Phylogeography and post-glacial dynamics in the clonal-sexual orchid *Cypripedium calceolus* L.

Phylogeography of *C. calceolus*

Roberta Gargiulo[1], Samuel Pironon[1], Ekaterina Zheleznaya[2,3], Michele D. Sanchez[1], Zoltan R. Balázs[4], Dorina Podar[4], Timothy Wilkinson[1], Anne Jäkäläniemi[5], Tiiu Kull[6], Henry Väre[7], Michael F. Fay[1,8]

[1]*Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK*
[2]*Peoples' Friendship University of Russia, Department of System Ecology, Podolskoye shosse 8/5, Moscow, 115093, Russia*
[3]*Timiryazev State Biological Museum, Malaya Gruzinskaya 15, Moscow, 123242, Russia*
[4]*Babeş-Bolyai University, Department of Molecular Biology and Biotechnology, 1 Kogălniceanu St, Cluj-Napoca, 400084, Romania*
[5]*Ecology and Genetics, University of Oulu, PO Box 4200, Oulu, 90014, Finland*
[6]*Estonian University of Life Sciences, Tartu, 51014, Estonia*
[7]*Finnish Museum of Natural History, Botanical Museum, University of Helsinki, PO Box 7, Helsinki, 00014, Finland*
[8]*School of Biological Sciences, University of Western Australia, Crawley, WA 6009, Australia*

Author for correspondence:
Roberta Gargiulo
Email: *r.gargiulo@kew.org*

Abstract

Aim: We investigated the phylogeographic history of a clonal-sexual orchid, to test the hypothesis that current patterns of genetic diversity and differentiation retain the traces of climatic fluctuations and of the species reproductive system.

Location: Europe, Siberia and Russian Far East.

Taxon: *Cypripedium calceolus* L. (Orchidaceae).

Methods: Samples (>900, from 56 locations) were genotyped at eleven nuclear microsatellite loci and plastid sequences were obtained for a subset of them. Analysis of genetic structure and approximate Bayesian computations were performed. Species distribution modelling was used to explore the effects of past climatic fluctuations on the species range.

Results: Analysis of genetic diversity reveals high heterozygosity and allele diversity, with no geographical trend. Three genetic clusters are identified with extant gene pools derived from ancestral demes in glacial refugia. Siberian populations exhibit different plastid haplotypes, supporting an early divergence for the Asian gene pool. Demographic results based on genetic data are compatible with an admixture event explaining differentiation in Estonia and Romania and they are consistent with past climatic dynamics inferred through species distribution modelling. Current population differentiation does not follow an isolation by distance model and is compatible with a model of isolation by colonisation.

Main conclusions: The genetic differentiation observed today in *C. calceolus* preserves the signature of climatic fluctuations in the historical distribution range of the species. Our findings support the central role of clonal reproduction in in reducing loss of diversity through genetic drift. The dynamics of the clonal-sexual reproduction are responsible for the persistence of ancestral variation and stability during glacial periods and post-glacial expansion.

Key words: approximate Bayesian computation, clonal propagation, glacial refugia, isolation by colonisation, population genetics, species distribution modelling, yellow lady's slipper orchid.

Introduction

Evolutionary benefits of clonality include both the reduction of the cost associated with sexual reproduction and, in general, survival in situations in which either mating, recruitment or the combination of both are limited (Barrett, 2015; Lin, Miriti, & Goodell, 2016; Vallejo-Marín, Dorken, & Barrett, 2010). Although clonality confers an advantage when conditions for sexual reproduction are adverse, in the long term, absence of recombination and potential accumulation of deleterious somatic mutations (Meselson effect) might be detrimental for fitness and survival (Barrett, 2015; Birky, 1996; Hojsgaard & Hörandl, 2015; Honnay & Bossuyt, 2005; Klekowski, 2003). Therefore, the combination of clonal and sexual reproduction constitutes a successful evolutionary strategy for overcoming environmental fluctuations (Eckert, 2002; Zhang & Zhang, 2007).

Asexual reproduction in plants essentially occurs with two different mechanisms: apomixis and vegetative propagation (Silvertown, 2008). For the purpose of the present study, we focus on vegetative propagation achieved through the growth of new ramets from different organs (e.g., leaves, roots). A controversial and crucial step in the methodological investigation of clonal plants is the definition of individual, as each ramet may be interpreted as an independent biological unit (Barrett, 2015; Douhovnikoff & Leventhal, 2016), although it is part of a larger clone. From a genetic perspective, genotypes of clonal individuals are not generated by random mating, leading to non-random

associations among loci (linkage disequilibrium) (De Meeûs & Balloux, 2004). Moreover, overlapping generations constitute a further violation of Hardy-Weinberg assumptions, potentially resulting in biased genetic indices (Balloux, Lehmann, & De Meeûs, 2003; De Meeûs, Lehmann, & Balloux, 2006).

Many studies have investigated the consequences of clonal reproduction on genetic diversity, especially on clonal and diploid animals (Halkett, Simon, & Balloux, 2005 and references therein), but also on plant systems (Eckert & Barrett, 1993; Ellstrand & Roose, 1987; Meloni et al., 2013; Stoeckel et al., 2006). The emerging picture is that diversity is context-specific and depends on the combination of environment and life-history traits, but clonal plants are generally as diverse as non-clonal plants. However, patterns of genetic differentiation affected by clonal reproduction remain relatively unexplored and few attempts have been made to relate clonality to the existing models (Oliva et al., 2014; Prati & Schmid, 2000).

Recent studies focused on different organisms have dismissed the Isolation by Distance paradigm (IBD; Wright, 1943), pointing out that population structure seldom depends on the connectivity within and among populations (Sexton, Hangartner & Hoffmann, 2014). The effect on genetic differentiation may emerge under the influence of adaptive variation, as a pattern of Isolation by Environment (IBE; Wang & Summers, 2010) or Isolation by Adaptation (IBA; Nosil, Egan, & Funk, 2008; Orsini, Vanoverbeke, Swillen, Mergeay, & Meester, 2013). Furthermore, intricate patterns of differentiation may depend on the concatenation of events following colonisation and they are usually more difficult to interpret, as predicted by the Isolation by Colonisation model (IBC; De Meester, Vanoverbeke, Kilsdonk, & Urban, 2016; Nadeau, Meirmans, Aitken, Ritland, & Isabel, 2016; Orsini et al., 2013).

With the exception of aquatic plants, where water facilitates vegetative dispersal (Oliva et al. 2014), clonal plants cannot rely on vegetative propagation for long-distance dispersal. Therefore, the expectation for mixed clonal-sexual system is that sexual reproduction acts in the establishment of new colonisers. In the next demographic stages, genetic differentiation depends on the trade-offs between sexual and asexual reproduction and the time elapsed since the colonisation event. In fact, although gene flow can homogenise the genetic structure arisen after founder events, it is subjected to the actual occurrence of sexual reproduction, and this depends on local conditions (Silvertown, 2008).

The perennial, diploid orchid *Cypripedium calceolus* L. is a suitable model species to investigate genetic diversity in clonal-sexual plants. The species is pollinated by insects with a restricted range (Antonelli, Dahlberg, Carlgren, & Appelqvist, 2009); outcrossing is predominant, but self-pollination by geitonogamy is also possible. Clonal propagation occurs via rhizome growth; each vegetative clump has been regarded as an active unit, capable of dormancy, rhizome fragmentation and genet admixture, also through recruitment within the clump (Brzosko, 2002; Davison, Nicolé, Jacquemyn, & Tuljapurkar, 2013; Kull, 1999; Shefferson, Kull, Tali, & Kellett, 2012). In contrast to this dynamism, the species also shows long-term stability, as the maximum longevity of the clumps has been reported as >100 years in Estonian populations (Kull, 1995) and between 110-350 years in Polish populations (Nicolé, Brzosko, & Till-Bottraud, 2005). Plants collected from the wild have survived in cultivation for >100 years (e.g., Fay & Taylor, 2015). Fruit-set and recruitment are generally low (Blinova, 2002; Brzosko, 2002; Kull, 1998; Khapugin, Chugunov, & Vargot, 2017; Zheleznaya; 2015), as the species does not offer any reward to pollinators (Neiland & Wilcock, 1998).

In addition to ecological observations, genetic investigations have shown low differentiation and high diversity. According to Brzosko, Wróblewska, & Ratkiewicz (2002), ancestral polymorphism persists in the species because of the predominance of vegetative reproduction, low recruitment and rhizome stability. Concerning differentiation, Fay et al. (2009), regarded the anemophilous dispersal of the light-weight seeds as a factor promoting multiple post-glacial dispersal events. However, all the evidence is supported by small-scale datasets or a limited number of markers, precluding conclusions on a global scale.

