BIODIVERSITY RESEARCH

Retrospective genetic monitoring of the threatened Yellow marsh saxifrage (*Saxifraga hirculus*) reveals genetic erosion but provides valuable insights for conservation strategies

Gemma E. Beatty[1], Neil Reid[2] and Jim Provan[1]*

[1] *School of Biological Sciences, Queen's University Belfast, 97 Lisburn Road, Belfast BT9 7BL*

[2] *Quercus, School of Biological Sciences, Queen's University Belfast, 97 Lisburn Road, Belfast BT9 7BL*

Correspondence: Dr Jim Provan

School of Biological Sciences,

Queen's University Belfast,

97 Lisburn Road,

Belfast BT9 7BL

E-mail: J.Provan@qub.ac.uk

Tel: +44 (0)28 9097 2280

Fax: +44 (0)28 9097 5588

Running title: Genetic erosion in a threatened plant species

Word count: 3,469 (excluding title page, abstract, references, tables and figures)

ABSTRACT

Aim Retrospective genetic monitoring, comparing genetic diversity of extant populations with historical samples, can provide valuable and often unique insights into evolutionary processes informing conservation strategies. The Yellow marsh saxifrage (*Saxifraga hirculus*) is listed as 'critically endangered' in Ireland with only two extant populations. We quantified genetic changes over time and identified genotypes in extant populations that could be used as founders for reintroductions to sites where the species is extinct.

Location Ireland.

Methods Samples were obtained from both locations where the species is currently found, including the most threatened site at the Garron Plateau, Co. Antrim, which held only 13 individuals during 2011. Herbarium samples covering the period from 1886 to 1957 were obtained including plants from the same area as the most threatened population, as well as three extinct populations. In total, 422 individuals (319 present-day and 103 historical) were genotyped at six microsatellite loci. Species distribution modelling was used to identify areas of potentially suitable habitat for reintroductions.

Results Level of phenotypic diversity within the most threatened population was significantly lower in the present-day compared to historical samples but levels of observed heterozygosity and number of alleles, whilst reduced, did not differ significantly. However, Bayesian Clustering Analysis suggested gradual lineage replacement over time. All three measures of genetic diversity were generally lower at the most threatened population compared to the more substantial extant populations in Co. Mayo. Species distribution

modelling suggested that habitat at one site where the species is extinct may be suitable for reintroduction.

Main conclusions The dominant genetic lineage in the most threatened population is rare elsewhere, thus care needs to be taken when formulating any potential reintroduction programme. Our findings highlight both the need for genetic monitoring of threatened populations, but also for its swift implementation before levels of diversity become critically low.

Keywords

Bayesian clustering analysis, microsatellites, polyploid, *Saxifraga hirculus*, Yellow marsh saxifrage.

INTRODUCTION

Knowledge of levels and patterns of intraspecific genetic diversity represents a fundamental aspect of modern conservation biology. Researchers and policy makers are now aware of the various implications of habitat loss and population extinction on diversity below the species level. Such information is crucial in the estimation of effective and minimum viable population sizes, as well as levels of inbreeding and adaptive potential (Allendorf & Luikart, 2007; Schwartz *et al.*, 2007). These factors are particularly relevant in populations comprising very low numbers of individuals, and typically those in immediate need of conservation.

Whilst there have been several recent attempts to predict the potential impacts of future habitat loss and/or population extinction on species' genetic diversity (Balint *et al.*, 2011; Beatty & Provan, 2011; Provan & Maggs, 2012), relatively few studies have directly quantified historical loss of diversity due to past (and ongoing) extinctions. These studies generally rely on the genetic analysis of museum or herbarium samples, and have often highlighted loss of genetic diversity in extant populations compared with their historical counterparts (reviewed in Wandeler *et al.*, 2007 and Leonard, 2008). Such retrospective genetic monitoring can provide valuable, and often unique, insights into evolutionary processes that can inform future conservation programmes (Schwartz *et al.*, 2007; Wandeler *et al.*, 2007; Jackson *et al.*, 2012).

