

Range expansion compromises adaptive evolution in an outcrossing plant

Santiago C. Gonzalez Martinez, Kate Ridout, John R. Pannell

▶ **To cite this version:**

Santiago C. Gonzalez Martinez, Kate Ridout, John R. Pannell. Range expansion compromises adaptive evolution in an outcrossing plant. Current Biology - CB, 2017, 27 (16), pp.2544-2551. 10.1016/j.cub.2017.07.007 . hal-01607144

HAL Id: hal-01607144
https://hal.science/hal-01607144

Submitted on 21 Oct 2021

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

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



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

Range expansion compromises adaptive evolution in an outcrossing plant

Santiago C. González-Martínez[1,2,4,*], Kate Ridout[1,3], John R. Pannell[1]

[1] Department of Ecology and Evolution, University of Lausanne, 1015, Lausanne, Switzerland

[2] BIOGECO, INRA, Univ. Bordeaux, 33610 Cestas, France

[3] Current address: RDM Nuffield Division of Clinical Laboratory Sciences, John Radcliffe Hospital, Oxford OX3 9DU, UK

[4] Lead Contact

[*] author for correspondence: santiago.gonzalez-martinez@inra.fr

Keywords: Colonisation; deleterious mutation; positive selection; selective sweep; site frequency spectrum; *Mercurialis annua*; population structure; dispersal

Running header: Selection during range expansion

Summary

Neutral genetic diversity gradients have long been used to infer the colonisation history of species [1, 2], but range expansion may also influence the efficacy of natural selection and patterns of non-synonymous polymorphism in different parts of a species' range [3]. Recent theory predicts both an accumulation of deleterious mutations and a reduction in the efficacy of positive selection as a result of range expansion [4-8]. These signatures have been sought in a number of studies of the human range expansion out of Africa, but with contradictory results [9-14]. We analysed the polymorphism patterns of 578,125 SNPs (17,648 genes) in the European diploid plant *Mercurialis annua*, which expanded its range from an eastern Mediterranean refugium into western habitats with contrasted climates [15]. Our results confirmed strong signatures of bottlenecks and revealed the accumulation of mildly to strongly deleterious mutations in range-front populations. A significantly higher number of these mutations were homozygous in individuals in range-front populations, pointing to increased genetic load and reduced fitness under a model of recessive deleterious effects. We also inferred a reduction in the number of selective sweeps in range-front versus core populations. These signatures have persisted even in a dioecious herb subject to substantial interpopulation gene flow [15]. Our results extend support from humans to plants for theory on the dynamics of mutations under selection during range expansion, showing that colonisation bottlenecks can compromise adaptive potential.

Results and Discussion

The annual wind-pollinated plant *Mercurialis annua* expanded its diploid range from an eastern Mediterranean core into central and western Europe, with range margins now established in the western Mediterranean Basin and the European north Atlantic coast [15] (Figure 1). To determine whether this range expansion has differentially affected the site frequency spectra (SFS) for putatively selected versus neutral nucleotide genetic variants between the core and range fronts of the species' distribution, as predicted by theory [4-7], we compared DNA sequence variation within

and among four localities in the eastern Mediterranean core and three localities in each of the Mediterranean and Atlantic range fronts of the species' range, with two males and two females sampled from each locality (Figure 1; *STAR Methods*). Using a gene-capture approach on the recently assembled and annotated genome of *M. annua* [16], we analysed polymorphism patterns for the sampled individuals at 578,125 single nucleotide polymorphic (SNP) sites after filtering from an initial 4 million SNPs (*STAR Methods*).

Genetic diversity and population structure within and among core and range-front populations

Overall, synonymous diversity (π_s = 0.0115) and the ratio of non-synonymous (π_a) to synonymous nucleotide diversity (π_a/π_s = 0.166; Table 1) were similar to those found in other outcrossing annual plants, such as *Capsella grandiflora* or *Zea mays* [reviewed in 17]. Our analysis further revealed strong genetic structure between core and the combined range-front regional populations (Figure S1a), consistent with the hypothesised range expansion from the east into western Europe [15]. We found only slight differences in nucleotide diversity or inbreeding between the core and the two range-front populations (Table 1). However, the total number of polymorphic sites was much lower in the range-front populations (Atlantic, 57.4%; Mediterranean, 56.3%) than in the core range (87.5%), a pattern consistent with that found in a wide range of species that have undergone colonisation bottlenecks and range expansion [12, and references therein, 18]. Statistics for tests of neutrality, such as Tajima's D, also suggested different demographic histories for range-front and core populations (Table 1). Results of demographic modelling using the software fastsimcoal2 v2.5.2.8 [19] point to a stable effective size (N_e) in the core population of *M. annua* in the eastern Mediterranean of c. 500,000, in contrast to much smaller N_e in both the Mediterranean range-front (c. 40,000 individuals) and the Atlantic range-front populations (c. 50,000 individuals), which were probably reduced by a moderate-intensity bottleneck (Table S1; *STAR Methods*).

Although geographically much closer to each other, the two range fronts were distinguished by strong genetic structure (Figure S1b; *STAR Methods*). It is possible that western Europe was initially colonised via a corridor around the Mediterranean, later branching north to occupy central Europe and the Atlantic Coast; alternatively, the western Mediterranean and the Atlantic coast might have been colonised via two relatively independent fronts, e.g., one south and one north of the Alps. Approximately 14% of the derived synonymous mutations found in the two range fronts were not shared with the core population, suggesting that they were sampled from low-frequency variation in the core [20], or arose as new mutations after the range expansion began. Of these front-specific mutations, approximately 50% were shared between the two range fronts, a scenario more consistent with the former of the two expansion scenarios above, i.e., a split between the Atlantic and the south-western Mediterranean populations from a population that had already expanded west. A well-supported phylogenetic network built using the NeighborNet method (*STAR Methods*, see [21], as implemented in SplitsTree4, www.splitstree.org), also suggests that range expansion likely forked in its later stages, so that the final colonisation routes into the western Mediterranean and Atlantic regions were different (see Figure S1d). Little further genetic structure was found among sampling localities within the core range and two range fronts (Figure S1c; *STAR Methods*); further analysis below is thus based on regional population comparisons, with sampling localities within each region grouped together.

Effect of range expansion on putatively neutral versus deleterious mutations

To test for effects of the range expansions on putative neutral mutations versus mutations under selection, we further filtered the variants to retain only those with a known ancestral state (396,450 SNPs) obtained by comparison with *M. huetti*, the sister species to diploid *M. annua* [22, 23]. Analysis of silent variation within this filtered dataset revealed the effects of a moderate-intensity bottleneck on the distribution of derived mutations in both range fronts but not the core range. In particular, whereas the observed silent derived SFS for sequences of individuals sampled in the core

range closely matched the distribution expected for a demographically stable population (Figure S2a), the SFS for both range fronts showed a significant and almost identical relative deficit in low-frequency and an excess in high-frequency or fixed variants (Figure S2b). Accordingly, the observed SFS for the two range-front populations conformed to that expected for a population expanded out of a bottleneck (Figure S2b) [reviewed in 24].

To assess the effects of range expansion on putatively deleterious mutations, we used the software SnpEff [25] to divide the observed derived mutations into three categories: synonymous mutations ('class A' mutations, which are likely to be largely neutral); non-synonymous mutations ('class B' mutations, which might have mildly deleterious effects on fitness); and mutations bringing about premature stop codons or translation frame shifts ('class C' mutations, which are likely to have more strongly deleterious mutations; see *STAR Methods*). Both range-front populations showed not only a relative deficit of low-frequency SNPs in all three mutation categories, but also a significant excess of fixed derived mutations (Figure S2c and S2d), including class B (Figure 2a) and class C mutations (Figure 2b). Specifically, while 0.72% of all derived mutations were fixed in the core range, this percentage rose approximately six-fold to 4.46% and 4.29% in the Mediterranean and Atlantic range-front populations, respectively. Of these, we observed only 271 class B and seven class C derived mutations in the core population, whereas these values rose respectively to 1,004 and 20 in the Mediterranean range front (involving 717 gene models), and to 989 and 22 in the Atlantic range front (involving 680 gene models; Figures 2a and 2b).