In the present work, we use plastid markers and nuclear microsatellites to verify the hypothesis that the current population genetic structure and differentiation preserve the signature of the mixed reproductive system. We infer evolutionary and demographic history of *C. calceolus* populations sampled across its entire range. By reconstructing the potential distribution of the species in the past

using species distribution modelling, we also interpret demographic and genetic results according to major historical fluctuations in climate.

Materials and Methods

Genetic diversity and differentiation

Populations were sampled from 56 localities (Table S1, Appendix S1 in Supporting Information; Fig. 1); a minimum sampling distance was kept, in order to avoid clones emerging from the same rhizome.

Eleven nuclear microsatellite loci were employed for genotyping, with the protocol described in Appendix S2.1. Microsatellite multilocus genotypes (MLGs) were determined in order to detect the presence of clonal individuals by employing the R package 'poppr'. We also constructed the genotype accumulation curve (Kamvar, Tabima, & Grünwald, 2014), in order to assess the minimum number of loci necessary to distinguish unique genotypes.

Computations of genetic diversity indices, including observed (H_O) and unbiased expected heterozygosity (uH_E) and allelic richness (A_R), are described in Appendix S2.2.

Phylogeographic and IBD patterns were evaluated with permutation tests (20,000 permutations) in SPAGeDi v1.4 (Hardy & Vekemans, 2002). Genetic structure and inbreeding coefficient were estimated by permuting gene copies among individuals and individuals among populations. Spatial locations were permuted among populations and the regression slopes were tested to detect IBD. The effect of mutation on genetic structure was evaluated by permuting allele sizes among allelic states. Permutation tests were conducted both under R-statistics and F-statistics.

Genetic differentiation among populations was determined by using both the Bayesian approach implemented in STRUCTURE v2.3.4 (Pritchard, Stephens, & Donnelly, 2000) and the Thermodynamic Integration (*TI*) implemented in MAVERICK v1.0.3 (Verity & Nichols, 2016), which is considered more accurate in estimating the number of demes (i.e., genetic clusters, *K*). Several exploratory analyses were conducted; both admixture and without-admixture models were evaluated (Appendix S2.3).

Molecular sequences were obtained for a subset of individuals, following the results of the genetic differentiation analysis (see next subsections). Sequences for *C. macranthos* Sw., overlapping with *C. calceolus* in Russia and Asia, and for *C. shanxiense* S.C.Chen, occurring in Asia and overlapping with *C. calceolus* in the Russian Far East (Averyanov, 1999; Andronova et al., 2017), were obtained from GenBank and compared. The nuclear Internal Transcriber Spacer (ITS) and four plastid regions (the genes *rbcL* and *trnL*, the intergenic spacers *trnL-trnF* and *trnH-psbA*) were obtained with the primers and the conditions reported in Appendix S2.1.

Demographic analyses

Different scenarios of colonisation were tested in an Approximate Bayesian Computation (ABC) framework as implemented in DIYABC v2.1.0 (Cornuet et al., 2014). Four gene pools were considered, according to the four main genetic clusters detected in the MAVERICK analysis. To avoid biases in the coalescent simulations induced by structure within the gene pools (Chikhi, Sousa, Luisi, Goossens, & Beaumont, 2010; Peter, Wegmann, & Excoffier, 2010), populations either suspected of cryptic structure (i.e., Kiiminki) or with high levels of allele fixation (i.e., Danish populations) were removed from the dataset. Four evolutionary scenarios were tested, assuming a possible divergence of the four gene pools from an ancestral population (*Na*). All scenarios assume a bottleneck, ideally corresponding to the Last Glacial Maximum (LGM) – when demes were isolated in refugia – followed by population expansions after the LGM.

Priors for simulations for each scenario are reported in Appendix S2.4; the final set of priors is reported in Table S4. A generalised stepwise mutation model was used to infer the best scenario;

posterior probabilities of each scenario were compared both with a direct estimate on the closest data points and a logistic regression on 1% of the simulated data points. Details of the analysis, model checking and estimation of bias and precision are described in Appendix S2.4.

Species Distribution Models

Occurrence records of *C. calceolus* were collected across its full Eurasian range from the GBIF database (https://www.gbif.org/), Averyanov (1999), and our sampled populations. After several data cleaning procedures, the final dataset consisted in 2854 occurrence points. At each of these presence points, we extracted different climatic variables at a 10 arc-min resolution from the version 1.4 of the Worldclim database (http://worldclim.org): annual mean temperature (BIO1), temperature annual range (BIO7), annual precipitation (BIO12), and precipitation seasonality (BIO15).

To model *C. calceolus* distribution, we used four different probabilistic models found in the 'biomod2' package in R (Thuiller, Lafourcade, Engler, & Araújo, 2009): Generalized Additive Model (GAM), Generalized Boosting Model (GBM), RandomForest (RF), and Multiple Adaptive Regression Splines (MARS). We ran the models with five different sets of 5000 pseudo-absence points sampled across the study area. As some parts of the species range were under-sampled, we decided not to select any pseudo-absence point within a two-degrees (~220 km) radius around each presence point (Barbet-Massin, Jiguet, Albert, & Thuiller, 2012; VanDerWal, Shoo, Graham, & Williams, 2009). Each model was then calibrated using 70% of each presence/pseudo-absence dataset and evaluated with the 30% remaining. We repeated this data-splitting strategy ten times and evaluated the model predictive accuracy using two different indices: the True Skill Statistics (TSS, Allouche, Tsoar, & Kadmon, 2006) and the area under the Receiver Operating Characteristic curve (ROC, Hanley & McNeil, 1982). In total, we therefore ran 400 species distribution models (four statistical models, five pseudo-absence selections, two evaluation indices, and ten repetitions).

We projected the distribution of *C. calceolus* under current, Mid-Holocene (~6,000 years ago; according to three Global Circulation Models), LGM (~22,000 years ago; according to three GCMs), and Last Inter-Glacial (LIG; ~120,000 – 140,000 years ago) climatic conditions generated by Braconnot et al., 2007 and Otto-Bliesner, Marshall, Overpeck, Miller, & Hu (2006) and extracted from Worldclim. Then, we computed a mean of the 400 maps obtained previously weighted by their respective evaluation score (i.e. TSS score), in order to obtain one consensus map for each time slice and GCM. More details about the species distribution modelling approach is given in Appendix S3.

Finally, we assessed the relationship between genetic diversity (A_R and H_O) and both current climatic suitability and stability in climatic suitability through-time (i.e., mean climatic suitability across the four time slices, according to the CCSM4 model) using linear regression and the *lm* function in R.

Results

Genetic diversity and differentiation

In total, 959 individuals were successfully genotyped at 11 microsatellite loci; no scoring errors were found in MICRO-CHECKER. Clones were detected as implemented in 'poppr' (see Fig. S3 in Appendix S2) and removed from the dataset, in order to avoid biases in the genetic estimations; 904 individuals were present in the final dataset. The genotype accumulation curve indicated that ten loci were necessary to capture the entire genotypic diversity (Fig. S4 in Appendix S2).

In total, we found 111 alleles across the 11 loci. H_O ranged from 0.04 (Skindbjerg-Denmark) to 0.75 (Simisti-Estonia) (Fig. 1a) and uHe from 0.09 (Skindbjerg-Denmark) to 0.73 (Sovata-Romania); however, in most populations both H_O and uHe were > 0.5 (Table S7 in Appendix S1). Ten private alleles were found in eight populations, mostly Russian and Italian (Table S7 in Appendix S1). The probability test detected deviations from random mating in some populations for some loci, although with no specific population- or locus-trend (except at "Lake Amur", for loci displaying a combination of missing data and homozygous individuals; Table S8 in Appendix S1). Evidence of non-random associations were also found for some loci (Table S9 in Appendix S1); however, they were mostly due to associations not constantly occurring and scattered across populations (data not shown). A_R was highest in Sovata (Romania) for almost all the loci and lowest in Skindbjerg (Denmark) (Table S10 in Appendix S1; Fig. 1b). A mixture of positive and negative F_{IS} for different loci (high variance over loci) was evident for most of the populations, except some populations in Russia, in which high positive values were detected (Table S11 in Appendix S1).

AMOVA analysis indicated that most of the variation is partitioned within individuals (86%) and only a small proportion among populations (13%) and among individuals within populations (0.7%) (Table S12 in Appendix S2). Population pairwise F_{ST} are reported in Table S13 in Appendix S1.