The Yellow Marsh Saxifrage (*Saxifraga hirculus*) is a perennial herbaceous plant with a circumpolar distribution (Hedberg, 1992). The species originated in central Asia (Hedberg, 1992), but the sole phylogeographic study carried out to date identified Alaska as the centre of genetic diversity, suggesting survival in an Alaskan / Beringian refugium during the Pleistocene glaciations (Oliver *et al*., 2006). The species suffered dramatic declines

63 throughout in Europe during the last *ca.* 200 years, primarily as a result of habitat loss,
64 principally wetlands, with remaining populations being small and widely scattered (Vittoz *et*
65 *al.*, 2006). The Irish Red Data Book for Vascular Plants lists the species under the
66 International Union for the Conservation of Nature (IUCN) category of 'critically
67 endangered' (Curis & McGough, 1988). We analysed herbarium samples spanning over 150
68 years, including extinct populations, as well as samples representing populations from the
69 only two locations where the species is currently extant, to determine if there has been loss of
70 genetic variation and to use this information to formulate species augmentation and
71 reintroduction programmes further informed by species distribution modelling, which was
72 used to identify potentially suitable habitat for any reintroductions.

METHODS

Study sites

Only two populations of *Saxifraga hirculus* occur in Ireland; the most threatened population consisted of only 13 individuals at the Garron Plateau, County Antrim during 2011 whilst the other is substantially larger occurring at 13 sites (each with *ca.* 100-200 individuals) near Bellacorcik, Co. Mayo. Herbarium samples show that the species occurred at much greater abundance at the Garron Plateau in the past and also occurred at another site near Rasharkin, County Antrim. Two further populations, one near Coleraine, County Derry and one near Lisclogher Co. Westmeath are now extinct.

Study species

Chromosome numbers for *S. hirculus* include $2n = 16$, $2n = 24$ and $2n = 32$, although all populations outside the Arctic Circle studied to date, including those in Europe, are tetraploid ($2n = 32$). Reproduction is both sexual and asexual. Flowers are markedly protandrous, but self-compatible, and are pollinated by a wide range of insects, with different pollinators in different parts of the species' range (Olesen & Warncke, 1989; Warncke *et al.*, 1993). Seeds lack special adaptations for dispersal, and are generally deposited close to the parent plant (Olesen & Warncke, 1989). Vegetative reproduction occurs via rhizomes, and is believed to be an important mechanism of propagation (Olesen & Warncke, 1990).

Sampling and DNA extraction

A total of 319 present-day samples were obtained during 2011. All 13 plants from the Garron Plateau, Co. Antrim were sampled and 306 samples were collected from all 13 sites included within the Co. Mayo population (24 individuals from each with the exception of site SHE-D,

98 for which 18 individuals were analysed). The species is protected under the Wildlife Acts
99 1976-2010 (Ireland) and Schedule 8 of the Wildlife (Northern Ireland) Order (1985), and it
100 was an offence to pick, uproot or destroy the plant. Consequently, a single leaf was taken
101 from each plant under Government licence. Samples were stored in silica gel for
102 transportation. A total of 103 historical samples were obtained from individual plants on
103 herbarium sheets representing both extant locations in Counties Antrim and Mayo, as well as
104 from extinct populations at Rasharkin, Co. Antrim, Coleraine, Co. Derry and Lisclogher, Co.
105 Westmeath. The sampling regime of herbarium samples and their distribution are given in
106 Table 1 and Figure 1 (for herbarium codes see Table S1).
107 DNA was extracted from all samples using the Qiagen DNeasy Plant Mini Kit, after an
108 initial 8 min grinding at 30 Hz using a Retsch MM300 mixer mill. DNA was quantified
109 visually on 1% agarose gels stained with ethidium bromide and diluted to a concentration of
110 50 ng μl^{-1} for subsequent PCR. All DNA extractions from herbarium samples were carried
111 out in a laboratory where no previous *S. hirculus* work had been performed.
112
113 **Microsatellite genotyping**
114 All samples were genotyped for six microsatellite markers developed for *S. hirculus* using the
115 ISSR cloning method outlined in Provan & Wilson (2007). Because of difficulties associated
116 with amplifying longer fragments from herbarium samples, primers were designed to amplify
117 products of less than *ca.* 200 bp (Table 2). The polyploid nature of *S. hirculus* in Ireland
118 represented a further problem in scoring allele sizes accurately where stutter bands were
119 present. Consequently, we were limited to using two trinucleotide microsatellites and two
120 tetranucleotides, which generally display minimal stuttering, and two dinucleotides with
121 relatively low stuttering. Forward primers were modified by the addition of a 19 bp M13 tail
122 (5'-CACGACGTTGTAAAACGAC-3') and reverse primers were modified by the addition

