



**HAL**  
open science

## Population genomic and historical analysis suggests a global invasion by bridgehead processes in *Mimulus guttatus*

Mario Vallejo-Marín, Jannice Friedman, Alex D Twyford, Olivier Lepais, Stefanie M Ickert-Bond, Matthew A Streisfeld, Levi Yant, Mark van Kleunen, Michael C Rotter, Joshua R Puzey

### ► To cite this version:

Mario Vallejo-Marín, Jannice Friedman, Alex D Twyford, Olivier Lepais, Stefanie M Ickert-Bond, et al.. Population genomic and historical analysis suggests a global invasion by bridgehead processes in *Mimulus guttatus*. *Communications Biology*, 2021, 4 (1), pp.1-12. 10.1038/s42003-021-01795-x . hal-03310344

HAL Id: hal-03310344

<https://hal.inrae.fr/hal-03310344>

Submitted on 30 Jul 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 4.0 International License

## Population genomic and historical analysis suggests a global invasion by bridgehead processes in *Mimulus guttatus*

Mario Vallejo-Marín <sup>1✉</sup>, Jannice Friedman <sup>2</sup>, Alex D. Twyford<sup>3,4</sup>, Olivier Lepais<sup>5</sup>, Stefanie M. Ickert-Bond<sup>6</sup>, Matthew A. Streisfeld<sup>7</sup>, Levi Yant <sup>8</sup>, Mark van Kleunen <sup>9,10</sup>, Michael C. Rotter<sup>11</sup> & Joshua R. Puzey<sup>12</sup>

Imperfect historical records and complex demographic histories present challenges for reconstructing the history of biological invasions. Here, we combine historical records, extensive worldwide and genome-wide sampling, and demographic analyses to investigate the global invasion of *Mimulus guttatus* from North America to Europe and the Southwest Pacific. By sampling 521 plants from 158 native and introduced populations genotyped at >44,000 loci, we determined that invasive *M. guttatus* was first likely introduced to the British Isles from the Aleutian Islands (Alaska), followed by admixture from multiple parts of the native range. We hypothesise that populations in the British Isles then served as a bridgehead for vanguard invasions worldwide. Our results emphasise the highly admixed nature of introduced *M. guttatus* and demonstrate the potential of introduced populations to serve as sources of secondary admixture, producing novel hybrids. Unravelling the history of biological invasions provides a starting point to understand how invasive populations adapt to novel environments.

<sup>1</sup> Biological and Environmental Sciences, University of Stirling, Stirling, Scotland, UK. <sup>2</sup> Biology Department, Queen's University, Kingston, ON, Canada. <sup>3</sup> School of Biological Sciences, University of Edinburgh, Edinburgh, Scotland, UK. <sup>4</sup> Royal Botanic Garden Edinburgh, Edinburgh, UK. <sup>5</sup> INRAE, Univ. Bordeaux, BIOGECO, Cestas, France. <sup>6</sup> Herbarium (ALA), University of Alaska Museum of the North, University of Alaska Fairbanks, Fairbanks, AK, USA. <sup>7</sup> Institute of Ecology and Evolution, University of Oregon, Eugene, OR, USA. <sup>8</sup> Future Food Beacon and School of Life Sciences, University of Nottingham, Nottingham, UK. <sup>9</sup> Department of Biology, University of Konstanz, Konstanz, Germany. <sup>10</sup> Zhejiang Provincial Key Laboratory of Plant Evolutionary Ecology and Conservation, Taizhou University, Taizhou, China. <sup>11</sup> Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, USA. <sup>12</sup> Biology Department, College of William and Mary, Williamsburg, VA, USA. ✉email: [mario.vallejo@stir.ac.uk](mailto:mario.vallejo@stir.ac.uk)

Increasing global connectivity is leading to widespread species translocations<sup>1</sup>. Most biological communities now include introduced members that have recently moved beyond their native ranges, often with negative impacts<sup>2–6</sup>. Finding the origins of invaders helps develop strategies for prevention, management and eradication<sup>7,8</sup>. It is also crucial for understanding to what extent invaders adapted to novel environments, along with the mechanisms of such adaptations<sup>9,10</sup>.

Tracing the migration and spread of invasives is typically very challenging. Inferring introduction histories is often accomplished using historical records, genetic analyses, or a combination of both<sup>11–13</sup>. In most cases, historical records of first introduction are unavailable or unreliable. Genetic data has greatly improved our ability to study the origins of invasions, and often uses information derived from extant populations<sup>10</sup>. However, genetic inferences are usually confounded by demographic processes that shape the introduced populations, including multiple introduction events, bottlenecks, evolution in the introduced range, admixture and hybridisation<sup>13–15</sup>.

Here, we use historical and genomic data to generate and test hypotheses in order to unravel the rapid worldwide invasion by the common yellow monkeyflower, *Mimulus guttatus* Fischer ex DC. (*Erythranthe* spp. (L.) G. L. Nesom; Phrymaceae), a herbaceous plant native to Western North America that was introduced across the world in the nineteenth century<sup>16–20</sup>. Unlike many invasive and non-native species, detailed historic botanical records<sup>21</sup> and travel diaries of early explorers<sup>22</sup> allow us to clearly retrace the history of the first introduction of *M. guttatus* into Europe. Historical records of *M. guttatus* reaching the United Kingdom (UK) paint a clear picture, but beyond this little is known. Here, we consider the hypothesis that the UK acted as a bridgehead (where invasive populations serve themselves as sources for further invasions<sup>12,23</sup>) for worldwide invasion.

The first European record of *M. guttatus* appears in *Curtis's Botanical Magazine*<sup>21</sup>, which presents a plate of Langsdorff's *Mimulus* (*Mimulus langsdorffii* Donn ex Sims), featuring a flowering individual of *M. guttatus*. The provenance of the depicted material is from Grigori von Langsdorff who "...brought it, as we are informed, from Unalaska, one of the Fox Islands" (Unalaska, Aleutian Islands)<sup>21</sup>, in his capacity as a naturalist on a Russian expedition to the Alaskan territories in 1805. Langsdorff describes how the expedition reaches Unalaska on 16 July 1805, and, after anchoring in Sea-Otters Bay (probably present-day Ugadaga Bay), they travelled on foot to Iluluk (Dutch Harbor). Here, Langsdorff first encounters *M. guttatus*: "splendid flowers were in blow upon the shore, among which a new *Mimulus* and *Potentilla*, which has never yet been described, were particularly to be distinguished"<sup>22</sup> (p. 329). Material brought by Langsdorff made its way to various Botanic Gardens, including Moscow (where it is listed as *M. guttatus* Fischer *nom. nudum*) and Montpellier (where De Candolle validly published the name *M. guttatus*). The seeds of *M. guttatus* also reached the Botanic Gardens at Cambridge in 1812, and it is therefore almost certain that the original species description included specimens collected by Langsdorff in Unalaska<sup>19</sup>.

Presciently, the *Botanical Magazine* recognised the potential for *M. guttatus* to become established outside western North America, and the 1812 entry states that because the taxon has showy flowers and is "easily propagated by seeds, and most probably by its runners, must soon be very common."<sup>21</sup> In fact, the first naturalised populations in the British Isles are recorded by 1830<sup>24</sup>, rapidly spreading throughout the UK<sup>25</sup>. The introduction history of *M. guttatus* outside of the UK is much less well understood. *Mimulus guttatus* seems to have reached New Zealand and become naturalised by 1878<sup>26</sup>, and the introduction of this taxon to eastern North America may have occurred much

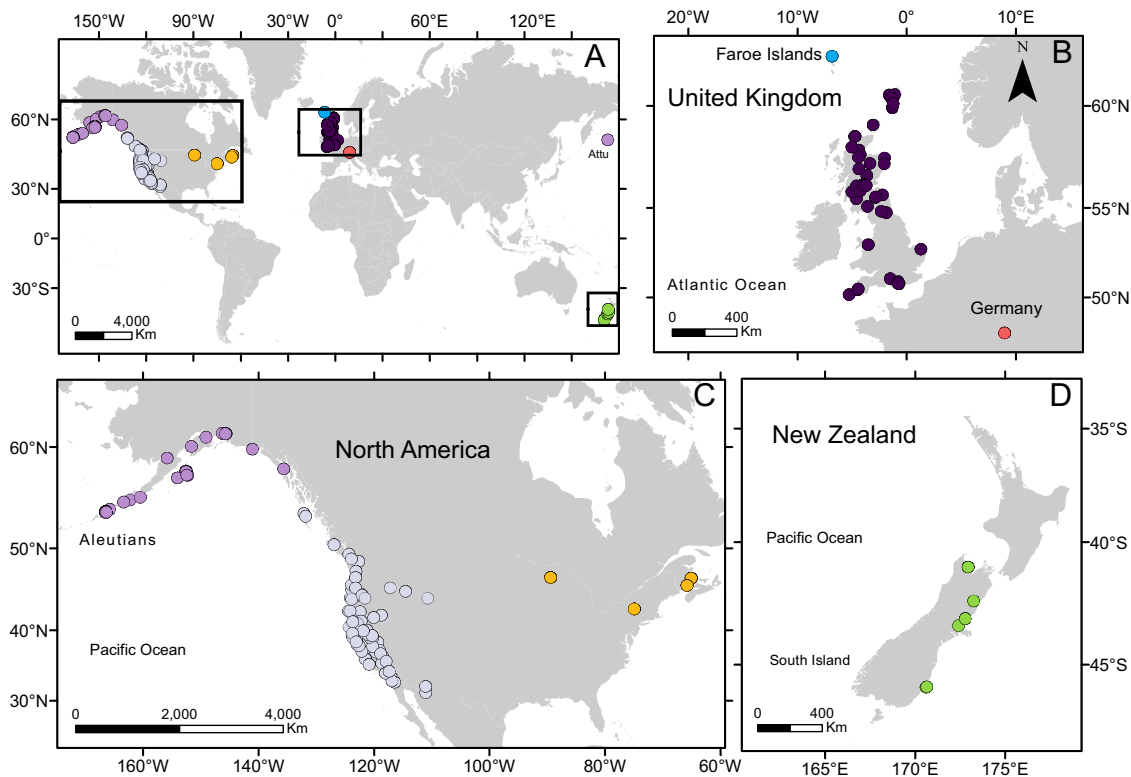
later in the second half of the twentieth century<sup>27</sup>. Therefore, the material brought in by Langsdorff represents the first introduction of *M. guttatus* outside its native range, and the subsequent arrival and naturalisation on the British Isles is the best documented, and currently most widespread, monkeyflower invasion<sup>17,20,24,25,28,29</sup>.