No IBD or phylogeographic signal was detected in SPAGEDI after permuting locations among populations, as regression slopes were not significant; permutation tests of allelic states were not significant (Table S14 in Appendix S1), denoting that either mutation is not predominant over drift or that mutation does not strictly follow a Stepwise Mutation Model (SMM). In this case, F-statistics is preferred over R-statistics (Hardy, Charbonnel, Fréville, & Heuertz, 2003). Significant observed value of F_{ST} was 0.126, corresponding to a moderate level of structure.

Genetic differentiation was better explained by using the without-admixture model (Fig. S1 in Appendix S2); the number of different genetic clusters was $K = 3$ according to STRUCTURE HARVESTER (Table S15 and Fig. S6 in Appendix S2) and $K = 5$ according to MAVERICK (Fig. S5 in Appendix S2). We report population differentiation according to MAVERICK in Figure 2, with K ranging from 3 to 5; Estonian and Romanian populations appear as the most admixed demes and belong to the same genetic cluster (i.e., gene pool A, Figs. 2-3). Gene pool B corresponds to populations in Finland, Western Russia and Siberia, whereas gene pool D corresponds to Italian and central European populations. Gene pool C, which occurs almost exclusively in Siberia and in the Russian Far East, might be identified as a potential descendant of an eastern Asian refugium with low levels of admixture in other populations. Furthermore, no relationship was observed between gene pool C cluster and plastid haplotypes, as individuals belonging to gene pool C exhibited both haplotypes 1 and 2 (Table S6 in Appendix S1). The least represented cluster for $K = 5$ defined two Estonian populations (Sarve and Kõrgessaare), the Finnish population Kiiminki and the Danish populations. This cluster is differentially absorbed by other clusters for smaller K values (Fig. 2).

The ITS alignment for 38 accessions consisted in 808 positions, 16 of which were variable in the accessions of *Cypripedium* analysed. One position was variable in *C. calceolus*, giving three variants occurring both in Europe and Asia (Table S5 in Appendix S1). ITS sequences of *C. calceolus* were distinct from both *C. macranthos* and *C. shanxiense*. The intermediate ITS variant (H3 in Table S5 in Appendix S1, indicated by the ambiguity Y) was also found in Italy and Finland, where other sympatric species do not occur.

Only three of the four plastid regions were successfully amplified. The alignment of the *trn*L intron and *trnL-trnF* intergenic spacer consisted of 983 positions and the *trnH-psbA* intergenic spacer

of 934 positions; two different haplotypes were detected in the plastid sequences (Table S6 in Appendix S1). Haplotype 1 was widespread, whereas haplotype 2 was only found in central Siberia (Slyud_Irkutsk) and in the Russian Far East (Amur).

Demographic analyses

Among the scenarios tested in DIYABC, shown in Fig. 3, the highest posterior probability 0.845 (95% CI: 0.827-0.864, logistic approach on 40,000 data points) was associated with the hypothesis of admixture between gene pools B and D, with consequent origin of gene pool A (Scenario 2). Posterior probabilities for Scenarios 1, 3 and 4 were respectively 0.0299 (95% CI: 0.056-0.034), 0.150 (95% CI: 0.0995-0.131) and 0.0098 (95% CI: 0.008-0.012) (Fig. S7 in Appendix S2). Model checking performed by using test quantities not previously used for model discrimination showed a good placement of the observed dataset in the posterior probability cloud (Fig. S8 in Appendix S2) with no summary statistics laying outside the confidence interval (Table S16 in Appendix S1).

Posterior distribution of parameters and RMAE values are shown in Table 1. If we consider a mean generation time of 100 years for *C. calceolus*, the divergence of the four gene pools, designated by t_2, appears to have started 24,400 BP (CI 95%: 9670-51,400) compatibly with the hypothesis of divergence during the LGM (31,000-16,000 BP). In scenario 2, the median values of t_1 places the admixture event during the Holocene, *c.* 6880 BP (CI 95%: 1880-20,500). RMAEs are usually low, indicating that parameter estimates are reliable. As usual with demographic analyses, confidence intervals are wide and caution is needed in strictly interpreting numerical values. However, broad intervals are mostly attributable to the limited information contained in the dataset rather than inaccurate choice of the priors or model misfit, as confirmed by model checking and RMAE values.

Species distribution modelling

Species distribution models performed well (mean ROC score: 0.99 ± 0.01, mean TSS score: 0.90 ± 0.03). Although over-prediction is suspected in many regions, we retrieved a good representation of the current distribution of *C. calceolus* across Eurasia with a particularly high climatic suitability in central, northern Europe and western Russia (Figs. 1, 2). The Mid-Holocene consensus maps report minor changes in climatic suitability across the distribution of *C. calceolus*, with a loss in the Black Sea region and a gain in central Siberian regions (Fig. 4a, Fig. S9 in Appendix S2). As expected, LGM maps denote a major shift in the distribution of climatically suitable areas with the Mediterranean, Black Sea and some East-Asian areas remaining suitable for *C. calceolus*, but with most northern Europe and Asian regions becoming unsuitable (Fig. 4b, Fig. S9 in Appendix S2). Finally, at the LIG, large and continuous regions are predicted to be suitable across Russia and Siberia, whereas many areas become unsuitable across Europe and the Mediterranean with only part of the Atlantic and mountainous areas remaining suitable (Fig. 4c). Regression analysis did not show any significant correlations between current climatic suitability/stability through time and genetic indices (Figures S10 and S11 in Appendix S2, p-value > 0.5).

Discussion

Clonal reproduction and implications on the genetic diversity and structure

Isolation by distance does not seem to play a role in the divergence of *C. calceolus* and spatial patterns of differentiation are not discrete (Table S14 in Appendix S1). Climatic suitability and stability through time are not correlated to genetic indices (Figs. S10-S11 in Appendix S1). Heterozygosity and allelic richness are uniformly high (Fig. 1a,b), therefore the nuclei of higher diversity normally expected in areas which have been more stable through time (e.g., southern Europe; Carnaval, Hickerson, Haddad, Rodrigues, & Moritz, 2009; Hewitt, 2000) are absent (Fig. 4). Moreover, there is no evidence for either longitudinal/latitudinal gradients or centre-periphery patterns of diversity (Fig. 4; Pironon, Villellas, Morris, Doak, & García, 2015; Pironon et al., 2017). Similarly, no decrease in performance of peripheral populations was observed in *C. calceolus* from the Spanish Pyrenees, based on demographic observations, flowering and recruitment factors (García, Goni, & Guzmán, 2010). We believe that global level of differentiation in *C. calceolus* is much more compatible with an IBC system, which assumes that genetic variation reflects the structure established after colonisation (DeWoody, Trewin, & Taylor, 2015; Orsini et al., 2013). According to the model, the absence of a cline in neutral genetic variation suggests that the first colonists determined population structure. In some cases, especially for long-lived species (Petit et al., 2003), the initial structure persists, whereas in other cases it develops as a consequence of other mechanisms (Nadeau et al., 2016). In *C. calceolus*, the establishment of the genetic structure by the first colonists may be particularly promoted by vegetative propagation. Different authors have proposed that, after colonisation, sexual reproduction is favoured by optimal conditions and no disturbance. As the habitat matures, clonal propagation gradually becomes more important (Piquot et al., 1998; Wilk, Kramer, & Ashley, 2009) and shapes fine-scale genetic structure of populations (Binks, Millar, & Byrne, 2015). From this perspective, genetic variation in *C. calceolus* might be ascribed to the persistence of ancestral polymorphisms and to the modifications induced by the clonal system dynamics (Brzosko et al., 2002; Gargiulo, Ilves, Kaart, Fay, & Kull, 2018). Although gene flow occurs, our results suggest that it is not a strong driver for the constitution of fine-scale genetic structure and it is more effective during the colonisation process and clearly at long distances. This is also supported by the observations of low recruitment from seeds in at least some populations (e.g., Khapugin et al., 2017). In the analysis of genetic differentiation, the without-admixture model is preferred over the admixture model (see Fig. S1 in Appendix S2), suggesting that each gene copy virtually comes from the same deme. Although this is clearly a simplification arising from high infra-population variation, it is in accordance with the IBC pattern and the persistence of ancestral variation.

