ORIGINAL ARTICLE

Phylogeographic pattern of range expansion provides evidence for cryptic species lineages in *Silene nutans* in Western Europe

H Martin[1], P Touzet[1], F Van Rossum[2,3,4], D Delalande[1] and J-F Arnaud[1]

As a result of recent or past evolutionary processes, a single species might consist of distinct Evolutionary Significant Units (ESUs), even corresponding to cryptic species. Determining the underlying mechanisms of range shifts and the processes at work in the build-up of divergent ESUs requires elucidating the factors that contribute to population genetic divergence across a species' range. We investigated the large-scale patterns of genetic structure in the perennial herbaceous plant species *Silene nutans* (Caryophyllaceae) in Western Europe. We sampled and genotyped 111 populations using 13 nuclear microsatellite loci and 6 plastid single-nucleotide polymorphisms. Broad-scale spatial population genetic structure was examined using Bayesian clustering, spatial multivariate analyses and measures of hierarchical genetic differentiation. The genotypic structure of *S. nutans* was typical of a predominantly allogamous mating system. We also identified plastid lineages with no intra-population polymorphism, mirroring two genetically differentiated nuclear lineages. No evidence of admixture was found. Spatial trends in genetic diversity further suggested independent leading-edge expansion associated with founding events and subsequent genetic erosion. Overall, our findings suggested speciation processes in *S. nutans* and highlighted striking patterns of distinct stepwise recolonisation of Western Europe shaped by Quaternary climate oscillations. Two main potential ESUs can be defined in Western Europe, corresponding to Eastern and Western nuclear-plastid lineages. *In situ* preservation of populations and genetic rescue implying *ex situ* conservation techniques should take the lineage identity into account. This is particularly true in Great Britain, northern France and Belgium, where *S. nutans* is rare and where distinct lineages co-occur in close contact.
Heredity (2016) **116,** 286–294; doi:10.1038/hdy.2015.100; published online 9 December 2015

INTRODUCTION

Most plant species in Europe have been extensively inventoried, described and documented. However, some of them show high intraspecific genetic, phenotypic and/or ecological variation leading to among-population differentiation that might ultimately result in speciation (Coyne and Orr, 2004; Gavin *et al.*, 2014). We may wonder whether one single species might consist of several Evolutionary Significant Units (ESUs; Fraser and Bernatchez, 2001) evolving in relative isolation from each other, or even corresponding to cryptic species, as a result of recent or past evolutionary processes. Identifying these ESUs has implications in terms of sustainable plant diversity conservation. Depending on their distribution (wide or restricted), abundance (common or rare), habitat requirements (more or less specialised) and patterns of genetic diversity and structure, the conservation status of ESUs may differ, and *in situ* and *ex situ* conservation strategies should be evaluated for each ESU separately (Palsbøll *et al.*, 2007). Besides, plant communities are facing strong and fast environmental changes owing to human activities (for example, climate change, habitat fragmentation), threatening their survival in the long term (Hautekèete *et al.*, 2015). Recent studies of the effects of contemporary climate warming on plant species distribution and

migration have emphasised the need for a better understanding of species response to past climatic changes (Petit *et al.*, 2008).

The study of population genetic structure and phylogeographic patterns across a species range, using a combination of plastid and nuclear molecular markers, allows retracing past and present species spread and identifying past and recent evolutionary processes responsible for the current patterns of genetic variation (Avise, 2000; Gavin *et al.*, 2014). Past climate events, especially the series of major ice ages that occurred during the Quaternary (2.4 million years ago (Mya) to the present), have driven the evolution of divergent genetic lineages by isolating populations in distant refugia, for instance, in the Iberian peninsula, Italy or the Balkans in Europe (Hewitt, 2000). Ultimately, these historical refugia may have promoted allopatric genetic divergence leading to different ESUs or to sibling species that may not be morphologically distinguishable, that is, cryptic species (Hewitt, 2000; Gavin *et al.*, 2014). Subsequently, postglacial migration and species range expansion from these different refugia have allowed divergent lineages to come into secondary contact in suture zones (for example, Taberlet *et al.*, 1998; Hewitt, 2011). Contemporary processes, such as gene flow, genetic drift and spatially heterogeneous selective pressures, in combination with species life-history traits, in particular the

[1]Unité Évolution, Écologie, Paléontologie, UMR CNRS 8198, Université de Lille 1—Sciences et Technologies, Villeneuve d'Ascq, France; [2]Meise Botanic Garden (formerly National Botanic Garden of Belgium), Meise, Belgium; [3]Écologie Végétale et Biogéochimie, Université Libre de Bruxelles, Brussels, Belgium and [4]Fédération Wallonie–Bruxelles, Brussels, Belgium
Correspondence: Professor J-F Arnaud, Unité Évolution, Écologie, Paléontologie, UMR CNRS 8198, Université de Lille 1, Bat. SN2, Villeneuve d'Ascq 59655, France.
E-mail: jean-francois.arnaud@univ-lille1.fr
Received 24 March 2015; revised 5 October 2015; accepted 6 October 2015; published online 9 December 2015

breeding system and seed and pollen dispersal capabilities, can also contribute to shape species' distribution, genetic diversity and population differentiation (Hamrick and Godt, 1996; Duminil *et al.*, 2007).

In this paper, we examine the genetic structure of the perennial herb *Silene nutans* in Western Europe. In Great Britain, northern France and Belgium, this species exhibits a disjunct distribution, is locally rare and vulnerable and populations show allozyme and/or morphological divergence (for example, Hepper, 1956; Fitter, 1978; Van Rossum *et al.*, 1997, 2003). In Great Britain, two varieties, *S. nutans* var. *salmoniana* Hepper and var. *smithiana* Moss, have been differentiated based on reproductive traits (Hepper, 1951), but this distinction is no longer considered in recent taxonomy (for example, Stace, 2010). In Belgium, two parapatric edaphic ecotypes (calcicolous *Ca* and silicicolous *Si*) with contrasting morphological traits and substantial genetic divergence have been described. Isolating reproductive barriers have been identified between them, suggesting incipient speciation processes (De Bilde, 1973; Van Rossum *et al.*, 1996, 1997). Our hypothesis is that populations of *S. nutans* in Western Europe may correspond to several ESUs or to cryptic biological species.

Using plastid single-nucleotide polymorphisms (SNPs) and nuclear microsatellite markers, we investigated the large-scale phylogeographic and population genetic structure patterns on a comprehensive sampling of populations of *S. nutans* from Western Europe. Using F-statistics, Bayesian clustering and spatial multivariate methods, we addressed the following questions:

1. What are the range-wide levels of genetic diversity and genetic differentiation in *S. nutans*? Can the patterns be related to historical gene flow, mating system and/or seed and pollen dispersal abilities?
2. If spatial genetic analyses reveal genetically differentiated lineages in *S. nutans* in Western Europe, what is the most plausible evolutionary process involved in the formation of these distinct lineages? We hypothesised that range shifts driven by the last Quaternary glaciations and leading-edge expansions associated with founding events may be the main processes explaining the large-scale patterns of genetic variability we observed.
3. Is there evidence for suture zones among genetically divergent lineages, and if so, can we detect hybridisation events mediated by seed and/or pollen flow in the contact zones?