The observed accumulation of potentially deleterious mutations in range-front populations is consistent with a hypothesis of relaxed purifying selection predicted for range-edge populations [5-7, 26]. Importantly, however, it was also accompanied by a proportionally similar accumulation of synonymous mutations, i.e., the ratio π_a/π_s was similar in the different parts of the species' range (core: 0.1654; Atlantic front: 0.1702; Mediterranean front: 0.1707; Table 1). These patterns of both

synonymous and non-synonymous variation in range-front populations of *M. annua* are strikingly similar to those found for humans [reviewed in 27]. For instance, the inferred accumulation of likely deleterious mutations is higher in populations out of Africa than in African populations [12, 28], as well as in populations that have been more recently colonised from Europe [10]. Interestingly, the relative proportion of fixed potentially deleterious alleles found for *M. annua* (approximately 3.6-fold) is also similar to that found for humans (approximately 4.4-fold, [9]). Notwithstanding evidence of genetic load at the individual level (see below), the less effective removal by selection of non-synonymous mutations hypothesised for range-front populations is nevertheless as unclear in *M. annua* as it is in humans [13, 14, but see 26]. A recent study failed to find any evidence for an increase in the number of deleterious mutations in the range front of each of six invasive plant species in the sunflower family, perhaps because of admixture among multiple source populations that may have countered the effects of drift during range expansion [29]. To our knowledge, the only clear indication of relaxed purifying selection in terms of the ratio π_a/π_s (from c. 0.2 to over 1) is provided by comparisons between northern and southern populations of wild tomato in western South America [3]. Compared to both *M. annua* and humans, the substantially lower nucleotide diversity (π) in the southern populations of wild tomato suggests that bottlenecks associated with its range expansions may have been much stronger.

The effects of range expansion on likely deleterious variants in *M. annua* are more evident at the level of individual genotypes. Although the average number of derived class B and class C mutations were similar among populations (core: 13,401 ± 571 mutations; Mediterranean front: 13,113 ± 327 mutations; Atlantic front: 13,245 ± 450; mean ± SD; Mann-Whitney test, n.s.), the recessive genetic load (i.e. the average number of derived class B and class C mutations in homozygous state) was substantially higher in individuals from range-front populations (Atlantic front: 5,197 ± 418; Mediterranean front: 5,381 ± 259) than in the core (4,331 ± 349; Mann-Whitney test: $P = 9 \times 10^{-06}$ and $P = 11 \times 10^{-06}$ for Atlantic and Mediterranean fronts, respectively). Importantly, even if we scale

these differences by class A (i.e., synonymous) mutations, they remain significant, albeit smaller (see Table S2).

We found further evidence for relaxed selection in range-front populations by comparing derived mutations that were found only in the range-front populations (those shared by the two range fronts as well as private mutations in each front) with those that were shared between the range fronts and the core. Again, we took a conservative approach by scaling class B or C by class A (BC/A) derived mutations in a homozygous state. For mutations observed only in the two range-front populations, the ratio BC/A = 0.86, whereas BC/A = 0.58 for (likely older) mutations found also in the core. Similarly, the ratio BC/A was > 1 for mutations private to each front compared to 0.79 for mutations private to the core. These results strongly suggest that the influx of new deleterious mutations in range-front populations may translate, at least temporally, into reduced individual fitness. We found similar patterns for the additive genetic load in range-front populations (i.e., the average number of derived deleterious alleles per individual, obtained by counting derived class B and C mutations in heterozygous and homozygous states once and twice, respectively; see Table S2).

The greater number of homozygous derived deleterious alleles observed in range-front populations of *M. annua* is predicted by theory [6]; it is also similar to the difference between human populations with African and European ancestry [12, 28]. If most deleterious mutations show largely recessive expression, as is common [30-32], the increased recessive genetic load in range-front populations of *M. annua* should confer reduced fitness on homozygous individuals, which should be more frequent than in the core. Even though the genetic load brought about by range expansions is expected to be transient, it can persist for thousands of generations [5]. This explanation has been suggested for the patterns observed in humans [reviewed in 27, 33], where selection may not yet have had sufficient time to reduce the frequency of new deleterious mutations accumulated in range-front populations during a previous range expansion. We are currently testing the fitness

consequences of increased recessive load in range-front populations of *M. annua* using reciprocal common gardens.

Effect of range expansion on positive selection

To determine how the range expansion of *M. annua* might have affected its potential for adaptive evolution, we sought evidence for selective sweeps [34, 35] in sequences from each of the three separate geographical regions sampled, including 490,964 SNPs in 14,060 gene models found in 7,375 scaffolds longer than 500 bp and with more than 10 SNPs per scaffold (*STAR Methods*). Here, we compared observed patterns with the overall SFS using a composite likelihood ratio test (CLRT) [36, 37], as well as used haplotype-based statistics (nS_L) that rely on the increase of haplotype homozygosity generated by positive selection [38, 39]. In order to preserve signatures of selection from depleted diversity, and to compare tests performed in populations with different levels of genetic variation, we included monomorphic sites in the CLRTs, following Pavlidis et al. [37], computing empirical *P* values using coalescence simulations without selection of population-specific (standard and bottleneck) demographic models for each geographic region [36]. The nS_L statistics were interpreted using an outlier approach [38, 39]; see *STAR Methods*.

Our CLRTs found evidence for a reduction in the number of selective sweeps in both range-front populations, with 146 and 150 sweeps in the Atlantic and Mediterranean fronts, respectively, versus 196 sweeps in the core population (based on CLRT at $P < 0.01$). Similarly, an analysis with $nS_L > 2.576$ revealed 240 and 224 sweeps in the two front populations compared to 297 in the core. The distribution of empirical *P* values for CLRTs also suggested a greater number of selective sweeps in the core population (Figure S3; *STAR Methods*). Most selective sweeps were restricted to only one of the three regions sampled, as shown by selective sweeps with CLRT values higher than the maximum obtained in the simulations, or by top nS_L outliers; however, two sweeps occurred in both the (Mediterranean) core and Mediterranean front populations (one supported by both analyses, the

other only by the CLRTs), and one was shared between the two (Atlantic and Mediterranean) range-front populations (supported by the CLRTs). It is possible that the two sweeps shared between the two Mediterranean regions (gm23646 and gm20866) relate to independent adaptation to a similar environment. Genetic differentiation (F_{ST}) was relatively high between the core and both range-front populations (average F_{ST} = 0.156 across 578,125 SNPs), suggesting that sweeps shared only between the core region and one front are unlikely to be the result of gene flow linking only those two regions and not the third. Moreover, when compared with the Atlantic range front, no particular similarity was found between the two Mediterranean regions in patterns of nucleotide diversity and divergence for gm23646 and gm20866 (Figure S4). The sweep shared between the two front populations (gm14172 and its 3' untranslated region; Figure 3a), however, seems to predate the divergence of the Atlantic and Mediterranean range fronts, as shown by reduced nucleotide diversity (π) and high genetic differentiation from the core population (F_{ST}) at the same sites (Figures 3b and 3c). We remain ignorant about the function of the genes involved in these putative selective sweeps (see Tables S3 and S4).