of a 7 bp tail (5'-GTGTCTT-3'). PCR was carried out in a total volume of 10 µl containing 100 ng genomic DNA, 10 pmol of HEX-labelled M13 primer, 1 pmol of M13-tailed forward primer, 10 pmol reverse primer, 1x PCR reaction buffer, 200 µM each dNTP, 2.5 mM $MgCl_2$ and 0.25 U GoTaq Flexi DNA polymerase (Promega). PCR was carried out on a MWG Primus thermal cycler using the following conditions for all loci with the exception of SH-3-B11: initial denaturation at 94 ^{o}C for 3 min followed by 10 touchdown cycles of denaturation at 94 ^{o}C for 30 s, annealing at 68 ^{o}C for 30 s (-1 ^{o}C per cycle), extension at 72 ^{o}C for 30 s followed by 30 cycles (40 for herbarium samples) of denaturation at 94 ^{o}C for 30 s, annealing at 58 ^{o}C for 30 s, extension at 72 ^{o}C for 30 s and a final extension at 72 ^{o}C for 5 min. For locus SH-3-B11, the following conditions were used: initial denaturation at 94 ^{o}C for 3 min followed by 40 cycles (50 for herbarium samples) of denaturation at 94 ^{o}C for 30 s, annealing at 52 ^{o}C for 30 s, extension at 72 ^{o}C for 30 s and a final extension at 72 ^{o}C for 5 min. Blank negative controls were routinely used, and approximately 25% of herbarium samples were genotyped twice to check for artifacts resulting from low DNA quality, which were not observed. Genotyping was carried out on an AB3730xl capillary genotyping system. Allele sizes were scored using LIZ-500 size standards and were checked by comparison with previously sized control samples.

Genetic analysis

No chromosome counts are available for Irish *S. hirculus*, but the species is believed to be polyploid in the majority of its non-Arctic range (Hedberg, 1992), and more than two bands were observed at all six loci analysed, suggesting polyploidy. Thus, it was not possible to score genotypes based on allele frequencies, and as a result we could not carry out many standard population genetic analyses (e.g. calculation of allelic richness, AMOVA). Consequently, the patterns of alleles observed for each locus in an individual were scored as

148 phenotypes. Within-population phenotype diversity was estimated for samples with $N \geq 5$
149 using the total number of alleles, observed heterozygosity (H_O), namely the proportion of
150 observed heterozygous individuals in a population, and the Gini-Simpson diversity index,
151 analogous to Nei's gene diversity, averaged over loci (Jost, 2006). Genetic clustering of
152 individuals was assessed using a Bayesian procedure implemented in the STRUCTURE
153 software package (V2.3.3; Pritchard *et al.*, 2000), which can accommodate ambiguous
154 codominant markers for polyploid species. The program was run using no prior knowledge
155 and the admixture ancestry model. Five independent runs were carried out for each value of
156 K, the number of genetic clusters, up to $K = 10$, since log-likelihood values reached a peak at
157 $K = 8$ and decreased thereafter. Each Markov chain Monte Carlo analysis used a burn-in
158 period of 10,000 followed by a further 100,000 iterations. The most likely value for K was
159 estimated using the ΔK statistic of Evanno *et al.* (2005) implemented in the STRUCTURE
160 HARVESTER software package (V0.6.1; Earl & vonHoldt, 2012).
161
162 **Species distribution modelling**
163 A presence-only maximum entropy approach was used to predict landscape suitability for the
164 *S. hirculus* throughout Ireland, from a sample set of known occurrences and spatially explicit
165 environmental parameters. Maximum entropy has been shown to frequently outperform other
166 presence-only modelling techniques particularly at very low sample sizes (Elith & Graham,
167 2009). Environmental parameters were described at a 500m cell resolution (Table S2). The
168 software package MAXENT was used (V3.3.3k; Philips *et al.,* 2010). To maximise model
169 flexibility, we considered linear, quadratic, threshold and hinged functions for all
170 environmental parameters (Phillips & Dudík, 2008). Due to the paucity of records it was not
171 possible to segregate the dataset into a training and test sets, thus only a training set was used
172 (Farren *et al.* 2010). Jackknife re-sampling analysis was used to determine a heuristic

173 estimate of the relative contribution of each variable based on the performance of the global
174 model (known as the regularized gain) without the variable of interest compared to the
175 influence of that variable in isolation (derived from a univariate model only). Global model
176 performance was judged using the area under the receiver operating characteristic (ROC)
177 curve (Liu *et al.*, 2005). Marginal response curves of the predicted probability of species
178 occurrence were graphed for each explanatory variable. A map of landscape favourability
179 was generated using ArcGIS 9.3 (ESRI, California, USA) and the 10^{th} percentile training
180 presence was used as the threshold.