Our results will be discussed in the light of biogeographical scenarios, climate refugia and past historical migration events since the Last Glacial Maxima, leading to divergent ESUs and possible cryptic speciation.

METHODS

Studied species and plant material

S. nutans L. (Caryophyllaceae) is a diploid, herbaceous, long-lived perennial rosette-forming plant species occurring in dry habitats, on rock outcrops, sand or shingle, such as xerothermophilous grasslands, open forests and forest edges. Its wide continental distribution range extends from western Europe to central Siberia and South Caucasus (Hepper, 1956; Fitter, 1978). Flowers are protandrous and insect-pollinated, mainly by moths (Jürgens *et al.*, 1996). *S. nutans* is self-compatible and expected to exhibit a mixed-mating system. Inbreeding depression in selfed progeny has been reported (Dufaÿ *et al.*, 2010).

Leaf material was sampled from a total of 1979 individuals from 111 populations located at the western border of *S. nutans* geographic range (Supplementary Table S1, Figure 1a). The sampling covered six countries, including 16 populations from the United Kingdom (populations 1–11 and 16 for *S. nutans* var. *smithiana*, and populations 12–15 for var. *salmoniana*; population assignment to varieties was based on fruit measurements;

Van Rossum, unpublished data), 50 populations originating from France, 10 populations from Luxemburg, 13 populations from Belgium (for which 5 have been identified as belonging to the *Ca* ecotype labelled 77, 78, 80, 87 and 88, and 5 to the *Si* ecotype labelled 79, 81, 82, 86 and 89; population assignment to ecotypes was based on De Bilde, 1973 and Van Rossum *et al.*, 1997), three populations from The Netherlands and 19 populations from Germany. Leaf samples were dried and conserved in silica gel until DNA extraction. Sample sizes ranged from 2 to 133, with a mean of 17.8 individuals per population (±14.8).

Molecular analyses

All sampled individuals were genotyped using both plastid and nuclear DNA markers. DNA was extracted from 15 to 20 mg of leaf tissue using Macherey-Nagel (Düren, Germany) NucleoSpin 96 Plant II Kits following the standard protocol outlined in the manufacturer's handbook.

Multilocus nuclear DNA genotypes were characterised at 13 recently isolated microsatellite loci named B09, E08, G01, H07, D10, Sil16, Sil19, Sil24, Sil31, Sil35, Sil36, Sil37 and Sil42 and described in Godé *et al.* (2014). Amplification procedures, multiplexing and genotyping were carried out following the standard protocols described in Godé *et al.* (2014).

Plastid diversity was investigated using SNPs recently developed from plastid sequences (pDNA) of *S. nutans* representative samples (Lahiani *et al.*, 2013). Plastid SNPs were defined based on six polymorphic pDNA nucleotides in the intergenic spacer sequences *psbA-trnH*, named Cp 42, and the *matK* gene fragment, named Cp 397, Cp 540, Cp 656, Cp 730 and Cp 804. The Kompetitive Allele Specific PCR (KASP) genotyping assays was used to detect the SNPs. This method is based on competitive allele-specific PCR and enables bi-allelic scoring of SNPs. The SNP-specific KASP assay mix (designed by LGC group, www.lgcgroup.com) and the universal KASP Master mix are added to DNA samples, and a thermal cycling reaction is then performed, followed by an end-point fluorescent read. Bi-allelic discrimination is achieved through the competitive binding of two allele-specific forward primers, each with a unique tail sequence that corresponds with two universal FRET (fluorescence resonant energy transfer) cassettes; one labelled with FAM dye and the other with HEX dye. The sequences flanking the SNPs that were used to define the primers are given in Supplementary Table S2.

Amplification reactions were carried out by PCR with 30 ng of DNA in 8.11 µl reaction volume containing 4 µl of universal KASP Master mix (2 ×) and 0.11 µl of the SNP-specific KASP Assay mix (composed of the two forward primers and the reverse primer). Cycling conditions for PCR amplification were 15 min at 94 °C, 10 cycles of 20 s at 94 °C and 60 s at 65 °C with a step-down of 0.8 °C per cycle, 26 cycles of 20 s at 94 °C and 60 s at 57 °C and finally 12 °C for 10 min. Fluorescence was detected using a LightCycler 480 (Roche Diagnostics, Basel, Switzerland). Individual haplotypes were defined as a combination of allelic states for all six SNPs.