The difference in the number of selective sweeps in the core relative to the range-front populations of *M. annua* contrasts with that found for humans, for which a number of studies have found similar (or lower) numbers of selective sweeps in African than European populations [38, 40, 41]. There is also growing evidence for strong selection associated with adaptation to novel challenges and opportunities in non-African populations [e.g., 42, 43, 44]. The observed patterns may correspond to differences in the power to detect selective sweeps in equilibrium and non-equilibrium populations (expected to be lower in the later). However, power comparisons under single hitchhiking models with intermediate selection ($2Ns$ = 100) show a substantial increase of power for this test when monomorphic sites are used in CLRTs [37, 45], as in our study, and comparable power under equilibrium and a wide range of bottleneck conditions [45]. In addition, haplotype-based neutrality tests, such as those based on the nS_L statistic, are generally more robust to bottlenecks than other

methods, as exemplified by only small (and similar) distances between nS_L distributions obtained under bottleneck models of different intensity and the standard neutral distribution [39]. In our study, between 17% and 25% of the selective sweeps, depending on population, were identified by both methods, suggesting reasonable overall power and no apparent differences in power across populations. Assuming adequate power, the lower number of selective sweeps observed in the range-front populations of *M. annua* suggests that the bottlenecks associated with its range expansion were stronger than those inferred for humans. In fact, while N_e underwent an eight- to 18-fold reduction in front relative to core populations in *M. annua* (see above), N_e for humans was reduced only three-fold with their expansion out of Africa [46]. Interestingly, previous work suggests that polyploid populations of *M. annua* underwent a similar range expansion to that studied in the present study for diploid *M. annua*, albeit over a shorter geographic distance, with the Iberian Peninsula colonised from a North African core [15, 47]. The polyploid *M. annua* expansion is thought to have been responsible for the strong cline in both additive genetic variance [48] and inbreeding depression across the species range [49], consistent with a scenario of a series of colonisation bottlenecks. Polyploid and diploid individuals of *M. annua* have similar life histories and occupy similar habitats [50] and might thus have expanded their ranges in similar ways. We would thus expect to find similar footprints of range expansion on gene sequences of polyploid *M. annua* as found here for its diploid relatives.

Concluding remarks

The field of phylogeography has been transformed in recent years both by the availability of genomic sequence data from samples across species ranges, and the development of new theory that predicts not only patterns of neutral genetic variation, but also variation at loci under both purifying and positive selection [4-7]. Some genetic signatures predicted by this theory have been detected in genomic sequence data of the human population, for which we have a good understanding of its history of range expansion out of Africa [9, 12]. Our study here extends

empirical tests for theory on range expansions to a non-model outcrossing plant, finding both a shift in the SFS towards higher frequency of deleterious mutations and a slightly greater genetic load in range-front populations compared to the species core, as well as fewer selective sweeps. The accumulation of putative deleterious mutations and a limited increase in the genetic load in range-front populations of *M. annua* is similar to corresponding patterns reported for humans [reviewed in 27, 33], whereas, to our knowledge, the inferred reduction in the number of selective sweeps, and thus positive selection, in range-edge regions in *M. annua* has not been found before. Indeed, while Böndel et al. [3] revealed a signature of a reduced efficacy of purifying selection in southern populations of wild tomato, as noted above, their study also identified signatures of local adaptation in the southern range margin, as evaluated by numbers of SNP outliers.

Whether range-front populations of *M. annua* have suffered materially from the potentially deleterious effects of their range expansion remains to be seen, e.g., through common garden or reciprocal transplant experiments [51]. In fact, the range expansion may actually have selected an increased capacity for dispersal and colonisation in *M. annua*, as has been observed in the cane toads of north-eastern Australia [52]. If so, it is plausible that the evolution of traits conferring a greater colonisation ability on *M. annua* has been able to counter any accompanying deleterious effects of range expansion, especially given that the species is subject to on-going population turnover through extinctions and colonisations in disturbed habitats [15, 47, 53, 54]. Recently, Gilbert et al. [8] have shown that range expansion that coincides with adaptation to the new environments encountered can also reduce its deleterious effects, largely because local adaptation provides time for migration from the core. Thus, just as the demographic consequences of evolution in a range expansion in which dispersal traits co-evolve with an expansion load deserve further exploration, so does the interaction between range expansion and local adaptation, too [7].

Author contributions

J.R.P. and S.C.G.-M. designed the study. K.R. and S.C.G.-M. conducted the experiments and data analyses. J.R.P. and S.C.G.-M. wrote the paper.

Acknowledgements

We are grateful to three anonymous reviewers whose comments greatly improved the manuscript. Thanks are extended to Dr. Myriam Heuertz for critical assistance in the lab and in the construction of phylogenetic networks. This work was supported by grants from the Swiss National Science Foundation (SNSF grant number 31003A_141052), a Marie Curie Intra European Fellowships within the 7th European Community Framework Programme (contract number 328146 to SCGM), and project POREXPAN (CGL2014-53120-P, Spanish Ministry of Economy and Competitiveness).

References

1. Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. Nature *405*, 907-913.
2. Avise, J.C. (2000). Phylogeography: The History and Formation of Species. (Cambridge MA, U.S.A: Harvard University Press).
3. Bondel, K.B., Lainer, H., Nosenko, T., Mboup, M., Tellier, A., and Stephan, W. (2015). North-south colonization associated with local adaptation of the wild tomato species *Solanum chilense*. Mol. Biol. Evol. *32*, 2932-2943.
4. Slatkin, M., and Excoffier, L. (2012). Serial founder effects during range expansion: a spatial analog of genetic drift. Genetics *191*, 171-181.
5. Peischl, S., Dupanloup, I., Kirkpatrick, M., and Excoffier, L. (2013). On the accumulation of deleterious mutations during range expansions. Mol. Ecol. *22*, 5972-5982.
6. Peischl, S., and Excoffier, L. (2015). Expansion load: recessive mutations and the role of standing genetic variation. Mol. Ecol. *24*, 2084-2094.
7. Peischl, S., Kirkpatrick, M., and Excoffier, L. (2015). Expansion load and the evolutionary dynamics of a species range. Amer. Nat. *185*, E81-E93.
8. Gilbert, K.J., Sharp, N.P., Angert, A.L., Conte, G.L., Draghi, J.A., Guillaume, F., Hargreaves, A.L., Matthey-Doret, R., and Whitlock, M.C. (2017). Local adaptation interacts with expansion load during range expansion: maladaptation reduces expansion load. Amer. Nat. *189*, 368-380.
9. Henn, B.M., Botigue, L.R., Peischl, S., Dupanloup, I., Lipatov, M., Maples, B.K., Martin, A.R., Musharoff, S., Cann, H., Snyder, M.P., et al. (2016). Distance from sub-Saharan Africa predicts mutational load in diverse human genomes. Proc. Nat. Acad. Sci. USA *113*, E440-E449.
10. Peischl, S., Dupanloup, I., Foucal, A., Jomphe, M., Bruat, V., Grenier, J.C., Gouy, A., Gbeha, E., Bosshard, L., Hip-Ki, E., et al. (2016). Relaxed selection during a recent human expansion. bioRxiv 064691; doi: http://dx.doi.org/10.1101/064691.