RESULTS

Between seven and 16 alleles were detected across the six loci analysed (average 10 alleles per locus; Table 2). The total number of alleles per population with sample numbers $N \geq 5$ ranged from 18 (GAR-1886 population) to 32 (AGH-2011 population). Observed heterozygosity (H_O) ranged from 0.389 (RAS-1884 population) to 0.882 (SHB-2011 population; Table 1). Levels of phenotype diversity (H) ranged from 0.387 (GAR-2011 population) to 0.821 (BEL-1968 population; Table 1). The extant Garron Plateau population had the third lowest level of H_O, after the GAR-1955 sample, and the lowest level of H. Both values (0.436 and 0.387, respectively) were lower than those in the extant Co. Mayo population, where H_O ranged from 0.500 to 0.882 (mean = 0.730) and H ranged from 0.512 to 0.808 (mean = 0.670). For the Garron Plateau population, the level of H (calculated over all individuals i.e. 24 individuals sampled across six time points vs. 13 samples from 2011) was significantly higher in the herbarium samples than in the extant samples, whereas levels of H_O and average number of alleles per individual (A), whilst higher for herbarium samples (again calculated over all individuals), were not significantly different (Table 3).

The results of the Bayesian clustering analysis indicated that the most likely number of genetic clusters was $K = 4$, followed by $K = 8$ (Figure 2). For $K = 4$ genetic clusters (Figure 3, top), the vast majority of plants from the Garron Plateau and the now extinct Rasharkin population were predominantly associated with the cluster shown in yellow, which was comparatively rare elsewhere in Ireland. The population from Co. Mayo displayed varying degrees of admixture, from being almost completely dominated by a single cluster (red at SHB) to being comprised of a substantial fraction of all four clusters (SHA). Assignment of populations to $K = 8$ clusters (Figure 3, middle and bottom) revealed some further subtle genetic substructuring, particularly in the Northern Ireland populations. This shows a distinct

206 temporal shift from the lineage represented by light yellow, which was initially the dominant
207 lineage at both the Garron Plateau and Rasharkin populations, to the lineage represented in
208 light blue, to the point where the light yellow lineage is now almost completely absent.
209 *S. hirculus* occurrence was strongly associated with landscapes dominated by bog, fen,
210 marsh and swamp typically on high altitude (>175m above sea level) plateaus (i.e. low
211 hilliness index) which experienced low maximum temperatures (<15^{o}C), high precipitation
212 (1,000 - 1,700mm of rain annually) and low seasonality i.e. consistently cool and wet all year
213 round (Figures S1 and S2). These regions were typically negatively correlated with
214 agricultural pastures. Model performance, defined as the area under the curve or AUC =
215 0.998. The model suggested that the extent of suitable habitats for *S. hirculus* is limited (Fig.
216 4a). Highest suitability and the greatest extent of habitat was predicted throughout north Co.
217 Mayo (Fig. 4b). The model also suggested that less optimal habitat was found throughout the
218 Co. Antrim including the Garron Plateau (Fig. 4c). A patch of potentially suitable habitat
219 (determined using the 10th percentile training presence = 0.692) was also identified near
220 Rasharkin which may represent the site of an extinct population. The habitat at locations of
221 other now extinct populations near Coleraine, Co. Derry and Lisclogher, Co. Westmeath was
222 predicted to have vanished.
223