Nuclear genetic variation within populations

Prior to further analyses, genotypic linkage disequilibrium among all pairwise locus combinations was checked with a log-likelihood ratio test, implemented in FSTAT version 2.9.3.2. (Goudet, 1995), using 10 000 randomisations of two-locus genotypes and Bonferroni correction. The following measures of genetic variation were calculated within and over all populations: allelic richness (A_r, El Mousadik and Petit, 1996) based on a minimum sample size of eight individuals (six populations with sample size below this threshold were not included, see Supplementary Table S1), observed heterozygosity (H_O), and expected heterozygosity (H_E) corrected for sample size. Some populations failed to be genotyped for the B09 locus (populations 59, 102 and 103) and the D10 locus (population 21); allelic richness was thus only estimated for the 11 remaining loci. Using GENEPOP version 4.3 (Rousset, 2008), deviations from Hardy–Weinberg equilibrium were also tested by estimating the intrapopulation fixation index (F_{IS}) following Weir and Cockerham (1984). Significance of F_{IS} estimates was tested using exact probability tests across loci and populations. Markov chain method provided unbiased estimates of the Fisher's exact test probability using the following parameters: 100 000 dememorisations, 1000 batches, and 100000 iterations per batch. To obtain corrected F_{IS} estimates (F_{IS_corr}) accounting for null allele occurrence, we used the Bayesian individual inbreeding model described in Chybicki and Burczyk (2009)
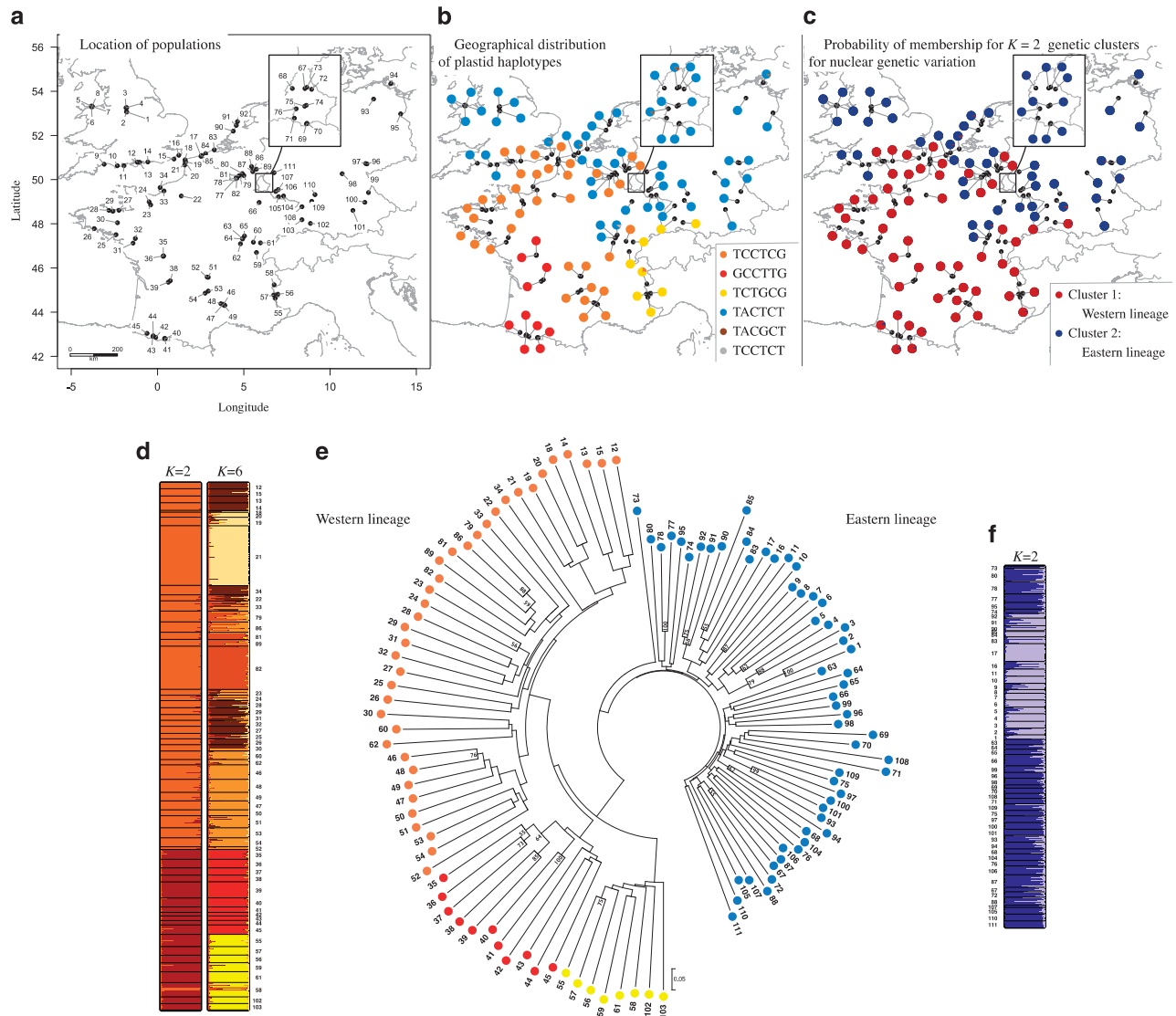
Figure 1 Map of the studied populations of *Silene nutans* in Western Europe. (**a**) Geographical location and population numbers. (**b**) Geographical distribution and within-population proportion of plastid haplotypes based on six SNPs. (**c**) Map of mean population membership probabilities of belonging to the two modal clusters based on nuclear microsatellite loci. (**d**) Assignment probabilities of individual membership into $K=2$ and $K=6$ clusters for Western lineage. Each individual is represented by a thin line partitioned into K coloured segments displaying the individual's estimated membership fractions in K clusters. (**e**) Neighbour-joining tree based on Cavalli-Sforza and Edwards' distance (1967), using nuclear data; only bootstraps >50% are indicated. Coloured circles next to population codes referred to plastid haplotype found in populations. (**f**) Assignment probabilities of individual membership into $K=2$ clusters for Eastern lineage. Each individual is represented by a thin line partitioned into two coloured segments displaying the individual's estimated membership fractions in two clusters. The inset box in panels (**a**), (**b**) and (**c**) zoom on a densely sampled area in greater detail.

and implemented in the software INEST version 2.0 available at http://genetyka. ukw.edu.pl. Using SPAGeDi V1.5 (Hardy and Vekemans, 2002), indirect and independent estimates of selfing rates (*s*) were derived from the multilocus correlation structure of population samples, following a multilocus estimate of the standardised identity disequilibrium and described in David *et al.* (2007). To examine geographic trends in the distribution of large-scale genetic diversity, we tested the relationship between allelic richness, longitude and latitude using a linear multiple regression model. Spatially interpolated values of allelic richness were then generated for a better visualisation using the thin plate spline method implemented in R version 3.0.2 (R Core Team, 2014) and represented on a geographical heat map.

Lineage distinctiveness and spatial patterns of genetic structure
Three complementary approaches were used on nuclear microsatellite data to investigate the genetic affinities among populations. First, genetic distance-

based population clustering was performed to infer the genetic relationships among *S. nutans* populations using an unrooted neighbour-joining tree. It was constructed using Cavalli-Sforza and Edwards (1967) genetic chord distance (D_{CE}) based on allele frequencies. Bootstrapped confidence values on obtained branches were determined using random resampling replications over loci (1000 replicates) implemented in the POPULATIONS software 1.2.32.

Second, a non-spatially explicit Bayesian clustering was carried out to infer the number of distinct gene pools in the whole data set, using the model-based Bayesian algorithm implemented in the STRUCTURE software (Pritchard *et al.*, 2000). The number of potential K clusters was assessed from 10 different runs of K ranging from 1 to 80. To ascertain adequate convergence of the Markov Chain Monte Carlo (MCMC) model, we allowed a burn-in of 100 000 iterations, followed by 2×10^6 MCMC replications without any prior geographic information on the putative affiliation of individuals. To identify the most likely number of K clusters, the *ad hoc* statistic ΔK was calculated as

described in Evanno *et al.* (2005). CLUMPP version 1.1.2. (Jakobsson and Rosenberg, 2007) accounted for label switching and was subsequently used to find the optimal alignment of independent runs by averaging the top runs. The resulting most likely grouping of populations was plotted using DISTRUCT version 1.1. (Rosenberg, 2004). This method was applied (i) to the whole data set to identify potential lineages, (ii) to each identified lineage separately to investigate more subtle lineage-specific genetic structure, and (iii) to populations from areas where different haplotypes (potential distinct lineages) co-occurred (potential contact zones), to detect potential admixture events. This kind of model-based Bayesian algorithm infers subtle population structure by minimising linkage disequilibrium and departures from Hardy–Weinberg equilibrium within each inferred cluster (Pritchard *et al.*, 2000). However, the assumption of panmixia does not necessarily hold in *S. nutans*, which can exhibit a mixed-mating system in natural populations (Dufaÿ *et al.*, 2010).

Third, to avoid this potential bias, we performed a multivariate ordination analysis that does not require any genetic assumptions: a spatial principal component analysis (sPCA, reviewed in Jombart *et al.*, 2009). We calculated independent synthetic variables that maximise the product of the genetic variance among populations and their spatial autocorrelation (based on Moran's *I*). The computation of was Moran's *I* carried out based on a Delaunay graph. Genetically distinguishable groups, clines in allele frequency and intermediate states can lead to what is called 'global structure' and are identified by positive components, that is, showing positive spatial autocorrelations (for example, DiLeo *et al.*, 2010). sPCA and tests for statistical significance of global structure were performed using the R ADEGENET package (Jombart, 2008). The main results of sPCA can be depicted on maps of population scores, allowing a visual assessment of spatial genetic structuring (Jombart *et al.*, 2008). Using the same strategy as that used for Bayesian clustering, we constructed maps for (i) the whole data set to investigate potential genetically distinct lineages, and (ii) each inferred lineage separately to identify specific evolutionary and phylogeographic processes.