The historical hypothesis of an Alaskan origin of European monkeyflowers is consistent with results from previous genetic analysis of *M. guttatus* in the United Kingdom<sup>30,31</sup>. However, these studies did not include material from the putative origin (Aleutian Islands), and due to their focus on UK populations, did not examine genetic relationships between native populations and introduced populations in other parts of the range such as in Eastern North America, the Faroe Islands, mainland Europe and New Zealand. Native *M. guttatus* presents an enormous breadth of ecological and genetic diversity<sup>32,33</sup>, and it remains unknown how much of this diversity is represented among introduced populations and the extent to which non-native populations have diverged. Recently, Da Re et al.<sup>20</sup> used climatic niche modelling to compare the climatic envelope of native and introduced *M. guttatus* populations, finding no evidence of niche shift in the introduced UK populations compared to the native ones. Moreover, the highest niche similarity of invasive UK populations occurred in the Aleutian Islands<sup>20</sup>, lending support to the historical hypothesis that traces their origin to Langsdorff.

Here, we provide the first global genetic analysis of native and introduced populations of *M. guttatus* by marrying historical information with genomic analyses. Specifically, we: (1) Resolve range-wide relationships at the population-level in the introduced range, as well as in the native range including the previously under-sampled regions of the Aleutian Islands and mainland Alaska; and (2) use genomic data to reconstruct the population genetic history of introduced UK populations and test the hypothesis that UK populations have a simple Aleutian origin or are the product of a more complex invasion history. Our results show that populations from the British Isles were first likely introduced from the Aleutian Islands in Alaska, followed by rapid admixture from other parts of the native range. We hypothesise that these British populations then served as a bridgehead for invasions worldwide into the rest of Europe, New Zealand and eastern North America. Our findings emphasise the highly admixed nature of introduced *M. guttatus* and raise the possibility that introduced populations might serve as sources of secondary admixture, producing novel hybrids.

## Results

**Population relationships in the native range.** The global sampling of *M. guttatus*, including populations sampled across ~5000 km of its distribution in North America (Fig. 1), allowed us to resolve population groupings in both native and introduced ranges. In the native range, including the newly sampled Alaskan region, strong geographic structure is evident from phylogenetic analysis (Fig. 2), with four well-resolved North, South, Coastal and North Pacific clades<sup>34</sup>. The newly sampled populations in Alaska and the Aleutian Islands form part of the North Pacific Clade (Fig. 2). This clade is sister to the Coastal clade and includes populations from northern Washington to the westernmost Aleutian Islands (Attu Island). Phylogenetic analysis revealed an unexpected placement of some populations from inland Oregon, including those from Iron Mountain, which conflicts with previous analyses and their expected relationships based on simple geography. The tetraploid *M. guttatus* population sampled in the Shetland Islands in the UK is nested among other geographically proximate populations, further supporting the local origin of this autopolyploid in the introduced range<sup>35</sup>.



**Fig. 1** Global sampling of *Mimulus guttatus* populations. **A** Overview of all sampled population. **B** European populations. **C** North American populations. Populations in the west are native, while orange symbols represent invasive populations in eastern North America. **D** New Zealand populations. The symbols for the two southernmost populations in the South Island are overlapping.

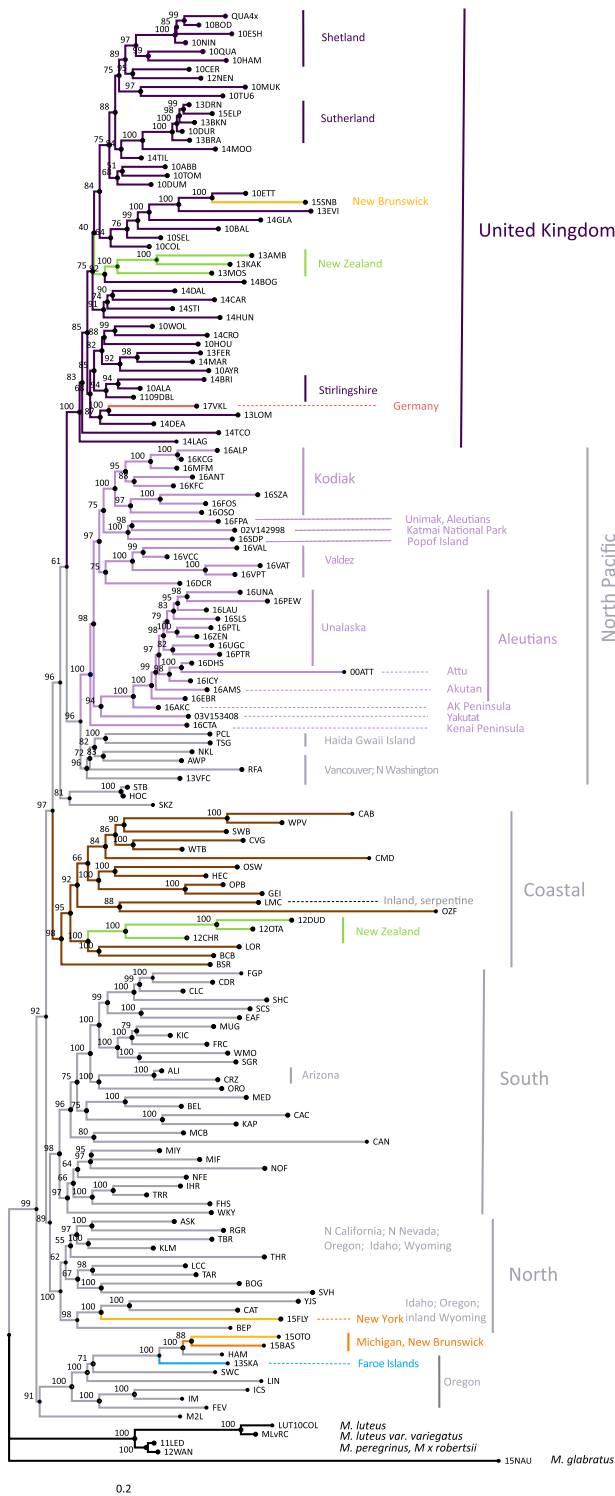
Finally, *M. luteus* formed a strongly supported clade, and the triploid and allohexaploid hybrids, *M. x robertsii* and *M. peregrinus* can be clearly distinguished from both parental taxa (*M. guttatus* and *M. luteus*).

**Global invasion of *Mimulus guttatus*.** At a global scale (Fig. 1), introduced *M. guttatus* populations are scattered across the phylogeny, indicating many independent introductions from across the native range (Fig. 2). In contrast, all UK *M. guttatus* populations form a sister group to the North Pacific clade. The UK group also includes other non-native populations from New Zealand, Canada and Germany, suggesting it may be the source for these. Other New Zealand populations are grouped within the Coastal clade, suggesting a potential second introduction. Moreover, interesting geographic discontinuities exist in North America, with a non-native New York population nested in the native North clade. Finally, two additional population populations from eastern North America, as well as the single sampled population from the Faroe Islands are grouped together with the native HAM-SWC group from Oregon (Fig. 2). Thus, the UK populations are genetically similar to each other and are closely related to some of the introduced populations of *M. guttatus* in New Zealand and eastern North America. However, the placement of other non-native populations within various native clades clearly indicates additional, independent introductions to New Zealand, eastern North America and the Faroe Islands, suggesting a complex history of colonisation.

Among native populations, those from the UK form a separate genetic cluster, as seen in principal component analysis (PCA) (Fig. 3). As in the phylogenetic reconstruction, the UK group is closely associated with non-native populations from New Zealand, Germany and eastern North America. The PCA is

also consistent with two separate introductions into New Zealand, one of them closely related to UK populations, and three independent origins of non-native populations in eastern North America. One of these origins of eastern North American populations is shared with the population from the Faroe Islands, forming a distinct group with two native populations from Oregon (SWC and HAM; Fig. 3). An interactive version of Fig. 3 with labelled individuals and populations is available at [https://mvallejo6.github.io/mimulus\\_voyage](https://mvallejo6.github.io/mimulus_voyage). Population structure in the native range is less clear from the worldwide PCA, although the North Pacific clade and particularly the Aleutian Islands populations are well differentiated along the first principal component (Fig. 3).