11. Gravel, S., Henn, B.M., Gutenkunst, R.N., Indap, A.R., Marth, G.T., Clark, A.G., Yu, F.L., Gibbs, R.A., The 1000 Genomes Project, Bustamante, C.D. (2011). Demographic history and rare allele sharing among human populations. Proc. Nat. Acad. Sci. USA *108*, 11983-11988.
12. Fu, W.Q., Gittelman, R.M., Bamshad, M.J., and Akey, J.M. (2014). Characteristics of neutral and deleterious protein-coding variation among individuals and populations. Am. J. Hum. Genet. *95*, 421-436.
13. Do, R., Balick, D., Li, H., Adzhubei, I., Sunyaev, S., and Reich, D. (2015). No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans. Nature Genet. *47*, 126-131.
14. Simons, Y.B., Turchin, M.C., Pritchard, J.K., and Sella, G. (2014). The deleterious mutation load is insensitive to recent population history. Nature Genet. *46*, 220-224.
15. Obbard, D.J., Harris, S.A., and Pannell, J.R. (2006). Sexual systems and population genetic structure in an annual plant: testing the metapopulation model. Amer. Nat. *167*, 354-366.
16. Ridout, K., Veltsos, P., Muyle, A., Emery, O., Rastas, P., Marais, G., Filatov, D., and Pannell, J.R. (2017). Hallmarks of early sex-chromosome evolution in the dioecious plant *Mercurialis annua* revealed by de novo genome assembly, genetic mapping and transcriptome analysis. bioRxiv *doi.org/10.1101/106120*.
17. Chen, C.X., Glémin, S., and Lascoux, M. (2017). Genetic diversity and the efficacy of purifying selection across plant and animal species. Mol. Biol. Evol. *34,* 1417-1428.
18. Dlugosch, K.M., and Parker, I.M. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Mol. Ecol. *17*, 431-449.
19. Excoffier, L., Dupanloup, I., Huerta-Sanchez, E., Sousa, V.C., and Foll, M. (2013). Robust demographic inference from genomic and SNP data. Plos Genet. *9*, e1003905.
20. Excoffier, L., Foll, M., and Petit, R.J. (2009). Genetic consequences of range expansions. Annu. Rev. Ecol. Evol. Syst. *40*, 481-501.
21. Huson, D.H., and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. *23*, 254-267.
22. Krahenbuhl, M., Yuan, Y.M., and Kupfer, P. (2002). Chromosome and breeding system evolution of the genus *Mercurialis* (Euphorbiaceae): implications of ITS molecular phylogeny. Pl. Syst. Evol. *234*, 155-170.
23. Obbard, D.J., Harris, S.A., Buggs, R.J.A., and Pannell, J.R. (2006). Hybridization, polyploidy, and the evolution of sexual systems in *Mercurialis* (Euphorbiaceae). Evolution *60*, 1801-1815.
24. Nielsen, R. (2005). Molecular signatures of natural selection. Annu. Rev. Genet. *39*, 197-218.
25. Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X.Y., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w(1118); iso-2; iso-3. Fly *6*, 80-92.
26. Gravel, S. (2016). When is selection effective? Genetics *203*, 451-462.
27. Lohmueller, K.E. (2014). The distribution of deleterious genetic variation in human populations. Curr. Opin. Genet. Dev. *29*, 139-146.
28. Lohmueller, K.E., Indap, A.R., Schmidt, S., Boyko, A.R., Hernandez, R.D., Hubisz, M.J., Sninsky, J.J., White, T.J., Sunyaev, S.R., Nielsen, R., et al. (2008). Proportionally more deleterious genetic variation in European than in African populations. Nature *451*, 994-U995.
29. Hodgins, K.A., Bock, D.G., Hahn, M.A., Heredia, S.M., Turner, K.G., and Rieseberg, L.H. (2015). Comparative genomics in the Asteraceae reveals little evidence for parallel evolutionary change in invasive taxa. Mol. Ecol. *24*, 2226-2240.
30. Assaf, Z.J., Petrov, D.A., and Blundell, J.R. (2015). Obstruction of adaptation in diploids by recessive, strongly deleterious alleles. Proc. Nat. Acad. Sci. USA *112*, E2658-E2666.
31. Garcia-Dorado, A., and Caballero, A. (2000). On the average coefficient of dominance of deleterious spontaneous mutations. Genetics *155*, 1991-2001.

32. Charlesworth, D., and Willis, J.H. (2009). The genetics of inbreeding depression. Nat. Rev. Genet. *10*, 783-796.
33. Peischl, S. (2016). Genetic surfing in human populations: from genes to genomes. Curr. Opin. Genet. Dev. *41*, 53-61.
34. Maynard Smith, J., and Haigh, J. (1974). The hitchhiking effect of a favourable gene. Genet. Res., Cambridge *219*, 23-35.
35. Barton, N.H. (2000). Genetic hitchhiking. Phil. Trans. Roy. Soc. Lond. B *355*, 1553-1562.
36. Nielsen, R., Williamson, S., Kim, Y., Hubisz, M.J., Clark, A.G., and Bustamante, C. (2005). Genomic scans for selective sweeps using SNP data. Genome Res. *15*, 1566-1575.
37. Pavlidis, P., Jensen, J.D., and Stephan, W. (2010). Searching for footprints of positive selection in whole-genome SNP data from nonequilibrium populations. Genetics *185*, 907-922.
38. Voight, B.F., Kudaravalli, S., Wen, X.Q., and Pritchard, J.K. (2006). A map of recent positive selection in the human genome. Plos Biol. *4*, 446-458.
39. Ferrer-Admetlla, A., Liang, M., Korneliussen, T., and Nielsen, R. (2014). On detecting incomplete soft or hard selective sweeps using haplotype structure. Mol. Biol. Evol. *31*, 1275-1291.
40. Fagny, M., Patin, E., Enard, D., Barreiro, L.B., Quintana-Murci, L., and Laval, G. (2014). Exploring the occurrence of classic selective sweeps in humans using whole-genome sequencing data sets. Mol. Biol. Evol. *31*, 1850-1868.
41. Fan, S.H., Hansen, M.E.B., Lo, Y., and Tishkoff, S.A. (2016). Going global by adapting local: A review of recent human adaptation. Science *354*, 54-59.
42. Beleza, S., Santos, A.M., McEvoy, B., Alves, I., Martinho, C., Cameron, E., Shriver, M.D., Parra, E.J., and Rocha, J. (2013). The timing of pigmentation lightening in Europeans. Mol. Biol. Evol. *30*, 24-35.
43. Colonna, V., Ayub, Q., Chen, Y., Pagani, L., Luisi, P., Pybus, M., Garrison, E., Xue, Y.L., Tyler-Smith, C., 1000 Genomes Project Consortium, et al. (2014). Human genomic regions with exceptionally high levels of population differentiation identified from 911 whole-genome sequences. Genome Biol. *15*, 14.
44. Malaspinas, A.S., Westaway, M.C., Muller, C., Sousa, V.C., Lao, O., Alves, I., Bergstrom, A., Athanasiadis, G., Cheng, J.Y., Crawford, J.E., et al. (2016). A genomic history of Aboriginal Australia. Nature *538*, 207-214.
45. Crisci, J.L., Poh, Y.P., Majajan, S., and Jensen, J.D. (2013). The impact of equilibrium assumptions on tests of selection. Frontiers Genet. *4*, 235.
46. Gattepaille, L., Günther, T., and Jakobsson, M. (2016). Inferring past effective population size from distributions of coalescent times. Genetics *204*, 1191-1206.
47. Pannell, J.R., Eppley, S.M., Dorken, M.E., and Berjano, R. (2014). Regional variation in sex ratios and sex allocation in androdioecious *Mercurialis annua*. J. Evol. Biol. *27*, 1467-1477.
48. Pujol, B., and Pannell, J.R. (2008). Reduced responses to selection after species range expansion. Science *321*, 96.
49. Pujol, B., Zhou, S.R., Sahchez-Vilas, J., and Pannell, J.R. (2009). Reduced inbreeding depression after species range expansion. Proc. Nat. Acad. Sci. USA *106*, 15379-15383.
50. Buggs, R.J.A., and Pannell, J.R. (2007). Ecological differentiation and diploid superiority across a moving ploidy contact zone. Evolution *61*, 125-140.
51. Keller, S.R., Fields, P.D., Berardi, A.E., and Taylor, D.R. (2014). Recent admixture generates heterozygosity-fitness correlations during the range expansion of an invading species. J. Evol. Biol. *27*, 616-627.
52. Rollins, L.A., Richardson, M.F., and Shine, R. (2015). A genetic perspective on rapid evolution in cane toads (*Rhinella marina*). Mol. Ecol. *24*, 2264-2276.
53. Eppley, S.M., and Pannell, J.R. (2007). Sexual systems and measures of occupancy and abundance in an annual plant: testing the metapopulation model. Amer. Nat. *169*, 20-28.