Genetic differentiation among nuclear–plastid lineages
Between-population genetic differentiation was investigated for nuclear data (i) over all populations for each locus and at the multilocus level, (ii) within the identified nuclear–plastid Western and Eastern lineages and (iii) between populations (over all loci) using the ANOVA procedure by Weir and Cockerham (1984). Pairwise F_{ST} values and their significance (*G* test) were calculated using GENEPOP. The partitioning of nuclear genetic diversity within and among genetically distinct lineages inferred from the spatial genetic structure analyses was estimated following the procedure proposed by Yang (1998) and implemented in the R package HIERFSTAT (Goudet, 2005). The among-plastid haplotype component is given by $F_{H/T}$, and its significance was tested by permuting individuals of each population across plastid haplotypes. The between-nuclear–plastid Western-Eastern lineage component is given by $F_{L/T}$ and its significance was tested by permuting individuals of each haplotype across nuclear–plastid lineages. Genetic differentiation among each pairwise comparison of plastid haplotypes was also performed. Variance in microsatellite allele size due to stepwise changes of repeat number can also account for the level of genetic differentiation and may give insights on the relative contribution of mutation versus migration rates in genetic divergence. Using SPAGeDi, we thus also calculated the microsatellite allele size-based estimates of genetic differentiation R_{ST} (Slatkin, 1995) and tested them for significance by allele-size permutations following Hardy *et al.* (2003). Finally, we estimated the proportion of specific alleles of each identified lineage for nuclear microsatellite loci.

Spatial genetic structure within nuclear–plastid lineages
To detect more subtle cryptic genetic structure, we reanalysed the nuclear data set by carrying out Bayesian clustering and sPCA separately for the Eastern and Western lineages. We also tested for isolation by distance by performing a spatial autocorrelation analysis on (i) the whole plastid data set and (ii) for the nuclear data within the two identified nuclear–plastid lineages. For plastid data, we used classic one-dimensional spatial correlograms based on Moran's *I*. Random permutations of population locations were carried out to test the statistical significance of each Moran's *I*. For nuclear data, one-dimensional Mantel correlograms were designed according to Oden and Sokal (1986). The normalised Mantel statistic *rz* (Smouse *et al.*, 1986) was used as described in Oden and Sokal (1986) to estimate the relationships between pairwise genetic distances (D_{CE} distance) and pairwise geographical distances among populations. All calculations were carried out using PASSAGE version 2 (Rosenberg and Anderson, 2011).

RESULTS
Nuclear and plastid genetic variation within populations
Exact tests for linkage disequilibrium between microsatellite loci only showed 38 significant *P*-values out of 7690 comparisons (0.49%, 384 expected from type 1 error at $\alpha = 0.05$). When analysed over all populations, 11 out of the 13 loci showed significant deviation from Hardy–Weinberg expectations, with mean single-locus F_{IS} values ranging from -0.007 to 0.284 (Table 1). The mean multilocus F_{IS} value was 0.146. At the population level, F_{IS} values ranged from -0.230 to 0.513, with 89 out of the 111 populations (80.1%) showing significantly positive multilocus F_{IS} values (Supplementary Table S1).

Table 1 Estimates of nuclear genetic variation for 13 microsatellite loci over 111 populations of *Silene nutans* from Western Europe: total number of alleles (A_n), the allelic richness (A_r), the observed heterozygosity (H_O), expected heterozygosity (H_E) and the mean fixation indices (F_{IS}, F_{IT}, F_{ST}, following Weir and Cockerham, 1984)

Locus	Size range	A_n	A_r	H_O	H_E	F_{IS}	F_{IT}	F_{ST}	% E-private	% W-private	% common
B09	162–198	29	6.1	0.487	0.585	0.140***	0.387***	0.287***	0	65.5	34.5
D10	175–328	38	6.3	0.450	0.594	0.246***	0.461***	0.285***	10.5	42.1	47.4
E08	214–370	72	6.3	0.437	0.512	0.150***	0.463***	0.369***	11.1	65.3	23.6
G01	233–357	79	8.7	0.779	0.854	0.099***	0.227***	0.141***	6.3	17.7	75.9
H07	158–313	46	6.4	0.551	0.685	0.206***	0.379***	0.218***	6.5	52.3	41.3
Sil16	114–204	29	6.5	0.524	0.666	0.227***	0.415***	0.244***	13.8	37.9	48.3
Sil31	129–234	49	5.3	0.447	0.547	0.208***	0.436***	0.287***	22.4	44.9	32.6
Sil35	84–225	49	7.8	0.701	0.757	0.094***	0.267***	0.191***	28.6	22.4	49.0
Sil37	111–252	47	6.9	0.709	0.725	0.028**	0.226***	0.204***	4.2	51.1	44.7
Sil19	111–165	26	6.0	0.577	0.580	0.012	0.365***	0.357***	23.1	26.9	50.0
Sil24	128–247	86	8.3	0.768	0.767	−0.007	0.181***	0.188***	8.1	54.6	37.2
Sil36	70–158	33	6.0	0.449	0.612	0.284***	0.494***	0.293***	42.4	21.2	36.4
Sil42	97–202	40	7.5	0.550	0.752	0.278***	0.406***	0.177***	35.0	7.5	57.5
Multilocus	—	47.9	6.8	0.571	0.664	0.146***	0.357***	0.247***	14.8	40.4	43.2

The proportion of diagnostic alleles of Eastern lineage (% E-private), Western lineage (% W-private) and common alleles (% common) in both nuclear-plastid lineages are also given. **$P < 0.01$, ***$P < 0.001$.

Corrected within-population multilocus F_{IS} estimates (F_{IS_corr}) accounting for null allele occurrence always yielded lower values with an arithmetic mean of 0.049. Using the classical relationship ($s = 2F_{IS}/(1+F_{IS})$) expected at genetic equilibrium for a mixed-mating system, estimated mean selfing rates were of 0.254 for F_{IS} and of 0.093 for F_{IS_corr}. Estimations of selfing rates based on standardised identity disequilibrium ranged from 0 to 0.379, with an arithmetic mean of 0.060 (Supplementary Table S1). It should be kept in mind that selfing rate estimations should be interpreted with caution. Indeed, populations with small sample sizes lead to large variance in s estimates and an upward bias is expected when propensity for selfing is low.

Nuclear microsatellite loci displayed high levels of polymorphism, with the number of alleles ranging from 26 (Sil19) to 86 (Sil24), for an average number of 47.9 alleles over all populations (Table 1). Mean H_O and H_E ranged from 0.437 to 0.779 and from 0.512 to 0.854, respectively (Table 1). The map interpolating levels of allelic richness (Figure 2) and the multiple regression analysis (adjusted $R^2 = 0.422$, $F_{(2,102)} = 38.9$, $P < 0.001$, Variance Inflation Factor = 1.003; Pearson's correlation coefficient between latitude and longitude $r = 0.054$) indicated a significant decrease in allelic richness along a southern–northern ($\beta = -0.150$, $P < 0.001$) and an eastern–western gradient ($\beta = 0.073$, $P < 0.001$).

