which both homozygotes were coded as 'AA', and all heterozygotes were coded as 'AB'. For significant tuber number associations, we calculated the proportion of phenotypic variance explained (R^2) by the top SNP by fitting linear models with tuber number as response variable and genotype as explanatory variable. Lastly, we generated Manhattan plots using the R package 'qqman'[65].

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability. Sequencing data used in this study have been deposited in the GenBank Sequence Read Archive (accession number PRJNA438356). We propagated, until 207, a subset of invasive *H. tuberosus* samples (12 genotypes) representative of all origins with phenotype data. These samples have been deposited and are being maintained at Plant Gene Resources of Canada (accession numbers CN 120051 to 120062). Phenotype data has been archived at Dryad (10.5061/dryad.4r1k05t).

Received: 13 November 2017; Accepted: 3 April 2018;
Published online: 7 May 2018

References

1. Ellstrand, N. C. & Schierenbeck, K. A. Hybridization as a stimulus for the evolution of invasiveness in plants? *Proc. Natl Acad. Sci. USA* **97**, 7043–7050 (2000).
2. Prentis, P. J., Wilson, J. R. U., Dormontt, E. E., Richardson, D. M. & Lowe, A. J. Adaptive evolution in invasive species. *Trends Plant Sci.* **13**, 288–294 (2008).
3. Bock, D. G. et al. What we still don't know about invasion genetics. *Mol. Ecol.* **24**, 2277–2297 (2015).
4. Colautti, R. I. & Lau, J. A. Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. *Mol. Ecol.* **24**, 1999–2017 (2015).
5. Tsutsui, N. D., Suarez, A. V., Holway, D. A. & Case, T. J. Reduced genetic variation and the success of an invasive species. *Proc. Natl Acad. Sci. USA* **97**, 5948–5953 (2000).
6. Perkins, T. A., Phillips, B. L., Baskett, M. L. & Hastings, A. Evolution of dispersal and life history interact to drive accelerating spread of an invasive species. *Ecol. Lett.* **16**, 1079–1087 (2013).
7. Dlugosch, K. M. & Parker, I. M. Invading populations of an ornamental shrub show rapid life history evolution despite genetic bottlenecks. *Ecol. Lett.* **11**, 701–709 (2008).
8. Colautti, R. I. & Barrett, S. C. Rapid adaptation to climate facilitates range expansion of an invasive plant. *Science* **342**, 364–366 (2013).
9. Krieger, M. J. B. & Ross, K. G. Identification of a major gene regulating complex social behavior. *Science* **295**, 328–332 (2002).
10. Gaines, T. A. et al. Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *Proc. Natl Acad. Sci. USA* **107**, 1029–1034 (2010).
11. Mueller, J. C. et al. Behaviour-related *DRD4* polymorphisms in invasive bird populations. *Mol. Ecol.* **23**, 2876–2885 (2014).
12. Baker, H. *The Genetics of Colonizing Species* 147–168 (Academic Press, New York, 1965).
13. Richards, C. L., Bossdorf, O., Muth, N. Z., Gurevitch, J. & Pigliucci, M. Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecol. Lett.* **9**, 981–993 (2006).
14. Davidson, A. M., Jennions, M. & Nicotra, A. B. Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. *Ecol. Lett.* **14**, 419–431 (2011).
15. West-Eberhard, M. J. *Developmental Plasticity and Evolution* (Oxford Univ. Press, Oxford, 2003).
16. Crispo, E. The Baldwin effect and genetic assimilation: revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution* **61**, 2469–2479 (2007).
17. Lande, R. Evolution of phenotypic plasticity in colonizing species. *Mol. Ecol.* **24**, 2038–2045 (2015).
18. Levis, N. A. & Pfennig, D. W. Evaluating 'plasticity-first' evolution in nature: key criteria and empirical approaches. *Trends Ecol. Evol.* **31**, 563–574 (2016).
19. Godoy, O., Valladares, F. & Pilar, C.-D. Multispecies comparison reveals that invasive and native plants differ in their traits but not in their plasticity. *Funct. Ecol.* **25**, 1248–1259 (2011).
20. Bock, D. G., Kane, N. C., Ebert, D. P. & Rieseberg, L. H. Genome skimming reveals the origin of the Jerusalem artichoke tuber crop species: neither from Jerusalem nor an artichoke. *New Phytol.* **201**, 1021–1030 (2014).
21. Baute, G. J., Owens, G. L., Bock, D. G. & Rieseberg, L. H. Genome-wide genotyping-by-sequencing data provide a high-resolution view of wild *Helianthus* diversity, genetic structure, and interspecies gene flow. *Am. J. Bot.* **103**, 2170–2177 (2016).
22. Kays, S. J. & Nottingham, S. F. *Biology and Chemistry of the Jerusalem Artichoke: Helianthus tuberosus L.* (CRC Press, Boca Raton, 2008).
23. Konvalinková, P. in *Plant Invasions: Ecological Threats and Management Solutions* (ed. Child, L.) 289–299 (Backhuys, Leiden, 2003).
24. Fehér, A. Historical reconstruction of expansion of non-native plants in the Nitra River basin (SW Slovakia). *Kanitzia* **15**, 47–62 (2007).
25. Balogh, L. in *The Most Important Invasive Plants in Hungary* (ed. Botta-Dukát, Z.) 227–255 (Hungarian Academy of Sciences Institute of Ecology and Botany, Vácrátót, 2008).
26. Filep, R., Balogh, L. & Csergö, A. M. Perennial *Helianthus* taxa in Târgu-Mureş city and its surroundings. *J. Plant Dev.* **17**, 69–74 (2010).