The signature of clonal reproduction is evident in the high heterozygosity values, in the negative F_{IS} and in the residual gametic disequilibrium. Contrary to departures from Hardy-Weinberg proportions, non-random associations among loci require many generations of random mating to disappear, therefore the equilibrium will never be reached if clonal individuals are constantly produced (De Meeûs et al., 2006). In most of the populations, F_{IS} values display an important variance among loci, with some remarkably negative values. Negative F_{IS} is considered as the ultimate signature of clonal reproduction; however, the occurrence of sexual reproduction and recombination guarantees the presence of positive F_{IS} values (Balloux et al., 2003; Reichel, Masson, Malrieu, Arnaud-Haond & Stoeckel, 2016). When sexual reproduction occurs, alleles are shuffled within and among populations. Compared to a strictly sexual system, however, the random sampling of allelic variants associated with sexual reproduction – genetic drift – has a minor impact on the genetic variation. In the long term, perennial rhizomes accumulate mutations and propagate polymorphisms to clonal offspring. Therefore, we suggest that allele diversity is predominantly due to the effect of mutations, rather than a real balance between mutation and drift. Genetic divergence with low level of drift occurs slowly, as (ancestral) polymorphism is protected by the clonal propagation, with all the consequences examined above. A similar effect on genetic diversity was found in the Canarian endemic *Ruta macrocarpa*, which is able to reproduce both sexually and clonally (Meloni et al., 2013).

Arnaud-Haond et al. (2014) explained the absence of IBD signal in clonal seagrass species as an artefact of the dominant role of mutation-drift in contrast to contemporary migration. We suggest that clonal propagation in *C. calceolus* can determine a similar imbalance among evolutionary forces, although the exact mechanisms need to be further explored. These mechanisms, together with post-glacial colonisation history, are behind the absence of IBD (see next subsection). In contrast with Arnaud-Haond et al. (2014), we believe that hypervariable markers are adequate in resolution, despite the saturation effect on the F_{ST}. Previous research on *C. calceolus* with more conservative markers such as allozymes (Brzosko et al., 2002; Case, 1994) and plastid length polymorphisms (Fay et al., 2009) showed the same lack of phylogeographical patterns, respectively on a regional and on a continental scale. Nevertheless, we are aware of the possible biases related to homoplasy, that we have tried to counterbalance by selecting microsatellite loci with a large amount of variation (Estoup, Jarne, & Cornuet, 2002).

Demographic history of *Cypripedium calceolus*

The scenario emerging from the ABC framework and the SDM results are compatible with the occurrence of bottlenecks in the history of *C. calceolus*, followed by gene pool divergence in glacial refugia and post-glacial expansion. The "Asian" gene pool C detected for populations in Siberia and in the Russian Far East (Fig. 2) was assumed as the most divergent (i.e., showing the lowest proportion of admixed individuals). An early divergence for Asian populations was corroborated by both nuclear data and the slowly mutating plastid DNA haplotypes, but also by the SDM results. In fact, climate was predicted to be highly suitable for *C. calceolus* across a large area in Russia-Siberia during the LIG (Fig. 4c, LIG), potentially guaranteeing continuity among the ancestral gene pools. In the following LGM, highly fragmented population(s) may have survived in some microrefugia across eastern Asia, where suitable macroclimatic conditions appear as significantly more reduced (Fig. 4b, LGM). Consequently, the relatively deeper divergence of the Asian gene pool might be due to stronger isolation. The core of most suitable areas in this period was southern and western Europe (Fig. 4b, LGM), corresponding to the gene pool D identified in the analysis of genetic structure (Fig. 3). In addition, two further gene pools were detected (A and B), supposedly related to different ancestral refugia in our demographic hypotheses. The scenario receiving the highest posterior probability after the ABC computations turned out to be the one assuming admixture between gene pools B and D, corresponding to southern-European and Finnish/Russian populations (Scenario 2); the median estimation for the beginning of the glacial bottleneck (i.e., time t_2 in Table 1) is 24,400 ya, broadly compatible with the LGM (~22,000 ya) and part of the last Ice Age (~110,000-11,000 ya). Moreover, t_2 (Table 1) formally corresponds to the period when the hypothetically panmictic population Na ceased to exist. This is consistent with the shift detected from LIG to LGM in the SDM predictions. Although we cannot exclude the occurrence of extant populations from further refugia related to the gene pool A (e.g., in the Balkans), its independent divergence from an ancestral deme, depicted both in Scenario 3 and 4, received less confidence after the ABC computations. Consequently, populations in Estonia and Romania appear to derive from the intersection of post-glacial colonisation routes from gene pools similar to the extant populations found in Italy and Finland/western Russia, as also supported by the range shift from the LGM to the Mid-Holocene climatic conditions (Fig. 4; Petit et al., 2003). According to the "southern refugia paradigm" (Hewitt, 2000), Italian peninsula is traditionally identified as an ancestral refugium. However, we cannot determine the exact location of the ancestral refugium from which gene pool B (Finland/western Russia) derives. The occurrence of cryptic refugia at central European latitudes has been reported in several studies on tree species (Tzedakis, Emerson, & Hewitt, 2013), especially cold and drought-tolerant, but also mesophilous species (Kuneš et al., 2008; Magri et al., 2006). Populations of *C. calceolus* occurring in Finland, western Russia and in some Siberian areas appear as potential derivatives of such a higher latitude refugium. The plant associations in which the species occurs nowadays, riparian mixed woodlands or boreal forests (e.g., Kull, 1999), might support this hypothesis. In terms of genetic structure, the existence of intermediate refugia (e.g., in central

Europe) is invoked as a potential explanation for the lack of phylogeographic structure (Mandák et al., 2016).

The difficulty in accurately predicting glacial refugia also depends on the limited availability of fossil pollen records. Entomophily in orchids, in fact, determines little pollen production, and pollen structure hampers an accurate distinction from pollen of other monocots (Pacini, 2009). From our results, a further difficulty is the absence of trends in heterozygosity and allelic richness, which are uniformly high. In fact, refugia are expected to have higher levels of genetic diversity (Comes & Kadereit, 1998; Pironon et al., 2017). In the same way, the SDMs identified large suitable areas, but the coarse resolution of the models may prevent from identifying micro-refugia. Moreover, the predictive accuracy of SDMs may be affected by non-analogous climatic conditions between the current and past time (Veloz et al., 2012) or by the potential of the species for niche shift (Pearman et al, 2008). However, the absence of *C. calceolus* in many of the predicted areas may also indicate that climatic factors are essential, though not sufficient requirements for the species establishment and survival. Probably the most important factors are the occurrence of the mycorrhizal symbionts necessary to seed germination, growth and, possibly, to adult survival (Fay, Feustel, Newlands, & Gebauer, 2018; Shefferson et al., 2007) and the occurrence of limestone soils (Kull, 1999). Persistence of populations has also been strongly influenced by humans, where entire populations have disappeared or become critically endangered after uprooting, collection or simply after indirect interventions of forest management (Fay & Taylor, 2015). Moreover, it has been predicted that changes in successional processes over centuries (i.e., canopy closure) are connected to stochasticity in the population dynamics of *C. calceolus* (Davison et al., 2013; Shefferson et al., 2012) and in general in orchid populations (Jacquemyn, Brys, Hermy, & Willems, 2007). In Polish and Estonian populations, light intensity and canopy shading play an important role in flowering and recruitment (Kull, 1999). Consequently, population survival is difficult to predict as it heavily relies on transient micro-conditions.

In conclusion, the present study reveals new perspectives on the demographic and biogeographic history of clonal orchids. Current genetic variation in *C. calceolus* can be interpreted in different spatio-temporal dimensions in which clonal propagation played a crucial role. Polymorphism is frozen as a result of minimal genetic drift, long generations and temporal stability. Occasional migration and gene flow keep population structure mostly dependent on the historical patterns of colonisation. Species distribution modelling reveals high climatic suitability in central Europe through time and a pattern of post-glacial colonisation compatible with genetic data, from refugial areas in southern Europe, eastern Asia and, possibly, central Europe. Future analysis of genome-scale data from *C. calceolus* may help clarify historical nuclei of diversity and corroborate the influence of clonal propagation to the observed genetic differentiation.

Data accessibility

Microsatellite scoring is available through DRYAD data repository (doi:10.5061/dryad.4674nn4). Internal transcribed spacers and plastid sequences have been submitted to Genbank and final accession numbers are indicated in Table S5 and Table S6 in Appendix S1.

Biosketch

Roberta Gargiulo is interested in population genetics and genomics. This work is part of the project focused on the conservation genetics of the lady's slipper orchid at the Royal Botanic Gardens, Kew.
Author contributions: R.G. and M.F.F. conceived the study. E.Z., D.P., Z.R.B., A.J., T.K. and H.V. collected field samples and provided ecological data. R.G. and M.D.S. conducted laboratory work. D.P. and Z.R.B. extracted DNA from Romanian samples. R.G. analysed molecular data. S.P. performed species distribution modelling. T.W. collected occurrence data, produced maps and graphical artworks. R.G. wrote the manuscript with significant input from S.P. and M.F.F. All authors reviewed the manuscript.