For plastid haplotype genetic diversity, we found a total of six different haplotypes, of which four corresponded to 99.5% of the overall plastid diversity observed (Figure 1b, Supplementary Table S1): the TACTCT SNP combination (hereafter called the blue haplotype), TCCTCG SNP combination (orange haplotype), GCCTTG SNP combination (red haplotype), and TCTGCG SNP combination (yellow haplotype). All populations were monomorphic, except seven populations that showed intra-population plastid polymorphism, with the second haplotype occurring at very low frequency (Supplementary Table S1). The Belgian *Ca* ecotype and the British var. *smithiana* showed the blue haplotype, whereas the *Si* ecotype and the British var. *salmoniana* showed the orange haplotype (Supplementary Figures S1b and c, Supplementary Table S1).
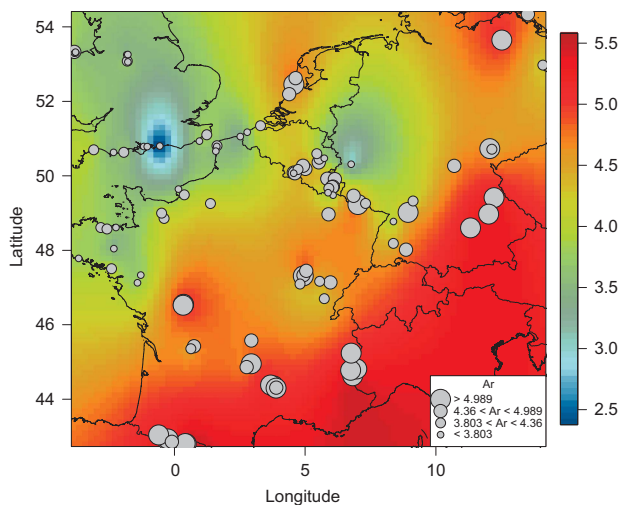
Lineage distinctiveness and spatial patterns of genetic structure

The four main plastid haplotypes showed geographic structuring (Figure 1b, Supplementary Figure S2) and several geographical areas in which two haplotypes co-occurred (Figure 1b, Supplementary Figures S1b and d). Given the lack of intra-population polymorphism and the strong spatial structure observed, these plastid haplotypes were used as a diagnostic criterion for population lineage identity. Based on nuclear microsatellite data, and taking plastid identity into account, a clear phylogeographic pattern and genetically differentiated nuclear–plastid lineages emerged from (i) the neighbour-joining tree based on allele frequencies using D_{CE} (Figure 1e); (ii) the non-spatially explicit Bayesian clustering method (Figures 1c–f); and (iii) the sPCA (Figure 3). Populations clearly clustered according to their plastid haplotype identity, with a Western lineage containing the red, orange and yellow plastid haplotypes, and an Eastern lineage characterised by the blue plastid haplotype. The highest likelihood found using the ΔK statistic corroborated the occurrence of the two main nuclear–plastid lineages (Supplementary Figure S3). The sPCA performed on the whole nuclear data set revealed significant global structure ($P < 0.001$). The first two sPCA axes accounted for most of the spatial genetic structure, covering a high proportion of the spatial autocorrelation and most of the variance in the genetic data ($I = 0.631$, var = 0.851 and $I = 0.610$, var = 0.223, respectively). The first two eigenvalues were 0.537 and 0.136, while the remaining eigenvalues were < 0.090. The first axis in the sPCA revealed a striking longitudinal cline, differentiating a Western and an Eastern nuclear–plastid lineage (Figure 3).

Genetic differentiation among nuclear–plastid lineages

The overall and pairwise genetic differentiation for nuclear data were high and highly significant, respectively, with single-locus F_{ST} estimates ranging from 0.177 to 0.369, and a mean multilocus F_{ST} value of 0.247 (all significant at $P < 0.001$, see Table 1). Pairwise F_{ST} estimates between the 111 studied populations ranged from 0 to 0.608, with 52.8% of significant estimates (Supplementary Figure S4).

The overall genetic distinctiveness among plastid haplotypes was supported by a significant hierarchical nuclear genetic differentiation
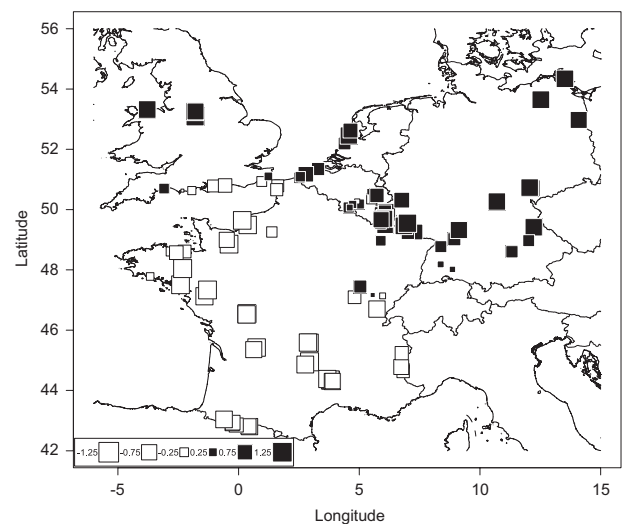

Figure 2 Map of the observed level of allelic richness (A_r) found in each population of *Silene nutans* (displayed as proportional circles). Spatially interpolated levels of allelic richness can also be visualised as colour plot gradient (from dark blue for low values to dark red for high values).

Figure 3 Geographic map of the first global scores of an sPCA performed on the whole data set in *Silene nutans* ($I = 0.631$, var = 0.851, $P < 0.001$). Each square represents a population. Populations that are more closely related in multivariate space share the same colour of squares. Sizes of squares indicate the magnitude of spatial positive autocorrelation.

($F_{H/T} = 0.143$, $P < 0.001$). The highest levels of nuclear genetic differentiation among haplotypes were found between blue and orange haplotypes ($F_{BO} = 0.157$, $P < 0.001$) and between blue and red haplotypes ($F_{BR} = 0.146$, $P < 0.001$). The differentiation was weaker but still significant between orange and red haplotypes ($F_{OR} = 0.078$, $P < 0.010$). We found no significant nuclear genetic differentiation between the yellow and any other haplotype, likely due to small sample size (the yellow haplotype occurred in only eight populations). Nuclear genetic differentiation between the Eastern and Western lineages was highly significant ($F_{L/T} = 0.126$, $P < 0.001$). Mean genetic differentiation (as measured by multilocus F_{ST} estimates) within these two main lineages was 0.187 and 0.195 for Eastern and Western lineages, respectively (all at $P < 0.001$). Mean multilocus allele size-based estimate of genetic differentiation R_{ST} were (i) 0.357 ($P < 0.01$) among the 111 sampled populations, (ii) 0.173 ($P > 0.05$) between the Western and Eastern lineages, (iii) 0.014 ($P > 0.05$) between the two main genetic clusters within the Eastern lineage, and (iv) 0.274 ($P < 0.01$) between the two main genetic clusters within the Western lineage. Over all loci, only 43% of the alleles were shared between the Eastern and Western nuclear–plastid lineages (Table 1).