27. van Kleunen, M., Dawson, W., Schlaepfer, D. R., Jeschke, J. M. & Fischer, M. Are invaders different? A conceptual framework of comparative approaches for assessing determinants of invasiveness. *Ecol. Lett.* **13**, 947–958 (2010).
28. Hollingsworth, M. L. & Bailey, J. P. Evidence for massive clonal growth in the invasive weed *Fallopia japonic*a (Japanese knotweed). *Bot. J. Linn. Soc.* **133**, 463–472 (2000).
29. Kliber, A. & Eckert, C. G. Interaction between founder effect and selection during biological invasion in an aquatic plant. *Evolution* **59**, 1900–1913 (2005).
30. Zhang, Y.-Y., Zhang, D.-Y. & Barrett, S. C. H. Genetic uniformity characterizes the invasive spread of water hyacinth (*Eichhornia crassipes*), a clonal aquatic plant. *Mol. Ecol.* **19**, 1774–1786 (2010).
31. Dorken, M. E. & Eckert, C. G. Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodon verticillatus* (Lythraceae). *J. Ecol.* **89**, 339–350 (2001).
32. Lippman, Z. B. & Zamir, D. Heterosis: revisiting the magic. *Trends Genet.* **23**, 60–66 (2007).
33. Keller, S. R. & Taylor, D. R. Genomic admixture increases fitness during a biological invasion. *J. Evol. Biol.* **23**, 1720–1731 (2010).
34. van Kleunen, M., Röckle, M. & Stift, M. Admixture between native and invasive populations may increase invasiveness of *Mimulus guttatus*. *Proc. R. Soc. B* **282**, 20151487 (2015).
35. Hahn, M. A. & Rieseberg, L. H. Genetic admixture and heterosis may enhance the invasiveness of common ragweed. *Evol. Appl.* **10**, 241–250 (2016).
36. Otto, S. P. The evolutionary consequences of polyploidy. *Cell* **131**, 452–462 (2007).
37. Flint-Garcia, S. A., Buckler, E. S., Tiffn, P., Ersoz, E. & Springer, N. M. Heterosis is prevalent for multiple traits in diverse maize germplasm. *PLoS ONE* **4**, e7433 (2009).
38. Kantar, M. B. et al. Evaluating an interspecific *Helianthus annuus* × *Helianthus tuberosus* population for use in a perennial sunflower breeding program. *Field Crops Res.* **155**, 254–264 (2014).
39. Ellstrand, N. C. et al. Crops gone wild: evolution of weeds and invasives from domesticated ancestors. *Evol. Appl.* **3**, 494–504 (2010).
40. Hodgins, K. A. et al. Comparative genomics in the Asteraceae reveals little evidence for parallel evolutionary change in invasive taxa. *Mol. Ecol.* **24**, 2226–2240 (2015).
41. Qi, X. et al. More than one way to evolve a weed: parallel evolution of US weedy rice through independent genetic mechanisms. *Mol. Ecol.* **24**, 3329–3344 (2015).
42. Li, L.-F. et al. Signatures of adaptation in the weedy rice genome. *Nat. Genet.* **49**, 811–814 (2017).
43. Kolbe, J. J. et al. Genetic variation increases during biological invasion by a Cuban lizard. *Nature* **431**, 177–181 (2004).
44. Jombart, T. & Ahmed, I. adegenet 1.3–1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**, 3070–3071 (2011).