Worldwide groupings by *K*-means cluster analysis (Fig. 4 and Supplementary Fig. 1) partition North American samples into three groups, New Zealand into two groups, and the single populations from the Faroe Islands and Germany in one group each, largely consistent with the results above. Non-native UK populations form two groups, one mixed with European and Eastern North American samples, and another with New Zealand samples. Native, non-Alaskan populations are distributed in five groups. Aleutian populations form a separate group not shared with other geographic regions. The *fastStructure* analysis with  $K^* = 7$  and  $K = 8$  (Supplementary Fig. 2) provides further support for these groupings ( $K^*$  = optimal *K* estimated in ref. <sup>36</sup>). UK populations form a separate group with multiple affinities to New Zealand and eastern North American samples. Furthermore, the distinctiveness of Aleutian populations relative to other native populations is also obvious (e.g., cluster 4 at  $K = 8$ , Supplementary Fig. 2). In terms of nucleotide diversity ( $\pi_{\text{GENOME}}$ ), UK populations show, on average, a relatively high level of diversity compared to Alaskan populations ( $\pi_{\text{GENOME}} = 0.0047$ ,  $n = 43$  populations vs.  $\pi_{\text{GENOME}} = 0.0037$ ,  $n = 32$ , respectively), but only slightly lower than the overall average



**Fig. 2 Phylogenetic reconstruction of studied *Mimulus*.** Maximum likelihood phylogenetic reconstruction of the relationship between studied *Mimulus guttatus* populations, and including populations from *M. luteus* (LUT10COL, UK), *M. luteus* var. *variegatus* (MLVRC, Chile) *M x robertsii* (12WAN) and *M. peregrinus* (11LED). The tree is rooted using a population of *M. glabratus* from Michigan (15NAU).

of other native North American populations ( $\pi_{\text{GENOME}} = 0.0050$ ,  $n = 68$  populations; Supplementary Fig. 5). Other invasive populations have similar or lower average levels of nucleotide diversity as the UK ( $\pi_{\text{GENOME}} = 0.0031\text{--}0.0047$ ), although the number of

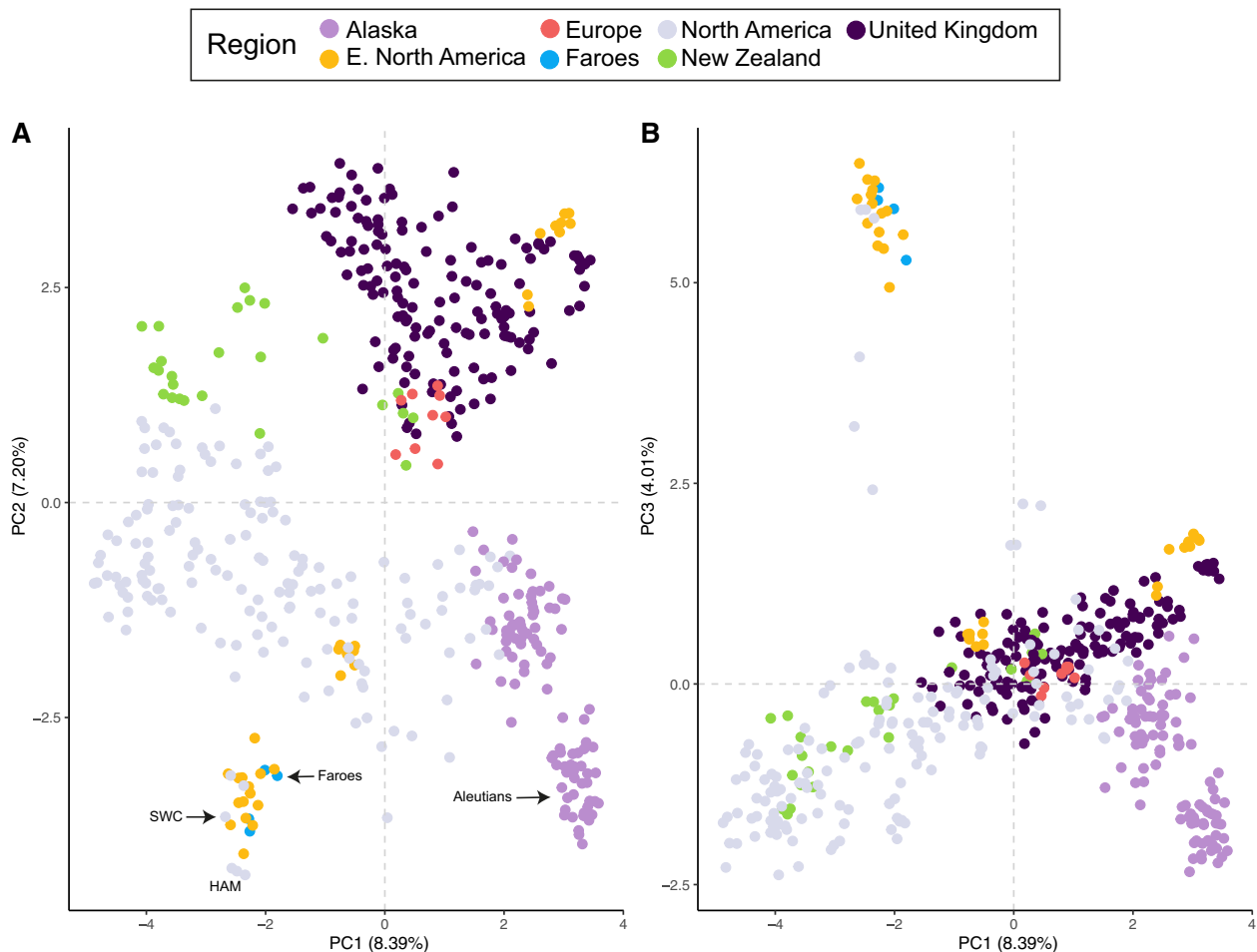
sampled populations in each of these regions is small (1–6; Supplementary Fig. 5).

**Introduction history in the UK.** To estimate a most likely scenario for the origin and history of introduction of UK populations, we next performed a coalescent analysis with ABC. Our analysis of demographic models allowed us to compare different scenarios for the origin and history of introduction of UK populations relative to five genetic groups in the native range: Aleutians (ALE) and Alaska-British Columbia (AKBC), both of which form part of the North Pacific clade, and the North (NORTH), South (SOUTH), and Coastal (COAST) clades (see Figs. 2 and 4)<sup>34,37</sup>. When assuming a single introduction event, the most likely source of UK individuals is the AKBC group (Table 1, posterior probability  $p = 0.89$ ). However, model comparisons favour scenarios with additional waves of introductions (Table 1). When we model two introductions, a first introduction from AKBC followed by a second introduction wave from NORTH has greatest support (Table 1,  $p = 0.48$ ) and is more likely than a single introduction scenario (237 votes against 32 votes, Table 1). Similarly, three introduction models result in selecting an introduction history with a first introduction from AKBC followed by additional introductions from NORTH and COAST ( $p = 0.53$ , Table 1) and then four introduction models identify a first introduction from SOUTH followed by additional introductions from AKBC, NORTH and COAST as the most likely scenario ( $p = 0.55$ , Table 1). Finally, when comparing all best one- to four-wave introduction models, with all possible five-wave introduction models, the most likely introduction history identified consisted of a first introduction from ALE followed by four subsequent waves from the AKBC, NORTH, SOUTH and COAST (E4 model;  $p = 0.55$ , Table 1). Full demographic parameters (e.g., estimated population sizes and introduction times per genetic group; E4 model) are presented in Supplementary Table 1.

Classification of the datasets simulated under the best one- to five-wave introduction scenario showed that 41.6% of the simulated datasets under E4 scenario were correctly classified, and a 50.6% probability that a simulation classified as E4 truly originated from this model (Table 2). Thus, the combination of the type and number of molecular markers and model prior specifications we used here contain enough information to confidently differentiate scenarios with different number of introductions (e.g., single or two-wave introductions vs. three- to five-wave introductions). Nevertheless, distinguishing the most likely scenario among these complex and sometimes very similar multiple introductions scenarios proved more difficult (Table 2, Supporting Material File 1). This is likely due to the wide prior distribution specification for the proportion of migrants that ranged from 0.001 to 0.999 (Supporting Material File 1). A scenario with a five-waves introduction but with very few migrants originating from one of the source (for instance 0.2% of the receiving population) would be very similar to a four-wave introduction. Conversely, a five-waves scenario with a very high migration rate from one source and very low migration rate from the other sources would underestimate the number of introductions. Given the scenario complexity, our ability to distinguish the order of introductions is even more limited. The combination of relatively low posterior probability of the best scenario (0.55, Table 1) and high prior error rate (0.41, estimated from out-of-bag procedure) indicates moderate levels of confidence in scenario choice that is typical for complex multiple introductions scenario<sup>38</sup>.

The posterior probability of 0.55 for the E4 model (Table 1), supports a first introduction from ALE followed by additional





**Fig. 3 Principal component analyses of native and introduced *Mimulus guttatus*.** Principal component analysis (PCA) of 474 individuals of *Mimulus guttatus* from both native and introduced populations genotyped at 1498 binary SNP loci. **A** Scatterplot of the first two principal components (PC2 vs. PC1). **B** Scatterplot of first and third principal components (PC3 vs. PC1). Colours indicate sample regions. An interactive 3D figure with individually labelled data points is available at: [https://mvallejo6.github.io/mimulus\\_voyage/](https://mvallejo6.github.io/mimulus_voyage/). A visualisation of the first 200 principal components (cumulative variance explained = 90.9%) projected in two dimensions using Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP)<sup>77</sup> is shown in Supplementary Fig. 4.

introductions from the other four origins (Supplementary Fig. 3). However, most of the posterior distributions of demographic parameters (e.g., effective population size, number of generations since introduction) for model E4 were nearly identical to the prior distributions (Supplementary Table 1 and Supplementary Notes 2), indicating limited information content of the genetic data set to estimate the demographic parameters of this complex introduction history.