Spatial genetic structure within nuclear–plastid lineages

When focussing on the Western lineage, the best mode suggested by the Bayesian clustering was $K = 2$ (Supplementary Figure S3), mirroring the nuclear genetic divergence between the red and the yellow plastid haplotype populations and the orange haplotype populations (Figure 1d). Additional hierarchical structure was detected at $K = 6$ (Supplementary Figure S3): red and yellow haplotype populations were distinct and the orange haplotype was partitioned into four further population clusters, corresponding to four geographical regions: southern Belgium (*Si* ecotype), northern France, western France together with southern England (var. *salmoniana*), and south-central France (Figure 1d, Supplementary Table S1). sPCA revealed a latitudinal (north–south) cline regardless of haplotype identity (Figure 4a). This global structure ($P = 0.001$) showed a strong signal of positive spatial autocorrelation ($I = 0.694$) and represented a high proportion of the genetic variability (var = 0.471). The Mantel correlogram displayed a clear continuous decrease in genetic similarity with increasing geographic distance (Figure 4a).

When analysing the Eastern nuclear–plastid lineage, a partition between (i) inland continental populations (Germany, Luxemburg, eastern France and the *Ca* ecotype from southern Belgium) and (ii) populations located near the Channel coast (northern France, Belgium and The Netherlands) and in the UK (var. *smithiana*) was underlined by both Bayesian clustering with the highest likelihood for $K = 2$ (Figure 1f, Supplementary Figure S3) and the plot of scores from the first global component of the sPCA (Figure 4b). The global structure ($P = 0.034$) showed a strong signal of positive spatial autocorrelation ($I = 0.784$). The Mantel correlogram pictured a typical pattern of long-distance differentiation owing to a substantial geographic break in allele frequencies across the range of the Eastern lineage (Figure 4b).

Detection of hybridisation events in contact zones

Although most populations were fixed for a single plastid haplotype, we tested for the possibility of hybridisation events mediated by pollen in contact zones. We focussed on three areas where populations of different haplotypes co-occurred and ran further Bayesian assignments on these subsets of populations (see Supplementary Figure S1). Admixture was absent: groups identified by Bayesian clustering always referred clearly to haplotype identity with no signature of admixed
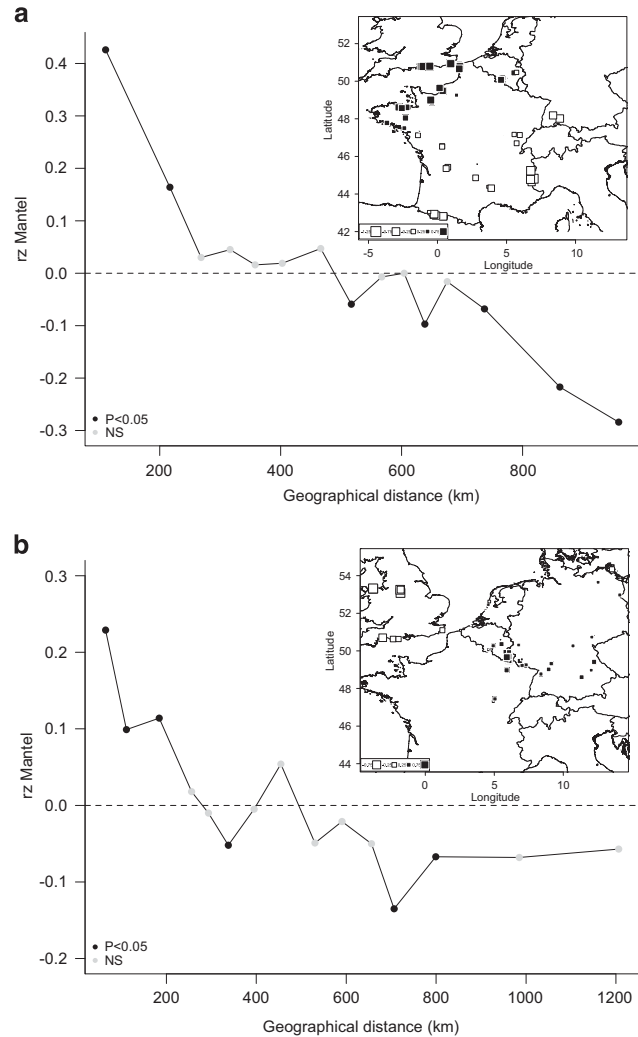

Figure 4 Spatial genetic structure within the two main Western and Eastern nuclear-plastid lineages in *Silene nutans*. One dimensional Mantel correlograms of (**a**) Western and (**b**) Eastern lineage. To reduce the random effects of sampling and ensure statistical supports for calculations, correlograms were designed with a minimum number of, respectively, 99 and 102 pairs of populations in each distance class. *rz*: normalised Mantel statistics. For each correlogram plot, the geographical map of the first global scores of sPCAs is also given. Each square represents a population. Populations that are more closely related in multivariate space share the same colour of squares. Sizes of squares indicate the magnitude of spatial positive autocorrelation.

nuclear genotypes or evidence of discordance between nuclear membership probability and plastid haplotype identity (Supplementary Figure S1).

DISCUSSION
Broad patterns of nuclear and plastid genetic diversity
S. nutans has been described as a self-compatible, gynomonoecious–gynodioecious species with a censer mechanism of seed dispersal (Hepper, 1956). Hermaphroditic plants are able to self by geitonogamy, with an extreme variation in selfing rates among individuals (Dufaÿ *et al.*, 2010). F_{IS} values were positive and significant for several populations and most loci, suggesting propensity for selfing even when accounting for null allele occurrence. Assuming that populations are at genetic equilibrium, the classic relationship between the

intra-population fixation index and the selfing rate $(s = 2F_{IS}/(1+F_{IS}))$ gave a mean selfing rate of 25.4% and 9.3% for F_{IS} and F_{IS}_corr, respectively. However, positive significant F_{IS} values may also result from biparental inbreeding owing to restricted pollen and/or seed dispersal. As a consequence, mean within-population selfing rate based on identity disequilibrium and less sensitive to biparental inbreeding (David *et al.*, 2007) gave a much lower rate of 6%. Overall, biparental inbreeding due to mating between relatives, along with a low rate of self-fertilisation, is expected to amplify the potential for isolation by distance pattern developing kin-structure at a local scale (Heywood, 1991; Vekemans and Hardy, 2004).

Inbreeding, along with restricted gene flow, leads to high genetic differentiation among populations. The mean level of genetic differentiation within Western ($F_{ST} = 0.195$) and Eastern ($F_{ST} = 0.187$) lineages of *S. nutans* was concordant with a mixed-mating system and moderate levels of gene flow among populations (Duminil *et al.*, 2009). The nearly complete fixation of plastid haplotypes and their strong geographical clustering may either confirm a spatially restricted seed dispersal and/or a lack of within-lineages plastid diversity with density-dependent processes or ecological factors that constraint lineages' distribution (Waters *et al.*, 2013). Moreover, a continuous decrease in genetic similarity with increasing geographical distance was shown, which suggested limited pollen/seed dispersal. Therefore, to conserve substantial levels of genetic diversity, *in situ* as well as *ex situ*, the preserved populations should reflect the nuclear–plastid differentiation patterns. Attention should particularly be paid to the southern populations, which have conserved the highest allelic richness.