54. Dorken, M.E., Freckleton, R.P., and Pannell, J.R. (2017). Small-scale and regional spatial dynamics of an annual plant with contrasting sexual systems. J. Ecol. *105,* 1044-1057.
55. Metzger, M.J., Bunce, R.G.H., Jongman, R.H.G., Mucher, C.A., and Watkins, J.W. (2005). A climatic stratification of the environment of Europe. Glob. Ecol. Biogeogr. *14*, 549-563.
56. Yang, J.A., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. Nature Genet. *42*, 565-569.
57. Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. arXiv:1207.3907 [q-bio.GN].
58. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156-2158.
59. Tajima, F. (1989). Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics *123*, 585-595.
60. Watterson, G.A. (1975). On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. *7*, 256-276.
61. Fay, J.C., and Wu, C.-I. (2000). Hitchhiking under positive Darwinian selection. Genetics *155*, 1405-1194.
62. Zeng, K., Fu, Y.-X., Shi, S., and Wu, C.-I. (2006). Statistical tests for detecting positive selection by utilizing high-frequency variants. Genetics *174*, 1431-1439.
63. Raj, A., Stephens, M., and Pritchard, J.K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics *197*, 573-589.
64. Ramos-Onsins, S.E., and Mitchell-Olds, T. (2007). mlcoalsim: multilocus coalescent simulations. Evolutionary Bioinformatics *2*, 41-44.
65. Jaramillo-Correa, J.P., Verdu, M., and González-Martínez, S.C. (2010). The contribution of recombination to heterozygosity differs among plant evolutionary lineages and life-forms. BMC Evol. Biol. *10*, 22.
66. Browning, S.R., and Browning, B.L. (2007). Rapid and accurate haplotype phasing and missing data inference for whole genome association studies by use of localized haplotype clustering. Am. J. Hum. Genet. *81*, 1084-1097.
67. Browning, S.R., and Browning, B.L. (2016). Genotype imputation with millions of reference samples. Am. J. Hum. Genet. *98*, 116-126.
68. Szpiech, Z.A., and Hernandez, R.D. (2014). selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. Mol. Biol. Evol. *31*, 2824-2827.

Figure titles and legends

Figure 1. **Sampled localities of diploid *M. annua* covering the core (eastern) Mediterranean range and both the Atlantic and (western) Mediterranean range fronts (see also Figure S1)**

Broken lines show the Atlantic North and Central, and the Mediterranean eco-regions, following the environmental stratification of Metzger et al. [55]. The straight line indicates the approximate frontier between diploid and polyploid (not included in this study) populations of *M. annua*, following Obbard et al. [15]. Base map and state borders from http://d-maps.com/carte.php?num_car=2232&lang=en.

Figure 2. **Site frequency distributions for derived mutations in core, and Atlantic and Mediterranean range-front populations (see also Table S1 and Figure S2)**

Pattern for (A) non-synonymous mutations and (B) high-effect mutations (e.g., premature stop codons or translation frame shifts). Accumulation of fixed derived mutations that are likely to have mildly to strongly deleterious effects was approximately 3.6 times greater in range-front than core populations; see text for details.

Figure 3. **Selective sweets shared between range-front populations (see also Figure S4)**

(A) The single selective sweep (in gm14172 and its 3' untranslated region, scaffold105331) that was found in common between the Atlantic and Mediterranean range-front populations using CLRTs. Broken lines represent likelihood significance thresholds, considering both a standard model for a demographically stable population and the best-fitting bottleneck model (Table S1; *STAR Methods*). (B) Detailed analyses of polymorphism (nucleotide diversity, π) and (C) divergence patterns (pairwise F_{ST}) by site for the selective sweep in scaffold105331, centred on a region at the 3' untranslated region of gm14172 of low polymorphism and increased divergence in the two range fronts.

Table 1. Genetic diversity, and neutrality-test and inbreeding statistics (see also *STAR Methods*, Table S1 and Figure S1)

Main statistics are also given separately for diploid *M. annua* range-front and core populations; *D*: Tajima's *D*; H_n: Fay and Wu's normalised *H*; E_z: Zeng et al.'s *E*; *F*: average inbreeding coefficient; *related*: average pairwise relatedness statistic (unadjusted A_{jk}) based on the method of Yang et al. [56]; values around zero indicate unrelated individuals.

Population	**N**	**Polymorphic sites**		**Nucleotide diversity[a]**						**Neutrality-test statistics**			**Inbreeding**	
				Tajima's π			**Watterson's Θ**							
		Number	**%**	**All**	**Syn**	**Non-syn**	**All**	**Syn**	**Non-syn**	***D***	***H_n***	***E_z***	***F***	***related***
Core population	16	505,650	87.46	2.196	11.579	1.916	2.168	10.881	2.098	-0.050	0.073	-0.373	0.0704	-0.0691
Atlantic front	12	331,997	57.43	2.201	10.709	1.823	1.847	8.793	1.564	0.780	-0.281	0.959	0.0490	-0.0913
Mediterranean front	12	325,321	55.44	2.137	10.454	1.785	1.791	8.598	1.529	0.785	-0.315	1.012	0.0775	-0.0938
Overall	*40*	*578,125*	*100.00*	*2.077*	*11.526*	*1.911*	*1.946*	*9.861*	*1.989*	*0.235*	*0.041*	*-0.076*	*0.0680*	*-0.0288*

[a]Nucleotide diversity per site × 10^{-3}; All: all sites, including coding regions (CDS), introns and intergenic sequence; Syn: synonymous sites; Non-syn: non-synonymous sites.

STAR Methods

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to, and will be fully fulfilled by, the Lead Contact, Santiago C. González-Martínez (santiago.gonzalez-martinez@inra.fr).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mercurialis annua L

Seeds from twenty to forty female plants of *Mercurialis annua* L. were collected in each of 10 localities covering the species core range in Turkey (Antalya, southern Mediterranean coast; Akçaabat, Trabzon) and Greece (Corinth, Peloponnese; Volos, Thessaly coast), and the two range fronts in the Atlantic (Brouck, northeastern France; Paris, northern France; Southampton, UK south coast) and western Mediterranean range (Barcelona, Sant Pere de Ribes and Tarragona, Mediterranean coast of Catalonia, Spain) (Figure 1). One seed per maternal family was planted in a cultivation tray and placed in a greenhouse under optimal germination and growing conditions. After about two weeks, leaves from two males and two females per sampling locality were collected, frozen in liquid nitrogen and stored at -80^{o}C until DNA extraction.

Mercurialis huetti Hanry

Seeds from one female plant of *Mercurialis huetti* Hanry from Catalonia (Spain) were collected to be used as outgroup. One seed from this maternal family was planted in a cultivation tray and placed in a greenhouse under optimal germination and growing conditions. After about two weeks, leaves from one female plant were collected, frozen in liquid nitrogen and stored at -80^{o}C until DNA extraction.

METHOD DETAILS

DNA extraction

High quality genomic DNA from the 40 *M. annua* plants and one outgroup of its sister species, *M. huetti*, was extracted from about 100 mg of frozen leave material using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany), following standard protocols.