DISCUSSION

S. hirculus genetic erosion over time

Maintaining genetic variation in threatened populations, which by their nature tend to be small and/or fragmented, is one of the central tenets of conservation genetics (Allendorf & Luikart, 2007; Schwartz *et al.*, 2007). The retrospective genetic monitoring afforded by the comparison of a critically threatened extant population with historical samples from the same area suggests that levels of population genetic diversity of *S. hirculus* in the critically endangered Garron Plateau, Co. Antrim are not only far lower than those in the other, larger extant populations in Co. Mayo, but also appreciably lower than the average values from historical samples (pre-1958) from the same area, although there is evidence of some fluctuation in these values through time. This loss of genetic diversity has been accompanied by the replacement of one lineage identified by the Bayesian clustering analysis (shown in Figure 2, bottom, in light blue) by another (shown in light yellow), to the extent that the new lineage accounts for over 60% of the total genetic diversity. These changes are probably the result of a combination of factors, foremost among which are stochastic fluctuations in allele frequencies due to the greatly exaggerated effects of genetic drift in very small populations. These changes may also have been accompanied by a founder effect, since the Garron Plataeu population was recorded as extinct after 1920, but subsequently "refound" in 1955 (Kertland, 1956). The extinction of the population at Rasharkin, Co. Antrim at the end of the 19th century would have contributed further to the loss of this lineage, both directly and indirectly via the cessation of possible gene flow between the two Co. Antrim populations. Nevertheless, although the levels of phenotypic diversity were significantly lower in the extant population than in the historical samples, the number of alleles was not. Given that the

number of alleles is correlated with the ability to respond to selection, this does not appear to be as serious a concern as potential inbreeding depression.

Conservation implications

Although information from population genetic studies can inform best-practice conservation strategies, the long decline in *S. hirculus* population numbers highlights a further conservation dilemma. Numbers of recorded individuals at the Garron Plateau have fallen from 130 during 1999 to 13 during 2011 (Georgina Thurgate, *pers. comm.*). Despite the importance of vegetative reproduction in the spread and persistence of *S. hirculus* (Olesen & Warncke, 1990), we detected only two potentially clonal individuals (i.e. 15% shared identical multi-locus phenotypes). Nevertheless, the extremely small number of individuals comprising the extant population at the Garron Plateau leaves it extremely vulnerable to sudden, stochastic extinction. Thus, it is clear that some sort of augmentation programme is necessary. One of the goals of conservation genetics is to ensure that the provenance of individuals used for such programmes is closely aligned with the population (Lesica & Allendorf, 1999), but the Garron Plateau individuals are generally associated with a genetically distinct group that is not found at any significant level elsewhere in the remaining populations in Ireland. Re-establishment or augmentation of a population using genetically divergent individuals results in a trade-off between increasing population numbers at the risk of outbreeding depression (Edmunds, 2007). The Bayesian clustering analysis suggests that if a source population is required for augmentation, it should be the SHA population in Co. Mayo, which not only has the highest percentage assignment to the predominant cluster found in the Garron Plateau, but also contains a mixture of several different lineages. This represents a further important aspect of strategic conservation based on molecular genetic approaches, namely the fact that data from a low number of putatively neutral loci do not

273 necessarily provide insights into the relative fitness of genotypes, particularly in differing
274 habitats (Ennos *et al.* 1997; Hollingsworth *et al.* 1999). Augmentation or introduction of a
275 range of genotypes, including those closest to genotypes found in the extant Garron
276 population, as is the case for the Co. Mayo SHA population, will provide a balance between
277 "genotype matching" and sufficient variation for natural selection to operate on. Of course,
278 reintroduction need not be limited to material from a single source population, and the results
279 of the genetic clustering analysis could be used to identify the widest possible range of
280 genetic diversity for reintroduction, if so desired.
281 *S. hirculus* occurrence was associated with bog, fen, marsh and swamp typically on high
282 altitude plateaus with low maximum temperatures, high precipitation and low seasonality i.e.
283 consistently cool and wet. The loss of the Coleraine, Co. Derry and Lisclogher, Co.
284 Westmeath populations could be attributed to loss of such habitat, since the species
285 distribution model did not identify these areas as currently suitable. A number of areas in the
286 Garron Plateau, Co. Antrim were identified as potentially suitable for establishing new
287 populations (via *ex-situ* conservation) to further supplement and expand the extant
288 population. A single area of potentially suitable habitat was identified near the site of the now
289 extinct population at Rasharkin, Co. Antrim. Insights into the historical genetic makeup of
290 the now extinct Rasharkin population afforded by the analysis of the herbarium samples
291 mean that a controlled reintroduction program, based on recreating the genetic make-up of
292 the original population from individuals extant elsewhere (i.e. Co. Mayo), could maximise
293 the potential success of re-establishing this lost population.
294 The genetic changes revealed by the retrospective genetic monitoring indicate the need to
295 implement such approaches as soon as possible. Regular censuses of the population at the
296 Garron Plateau began during 1999 when there were 130 plants If genetic monitoring had
297 commenced at the same time there would have been more chance of developing a successful