45. Huson, D. H. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254–267 (2005).
46. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611–2620 (2005).
47. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
48. Earl, D. A. & von Holdt, B. M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359–361 (2012).
49. Francis, R. M. POPHELPER: an R package and web app to analyse and visualise population structure. *Mol. Ecol. Resour.* **17**, 27–32 (2017).
50. R Development Core Team *R: A language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2015).
51. Aulchenko, Y. S., Ripke, S., Isaacs, A. & van Duijn, C. M. GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23**, 1294–1296 (2007).
52. Lenth, R. V. Least-squares means: the R package lsmeans. *J. Stat. Softw.* **69**, 1–33 (2016).
53. Pinheiro, J., Bates, D., DebRoy, S. & Sarkar, D. *nlme: Linear and Nonlinear Mixed Effects Models* v. 3.1-131 (Comprehensive R Archive Network, 2017); http://CRAN.R-project.org/package=nlme
54. Hadfield, J. D. MCMC methods for multi-response generalised linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* **33**, 1–22 (2010).
55. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **6**, 65–70 (1979).
56. Hodgins, K. A. & Rieseberg, L. H. Genetic differentiation in life-history traits of introduced and native common ragweed (*Ambrosia artemisiifolia*) populations. *J. Evol. Biol.* **24**, 2731–2749 (2011).
57. Mayrose, M., Kane, N. C., Mayrose, I., Dlugosch, K. M. & Rieseberg, L. H. Increased growth in sunflower correlates with reduced defences and altered gene expression in response to biotic and abiotic stress. *Mol. Ecol.* **20**, 4683–4694 (2011).
58. Dlugosch, K. M. et al. Evolution of invasiveness through increased resource use in a vacant niche. *Nat. Plants* **1**, 15066 (2015).
59. Kantar, M. B. et al. Ecogeography and utility to plant breeding of the crop wild relatives of sunflower (*Helianthus annuus* L.). *Front. Plant Sci.* **6**, 841 (2015).
60. Walter, J., Essl, F., Englisch, T. & Kiehn, M. Neophytes in Austria: habitat preferences and ecological effects. *Neobiota* **6**, 13–25 (2005).
61. Descombes, P. et al. Monitoring and distribution modelling of invasive species along riverine habitats at very high resolution. *Biol. Inv.* **18**, 3665–3679 (2016).
62. Petitpierre, B. et al. Climatic niche shifts are rare among terrestrial plant invaders. *Science* **335**, 1344–1348 (2012).
63. Champely, S. *pwr: Basic Functions for Power Analysis* v. 1.2-0 (Comprehensive R Archive Network, 2012); http://CRAN.R-project.org/package=pwr
64. Rosyara, U. R., De Jong, W. S., Douches, D. S. & Endelman, J. B. Software for genome-wide association studies in autopolyploids and its application to potato. *Plant Genome* **9**, 73 (2016).
65. Turner, S. D. qqman: an R package for visualizing GWAS results using QQ and Manhattan plots. Preprint at https://doi.org/10.1101/005165 (2014).