Gene-capture experiment

The annotated *M. annua* reference genome v. 1.3 [16] was used to design a gene capture experiment that included all exomic regions larger than 30 bp (24.26 Mbp from 103,478 exons) as well as introns and intergenic sequence from scaffolds over 1,000 bp in length. In total, this assay covered 44.71 Mbp accounting for approximately 7% of the estimated *M. annua* genome size. Sequence capture using SureSelect DNA Capture technology (Agilent Technologies, Santa Clara, CA) and next generation sequencing (Illumina HiSeq 2500 1x100bp) of captured regions was outsourced to Rapid Genomics (Gainesville, FL). Briefly, 400,000 120-mer probes covering the targeted sequence were designed, avoiding repetitive regions and organelle genomes. About 90% of the probes worked in each individual. Captured sequences were filtered, based on sequencing quality, and aligned against the *M. annua* reference genome with alignment rates from 77.1 % (56.6 % unique) to 55.5 % (37.2% unique). SNP calling was based on captured sequence with unique alignments to the reference genome, using FreeBayes v0.9.18 [57], which allowed the initial identification of 4,300,049 SNPs.

Raw SNP data were filtered using VCFtools v0.1.14 [58] and vcffilter (part of vcflib C++ library). SNPs were first filtered based on their quality score (Q>20) and individual (>8x) and mean population coverage (between 20-250x). For the heterozygous calls, only the SNPs were kept that had an allele balance between 0.3 and 0.7. Second, only biallelic SNPs were kept with both strands fairly well represented (i.e., no strand bias) and fewer than 15% of the individuals missing. This stringent filtering reduced the SNP dataset about 7.5-fold to 578,125 SNPs. Key results were also validated

with a SNP dataset that was additionally filtered by Hardy-Weinberg disequilibrium at each population and polymorphism excess (529,793 SNPs; 120-mer capture probes with more than nine polymorphic sites were considered unreliable given the average number of polymorphic sites in *M. annua* and Poisson distribution assumptions). Finally, SNPs were annotated based on the *M. annua* reference genome v. 1.3 [16] using SnpEff v4.1e [25], and ancestral states were identified by comparison with the *M. huetti* outgroup, resulting in an annotated SNP dataset of 396,450 SNPs.

QUANTIFICATION AND STATISTICAL ANALYSIS

Genetic diversity and population structure

Nucleotide diversity (π [59]; θ_w [60]) and neutrality-test (Tajima's *D* [59]; and Fay and Wu's normalised *H* and Zeng et al.'s *E* [61, 62]) statistics were computed both for all populations together and separately for range-front and core populations, using mstatspop v0.1 (https://bioinformatics.cragenomica.es/numgenomics/people/sebas/software/software.html) on concatenated sequence files. Average inbreeding coefficients (*F*) and pairwise relatedness (as evaluated by unadjusted A_{jk} statistic [56]) per population were computed using VCFtools v0.1.14 [58].

Population genetic structure was investigated using fastStructure [63] with *K* groups from *K*= 1 (i.e., no structure) to *K* = 10 (the number of localities in our sample). Runs for *K* = 2 (separating the core range from the two range fronts, see Figure S1a) and for *K*= 3 (separating the three study regions; Figure S1b) were repeated 10 times, and averaged *Q* values (i.e., the individual assignment probability to each of the *K* groups) were used to draw bar plots. Runs including only sampling localities from the core or each one of the two range fronts separately showed no structure within regional populations (Figure S1c). A phylogenetic network was built using the NeighborNet method, based on the Uncorrected_P distance between sequences, as implemented in the program SplitsTree4 [21]; program available at www.splitstree.org. These computations were performed on

two random subsets (without replacement) of 100,000 variable sites, and the support of the phylogenetic relationships was assessed with 1,000 bootstrap replicates. The network had high support (see Figure S1d) and suggested that western Europe was colonised via an initial corridor that later branched to occupy the current Atlantic and Mediterranean range fronts.

SFS and demographic modelling

The unfolded site frequency spectrum (SFS) for each population (and mutation class) was estimated by maximum-likelihood, following Nielsen et al. [36]. Males and females from the same populations were considered together for this analysis. First, sites from potentially sex-linked regions were removed. Then, the unfolded SFS based on the remaining annotated silent sites (338,497 SNPs) was used to fit specific demographic models for the core and range-front populations by maximum likelihood, using fastsimcoal2 v2.5.2.8 [19]. Both the standard model for a demographically stable population and a wide variety (in terms of time of the bottleneck, duration and intensity) of bottleneck models were assayed (see Table S1 for search ranges). Demographic models with best-fit to observed SFS were then retained for each population. *M. annua* is an annual herb and thus number of generations can be directly translated to years, ignoring effects of a possible seedbank.

Genetic load

Genetic load in range-front and core populations was evaluated by computing the average recessive and additive genetic load per individual. The recessive genetic load was obtained by counting the average number of derived class B or class C mutations (i.e., non-synonymous mutations and mutations bringing about premature stop codons or translation frame shifts) in homozygous state. The additive genetic load, defined as the average number of derived deleterious alleles per individual, was obtained by counting derived class B or class C mutations in homozygous state twice and derived class B or class C mutations in heterozygous state once. Scaled values by class A (i.e., synonymous) derived mutations for these quantities were also computed in order to evaluate

differences in the efficacy of purifying selection at individual level. The average recessive and additive genetic load per individual in each population was obtained for the full SNP dataset (396,450 SNPs with known ancestral allele), for only shared mutations between range-front and core populations (i.e., presumably older mutations; 157,088 SNPs), for mutations private to the two fronts (18,652 SNPs) and for private mutations in each population (i.e., presumably newer mutations; 125,266 SNPs in the core population, 16,885 SNPs in the Atlantic front and 15,661 SNPs in the Mediterranean front). Significant differences between the core and range-front populations were identified using Mann-Whitney nonparametric tests.

Selective sweeps

Evidence for selective sweeps in core and range-front populations was obtained by computing both parametric tests based on maximum composite likelihood ratio (CLRTs) and haplotype-based statistics (nS_L). To increase power to detect selective sweeps, we only applied the tests on 7,375 scaffolds that were longer than 500 bp and had more than 10 SNPs, which involved 490,964 SNPs (14,060 gene models).

CLRTs included monomorphic sites (i.e., equal SNP density across populations) and considered both folded and unfolded sites [36, 37], as implemented in SweeD v3.3.2 (http://pop-gen.eu/wordpress/software/sweed). Runs used a 60-bp grid to compute CLRTs, although denser grids (10-bp and 30-bp) produced similar results (and are thus not shown). Significance thresholds for selective sweeps detected in each population were obtained by running SweeD on files obtained by coalescence simulations from best-fitting demographic models (a standard model of stable population size for the core population and a bottleneck model for both range-front populations, see Table S1) using mlcoalsim v1.40 [64]. Coalescence simulations considered recombination rates from $\rho = 0$, to $\rho = 0.031$ (the average recombination rate in a wide range of plants; [65]), to $\rho = 0.155$ (very high recombination), and the most conservative parameters were retained (i.e., those

resulting in higher significance thresholds for CLRTs). Empirical *P* values were obtained from the simulations, and likelihood thresholds at the 0.01 significance level were used for comparison between core (likelihood threshold of 4.550) and range-front populations (likelihood threshold of 4.639). Moreover, selective sweeps with higher likelihood than the maximum produced in the simulations (likelihood threshold of 6.725 and 8.177 for core and range-front populations, respectively) were considered to have lower Type I error and were used in annotation tables (Table S3) and as examples in Figures 3 and S4.