Table 1 Demographic approximate Bayesian computations for *C. calceolus*, as obtained in the most likely scenario (Scenario 2). See also Appendix S2, Table S4.

Parameter	Explanation	Median	q050	q950	RMAE
N1	N_e of Deme N1	3490	1590	4790	0.189 (0.007)
N2	N_e of Deme N2	3500	1910	4810	0.168 (0.099)
N3	N_e of Deme N3	1560	537	4260	0.144 (0.096)
N4	N_e of Deme N4	3720	2100	4870	0.153 (0.102)
t1	time to the admixture event	68.8	18.8	205	0.253 (0.214)
ra	admixture rate	0.522	0.172	0.853	0.154 (0.006)
t2	time to the divergence from the ancestral deme Na	244	96.7	514	0.220 (0.243)
dbx	time interval related to the change in population size	49.5	9.29	201	0.109 (0.006)
N2b	size of the bottlenecked deme N2b	693	266	968	0.183 (0.051)
N3b	size of the bottlenecked deme N3b	405	103	888	0.186 (0.057)
N4b	size of the bottlenecked deme N4b	599	191	954	0.190 (0.061)
Na	Ne of the ancestral deme N1	2221	899	7710	0.188 (0.015)
μ	Mean mutation rate	5.09×10^{-4}	1.92×10^{-4}	9.21×10^{-4}	0.150 (0.007)
P	Proportion of multiple step mutation in the generalised stepwise model	0.24	0.140	0.3	0.108 (0.004)
μ$_{SNI}$	Mean single nucleotide insertion/deletion rate	3.72×10^{-6}	3.60×10^{-7}	9.00×10^{-6}	0.229 (0.007)