298 *ex-situ* conservation programme to maximise genetic diversity than at present. As it is, the
299 current scenario further highlights the need for conservation practitioners to move away from
300 a 'fire-fighting' mentality (Mace and Purvis, 2008). Nevertheless, the findings of our study
301 can be used to inform any potential reintroduction / augmentation programmes. Results from
302 a previous re-establishment program in Scotland indicate that *ex-situ* propagation of seedlings
303 followed by transplantation is a more successful method than simply sowing seeds directly
304 onto potential recovery sites (Welch 2002). Based on the information from the current study,
305 genetic analysis of *ex-situ* individuals could be used to select individuals most representative
306 of the current extant gene pool, whilst aiming to maximise genetic diversity.
307

ACKNOWLEDGEMENTS

This study was commissioned and funded by the Natural Heritage Research Partnership between the Northern Ireland Environment Agency (NIEA), Department of the Environment Northern Ireland and *Quercus*, Queen's University Belfast. We are grateful to Kyle Hunter and Mark Wright (NIEA) and James Kilroy, Cameron Clothworthy, Leonard Floyd and Deirdre Lynn from the National Parks & Wildlife Service, Department of Arts, Heritage and the Gaeltacht, Republic of Ireland for sample collection. Herbarium samples were provided by Colin Kelleher, Noeleen Smyth and Mathew Jebb (National Botanic Garden of Ireland) and Peter Crowther (Ulster Museum, National Museums Northern Ireland). We are also grateful to Associate Editor Andy Lowe and two anonymous reviewers for constructive comments that greatly improved the manuscript. Many thanks to Georgina Thurgate who acted as NIEA Client Officer.

REFERENCES

Allendorf, F.W. & Luikart, G. (2007) *Conservation and the Genetics of Populations*, Oxford: Blackwell.

Bálint, M., Domisch, S., Engelhardt, C.H.M., Haase, P., Lehrian, S., Sauer, J., Theissinger, K., Pauls, S.U. & Nowak, C. (2011) Cryptic biodiversity loss linked to global climate change. *Nature Climate Change* **1**, 313-318.

Beatty, G.E. & Provan, J. (2011) Comparative phylogeography of two related plant species with overlapping ranges in Europe, and the potential effects of climate change on their intraspecific genetic diversity. *BMC Evolutionary Biology* **11**, 29.

Earl, D.A. & vonHoldt, B.M. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**, 359-361.

Edmands, S. (2007) Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Molecular Ecology* **16**, 463-475.

Elith, J. & Graham, C. H. (2009) Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography* **32**, 66-77.

Evanno, G., Regnaud, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Farren A, Prodöhl, P. A., Laming, P. Reid N. (2010) Distribution of the common lizard (*Zootoca vivipara*) and landscape favourability for the species in Northern Ireland. *Amphibia-Reptilia*, **31**, 387-394.

344 Hedberg, K.O. (1992) Taxonomic differentiation in *Saxifraga hirculus* L. (Saxifragaceae) – a
345 circumpolar Arctic-Boreal species of Central-Asiatic origin. *Botanical Journal of the*
346 *Linnean Society* **109**, 377-393.
347 Hollingsworth, P. M., Tebbitt, M, Watson, K.J. & Gornall, R.J. (1998) Conservation genetics
348 of an arctic species, *Saxifraga rivularis* L., in Britain. *Botanical Journal of the Linnean*
349 *Society* **128**, 1-14.
350 Jackson, J.A., Laikre, L., Baker, C.S. & Kendall, K.C. (2012) Guidelines for collecting and
351 maintaining archives for genetic monitoring. *Conservation Genetics Resources* **4**, 527-
352 536.
353 Jost, L. (2006) Entropy and diversity. *Oikos* **113**, 363-375.
354 Kertland, M.P.H. (1956) *Subularia aquatic* and *Saxifraga hirculus* in Co. Antrim. *Irish*
355 *Naturalists' Journal* **12**, 51-52.
356 Leonard, J.A. (2008) Ancient DNA: applications for wildlife conservation. *Molecular*
357 *Ecology* **17**, 4186-4196.
358 Lesica, P. & Allendorf, F.W. (1999) Ecological genetics and the restoration of plant
359 communities: mix or match? *Restoration Ecology* **7**, 42-50.
360 Liu, C., Berry, P. M., Dawson, T. P. & Pearson, R. G. (2005) Selecting thresholds of
361 occurrence in the prediction of species distributions. *Ecography*, **28**, 385-393.
362 Mace, G.M. & Purvis, A. (2008) Evolutionary biology and practical conservation: bridging a
363 widening gap. *Molecular Ecology* **17**, 9-19.
364 Olesen, J.M. & Warncke, E. (1989) Flowering and seasonal changes in flower sex ratio and
365 frequency of flower visitors in a population of *Saxifraga hirculus. Holarctic Ecology* **12**,
366 21-30.