Phylogeographic patterns in *S. nutans*

At a continental scale, leading-edge expansions following the Last Glacial Maximum may result in multiple phylogeographic patterns that are species specific as each taxon responded independently to Quaternary cold periods (Taberlet *et al.*, 1998). Rapid postglacial recolonisation through repeated population bottlenecks of founding lineages, blocking establishment of latecomers, is expected to reduce within-lineage genetic diversity with increasing distance from its refugium (Waters *et al.*, 2013).

In Western Europe, *S. nutans* showed a phylogeographic pattern with two main, strongly differentiated Western and Eastern nuclear–plastid lineages ($F_{L/T} = 0.126$). Patterns of long-distance differentiation and spatial trends in allelic richness suggested a Western lineage originating from the western Mediterranean Basin and an Eastern lineage, probably originating from Eastern Europe. The Eastern lineage exhibited a north-westward expansion, from eastern Germany to northern Wales with a genetic break found along the continental European coastline. The Western lineage harboured three plastid haplotypes, spatially structured, consistent with a northward expansion from distinct potential refugia. These hypotheses are supported by the asymmetrical patterns of private alleles between the two main nuclear–plastid lineages. Most private alleles were found in the Western lineage. Moreover, allele size-based measure of genetic differentiation yielded a highly significant R_{ST} estimate of 0.274 between the two main genetic clusters found in this Western lineage. Overall, this could suggest (i) a predominant role of stepwise-like mutation processes over migration and relative cluster isolation, and (ii) different glacial refugia composed of distinct genetic pools. We should keep in mind that our microsatellite loci may not follow a strict stepwise mutation process and that homoplasy is likely to occur, which may explain why no significant R_{ST} estimate was detected between Western and Eastern lineages.

Secondary contacts: from diverging lineages to cryptic species

Beyond shaping the distribution of present genetic diversity, Quaternary climate oscillations have been shown to favour speciation processes (Hewitt, 2000). Genetic drift and selection operating in different, isolated glacial refugia can lead to genetic divergence. Patterns of postglacial expansion from different refugia found in animals, plants, fungal taxa and described in *S. nutans* here produce suture zones in which lineages come into secondary contact (Taberlet *et al.*, 1998; Hewitt, 2000, 2011). Many of these suture zones have been described across Europe, such as in the Alps, the Pyrenees or in eastern Central Europe along the eastern borders of Germany (Hewitt, 2000). Accordingly, our results in Western Europe identified three suture zones where the Eastern and Western nuclear–plastid lineages of *S. nutans* co-occur: in southern Belgium, eastern France, and surprisingly, in southern England. Although previous phylogeographic studies have reported that only one lineage colonised Great Britain (for example, oaks expanding from Spain or grasshoppers from the Balkans, reviewed in Hewitt, 2011), our study on *S. nutans* is, to our knowledge, the first that shows the existence of a contact zone in Great Britain shaped by postglacial expansion.

Our study shows that the two morphological varieties previously described in Hepper (1951) in Great Britain, *S. nutans* var. *salmoniana* and var. *smithiana*, belong to separate nuclear–plastid lineages, Western and Eastern, respectively. Morphological differences, for example, differences in fruit size, have been also found elsewhere, in particular between blue and orange/red haplotype populations, whereas the continental populations belonging to the yellow and blue haplotypes did not show such contrasting differences (Van Rossum, unpublished data). Moreover, the two reproductively isolated Belgian *Ca* and *Si* parapatric edaphic ecotypes also correspond to the Eastern and Western nuclear–plastid lineages, respectively. Bayesian clustering performed on the three contact zones identified in *S. nutans* did not reveal any signatures of admixture between haplotype lineages, suggesting that no hybridisation events occur in natural populations. Such result was somewhat unexpected given that range shifts and expansions of a colonising species into the range of a local species may lead to haplotype sharing owing to longer pollen than seed dispersal (for example, Petit and Excoffier, 2009). Altogether, this strict association between nuclear gene pools and plastid haplotype suggests cryptic speciation in *S. nutans*, at least for Eastern and Western nuclear–plastid lineages for which a strong reproductive barrier can be suspected. From an evolutionary point of view, suture zones resulting from endogenous selection are not stabilised geographically and their dynamics are thought to be maintained by a balance between dispersal and selection against hybrids (Barton, 1979). These tension zones may be associated with an environmental transition where different genotypic combinations allow differential adaptive values, resulting in a genetic–environment association and ecotypic differentiation (Barton and Hewitt, 1985; Hewitt, 1988; Bierne *et al.*, 2011). Although in southern Belgium the genetic lineages are clearly associated with edaphic properties (for example, Van Rossum *et al.*, 1997), to date, no evidence of such association between genetic variation and soil characteristics have been found elsewhere (Van Rossum *et al.*, 2003; Van Rossum and Prentice, 2004). This inconsistency suggests that the divergence of the two main lineages precedes the specific adaptation to the nature of soil and indicates locally restricted association between edaphic adaptation and lineage identity.

CONCLUSION

Based on our findings on genetic structure in *S. nutans*, two strongly differentiated evolutionary lineages were identified with no

hybridisation events detected using nuclear markers. Besides, no plastid haplotypes were shared between these two lineages. This suggests the occurrence of at least two cryptic species within *S. nutans*, with no introgression events affecting species integrity. According to Mayr's biological species concept (Coyne and Orr, 2004), two main ESUs can therefore be defined in Western Europe, corresponding to Eastern and Western nuclear–plastid lineages. Moreover, within the Western lineage, three potential ESUs might also be considered, corresponding to the three sub-lineages. Given the differences in their distribution range, the conservation status of these distinct ESUs should be evaluated separately. *In situ* preservation of populations and genetic rescue implying *ex situ* conservation techniques should take the lineage identity into account. Applying such conservation approaches are particularly true in Great Britain, northern France and Belgium, where *S. nutans* is rare and where distinct lineages co-occur in close contact.

DATA ARCHIVING

1. DNA sequences: Genbank accession numbers KJ671551–KJ671562.
2. Coordinates of sample locations and nuclear genetic data are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.vq4j4.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Avise JC (2000). *Phylogeography: the History and Formation of Species*. Harvard University Press: Cambridge, MA, USA.

Barton NH (1979). The dynamics of hybrid zones. *Heredity* **43**: 341–359.

Barton NH, Hewitt GM (1985). Analysis of hybrid zones. *Annu Rev Ecol Syst* **16**: 113–148.

Bierne N, Welch J, Loire E, Bonhomme F, David P (2011). The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol Ecol* **20**: 2044–2072.

Cavalli-Sforza LL, Edwards AWF (1967). Phylogenetic analysis. Models and estimation procedures. *Am J Hum Genet* **19**: 233–257.

Chybicki IJ, Burczyk J (2009). Simultaneous estimation of null alleles and inbreeding coefficients. *J Hered* **100**: 106–113.

Coyne JA, Orr HA (2004). *Speciation*. Sinauer associates: Sunderland, Massachusetts, USA.