N_e is the effective population size. q050-q950 indicate the 95% credibility values. RMAE: relative median of the absolute error. See supporting information for explanations about the parameters
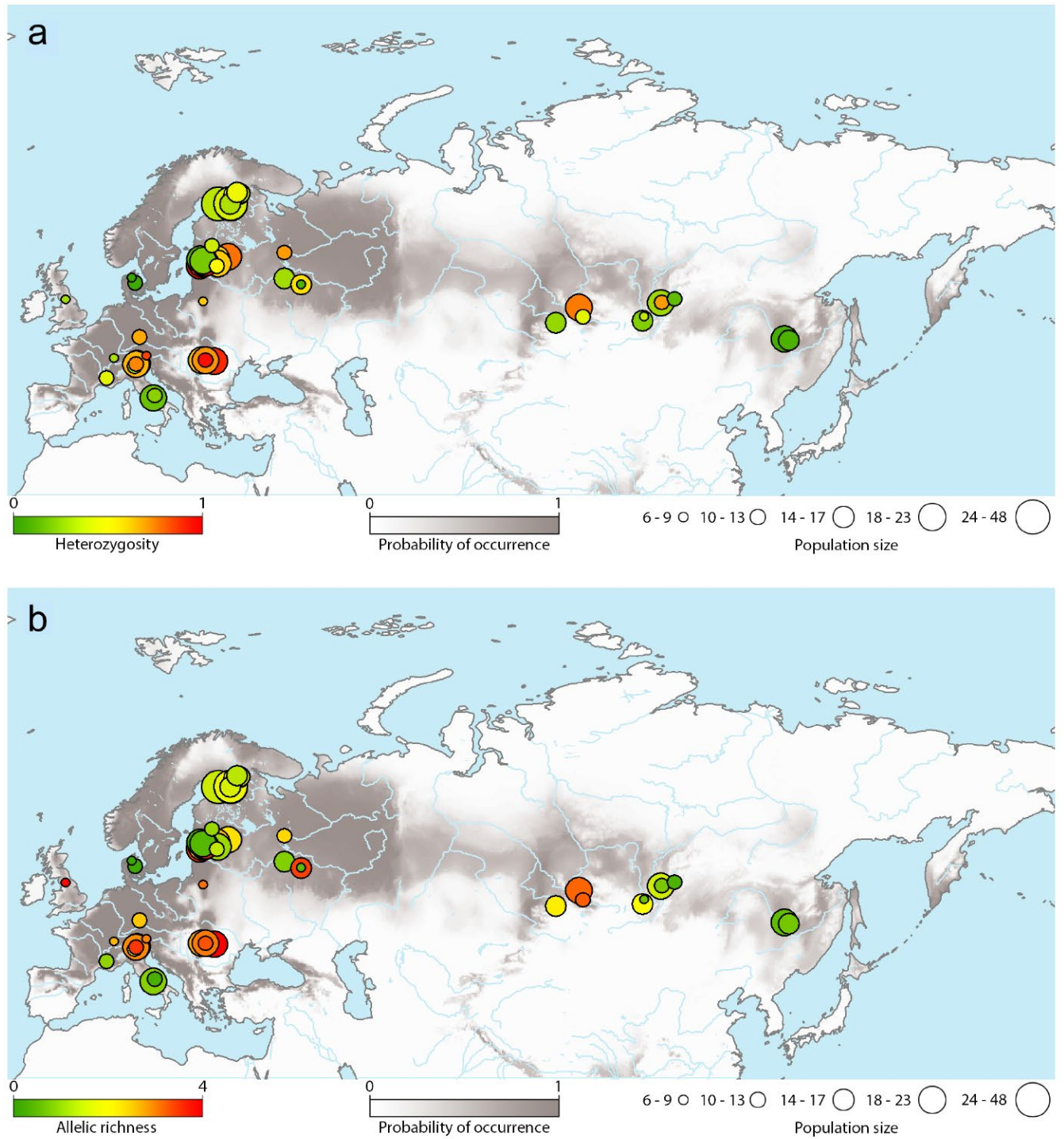
Fig. 1 Map of the current climatic suitability for the populations of *C. calceolus*. Pie charts are proportional to sample sizes after clone exclusion. Spatial patterns of genetic diversity in terms of (a) observed heterozygosity, H_O, and (b) allelic richness, A_R, are represented by colour gradients.
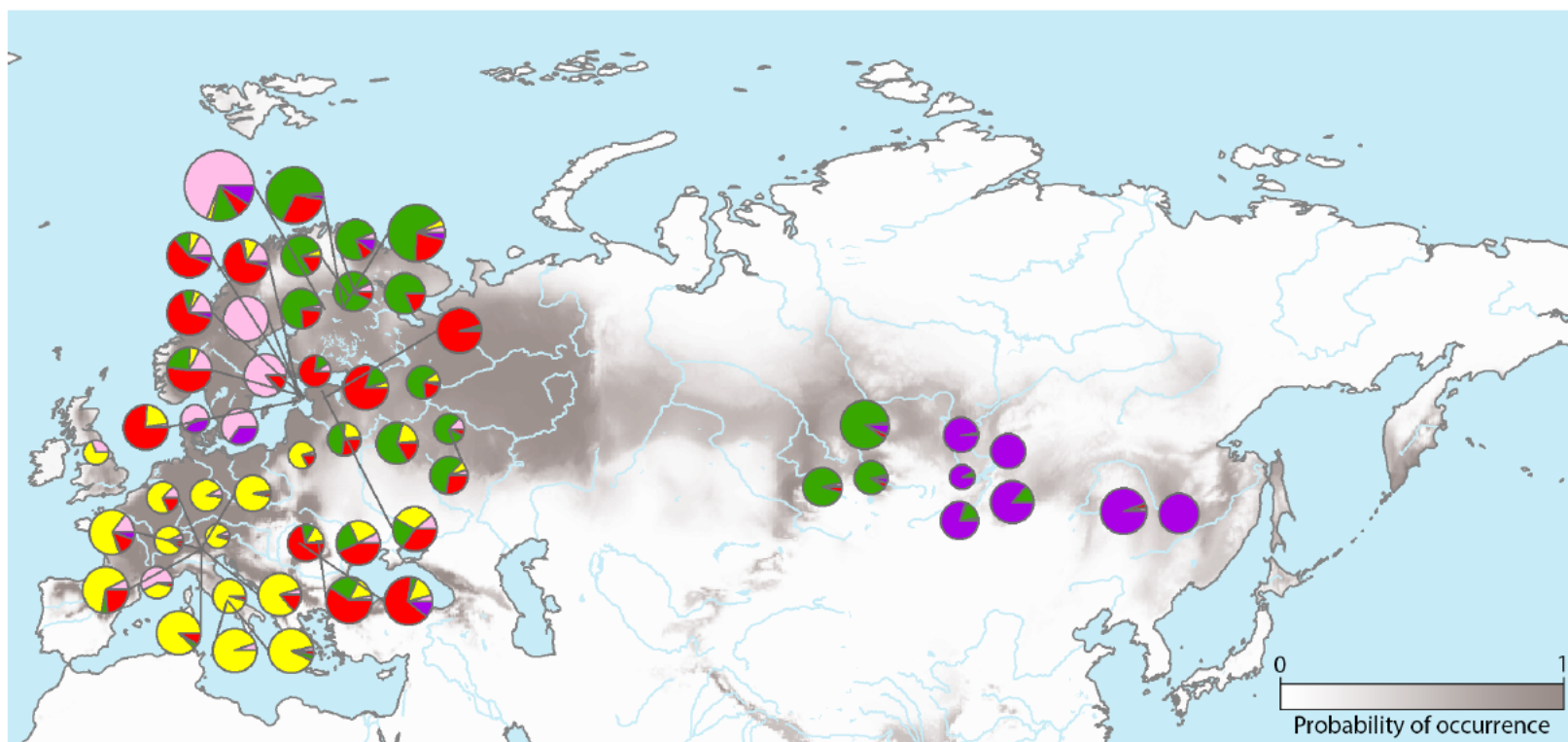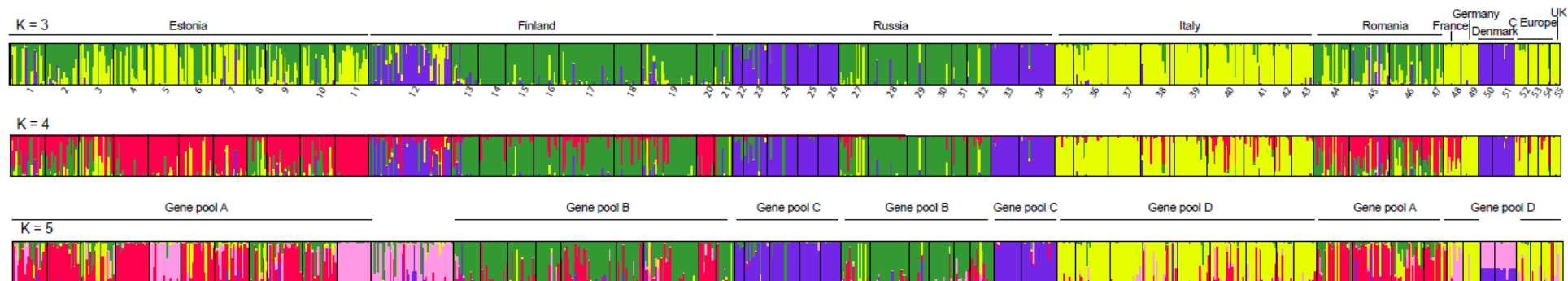
Fig. 2 Genetic structure of *C. calceolus* as implemented in MAVERICK. Pie charts in the map represent the genetic structure for the most likely $K = 5$, with circles proportional to sample size (see Table S1 for population codes). Bar plots for $K = 3$ and $K = 4$ are shown for comparison.
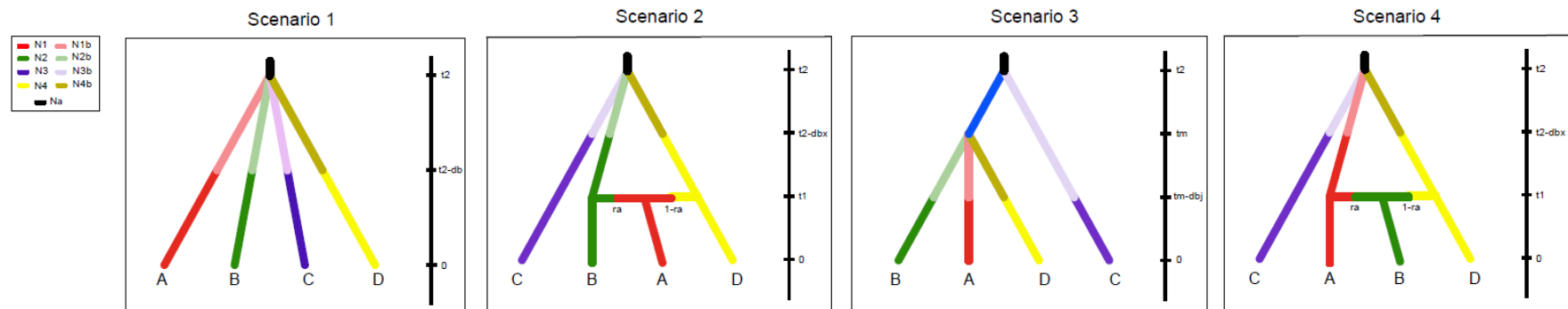
Fig. 3. Hypothetical coalescent scenarios tested in DIYABC. Letters correspond to the gene pools detected in the analysis of genetic structure (Fig. 2). Scenario 1 assumes independent divergence of the four gene pools from ancestral demes during post-glacial expansion. Scenario 2 assumes the gene pool A as deriving from an admixture event of parental demes B and D. In Scenario 3, an independent divergence is considered rather than an admixture event, in which gene pools A, B and D evolved independently from a progenitor deme. Scenario 4 assumes that gene pool B derives from an admixture event of parental demes A and D. Scenario 2 is the most plausible one, according to the DIYABC analysis (see details in Appendix S2).
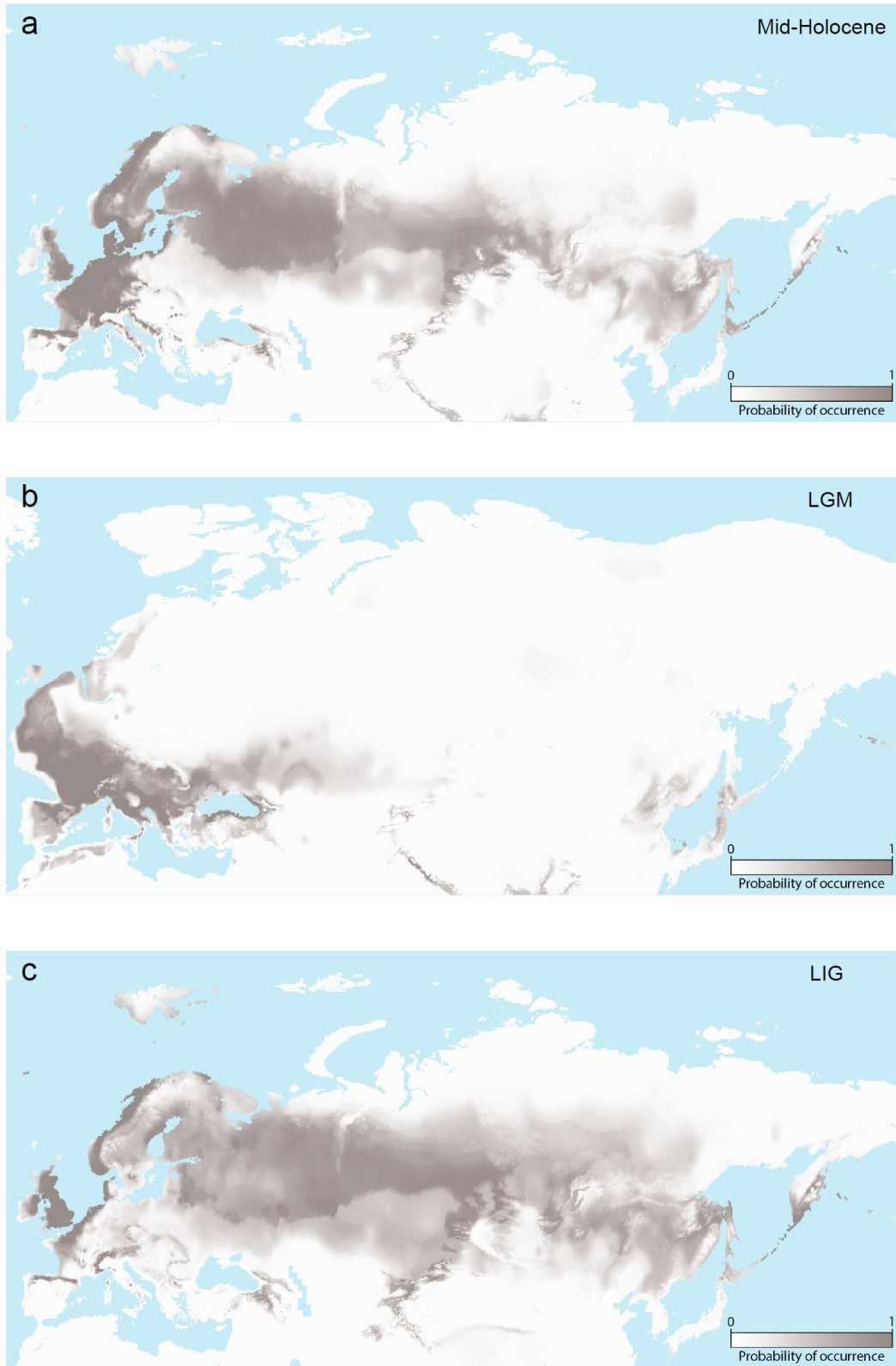
Fig. 4 Maps showing predicted suitability through time according to the species distribution modelling (CCSM4 model). In (a) Mid-Holocene (~6,000 years ago) suitability, (b) Last Glacial Maximum (LGM; ~22,000 years ago) suitability and (c) Last Inter-Glacial (LIG; ~120,000 – 140,000 years ago) suitability.