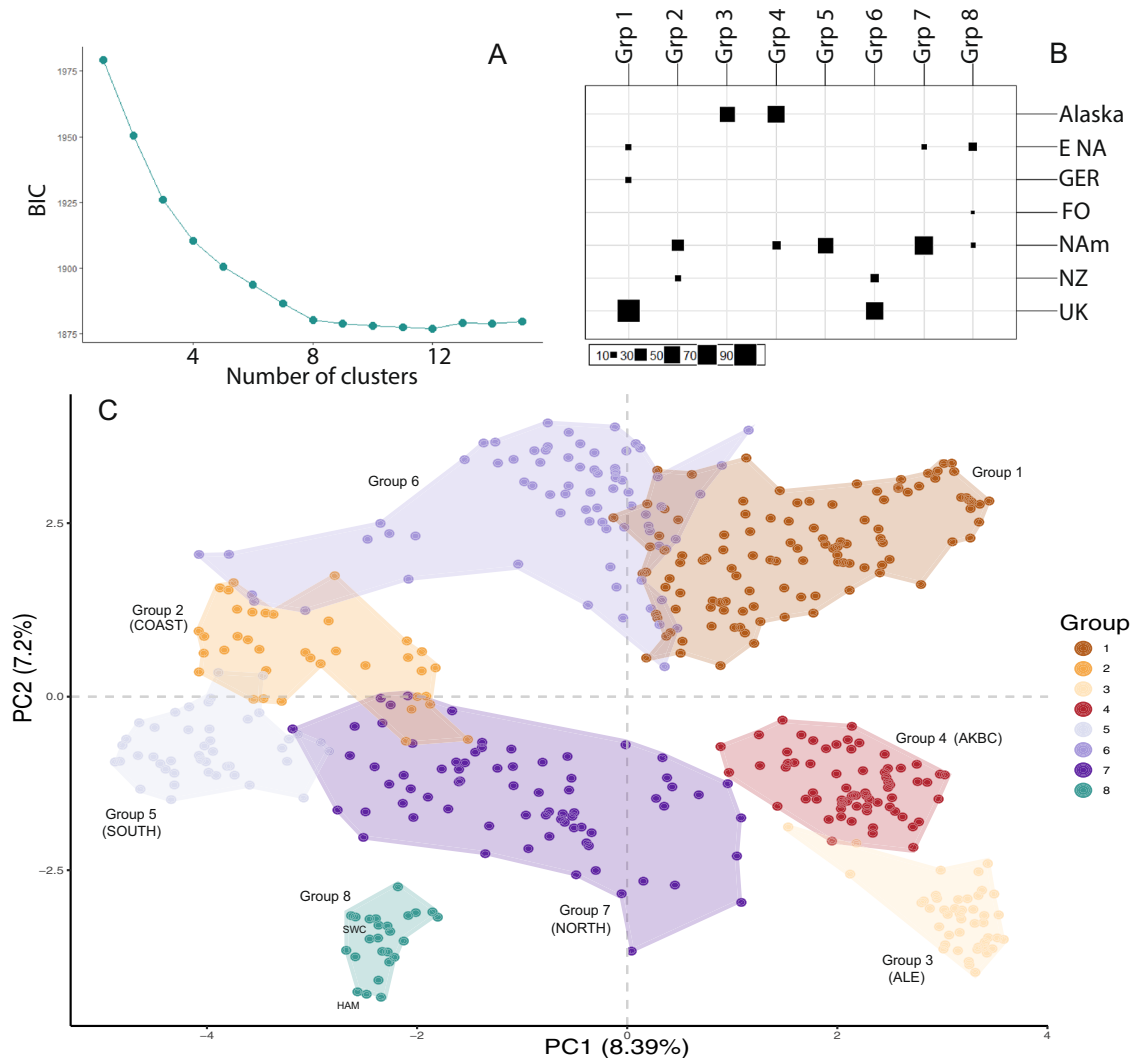
## Discussion

Here, we provide the first global picture of the genetic relationships between native and introduced populations of *Mimulus guttatus*, including targeted sampling of a historically indicated origin for the UK populations. Our results can be summarised in three main findings: (1) *Mimulus guttatus* achieved a broad distribution across geographic boundaries through multiple introductions from genetically distinct source populations. (2) In some cases, the establishment of *M. guttatus* in the invasive range might have been achieved via a bridgehead process, where invasive populations serve themselves as sources for further, more distant vanguard invasions. This is well illustrated in our discovery of the establishment of invasive populations in New Zealand and eastern North America by way of UK invasive populations. (3) Admixture in the introduced range has given rise to genetically distinct

populations generating novel genetic, and therefore phenotypic, combinations (e.g., invasive phenotypes that produce both large numbers of flowers and are highly clonal<sup>39</sup>).

Widely distributed taxa that serve as a source of invasive populations pose a particular challenge for molecular studies aiming to reconstruct the history of biological invasions. The distribution of *M. guttatus* spans from Mexico to the Aleutians and covers >6000 km of coastline<sup>33</sup>. To identify potential sources of specific invasion events, sampling large geographic regions is required. *Mimulus guttatus* has been the subject of continuous study for the last 60 years<sup>32</sup>, and previous work has collected population samples across nearly its entire native range<sup>40–42</sup>. Our analyses of large-scale population samples from the native range builds on the recent finding of geographic genetic structure corresponding to separate coastal and northern colonisation events in North America<sup>37</sup>. Here, we fill in crucial gaps with sampling from Alaska and the Aleutian Islands, which reveals strong geographic structure in the far north west of the species range, with genetic clusters by islands in the Aleutians. This extensive sampling in the native range allows us to show that Aleutian populations have acted as important conduits to the invasion of *Mimulus* in Europe and beyond.

Many biological invasions by both plants and animals are associated with multiple introductions, to the extent that single



**Fig. 4** Clustering analysis of native and introduced populations of *Mimulus guttatus*. K-means clustering analysis of native and introduced populations of *Mimulus guttatus*. The analysis is based on the first 300 principal components. **A** Bayesian Information Criterion values for models ranging from 2 to 15 clusters. **B** Group membership of each geographic group for the optimal number of clusters ( $K = 8$ ). **C** Principal component analysis depicted in Fig. 3A but coloured by the groups identified in the K-means cluster analysis ( $K = 8$ ). To improve clarity a polygon was drawn around each of the eight genetic groups. The genetic groups of native populations used for the ABC analysis are named in parenthesis: AKBC = North Pacific Clade.

introduction invasions are considered the exception<sup>9</sup>. Here, we found clear evidence that introduction of *M. guttatus* into various geographic regions has occurred by colonisation from multiple genetically distinct sources. For example, among the four populations we sampled in eastern North America, where *M. guttatus* was introduced in the last century, there is evidence of three genetically distinct groups, one of which also occurs in the Faroe Islands (Fig. 3). Similarly, introduced populations in New Zealand have at least two separate genetic origins, including a close affinity with native populations (near Santa Cruz, California) located 11,000 km away and with non-native populations in the UK. The multiple origins of invasive populations found in the same geographic region is important for several reasons. From a management perspective, multiple introductions can help identify locations of transport routes that are susceptible for further invasions. Moreover, multiple introductions may help invasive populations overcome demographic and genetic bottlenecks associated with introduction events<sup>9</sup>, as demonstrated by the relatively high levels of nucleotide diversity observed in introduced populations of *M. guttatus* in the British Isles (Supplementary Fig. 5). In species that are introduced via the ornamental

trade, as was probably the case for monkeyflowers, repeated introductions may not be unusual. To date it is still possible to freely purchase monkeyflowers in UK garden centres. However, because the type sold is no longer *M. guttatus* but horticultural varieties of its close relative *M. luteus*, we speculate that the multiple introductions detected in the invasive range of *M. guttatus* reflect historical events (nineteenth and twentieth centuries) rather than recent reintroductions. In addition, we did not find evidence of large-scale admixture from *M. luteus* shaping genetic variation in *M. guttatus*, consistent with the strong reproductive barriers imposed by differences in ploidy level between these *Mimulus* taxa<sup>43</sup>.

The genetic history of these invasions reveals a complex series of introduction events associated with early establishment (nineteenth century). Our ABC analyses reconstruct this history and show that extant populations are composed of a combination of multiple genetic groups from across the native range. Reconstruction of demographic events during introduction (Supplementary Fig. 3) supports an initial introduction of *M. guttatus* from the Aleutian Islands, which is consistent with the historical records of Langsdorff's expedition and subsequent transfer of