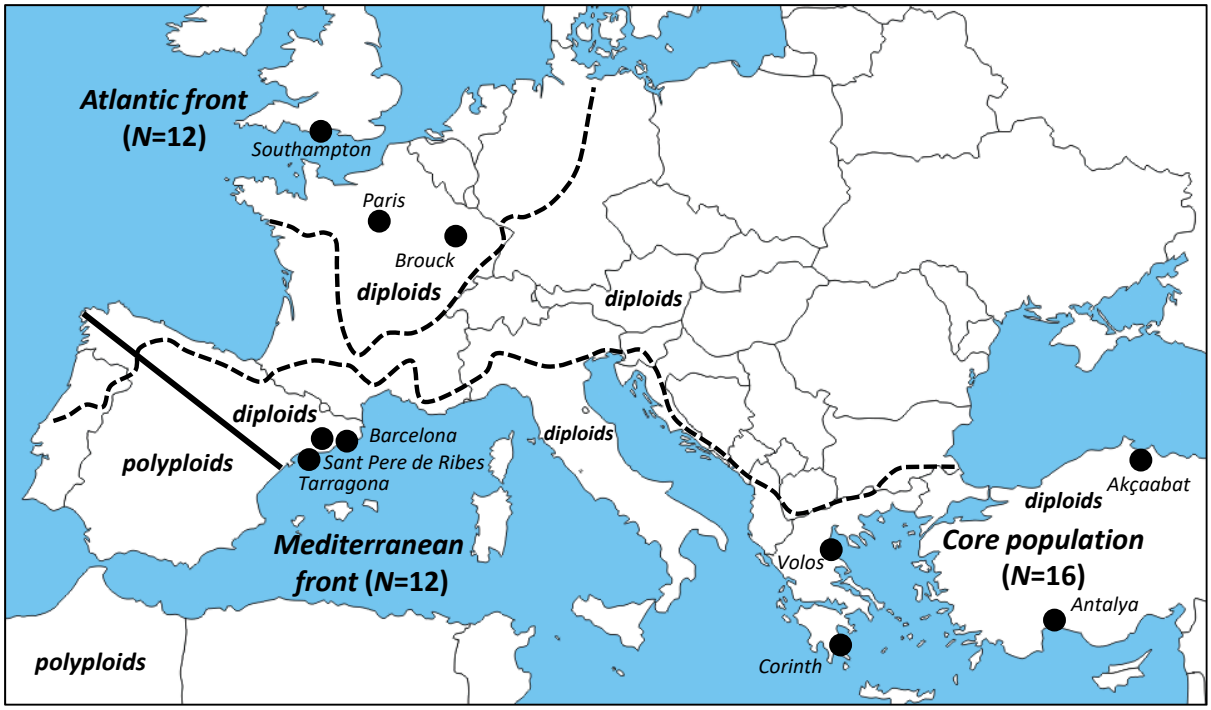
Haplotype-based tests have shown increased power to detect selective sweeps, in particular under non-equilibrium situations [36, 39, 45]. VCF files were phased and missing genotypes were imputed, using Beagle v4.1 [66, 67]; updated version available from http://faculty.washington.edu/browning/beagle/beagle.html. Phased VCFs without missing data were used to compute the nS_L statistic using selscan v1.1.0b, subsequently normalised across allele frequency bins using norm v1.1.0a [68]. The nS_L statistic compares the decay of haplotype homozygosity between the ancestral and derived haplotypes extending from a target polymorphic site, and differs from the well-known iHS statistic [38] in that it measures the length of haplotype homozygosity between a pair of haplotypes in terms of the number of mutations in the rest of the haplotypes, and thus it does not require a genetic map [39]. The nS_L statistic is generally interpreted via an outlier approach, considering both tails of the distribution, with evidence of selection being recognised when $|nS_L| > 2$. Here, however, we focused on positive values (i.e., only selective sweeps on derived haplotypes) greater than 2.576, which should have resulted in a more robust test.

The three high-confidence selective sweeps that were shared between two regional populations (see main text) were further analysed for patterns of polymorphism (nucleotide diversity, π) and divergence (pairwise F_{ST}) per site using VCFtools v0.1.14 [58]. Annotation of genomic regions subject to selective sweeps were obtained from [16] and confirmed by blast against NCBI's GenBank public

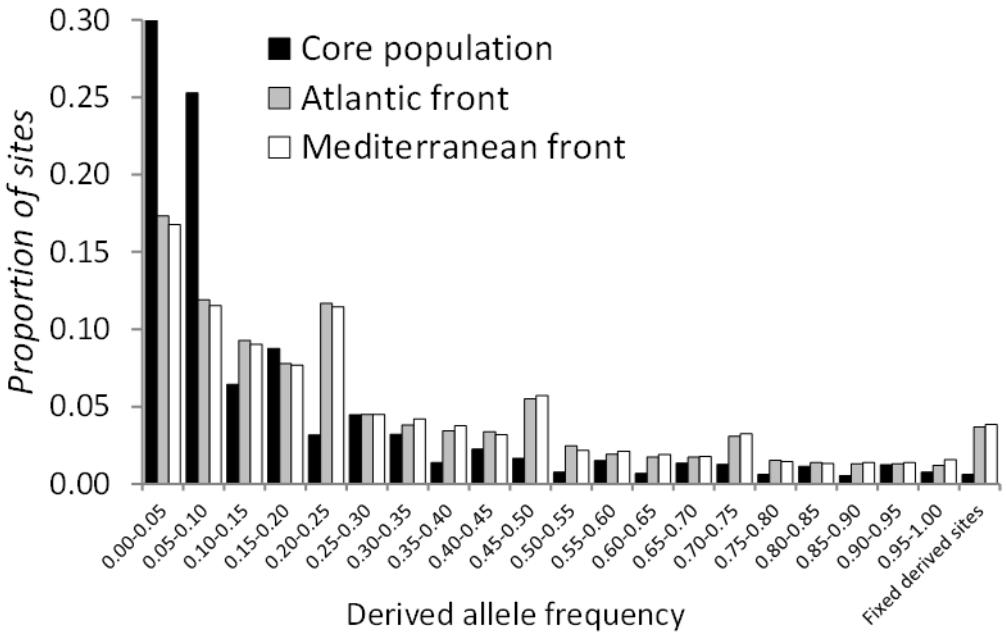
database.

DATA AND SOFTWARE AVAILABILITY

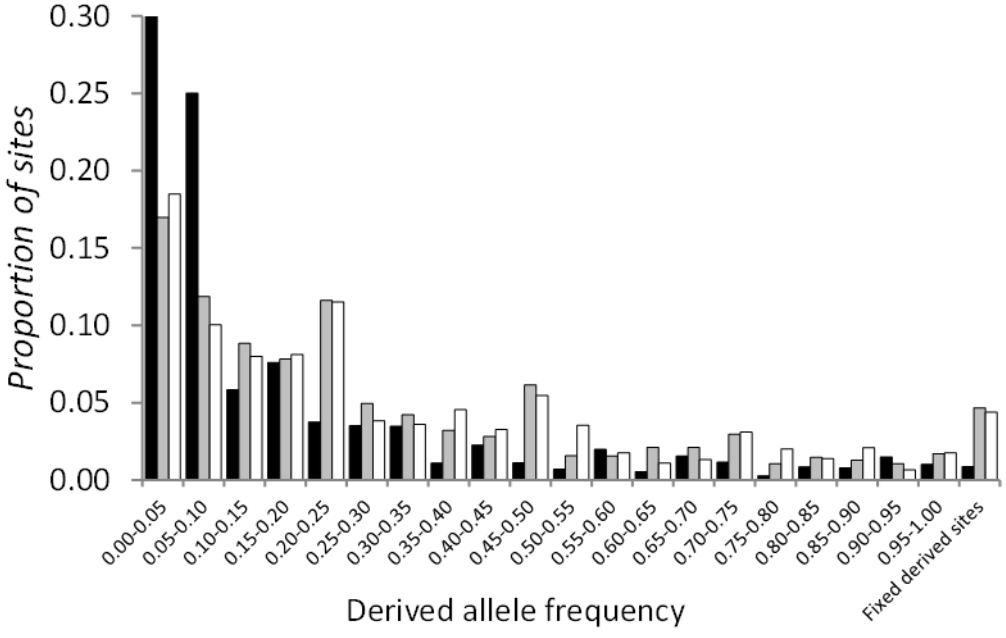
SNP data (VCF v4.1 format) were deposited in the Dryad repository with doi:10.5061/dryad.74631. This VCF file includes the filtered SNP dataset for all 40 *Mercurialis annua* samples (578,125 SNPs). Information on gene capture probe, ancestral state (by comparison with *M. huetti* outgroup), and SnpEff annotation is provided in the VCF INFO field.