References

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology, 43*, 1223–1232.

Andronova, E. V., Machs, E. M., Filippov, E. G., Raiko, M. P., Lee, Y. I., & Averyanov, L. V. (2017). Phylogeography of the genus *Cypripedium* (Orchidaceae) taxa in Russia. *Botanical Journal (Ботанический журнал), 102*, 1027–1059.

Antonelli, A., Dahlberg, C.J., Carlgren, K.H., & Appelqvist, T. (2009). Pollination of the Lady's slipper orchid (*Cypripedium calceolus*) in Scandinavia–taxonomic and conservational aspects. *Nordic Journal of Botany*, *27*, 266–273.

Arnaud-Haond, S., Moalic, Y., Hernández-García, E., Eguiluz, V.M., Alberto, F. … Duarte, C.M. (2014). Disentangling the influence of mutation and migration in clonal seagrasses using the genetic diversity spectrum for microsatellites. *Journal of Heredity*, *105*, 532–541.

Averyanov, L.V. (1999). Genus *Cypripedium* (Orchidaceae) in the Russia. *Turczaninowia, 2,* 5–40.

Balloux, F., Lehmann, L., & De Meeûs, T. (2003). The population genetics of clonal and partially clonal diploids. *Genetics*, *164*, 1635–1644.

Barbet-Massin, M., Jiguet, F., Albert, C.H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, *3*, 327–338.

Barrett, S.C. (2015). Influences of clonality on plant sexual reproduction. *Proceedings of the National Academy of Sciences*, *112*, 8859–8866.

Binks, R.M., Millar, M.A., & Byrne, M. (2015). Contrasting patterns of clonality and fine-scale genetic structure in two rare sedges with differing geographic distributions. *Heredity*, *115*, 235.

Birky, C.W. (1996). Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics*, *144*, 427–437.

Blinova, I. (2002). A Northernmost Population of *Cypripedium calceolos* L.(Orchidaceae): Demography, Flowering, and Pollination. *Selbyana*, *23*, 111–120.

Braconnot, P., Otto-Bliesner, B., Harrison, S., Joussaume, S., Peterchmitt, J.Y., Abe-Ouchi, A. ... Kageyama, M. (2007). Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum–Part 1: experiments and large-scale features. *Climate of the Past*, *3*, 261–277.

Brzosko, E. (2002). Dynamics of island populations of *Cypripedium calceolus* in the Biebrza river valley (north-east Poland). *Botanical Journal of the Linnean Society*, *139*, 67–77.

Brzosko, E., Wróblewska, A., & Ratkiewicz, M. (2002). Spatial genetic structure and clonal diversity of island populations of lady's slipper (*Cypripedium calceolus*) from the Biebrza National Park (northeast Poland). *Molecular Ecology*, *11*, 2499–2509.

Case, M.A. (1994). Extensive variation in the levels of genetic diversity and degree of relatedness among five species of *Cypripedium* (Orchidaceae). *American Journal of Botany*, *81*, 175–184.

Carnaval, A.C., Hickerson, M. J., Haddad, C. F., Rodrigues, M. T., & Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, *323*, 785–789.

Chikhi L., Sousa V.C., Luisi, P., Goossens, B., & Beaumont, M.A. (2010). The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. *Genetics, 186,* 983–995.

Comes, H.P., & Kadereit, J.W. (1998). The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science*, *3*, 432–438.

Cornuet, J.M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R. ... Estoup, A. (2014). DIYABC v2. 0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, *30*, 1187–1189.

Davison, R., Nicolé, F., Jacquemyn, H., & Tuljapurkar, S. (2013). Contributions of covariance: decomposing the components of stochastic population growth in *Cypripedium calceolus*. *The American Naturalist*, *181*, 410–420.

De Meester, L., Vanoverbeke, J., Kilsdonk, L.J., & Urban, M.C. (2016). Evolving perspectives on monopolization and priority effects. *Trends in Ecology & Evolution*, *31*, 136–146.

De Meeûs, T., & Balloux, F. (2004). Clonal reproduction and linkage disequilibrium in diploids: a simulation study. *Infection, Genetics and Evolution*, *4*, 345–351.

De Meeûs, T., Lehmann, L., & Balloux, F. (2006). Molecular epidemiology of clonal diploids: a quick overview and a short DIY (do it yourself) notice. *Infection, Genetics and Evolution*, *6*, 163–170.

DeWoody, J., Trewin, H., & Taylor, G. (2015). Genetic and morphological differentiation in *Populus nigra* L.: isolation by colonization or isolation by adaptation? *Molecular Ecology*, *24*, 2641–2655.

Douhovnikoff, V., & Leventhal, M. (2016). The use of Hardy–Weinberg Equilibrium in clonal plant systems. *Ecology and Evolution*, *6*, 1173–1180.

Eckert, C.G. (2002). The loss of sex in clonal plants. In *Ecology and Evolutionary Biology of Clonal Plants* (pp. 279–298). Springer, Dordrecht.

Eckert, C.G., & Barrett, S.C. (1993). Clonal reproduction and patterns of genotypic diversity in *Decodon verticillatus* (Lythraceae). *American Journal of Botany*, *80*, 1175–1182.

Ellstrand, N.C., & Roose, M.L. (1987). Patterns of genotypic diversity in clonal plant species. *American Journal of Botany*, *74*, 123–131.

Estoup, A., Jarne, P., & Cornuet, J.M. (2002). Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*, *11*, 1591–1604.

Fay, M.F., Bone, R., Cook, P., Kahandawala, I., Greensmith, J., Harris, S. ... Lexer, C. (2009). Genetic diversity in *Cypripedium calceolus* (Orchidaceae) with a focus on north-western Europe, as revealed by plastid DNA length polymorphisms. *Annals of Botany*, *104*, 517–525.

Fay, M.F., Feustel, M., Newlands, C., & Gebauer, G. (2018). Inferring the mycorrhizal status of introduced plants of *Cypripedium calceolus* (Orchidaceae) in northern England using stable isotope analysis. *Botanical Journal of the Linnean Society*, *186*, 587–590.

Fay, M.F., & Taylor, I. (2015). 801. *Cypripedium calceolus*. *Curtis's botanical magazine*, *32*, 24–32.

García, M.B., Goni, D., & Guzmán, D. (2010). Living at the Edge: Local versus Positional Factors in the Long-Term Population Dynamics of an Endangered Orchid. *Conservation Biology*, *24*(5), 1219–1229.

Gargiulo, R., Ilves, A., Kaart, T., Fay, M.F., & Kull, T. (2018). High genetic diversity in a threatened clonal species, *Cypripedium calceolus* (Orchidaceae), enables long-term stability of the species in different biogeographical regions in Estonia. *Botanical Journal of the Linnean Society*, *186*, 560–571.

Halkett, F., Simon, J.C., & Balloux, F. (2005). Tackling the population genetics of clonal and partially clonal organisms. *Trends in Ecology & Evolution*, *20*, 194–201.

Hanley, J.A., & McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*, 29–36.

Hardy, O.J., Charbonnel, N., Fréville, H., & Heuertz, M. (2003). Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics*, *163*, 1467–1482.

Hardy, O.J., & Vekemans, X. (2002). SPAGeDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Resources*, *2*, 618–620.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*, 907–913.

Hojsgaard, D., & Hörandl, E. (2015). A little bit of sex matters for genome evolution in asexual plants. *Frontiers in Plant Science*, *6*, 82.

Honnay, O., & Bossuyt, B. (2005). Prolonged clonal growth: escape route or route to extinction? *Oikos, 108*, 427-432.

Jacquemyn, H., Brys, R., Hermy, M., & Willems, J.H. (2007). Long-term dynamics and population viability in one of the last populations of the endangered *Spiranthes spiralis* (Orchidaceae) in the Netherlands. *Biological Conservation*, *134*, 14–21.

Kamvar, Z.N., Tabima, J.F., & Grünwald, N.J. (2014). PoPPR: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281.