**Table 1 Comparison of demographic models of the invasion of *Mimulus guttatus* into the United Kingdom.**

Model group	Number of introduced origins	Model	First introduced origin	Following introduced origins	Votes (posterior probability of best model)
A	1	A1	ALE		121
	1	<b>A2</b>	<b>AKBC</b>		<b>276 (0.89)</b>
	1	A3	COAST		219
	1	A4	NORTH		206
	1	A5	SOUTH		178
B	1	A2	AKBC		32
	2	B1	AKBC	ALE	45
	2	B2	AKBC	COAST	78
	2	<b>B3</b>	<b>AKBC</b>	<b>NORTH</b>	<b>237 (0.48)</b>
	2	B4	AKBC	SOUTH	172
	2	B5	ALE	AKBC	30
	2	B6	COAST	AKBC	92
	2	B7	NORTH	AKBC	183
C	2	B8	SOUTH	AKBC	131
	1	A2	AKBC		28
	2	B3	AKBC	NORTH	30
	3	C1	AKBC	NORTH,ALE	60
	3	<b>C2</b>	<b>AKBC</b>	<b>NORTH,COAST</b>	<b>160 (0.53)</b>
	3	C3	AKBC	NORTH,SOUTH	74
	3	C4	NORTH	AKBC,ALE	98
	3	C5	NORTH	AKBC,COAST	118
	3	C6	NORTH	AKBC,SOUTH	114
D	3	C7	ALE	AKBC,NORTH	96
	3	C8	COAST	AKBC,NORTH	136
	3	C9	SOUTH	AKBC,NORTH	86
	1	A2	AKBC		22
	2	B3	AKBC	NORTH	34
	3	C2	AKBC	NORTH,COAST	106
	4	D1	AKBC	NORTH,COAST,ALE	116
	4	D2	AKBC	NORTH,COAST,SOUTH	98
	4	D3	NORTH	AKBC,ALE,COAST	86
	4	D4	NORTH	AKBC,COAST,SOUTH	122
E	4	D5	COAST	AKBC,NORTH,ALE	92
	4	D6	COAST	AKBC,NORTH,SOUTH	78
	4	D7	ALE	AKBC,NORTH,COAST	110
	4	<b>D8</b>	<b>SOUTH</b>	<b>AKBC,NORTH,COAST</b>	<b>136 (0.55)</b>
	1	A2	AKBC		46
	2	B3	AKBC	NORTH	42
	3	C2	AKBC	NORTH,COAST	127
	4	D8	SOUTH	AKBC,NORTH,COAST	106
E	5	E1	SOUTH	AKBC,NORTH,COAST,ALE	133
	5	E2	AKBC	NORTH,COAST,SOUTH,ALE	134
	5	E3	NORTH	AKBC,SOUTH,COAST,ALE	120
	5	<b>E4</b>	<b>ALE</b>	<b>AKBC,NORTH,SOUTH,COAST</b>	<b>151 (0.55)</b>
	5	E5	COAST	AKBC,NORTH,SOUTH,ALE	141

Stepwise comparison of demographic models of the invasion of *Mimulus guttatus* into the United Kingdom using 10,000 simulations for each of the model and random forest ABC model selection approach. At each step (model groups A-E), more complex introduction histories are considered while keeping the most likely models selected in previous comparison steps. The most likely model at each step is indicated in bold.

material to Russian, European and British collections. The colonisation of the UK by these exotic Aleutian monkeyflowers may have been facilitated by the close similarity of the ecological niche of *M. guttatus* in the British Isles and the Aleutian Islands<sup>20</sup>. Climatic pre-adaptation<sup>44</sup> of Aleutian monkeyflowers provided early arrivals with an opportunity for initial establishment. It is also clear that an initial introduction from the Aleutian Islands was accompanied or quickly followed by multiple introductions from other parts of the range. The UK seems to have become a melting pot for *M. guttatus* resulting in admixture of previously differentiated populations, which resulted in the creation of a unique set of genotypes that are now characteristic of UK populations (Fig. 4).

Invasive populations can themselves become sources for subsequent invasions, a phenomenon termed the “bridgehead effect”<sup>12</sup>. For example, the invasion of Australia by ragweed (*Ambrosia artemisiifolia*, Asteraceae) occurred not from native North American populations, but from populations in the introduced European range<sup>23</sup>. Our results indicate that UK populations served as a stepping-stone for secondary invasions in other parts of the non-native range. This potential bridgehead effect in invasive monkeyflowers is most clearly illustrated in the invasion of New Zealand. Some invasive populations in New Zealand share a close genetic affinity to UK populations. The genetic similarity is consistent with the exchange of biological material, including horticultural taxa, in the nineteenth century,



**Table 2 . (A) Power to discriminate between alternative demographic models using an “out-of-bag” procedure given the parameter model specification.**

A.	Classified models	A2	B3	C2	D8	E1	E2	E3	E4	E5	Total	Classification error	Probability that E4 is selected
	<b>Simulated models</b>												
	A2	<b>8902</b>	1043	24	12	2	9	0	1	7	10000	11.0%	0.0%
	B3	<u>2151</u>	<b>7616</b>	26	15	4	73	11	25	79	10000	23.8%	0.3%
	C2	<u>210</u>	463	<b>4844</b>	<u>1576</u>	675	642	534	469	587	10000	51.6%	5.5%
	D8	447	330	<u>3067</u>	<b>1905</b>	<u>1039</u>	662	908	911	731	10000	81.0%	10.7%
	E1	355	336	<u>2094</u>	<u>1342</u>	<b>1462</b>	735	<u>1323</u>	<u>1339</u>	<u>1014</u>	10000	85.4%	15.8%
	E2	400	<u>1317</u>	<u>2161</u>	<u>1039</u>	643	<b>1473</b>	<u>756</u>	<u>1062</u>	<u>1149</u>	10000	85.3%	12.5%
	E3	28	<u>625</u>	<u>1894</u>	<u>1173</u>	<u>1277</u>	857	<b>1641</b>	<u>1438</u>	<u>1067</u>	10000	83.6%	17.0%
	E4	344	<u>1096</u>	<u>1024</u>	993	<u>1219</u>	918	1313	<b>2009</b>	<u>1084</u>	10000	79.9%	<b>23.7%</b>
	E5	435	<u>1291</u>	<u>1491</u>	777	963	<u>1170</u>	<u>1006</u>	<u>1228</u>	<b>1639</b>	10000	83.6%	14.5%
	Total	<b>13272</b>	<b>14117</b>	<b>16625</b>	<b>8832</b>	<b>7284</b>	<b>6539</b>	<b>7492</b>	<b>8482</b>	<b>7357</b>			

The comparisons are made at the final selection step between the most likely one- to four-wave introduction models and all possible five-wave introduction models. The table shows how many of the 10,000 simulated datasets generated under a given scenario (A2 to E5, rows) were classified into each demographic scenario (A2 to E5 columns). The number of incorrect classifications is then used to compute the overall classification error. The last column shows the percentage of simulated models classified as E4 (which was the most likely scenario for the observed genetic dataset). Bold numbers indicate correct classification, and underlined numbers indicate >10% incorrect classification. (B) Probability of a given number of origins given that the E4 model is selected.

as British people migrated to New Zealand<sup>45</sup>. The single sampled population in continental Europe (Germany) also shows a close relationship to UK populations. Unfortunately, without further sampling it is difficult to establish whether UK populations contribute to the extant populations of *M. guttatus* in Europe. The hypothesis that UK populations have served as bridgeheads to other invasions remains to be further investigated. Morphologically, *M. guttatus* populations in Russia, Germany and the Czech Republic resemble UK material (Vallejo-Marin, *pers. obs.*) but the genetic identity of continental Europe populations remains to be investigated. In this regard, genomic analyses of herbarium specimens could provide important additional insights<sup>46</sup>. Particularly tantalising would be to compare specimens from herbaria in Russia, France and the UK, where historical links connect early *Mimulus* collections with Langsdorff's expedition to Alaska in the early 19th century. Finally, we also detected a close affinity between UK populations and a population in the non-native range in eastern North America. Populations of *M. guttatus* in eastern North America are generally small, occurring in the states of Michigan, New York, USA and in New Brunswick, Canada<sup>27</sup>. These small and sparsely distributed populations show diverse genetic origins and seem to be much more recently established (second half of the twentieth century). The mechanism of introduction of UK material into eastern North America is unknown but it could be associated with horticultural exchanges<sup>1,5,47</sup>. Our results combined with previous work on other systems (e.g., refs. 12,23), highlight the importance that bridgehead populations can have on biological invasions. Bridgehead populations may have already been selected for their ability to colonise beyond their native range, making them particularly good candidates for additional range expansions.

Multiple introductions and admixture can, in principle, both increase or decrease the performance and adaptive potential of invasive populations<sup>48–50</sup>. Multiple introductions from genetically distinct sources introduce variation and alleviate the negative effects of demographic bottlenecks associated with colonisation. Moreover, genetically diverse populations are less likely to experience the deleterious effects of inbreeding depression<sup>50,51</sup> and can increase individual fitness through heterosis<sup>49</sup>, which can potentially make invasive populations more difficult to control. In contrast, admixture may reduce overall fitness if gene flow results in outbreeding depression<sup>52</sup>, a phenomenon that can occur due to epistatic interactions or, for

example, the breakdown of locally adapted genotypes. In *M. guttatus*, experimental work indicates that both positive and negative effects of admixture can be observed in invasive populations. For example, crossing native and introduced populations results in an increase in biomass, and both clonal and sexual reproduction in greenhouse conditions<sup>53,54</sup>. In field conditions, the effects of admixture can be reversed, and a common garden study shows that admixture between UK *M. guttatus* and both annual and perennial populations from the native range result in lower fitness as estimated using population growth rates<sup>39</sup>. The effects of admixture may be particularly strong on invasive species with a widespread, highly diverse native distribution, such as *M. guttatus*. Native populations that occur over large, biogeographically diverse areas may serve as reservoirs of genetic and ecological variation. This wide range of ecogeographic variation may facilitate the colonisation of new regions in the introduced range and potentiate the effects of subsequent introductions and admixture on the performance and adaptive potential of invasive populations.

## Methods

**Study system and population sampling.** *Mimulus guttatus* Fischer ex DC (section *Simiolus*, Phrymaceae), the common monkeyflower, is a widespread species with a native range extending across western North America from northern Mexico to the farthest reaches of the Aleutian Island chain in Alaska<sup>20,33</sup>. The invasive range includes much of the UK, the Faroe Islands, parts of mainland Europe, New Zealand, and Eastern North America<sup>20</sup>. The species is self-compatible and predominantly outcrossing<sup>55</sup>. Most populations are diploid, although tetraploid populations occur throughout the native range<sup>56</sup> and tetraploid populations have also evolved in the introduced range<sup>35,56</sup>. In the native range, populations comprise either small annual plants that reproduce exclusively by seed or perennial plants that reproduce by both seed and vegetative stolons. Only perennial plants are documented in the invasive range.