David P, Pujol B, Viard F, Castella V, Goudet J (2007). Reliable selfing rate estimates from imperfect population genetic data. *Mol Ecol* **16**: 2474–2487.

De Bilde J (1973). Etude génécologique du *Silene nutans* L. en Belgique: populations du *Silene nutans* L. sur substrats siliceux et calcaires. *Rev Gén Bot* **80**: 161–176.

DiLeo MF, Row JR, Lougheed SC (2010). Discordant patterns of population structure for two co-distributed snake species across a fragmented Ontario landscape. *Divers Distrib* **16**: 571–581.

Dufaÿ M, Lahiani E, Brachi B (2010). Gender variation and inbreeding depression in gynodioecious-gynomonoecious *Silene nutans* (Caryophyllaceae). *Int J Plant Sci* **171**: 53–62.

Duminil J, Fineschi S, Hampe A, Jordano P, Salvini D, Vendramin GG et al. (2007). Can population genetic structure be predicted from life-history traits? *Am Nat* **169**: 662–672.

Duminil J, Hardy OJ, Petit RJ (2009). Plant traits correlated with generation time directly affect inbreeding depression and mating system and indirectly genetic structure. *BMC Evol Biol* **9**: 177–190.

El Mousadik A, Petit RJ (1996). High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinos* (L.) Skeels] endemic to Morocco. *Theor Appl Genet* **92**: 832–839.

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**: 2611–2620.

Fitter A (1978). *An Atlas of the Wild Flowers of Britain and Northern Europe*. Collins: Glasgow, UK.

Fraser DJ, Bernatchez L (2001). Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Mol Ecol* **10**: 2741–2752.

Gavin DG, Fitzpatrick MC, Gugger PF, Heath KD, Dobrowski SZ, Hampe A et al. (2014). Climate refugia : joint inference from fossil records, species distribution models and phylogeography. *New Phytol* **204**: 37–54.

Godé C, Touzet P, Martin H, Lahiani E, Delph LF, Arnaud J-F (2014). Characterization of 24 polymorphic microsatellite markers for *Silene nutans*, a gynodioecious-gynomonoecious species, and cross-species amplification in other *Silene* species. *Conserv Genet Resour* **6**: 915–918.

Goudet J (1995). FSTAT (Version 1.2): a computer program to calculate *F*-statistics. *J Hered* **86**: 485–486.

Goudet J (2005). HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics. *Mol Ecol Notes* **2**: 184–186.

Hamrick JL, Godt MJW (1996). Effects of life history traits on genetic diversity in plant species. *Philos Trans R Soc Lond B Biol Sci* **351**: 1291–1298.

Hardy OJ, 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 OJ, Vekemans X (2002). SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* **2**: 618–620.

Hautekèete N-C, Frachon L, Luczak C, Toussaint B, Van Landuyt W, Van Rossum F et al. (2015). Habitat type shapes long-term plant biodiversity budgets in two densely populated regions in north-western Europe. *Divers Distrib* **21**: 631–642.

Hepper FN (1951). The variations of *Silene nutans* L. in Great Britain. *Watsonia* **2**: 80–90.

Hepper FN (1956). Biological flora of the British Isles: *Silene nutans* L. *J Ecol* **44**: 693–700.

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

Hewitt GM (1988). Hybrid zones—natural laboratories for evolutionary studies. *Trends Ecol Evol* **3**: 158–167.

Hewitt GM (2011). Quaternary phylogeography: the roots of hybrid zones. *Genetica* **139**: 617–638.

Heywood JS (1991). Spatial analysis of genetic variation in plant populations. *Annu Rev Ecol Syst* **22**: 335–355.

Jakobsson M, Rosenberg NA (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**: 1801–1806.

Jombart T (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403–1405.

Jombart T, Devillard S, Dufour A-B, Pontier D (2008). Revealing cryptic spatial patterns in genetic variability with a new multivariate method. *Heredity* **101**: 92–103.

Jombart T, Pontier D, Dufour A-B (2009). Genetic markers in the playground of multivariate analysis. *Heredity* **102**: 330–341.

Jürgens A, Witt T, Gottsberger G (1996). Reproduction and pollination in Central European populations of *Silene* and *Saponaria* species. *Bot Acta* **109**: 316–324.

Lahiani E, Dufaÿ M, Castric V, Le Cadre S, Charlesworth D, Van Rossum F et al. (2013). Disentangling the effects of mating systems and mutation rates on cytoplasmic diversity in gynodioecious *Silene nutans* and dioecious *Silene otites*. *Heredity* **111**: 157–164.

Oden NL, Sokal RR (1986). Directional autocorrelation an extension of spatial correlograms to two dimensions. *Syst Zool* **35**: 608–617.

Palsbøll PJ, Bérubé M, Allendorf FW (2007). Identification of management units using population genetic data. *Trends Ecol Evol* **22**: 11–16.

Petit RJ, Excoffier L (2009). Gene flow and species delimitation. *Trends Ecol Evol* **24**: 386–393.

Petit RJ, Hu FS, Dick CW (2008). Forests of the past: a window to future changes. *Science* **320**: 1450–1452.

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

R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, Austria. Available from http://www.R-project.org/.

Rosenberg MS, Anderson CD (2011). PASSaGE: Pattern Analysis, Spatial Statistics and Geographic Exegesis. Version 2. *Methods Ecol Evol* **2**: 229–232.

Rosenberg NA (2004). DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes* **4**: 137–138.

Rousset F (2008). GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Res* **8**: 103–106.

Slatkin M (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**: 457–462.

Smouse PE, Long JC, Sokal RR (1986). Multiple regression and correlation mantel test of matrix correspondence. *Syst Zool* **35**: 627–632.

Stace C (2010). *New Flora of the British Isles*, 3rd edn. Cambridge University Press: Cambridge, UK.

Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F (1998). Comparative phylogeography and postglacial colonization. *Mol Ecol* **7**: 453–464.

Van Rossum F, De Bilde J, Lefèbvre C (1996). Barriers to hybridization in calcicolous and silicolous populations of *Silene nutans* from Belgium. *Belgian J Bot* **129**: 13–18.

Van Rossum F, Prentice HC (2004). Structure of allozyme variation in Nordic *Silene nutans* (Caryophyllaceae): population size, geographical position and immigration history. *Biol J Linn Soc* **81**: 357–371.

Van Rossum F, Vekemans X, Gratia E, Meerts P (2003). A comparative study of allozyme variation of peripheral and central populations of *Silene nutans* L. (Caryophyllaceae) from Western Europe: implications for conservation. *Plant Syst Evol* **242**: 49–61.

Van Rossum F, Vekemans X, Meerts P, Lefèbvre C (1997). Allozyme variation in relation to ecotypic differentiation and population size in marginal populations of *Silene nutans*. *Heredity* **78**: 552–560.

Vekemans X, Hardy OJ (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol Ecol* **13**: 921–935.

Waters JM, Fraser CI, Hewitt GM (2013). Founder takes all: density-dependent processes structure biodiversity. *Trends Ecol Evol* **28**: 78–85.

Weir BS, Cockerham CC (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.

Yang R-C (1998). Estimating hierarchical *F*-statistics. *Evolution* **52**: 950–956.

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)