Khapugin, A.A., Chugunov, G.G., & Vargot, E.V. (2017). *Cypripedium calceolus* (Orchidaceae) in Central Russia: a case study for its populations in two protected areas in the Republic of Mordovia (Russia). *Lankesteriana*, *17*(3).

Klekowski, E.J. (2003). Plant clonality, mutation, diplontic selection and mutational meltdown. *Biological Journal of the Linnean Society*, *79*, 61–67.

Kull, T. (1995). Genet and ramet dynamics of *Cypripedium calceolus* in different habitats. *Abstracta Botanica*, *19*, 95–104.

Kull, T. (1998). Fruit-set and recruitment in populations of *Cypripedium calceolus* L. in Estonia. *Botanical Journal of the Linnean Society*, *126*, 27–38.

Kull, T. (1999). Biological Flora of the British Isles. No. 208 *Cypripedium calceolus* L. *Journal of Ecology, 87,* 913–924.

Kuneš, P., Pelánková, B., Chytrý, M., Jankovská, V., Pokorný, P., & Petr, L. (2008). Interpretation of the last-glacial vegetation of eastern-central Europe using modern analogues from southern Siberia. *Journal of Biogeography*, *35*, 2223–2236.

Lin, C.H., Miriti, M.N., & Goodell, K. (2016). Demographic consequences of greater clonal than sexual reproduction in *Dicentra canadensis*. *Ecology and evolution*, *6*, 3871–3883.

Magri, D., Vendramin, G.G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D. ... Tantau, I. (2006). A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist*, *171*, 199–221.

Mandák, B., Havrdová, A., Krak, K., Hadincová, V., Vít, P., Zákravský, P., & Douda, J. (2016). Recent similarity in distribution ranges does not mean a similar postglacial history: a phylogeographical study of the boreal tree species *Alnus incana* based on microsatellite and chloroplast DNA variation. *New Phytologist*, *210*, 1395-1407.

Meloni, M., Reid, A., Caujapé-Castells, J., Marrero, A., Fernández-Palacios, J.M. … Conti, E. (2013). Effects of clonality on the genetic variability of rare, insular species: the case of *Ruta microcarpa* from the Canary Islands. *Ecology and Evolution*, *3*, 1569–1579.

Nadeau, S., Meirmans, P.G., Aitken, S.N., Ritland, K., & Isabel, N. (2016). The challenge of separating signatures of local adaptation from those of isolation by distance and colonization history: The case of two white pines. *Ecology and Evolution*, *6*, 8649–8664.

Neiland, M.R.M., & Wilcock, C.C. (1998). Fruit set, nectar reward, and rarity in the Orchidaceae. *American Journal of Botany*, *85*, 1657–1671.

Nicolé, F., Brzosko, E., & Till-Bottraud, I. (2005). Population viability analysis of *Cypripedium calceolus* in a protected area: longevity, stability and persistence. *Journal of Ecology*, *93*, 716–726.

Nosil, P., Egan, S.P., & Funk, D.J. (2008). Heterogeneous genomic differentiation between walking-stick ecotypes: "isolation by adaptation" and multiple roles for divergent selection. *Evolution*, *62*, 316–336.

Oliva, S., Romero, J., Pérez, M., Manent, P., Mascaró, O., Serrão, E. A. ... Alberto, F. (2014). Reproductive strategies and isolation-by-demography in a marine clonal plant along an eutrophication gradient. *Molecular Ecology*, *23*, 5698–5711.

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J., & Meester, L. (2013). Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, *22*, 5983–5999.

Otto-Bliesner, B.L., Marshall, S.J., Overpeck, J.T., Miller, G.H., & Hu, A. (2006). Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *Science*, *311*, 1751–1753.

Pacini, E, (2009). Orchid pollen dispersal units and reproductive consequences. In: Kull, T., Arditti, J., Wong, S.M., (Eds.) *Orchid biology: reviews and perspectives, 10* (pp. 185–218)*.* Berlin, Germany: Springer.

Pearman, P.B., Guisan, A., Broennimann, O., & Randin, C.F. (2008). Niche dynamics in space and time. *Trends in Ecology & Evolution*, *23*, 149–158.

Peter, B.M., Wegmann, D., & Excoffier, L. (2010). Distinguishing between population bottleneck and population subdivision by a Bayesian model choice procedure. *Molecular ecology, 19(21),* 4648-4660.Petit, R.J., Aguinagalde, I., de Beaulieu, J.L., Bittkau, C., Brewer, S., Cheddadi, R. ... Mohanty, A. (2003). Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, *300*, 1563–1565.

Piquot, Y., Petit, D., Valero, M., Cuguen, J., de Laguerie, P., & Vernet, P. (1998). Variation in sexual and asexual reproduction among young and old populations of the perennial macrophyte *Sparganium erectum*. *Oikos*, *82*, 139–148.

Pironon, S., Papuga, G., Villellas, J., Angert, A.L., García, M.B., & Thompson, J.D. (2017). Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. *Biological Reviews*, *92*, 1877–1909.

Pironon, S., Villellas, J., Morris, W.F., Doak, D.F., & García, M.B. (2015). Do geographic, climatic or historical ranges differentiate the performance of central versus peripheral populations? *Global Ecology and Biogeography*, *24*, 611–620.

Prati, D., & Schmid, B. (2000). Genetic differentiation of life-history traits within populations of the clonal plant *Ranunculus reptans*. *Oikos*, *90*, 442–456.

Pritchard, J.K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Reichel, K., Masson, J.P., Malrieu, F., Arnaud-Haond, S., & Stoeckel, S. (2016). Rare sex or out of reach equilibrium? The dynamics of F_{IS} in partially clonal organisms. *BMC genetics*, *17*, 76.

Sexton, J.P., Hangartner, S.B., & Hoffmann, A.A. (2014). Genetic isolation by environment or distance: which pattern of gene flow is most common? *Evolution*, *68*(1), 1–15.

Shefferson, R.P., Kull, T., Tali, K., & Kellett, K.M. (2012). Linking vegetative dormancy to fitness in two long-lived herbaceous perennials. *Ecosphere*, *3*(2), 1–19.

Shefferson, R.P., Taylor, D.L., Weiß, M., Garnica, S., McCormick, M.K., Adams, S. ... Yukawa, T. (2007). The evolutionary history of mycorrhizal specificity among lady's slipper orchids. *Evolution*, *61*, 1380–1390.

Silvertown, J. (2008). The evolutionary maintenance of sexual reproduction: evidence from the ecological distribution of asexual reproduction in clonal plants. *International Journal of Plant Sciences*, *169*, 157–168.

Stoeckel, S., Grange, J., Fernández-Manjarres, J.F., Bilger, I., Frascaria-Lacoste, N., & Mariette, S. (2006). Heterozygote excess in a self-incompatible and partially clonal forest tree species—*Prunus avium* L. *Molecular Ecology*, *15*, 2109–2118.

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M.B. (2009). BIOMOD–a platform for ensemble forecasting of species distributions. *Ecography*, *32*, 369–373.

Tzedakis, P.C., Emerson, B.C., & Hewitt, G.M. (2013). Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution*, *28*, 696–704.

Vallejo-Marín, M., Dorken, M.E., & Barrett, S.C. (2010). The ecological and evolutionary consequences of clonality for plant mating. *Annual Review of Ecology, Evolution, and Systematics*, *41,* 193–213.

VanDerWal, J., Shoo, L.P., Graham, C., & Williams, S.E. (2009). Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecological Modelling*, *220*, 589–594.

Veloz, S.D., Williams, J.W., Blois, J.L., He, F., Otto-Bliesner, B., & Liu, Z. (2012). No-analog climates and shifting realized niches during the late quaternary: implications for 21st-century predictions by species distribution models. *Global Change Biology*, *18*, 1698–1713.

Verity, R., & Nichols, R.A. (2016). Estimating the number of subpopulations (K) in structured populations. *Genetics*, *203*(4), 1827–1839.

Wang, I.J., & Summers, K. (2010). Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular Ecology*, *19*, 447–458.

Wilk, J.A., Kramer, A.T., & Ashley, M.V. (2009). High variation in clonal vs. sexual reproduction in populations of the wild strawberry, *Fragaria virginiana* (Rosaceae). *Annals of Botany*, *104*, 1413–1419.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*, 114–138.

Zhang, Y., & Zhang, D. (2007). Asexual and sexual reproductive strategies in clonal plants. *Frontiers of Biology in China*, *2*, 256–262.

Zheleznaya, E. (2015). Results of a study of *Cypripedium* in several regions of Siberia (Russia). *European Journal of Environmental Sciences*, *5*(2).