We sampled populations of *M. guttatus* in the native range of western North America and the main areas of introduction in eastern North America, Europe and New Zealand for a total of 521 individuals from 158 populations (Fig. 1 and Table 3). In the native range, the samples included 70 previously genotyped populations<sup>34</sup>, spanning Arizona to British Columbia, plus an additional population from Vancouver Island. To fill the gap of previous studies, and to specifically address the hypothesis of an Alaskan origin of introduced UK populations, we collected samples from 32 populations in Alaska, including 14 populations from the Aleutian Islands (Attu, Unalaska, Akutan and Unimak) (Supplementary Data 1). Voucher specimens of the newly sampled populations are deposited in the University of Alaska herbarium (ALA). In the introduced range, we sampled four populations in eastern North America, one from the Faroe Islands, one from Germany, six from New Zealand, and 43 from UK populations from Cornwall to the Shetland Islands. As an outgroup we included three diploid individuals from a population of *M. glaberrimus* from Michigan, USA. We also

**Table 3 Number of populations and individuals sampled and sequenced.**

	Region	Number of populations	Number of individuals
Native	Western North America (excluding Alaska)	71	182
	Western North America (Alaska only)	32	106
Introduced	Eastern North America	4	34
	Faroe Islands	1	4
	United Kingdom	43	161
	Germany	1	9
	New Zealand	6	25
Total		158	521

Summary of the number of populations and individuals sampled and sequenced. A detailed breakdown by population is shown in Supplementary Data 1.

sampled three tetraploid UK *M. guttatus*, 19 individuals of *M. leuteus* from both native and introduced ranges (with which *M. guttatus* hybridises in the introduced range to produce a sterile but widespread triploid, *M. x robertsii*), three *M. x robertsii*, and three *M. peregrinus* (the allohexaploid species derived by whole genome duplication from *M. x robertsii*<sup>57</sup>; (Supplementary Data 1). In total, we had samples from 103 populations of *M. guttatus* from the native range, and 55 populations from the introduced range (Table 3). Full sample details are provided in Supplementary Data 1.

**Genotyping.** To obtain DNA for genotyping, we germinated field-collected seeds from all new populations in a controlled environment facility at the University of Stirling. We extracted genomic DNA from fresh leaves or flower buds using the DNeasy Plant Kit (Qiagen, Germantown, MD), with samples standardised to 100 ng DNA for library preparation. We used genotyping by sequencing (GBS) to generate genome-wide polymorphism data<sup>58</sup>. For GBS library preparation, we used the same protocol as Twyford and Friedman<sup>34</sup>, using the enzyme PstI, and pooling samples in a 95-plex (plus one blank water control) for 100 bp single-end sequencing on the Illumina HiSeq 4000 at the University of Oregon. We analysed raw sequence reads using the Tassel5-GBSv2Pipeline<sup>59</sup>, using the *M. guttatus* v2 genome<sup>60</sup> as a reference. Alignment with the reference genome was done with BWA with default parameters in Tassel5-GBSv2Pipeline. To obtain variant sites we used a minimum allele frequency of  $\text{mnMAF} = 0.01$ . For the genotyping step (SNP CallerPluginV2) we used the following options:  $\text{kmerLength} = 64$ ,  $\text{mnQs} = 20$ ,  $\text{minPosQS} = 20$ . For population genomic analyses, we retained only variable sites (SNPs), but for estimating nucleotide diversity across the genome ( $\pi_{\text{GENOME}}$ ) and for tree reconstruction, we generated a sequence matrix with both SNPs and invariant sites (setting  $\text{mnMAF} = 0$ ).

To estimate pairwise nucleotide diversity across the genome of *M. guttatus* populations ( $\pi_{\text{GENOME}}$ ), we filtered the data set containing both variant and invariant sites (123,141 loci) using VCFtools version 0.1.15<sup>61</sup>. We kept only biallelic loci that were genotyped in at least 75% of all individuals, which reduced the data set to 26,012 loci. We then removed individuals with <50% genotyped loci, reducing the number of individuals from 521 to 475. Nucleotide diversity per site was calculated separately for each of 155 populations, both native and invasive, using the option `--site-pi` in VCFtools. An overall pairwise nucleotide diversity across the ( $\pi_{\text{GENOME}}$ ) for each population was then obtained by averaging across all sites.

For population genetic analyses in *M. guttatus*, we filtered the SNP data (44,552 loci from 521 *M. guttatus* individuals) using VCFtools version 0.1.15<sup>61</sup> and kept only biallelic loci that were genotyped in at least 75% of all individuals, which reduced the number of genotyped SNPs to 1820 loci. The parameters we used in this filtering step were:  $\text{max-missing} = 0.75$ ,  $\text{mac} = 3$ ,  $\text{min-alleles} = 1$ ,  $\text{max-alleles} = 2$ . We then removed individuals with <50% genotyped loci, reducing the number of individuals from 521 to 474. Finally, we used PLINK version 1.9<sup>62</sup> to thin the data set to reduce linkage disequilibrium among SNPs using a pairwise correlation coefficient of 0.5 (`--indep-pairwise 50 5 0.5`). The final *M. guttatus* data set contained 1498 SNPs from 474 individuals in 155 populations (Supplementary Table 2).

**Tree building.** We sought to resolve evolutionary relationships between populations and species using polymorphism-aware phylogenetic models implemented in IQ-TREE<sup>63</sup>. These models use population site frequency data, and therefore account for incomplete lineage sorting<sup>64</sup>. This phylogeographic approach generates an initial visualisation of population history and broad scale geographic genetic structure from the genome-wide signal, prior to more detailed characterisation with population-level approaches (described below). We analysed two datasets, one for all sampled *Mimulus* taxa, and one for *M. guttatus*, with both datasets, including *M. glabratus* as an outgroup. Each analysis used the full GBS sequences with invariant sites (123,141 loci including both variant and invariant sites), filtered to include 8798 sites with <50% missing data. We calculated population allele frequencies using the counts file library (cflib) python scripts that accompany<sup>64</sup>. Model-fitting was performed with ModelFinder<sup>65</sup>. IQ-TREE analyses subsequently used the best-fitting model (TVM + F + G4) allowing for excess polymorphism

(+P) and with five chromosome sets per population (+N5). Tree searches were performed with settings recommended for short sequences, including a small perturbation strength (-pers 0.2) and large number of stop iterations (-nstop 500). Topological support was assessed using an ultrafast bootstrap approximation approach<sup>66</sup>, with 1000 bootstrap replicates. Trees were visualised with FigTree<sup>67</sup>.

**Population genetic structure.** To analyse population genetic structure, we conducted a principal component analysis using the *glPca* function in *adegenet*<sup>68</sup> in R ver. 4.0.0<sup>69</sup>. We used *K*-means grouping implemented with the function *find.clusters* in *adegenet* to identify clusters of individuals in the data without using a priori groupings. For this analysis, we used 100 randomly chosen centroids for each run, and calculated the goodness of fit for each model for values of *K* between two and 15. For the selected *K*-value, we also ran a discriminant analysis of principal components (DAPC)<sup>70</sup> using the inferred groups for assigning individual membership. We further used *fastStructure* version 1.0<sup>71</sup> to infer population structure across *M. guttatus* populations using a Bayesian framework. For this analysis, we randomly subsampled the data to include a maximum of three individuals per population (408 individuals in total) from both native and introduced ranges, and analysed values of *K* from 2 to 8.

**Introduction history reconstruction by ABC.** Our preliminary analyses indicated that introduced *M. guttatus* had a complex origin with multiple introductions in different non-native regions. In order to gain a more detailed understanding of the demographic history of non-native populations, we focused on the introduction of *M. guttatus* to the UK, which has been best studied both historically and genetically<sup>30,31</sup>. Therefore, we implemented an approximate Bayesian computation (ABC) approach to determine the most likely *M. guttatus* introduction history in the UK. For this analysis, we used the pruned data set consisting of 1498 SNPs but included only individuals from the native range or the UK (399 individuals). Individuals from the native range were grouped into one of five groups ("genetic group"). Four of these groups (AKBC, NORTH, SOUTH and COAST) have been previously identified in analyses using native populations only<sup>34,37</sup>. The fifth group (Aleutian Islands, ALE) is identified in the present study (see Results section). To assign native individuals to each of these genetic groups, we used the group membership obtained in the *K*-means cluster analysis, which produced genetic groups that closely correspond to the native clades previously identified in Twyford and Friedman<sup>34</sup> and Twyford et al.<sup>37</sup> (see "Results" section). The number of individual in each group is as follows: NORTH,  $N = 62$ ; SOUTH,  $N = 42$ ; COAST,  $N = 30$ ; AKBC (Alaska-British Columbia),  $N = 70$ ; and ALE,  $N = 45$ . Six individuals from two populations (SWC and HAM) that formed a separate genetic group in the native range were not included in this analysis. Individuals from the UK were considered to belong to a single population (UK;  $N = 150$ ).

As all possible scenarios of divergence between the five native groups would have been computationally impossible to test, native group genetic relationships were determined from the phylogenetic tree topology (see "Results" section). All the simulations assumed that the North population diverged from an ancestral population at time  $t_4$ , from which the South population diverged at time  $t_5$ . In addition, the Coastal population diverged from the ancestral population at time  $t_3$  from which the Alaska-British Columbia population diverged at time  $t_2$ , and the Aleutian population diverged from there at time  $t_1$ . The simulated demographic models share this native population divergence history and only differed by their introduction history into the UK.