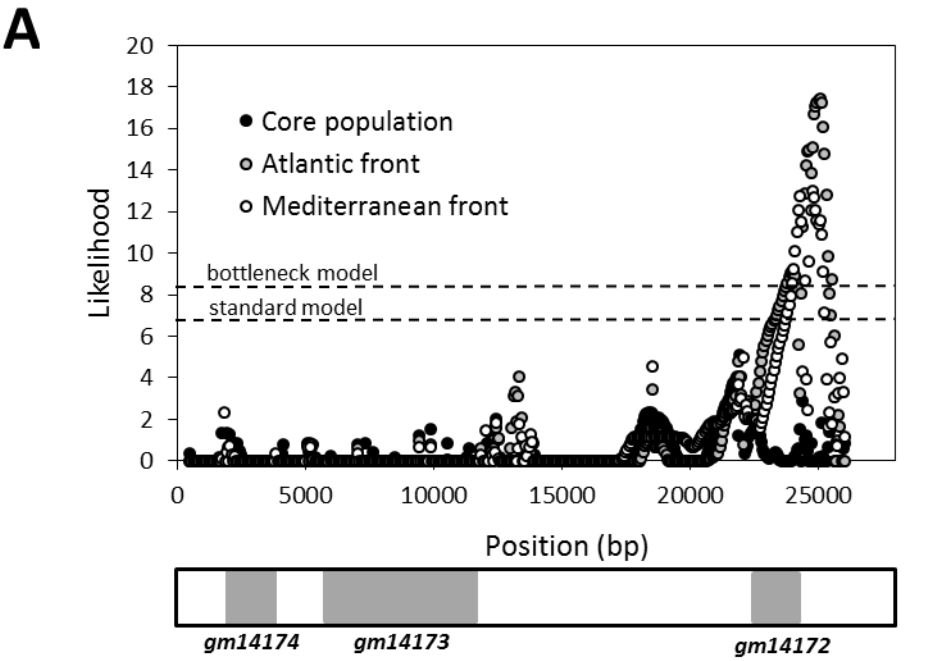

Atlantic front
(N=12)
Southampton
Paris
Brouck
diploids
diploids
diploids
Barcelona
Sant Pere de Ribes
Tarragona
polyploids
diploids
Mediterranean
front (N=12)
polyploids
Akçaabat
diploids
Core population
(N=16)
Volos
Antalya
Corinth

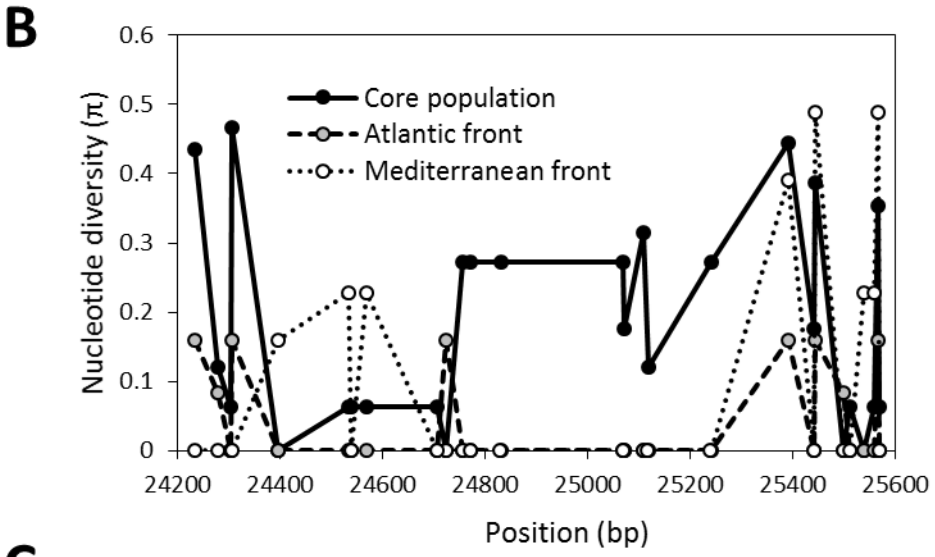
A

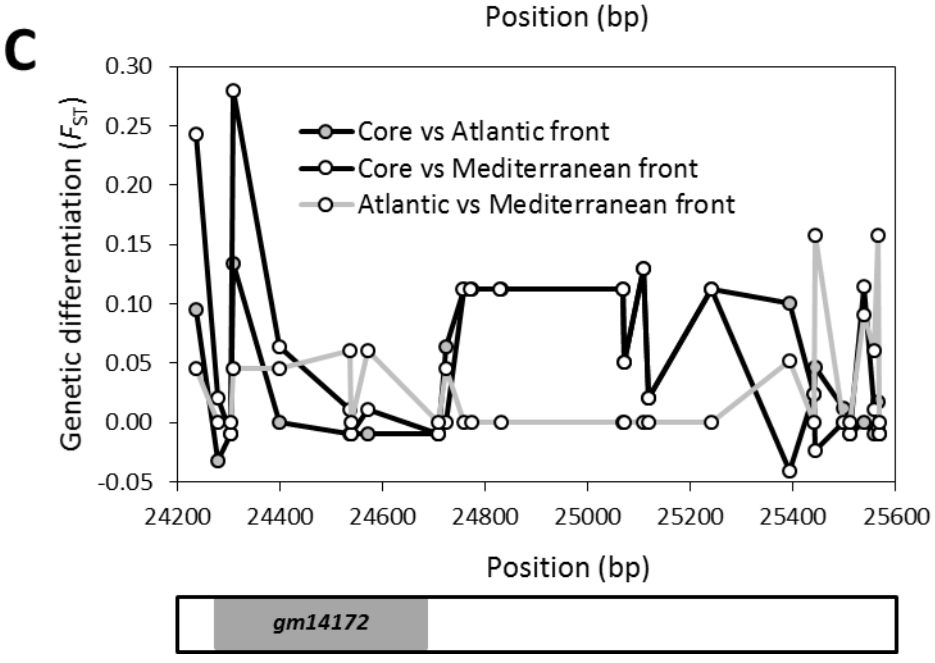

Non-synonymous mutations
Core population
Atlantic front
Mediterranean front
Proportion of sites
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.00-0.05
0.05-0.10
0.10-0.15
0.15-0.20
0.20-0.25
0.25-0.30
0.30-0.35
0.35-0.40
0.40-0.45
0.45-0.50
0.50-0.55
0.55-0.60
0.60-0.65
0.65-0.70
0.70-0.75
0.75-0.80
0.80-0.85
0.85-0.90
0.90-0.95
0.95-1.00
Fixed derived sites
Derived allele frequency

B


Mutations with high potential effect size
Proportion of sites
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.00-0.05
0.05-0.10
0.10-0.15
0.15-0.20
0.20-0.25
0.25-0.30
0.30-0.35
0.35-0.40
0.40-0.45
0.45-0.50
0.50-0.55
0.55-0.60
0.60-0.65
0.65-0.70
0.70-0.75
0.75-0.80
0.80-0.85
0.85-0.90
0.90-0.95
0.95-1.00
Fixed derived sites
Derived allele frequency


A
Likelihood
Core population
Atlantic front
Mediterranean front
bottleneck model
standard model
Position (bp)
gm14174
gm14173
gm14172


B
Nucleotide diversity (π)
Core population
Atlantic front
Mediterranean front
Position (bp)


C
Genetic differentiation (F_{ST})
Core vs Atlantic front
Core vs Mediterranean front
Atlantic vs Mediterranean front
Position (bp)
gm14172