Acknowledgements

Plant material used in this study was provided by L. Marek at the United States Department of Agriculture; D. Kessler and A. Diederichsen at Plant Gene Resources of Canada; M. Oppermann and H. Knuepffer at the Leibniz Institute of Plant Genetics and Crop Plant Research; and by N. Langlade and A. Zanetto at the French National Institute for Agricultural Research. We thank C. Bock, I. Bock, A. Bock, M. Bock, R. Filep and A. Stermin for help with sampling of invasive populations. We also thank C. Phillips, C. Tashlin Fluegel and C. Ramsay for assistance with DNA extractions and greenhouse work. The allelopathy bioassays were performed with the help of P. Qiao, A. Poon, S. Li, K. MacDonald, K. Borkowski and M. Iseminger. Also, we acknowledge M. Biron for assistance with greenhouse work, and A. Kuzmin for assistance with sequencing. The set-up and planting of the common garden was performed with help from M. Bock, S. Hubner, D. Burge, G. Baute, K. Stepien and S. Trehearne. The harvest of the common garden was performed with the help of M. Bock, O. Helmy, J. Brodie, S. Heredia, J. Henry, W. Cheung, K. Moran, M. Hahn, K. Ostevik, M. Todesco, E. Drummond, M. Pascual, J. Lee, K. Baute, A. Lalji, W. Choi, D. Skonieczny, I. Hamadi, V. Cha and C. Jewell. This work was supported by a Natural Sciences and Engineering Research Council (NSERC) Vanier CGS and a Killam Doctoral Fellowship to D.G.B., by the Swiss National Science Foundation (SNSF) Postdoctoral Fellowship P2FRP3_151662 to C.C., by the SNSF Doc. mobility Fellowship P1SKP3_168393 to R.M.-D. and by NSERC grant 327475 to L.H.R.

Author contributions

D.G.B. and L.H.R. conceived and planned the study. D.G.B. performed the sampling of invasive populations. D.G.B. and M.B.K. coordinated and conducted greenhouse plant propagations, common-garden fieldwork, and measurements of common-garden traits. D.G.B. and C.C. planned and conducted the bioassays that were used to quantify allelopathy. D.G.B. generated the GBS libraries and analysed the GBS data. R.M.-D. wrote the code and performed the analyses for trait differentiation among sample categories. D.G.B. performed the GWA mapping and the greenhouse drought experiment. All authors interpreted the results. D.G.B. wrote the paper with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-018-0553-z.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to D.G.B.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Dan G. Bock

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

We collected GBS data for 691 H. tuberosus samples and for 175 samples of the progenitor species of H. tuberosus, reasoning that these sample sizes would be sufficient to obtain an accurate view of genetic variation. For the common garden experiment, we selected all non-clonal genotypes in our dataset for which plants were available for propagation.

2. Data exclusions

Describe any data exclusions.

Sequence data analyses included all samples that passed the filtering criteria. Analyses of trait data based on sample categories did not include samples labeled as "unknown" (as described in the paper).

3. Replication

Describe whether the experimental findings were reliably reproduced.

N/A

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

N/A

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

N/A

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	pyRAD v2.17, cutadapt v1.8, BWA v0.7.10, GATK v3.4, Picard tools v1.79, FreeBayes v1.0.1, vcflib, R v3.1.2, SNPRelate, adegenet v2.0.1, STRUCTURE v2.3.4, Structure Harvester, pophelper v1.2.1, SNPhylo v20140701, Splitstree4, IQ-TREE v1.4.2, FigTree v1.4.2, Tomato Analyzer, GenABEL, lsmeans, nlme, MCMCglmm, pwr v1.2, compute.es, GWASpoly, qqman.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	Sequence data are available from the Sequence Read Archive, as described in the Data availability section of the paper. As well, the plant material obtained from USDA, PGRC, INRA and IPK is available upon request from these repositories. Lastly, 12 genotypes collected by the Rieseberg Lab, that were chosen to represent all origins with phenotype data in the paper, and that were propagated at UBC until 2017, have been archived at PGRC (see Data availability section).

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	N/A

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	N/A
b. Describe the method of cell line authentication used.	N/A
c. Report whether the cell lines were tested for mycoplasma contamination.	N/A
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	N/A

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	N/A

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	N/A