We first considered simple introduction models where the UK population was derived from a single native origin at time  $t_{0a}$  (models A1 to A5, Supplementary Notes 1). We then simulated UK introduction from a single origin at time  $t_{0a}$  followed by a second introduction at time  $t_{0b}$  (two-waves introduction models; models B). This strategy resulted in the definition of eight different two-waves introduction models (models B1 to B8, Supplementary Notes 1). We then tested more complex introduction models using a similar logic, modelling three-waves (models C1 to C9), four-waves (models D1 to D8) and five-waves (models E1 to E5) introduction models by integrating the most likely origins identified in

previous sets of models to define a restricted number of models to compare. A full version of other assumptions and simulation parameters is given in Supplementary Notes 1.

For each demographic model, we simulated 10,000 genetic datasets consisting of 1435 independent SNP genotypes for 798 haploid individuals distributed following the sample size of all six populations in the real data set using *Fastsimcoal2* version 2.6.0.3<sup>72</sup> called by *ABCtoolbox* version 1<sup>73</sup>. We passed a custom bash script to *ABCtoolbox* to add missing genotypes to the simulated data set at an identical rate to the observed level in the real data. Then, we used *ABCtoolbox* to call the *arlsunstat* programme<sup>74</sup> to compute summary statistics from the simulated genotypes that were combined into diploid genotypes by randomly pairing haploid genotypes assuming Hardy-Weinberg equilibrium<sup>75</sup>. We computed all available statistics within and between populations for biallelic loci (67 summary statistics). In addition, we computed summary statistics within and between three defined regional groups (NORTH and SOUTH in one group; COASTAL, AKBC and ALE in a second group; and UK in a third group) representing an additional set of 29 summary statistics.

**ABC model comparisons.** We performed iterative model comparisons by comparing increasingly complex models (Table 1). In the first round, the introduction models assume a single introduction from one of the five native genetic groups. Then in round two, we considered two introductions models that necessarily involved the population origin from round one. This allowed us to define two sets of two-waves introduction models: one set consisting of four models with the most likely origin in previous rounds as the first introduction origin, followed by a second introduction from one of the four other native populations. Moreover, a second set of four models, which assume that the most likely origin in the previous round constitutes the second introduction, while the first introduction originated from one of the four other native populations (Table 1). We compared the most likely single introduction model and the eight two-waves introduction models. We then considered more complex models, comparing nine three-waves introduction models and the most likely single and two-waves introduction models (Table 1). We subsequently compared models assuming four-waves and five-waves of introduction while still including more simple models in the comparisons (Table 1). Demographic models were compared using a random forest approach implemented in the R package *abcrf*<sup>6</sup>.

We built a classification random forest model using 1000 trees and a training data set consisting of the summary statistics computed for the 10,000 simulated genetic datasets for each model. We estimated the classification error rate for each model using an “out-of-bag” procedure to quantify the power of the genetic data given the models and prior distribution specifications to differentiate the different demographic models. Then, we used the summary statistics computed based on the observed genotypic data to predict the demographic model that best fit the data using a regression forest with 1000 trees. We report the number of “votes” for each demographic scenario and the approximation of the posterior probability of the most likely model. We used the overall most likely scenario to simulate 100,000 genetic datasets using parameters and prior distributions described above to estimate demographic model parameters. We built a regression random forest model implemented in *abcrf* based on the summary statistics using 1000 trees. We estimated the posterior median, 0.05 and 0.95 quantiles of the model parameters by random forest regression model based on the summary statistics of the observed genotypic composition.

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

## Data availability

Genotype data as a VCF file is publicly available at <http://hdl.handle.net/11667/168> (DATAStorre, U. Stirling). Location data of sampled populations is available in the Supplementary Materials (Supplementary Data 1). Source data of Figs. 3 and 4 are provided in Supplementary Data 2. Herbarium specimens of newly collected material in Alaska is deposited at the ALA herbarium.

## Code availability

Code for the Approximate Bayesian Computation Analysis is available at <https://doi.org/10.15454/VUMC1P>.

Received: 15 July 2020; Accepted: 5 February 2021;

Published online: 12 March 2021

## References

- Chapman, D., Purse, B. V., Roy, H. E. & Bullock, J. M. Global trade networks determine the distribution of invasive non-native species. *Glob. Ecol. Biogeogr.* **26**, 907–917 (2017).
- Pysek, P. et al. A global assessment of invasive plant impacts on resident species, communities and ecosystems: the interaction of impact measures, invading species’ traits and environment. *Glob. Change Biol.* **18**, 1725–1737 (2012).
- Vila, M. et al. Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecol. Lett.* **14**, 702–708 (2011).
- Seebens, H. et al. No saturation in the accumulation of alien species worldwide. *Nat. Commun.* **8**, <https://doi.org/10.1038/ncomms14435> (2017).
- Seebens, H. et al. Global trade will accelerate plant invasions in emerging economies under climate change. *Glob. Change Biol.* **21**, 4128–4140 (2015).
- van Kleunen, M. et al. Global exchange and accumulation of non-native plants. *Nature* **525**, 100–103 (2015).
- Hulme, P. E. et al. Grasping at the routes of biological invasions: a framework for integrating pathways into policy. *J. Appl. Ecol.* **45**, 403–414 (2008).
- Hufbauer, R. A. Population genetics of invasions: can we link neutral markers to management? *Weed Technol.* **18**, 1522–1527 (2004).
- Dlugosch, K. M. & Parker, I. M. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Mol. Ecol.* **17**, 431–449 (2008).
- Welles, S. R. & Dlugosch, K. M. Population genomics of colonization and invasion. *Population Genomics: Concepts, Approaches and Applications*, (ed. Rajara, O. P.) 655–683 (Springer, 2019).
- van Boheemen, L. A., Atwater, D. Z. & Hodgins, K. A. Rapid and repeated local adaptation to climate in an invasive plant. *N. Phytologist* **222**, 614–627 (2019).
- Lombaert, E. et al. Bridgehead effect in the worldwide invasion of the biocontrol harlequin ladybird. *PLoS ONE* **5**, e9743–e9743 (2010).
- Estoup, A. & Guillemaud, T. Reconstructing routes of invasion using genetic data: why, how and so what? *Mol. Ecol.* **19**, 4113–4130 (2010).
- Dlugosch, K. M., Anderson, S. R., Braasch, J., Cang, F. A. & Gillette, H. D. The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Mol. Ecol.* **24**, 2095–2111 (2015).
- Bock, D. G. et al. What we still don’t know about invasion genetics. *Mol. Ecol.* **24**, 2277–2297 (2015).
- Vallejo-Marin, M. & Lye, G. C. Hybridisation and genetic diversity in introduced *Mimulus* (Phrymaceae). *Heredity* **110**, 111–122 (2013).
- Stace, C. A. *New Flora of the British Isles*. 3rd edn, (Cambridge University Press, 2010).
- Tokarska-Guzik, B. & Dajdok, Z. NOBANIS—Invasive alien species fact sheet: *Mimulus guttatus*. (Online Database of the North European and Baltic Network on Invasive Alien Species, NOBANIS. [www.nobanis.org](http://www.nobanis.org), date of access 2/2/2012, 2010).
- Grant, A. L. A monograph of the genus *Mimulus*. *Ann. Mo. Botanical Gard.* **11**, 99–380 (1924).
- Da Re, D., Olivares, A. P., Smith, W. & Vallejo-Marin, M. Global analysis of ecological niche conservation and niche shift in invasive and hybrid populations of monkeyflowers (*Mimulus guttatus*, *M. luteus* and *M. × robertsii*). *Plant Ecol. Divers.*, <https://doi.org/10.1080/17550874.2020.1750721> (2020).
- Sims, J. *Curtis’s Botanical Magazine. Flower-Garden Displayed*. Vol. 35 (Sherwood, Neeley & Jones, 1812).
- von Langsdorff, G. H. *Voyages and Travels in Various Parts of the World*. (George Philips, 1817).
- van Boheemen, L. A. et al. Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia. *Mol. Ecol.* **26**, 5421–5434 (2017).
- Roberts, R. *Mimulus* hybrids in Britain. *Watsonia* **6**, 70–75 (1964).
- Preston, C. D., Pearman, D. A. & Dines, T. D. *New Atlas of the British and Irish Flora*. (Oxford University Press, 2002).
- Owen, S. J. (Department of Conservation, Wellington, 1996).
- Murren, C. J., Chang, C. C. & Dudash, M. R. Patterns of selection of two North American native and nonnative populations of monkeyflower (Phrymaceae). *N. Phytologist* **183**, 691–701 (2009).
- McArthur, E. D. The cytotaxonomy of naturalized British *Mimulus*. *Watsonia* **10**, 155–158 (1974).
- Stace, C. A. & Crawley, M. J. *Alien Plants*. (William Collins, 2015).
- Puzey, J. R. & Vallejo-Marin, M. Genomics of invasion: diversity and selection in introduced populations of monkeyflowers (*Mimulus guttatus*). *Mol. Ecol.* **23**, 4472–4485 (2014).
- Pantoja, P. O., Simón-Porcar, V. I., Puzey, J. R. & Vallejo-Marin, M. Genetic variation and clonal diversity in introduced populations of *Mimulus guttatus* assessed by genotyping at 62 single nucleotide polymorphism loci. *Plant Ecol. Diversity* **10**, 5–15 (2017).
- Wu, C. A. et al. *Mimulus* is an emerging model system for the integration of ecological and genomic studies. *Heredity* **100**, 220–230 (2008).
- Vickery, R. K. *Case Studies in the Evolution of Species Complexes in Mimulus*. (Springer, 1978).



34. Twyford, A. D. & Friedman, J. Adaptive divergence in the monkey flower *Mimulus guttatus* is maintained by a chromosomal inversion. *Evolution* **69**, 1476–1486 (2015).
35. Simón-Porcar, V. I., Silva, J. L., Meeus, S., Higgins, J. D. & Vallejo-Marín, M. Recent autopolyploidization in a naturalized population of *Mimulus guttatus* (Phrymaceae). *Botanical J. Linn. Soc.* **185**, 189–207 (2017).
36. Li, Y.-L. & Liu, J.-X. StructureSelector: a web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol. Ecol. Resour.* **18**, 176–177 (2018).
37. Twyford, A. D., Wong, E. L. Y. & Friedman, J. Multi-level patterns of genetic structure and isolation by distance in the widespread plant *Mimulus guttatus*. *Heredity* **125**, 227–239 (2020).
38. Fraimout, A. et al. Deciphering the routes of invasion of *Drosophila suzukii* by means of ABC random forest. *Mol. Biol. Evol.* msx050, <https://doi.org/10.1093/molbev/msx050> (2017).
39. Pantoja, P. O., Paine, C. E. T. & Vallejo-Marín, M. Natural selection and outbreeding depression suggest adaptive differentiation in the invasive range of a clonal plant. *Proc. R. Soc. B: Biol. Sci.* **285**, 20181091 (2018).
40. Lowry, D. B., Hall, M. C., Salt, D. E. & Willis, J. H. Genetic and physiological basis of adaptive salt tolerance divergence between coastal and inland *Mimulus guttatus*. *N. Phytologist* **183**, 776–788 (2009).
41. Oneal, E., Lowry, D. B., Wright, K. M., Zhu, Z. & Willis, J. H. Divergent population structure and climate associations of a chromosomal inversion polymorphism across the *Mimulus guttatus* species complex. *Mol. Ecol.* **23**, 2844–2860 (2014).
42. Friedman, J., Twyford, A. D., Willis, J. H. & Blackman, B. K. The extent and genetic basis of phenotypic divergence in life history traits in *Mimulus guttatus*. *Mol. Ecol.* **24**, 111–122 (2015).
43. Meeus, S., Šemberová, K., De Storme, N., Geelen, D. & Vallejo-Marín, M. Effect of whole-genome duplication on the evolutionary rescue of sterile hybrid monkeyflowers. *Plant Commun.* <https://doi.org/10.1016/j.xplc.2020.100093> (2020).
44. Mack, R. N. Phylogenetic constraint, absent life forms, and preadapted alien plants: a prescription for biological invasions. *Int. J. Plant Sci.* **164**, S185–S196 (2003).
45. Bridge, C. & Fedorowich, K. *The British World: Diaspora, Culture and Identity*. (Routledge, 2004).
46. Gutaker, R. M., Reiter, E., Furtwangler, A., Schuenemann, V. J. & Burbano, H. A. Extraction of ultrashort DNA molecules from herbarium specimens. *Biotechniques* **62**, 76–79 (2017).
47. Haeuser, E. et al. European ornamental garden flora as an invasion debt under climate change. *J. Appl. Ecol.* **55**, 2386–2395 (2018).
48. Barker, B. S. et al. Potential limits to the benefits of admixture during biological invasion. *Mol. Ecol.* **28**, 100–113 (2019).
49. Rius, M. & Darling, J. A. How important is intraspecific genetic admixture to the success of colonising populations? *Trends Ecol. Evolution* **29**, 233–242 (2014).
50. Verhoeven, K. J. F., Macel, M., Wolfe, L. M. & Biere, A. Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. *Proc. R. Soc. B-Biol. Sci.* **278**, 2–8 (2011).
51. Dudash, M. R., Murren, C. J. & Carr, D. E. Using *Mimulus* as a model system to understand the role of inbreeding in conservation: genetic and ecological approaches. *Ann. Mo. Botanical Gard.* **92**, 36–51 (2005).
52. Frankham, R. et al. Predicting the probability of outbreeding depression. *Conserv. Biol.* **25**, 465–475 (2011).
53. Li, Y., Stift, M. & van Kleunen, M. Admixture increases performance of an invasive plant beyond first-generation heterosis. *J. Ecol.* **106**, 1595–1606 (2018).
54. van Kleunen, M., Rockle, M. & Stift, M. Admixture between native and invasive populations may increase invasiveness of *Mimulus guttatus*. *Proc. R. Soc. B-Biol. Sci.* **282**, 20151487 (2015).
55. Ritland, K. Correlated matings in the partial selfer *Mimulus guttatus*. *Evolution* **43**, 848–859 (1989).
56. Vickery, R. K., Crook, K., Lindsay, D., Mia, M. & Tai, W. Chromosome counts in section *Simiolus* of the genus *Mimulus* (Scrophulariaceae). VII. New numbers for *M. guttatus*, *M. cupreus*, *M. tilingii*. *Madroño*. Vol. 19, 211–218 (California Botanical Society, 1968).
57. Vallejo-Marín, M., Buggs, R. J. A., Cooley, A. M. & Puzey, J. R. Speciation by genome duplication: Repeated origins and genomic composition of the recently formed allopolyploid species *Mimulus peregrinus*. *Evolution* **69**, 1487–1500 (2015).
58. Elshire, R. J. et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **6**, e19379 (2011).
59. Glaubitz, J. C. et al. TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PLoS ONE* **9** (2014).
60. Hellsten, U. et al. Fine-scale variation in meiotic recombination in *Mimulus* inferred from population shotgun sequencing. *Proc. Natl Acad. Sci. USA* **110**, 19478–19482 (2013).
61. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
62. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 10047–13748 (2015). s13742-13015-.
63. Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evolution* **32**, 268–274 (2015).
64. Schrempf, D., Minh, B. Q., De Maio, N., von Haeseler, A. & Kosiol, C. Reversible polymorphism-aware phylogenetic models and their application to tree inference. *J. Theor. Biol.* **407**, 362–370 (2016).
65. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–+ (2017).
66. Minh, B. Q., Nguyen, M. A. T. & von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evolution* **30**, 1188–1195 (2013).
67. Rambaut, A. <http://tree.bio.ed.ac.uk/software/figtree/> (2014).
68. Jombart, T. & Ahmed, I. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**, 3070–3071 (2011).
69. R Development Core. *R A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>, 2020).
70. Jombart, T., Devillard, S. & Balloux, F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* **11**, 94–94 (2010).
71. Raj, A., Stephens, M. & Pritchard, J. K. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* **197**, 573–U207 (2014).
72. Excoffier, L., Dupanloup, I., Huerta-Sanchez, E., Sousa, V. C. & Foll, M. Robust demographic inference from genomic and SNP data. *PLoS Genet.* **9** (2013).
73. Wegmann, D., Leuenberger, C., Neuenschwander, S. & Excoffier, L. ABCtoolbox: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics* **11**, 116 (2010).
74. Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
75. Ray, N., Currat, M., Foll, M. & Excoffier, L. SPLATCHE2: a spatially explicit simulation framework for complex demography, genetic admixture and recombination. *Bioinformatics* **26**, 2993–2994 (2010).
76. Pudlo, P. et al. Reliable ABC model choice via random forests. *Bioinformatics* **32**, 859–866 (2016).
77. McInnes, L., Healy, J. & Melville, J. UMAP: uniform manifold approximation and projection for dimension reduction. Preprint at <https://arxiv.org/abs/1802.03426> (2018).

## Acknowledgements

We thank John Willis and current and former members of his lab, including David Lowry and Kevin Wright, for providing access to North American seed material collected over many years, and to the Botanical Society of Britain and Ireland for their continued support locating UK *Mimulus* populations. Arielle Cooley kindly provided seed material from Chilean populations of *M. luteus* var. *variegatus*. We are very grateful to Claudia Buser and John Bailey for providing the New Zealand material, and Nils Bunnefeld, Anna Maria Fosaa and Simun Arge for their help while collecting *Mimulus* in the Faroe Islands. We thank Oregon Genomics (University of Oregon) for sequencing services, the University of Stirling Controlled Environment Facility for access to plant growth facilities, and Sophie Webster for help in the laboratory. Computer time for the ABC analysis was provided by the computing facilities MCIA (Mésocentre de Calcul Intensif Aquitain) of the Université de Bordeaux and of the Université de Pau et des Pays de l'Adour. L.Y. was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme [grant number ERC-StG 679056 HOTSPOT], via a grant to L.Y. This work was supported by NSF 1754080 to J.R.P. This project was made possible by a grant from the Global Exploration Fund, Northern Europe from National Geographic (GEFNE164-15) to M.V.M., J.R.P. and S.M.I.-B., and a grant from the Natural Environment Research Council (NERC; NE/J012645/1) to M.V.M. We thank all the people who helped us during fieldwork in Alaska, particularly Suzi Golodoff (Unalaska Island) and Stacy Studebaker (Kodiak Island) for providing their exceptional knowledge of the local flora, and Roger Topp (U. Alaska, Fairbanks/Museum of the North), who documented the expedition with his outstanding photographs and video.

## Author contributions

M.V.M., J.R.P., S.M.I.B., J.F. and A.D.T. designed the research. M.V.M., J.R.P., J.F., S.M.I.B., M.C.R. and M.v.K. collected material. M.V.M., A.D.T., and O.L. analysed the data. M.V.M., J.F., L.Y., A.D.T. and J.R.P. wrote the manuscript with input from all the authors.

**Competing interests**

The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-021-01795-x>.

**Correspondence** and requests for materials should be addressed to M.V.-M.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



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

© The Author(s) 2021