Olesen, J.M. & Warncke, E. (1990) Morphological, phenological and biochemical differentiation in relation to gene flow in a population of *Saxifraga hirculus*. *Sommerfeltia* **11**, 159-171.

Oliver, C., Hollingsworth, P.M. & Gornall, R.J. (2006) Chloroplast DNA phylogeography of the arctic-montane species *Saxifraga hirculus* (Saxifragaceae). *Heredity* **96**, 222-231.

Phillips, S. J., Anderson, R. P. & Schapire, R. E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231-259.

Phillips, S. J. & Dudík, M. (2008) Modelling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161-175.

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

Provan, J. & Wilson, P.J. (2007) Development of microsatellites for the peat moss *Sphagnum capillifolium* using ISSR cloning. *Molecular Ecology Notes* **7**, 254-256.

Provan, J. & Maggs, C.A. (2012) Unique genetic variation at a species' rear edge is under threat from global climate change. *Proceedings of the Royal Society of London Series B* **279**, 39-47.

Schwartz, M.K., Luikart, G. & Waples, R.S. (2007) Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology and Evolution* **22**, 25-33.

Vittoz, P., Wyss, T. & Gobat, J-M. (2006) Ecological conditions for *Saxifraga hirculus* in Central Europe: a better understanding for a good protection. *Biological Conservation* **131**, 594-608.

Wandeler, P., Hoeck, P.A.E. & Keller, L.F. (2007) Back to the future: museum specimens in population genetics. *Trends in Ecology and Evolution* **22**, 634-642.

390 Warncke, E., Terndrup, U., Michelsen, V. & Erhardt, A. (1993) Flower visitors to *Saxifraga*
391 *hirculus* in Switzerland and Denmark, a comparative study. *Botanica Helvetica* **103**, 141-
392 147.
393 Welch, D. (2002) The establishment of recovery sites for *Saxifraga hirculus* L. in NE
394 Scotland. *Botanical Journal of Scotland* **54**, 75-88.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Table S1 Herbarium codes of samples used.

Table S2 Description of variables used to describe landscape suitability for the Yellow marsh saxifrage.

Figure S1 Jackknife analyses of the importance of environmental variables in maximum entropy modelling of yellow marsh saxifrage distribution.

Figure S2 Marginal response curves of the predicted probability of yellow marsh saxifrage occurrence for each explanatory variable that contributed to 95% of the cumulative variance.

BIOSKETCHES

Gemma Beatty is a Postdoctoral Research Fellow at Queen's University Belfast. Her PhD research compared how postglacial recolonization and range-edge effects have shaped the genetic diversity of several Monotropoideae species. She is interested in using genetic approaches to study the effects of past and present climate change on the distribution ranges of natural populations, and the various factors that determine these ranges.

Neil Reid is Manager of *Quercus*, Northern Ireland's Centre for Biodiversity and Conservation Biology. Has has a background in species distribution modelling as a tool for identifying high conservation value areas for endangered species.

Jim Provan is a Reader in Evolutionary Genetics at Queen's University Belfast. His research interests focus on how genetic variation is distributed across species ranges, and on the effects of past, present and future climate change on levels and patterns of intraspecific diversity.

Author contributions: N.R. and J.P. conceived the study; G.E.B. and J.P. collected the data; All three authors carried out the analyses and wrote the manuscript.

Table 1 Details of populations analysed in the present study. *N* – number of individuals studied; *A* –number of alleles; H_O – observed heterozygosity; *H* – phenotype diversity (Gini-Simpson index).

County	Population	Year	Code	Lat	Long	*N*	*A* [a]	H_O [a]	*H* [a]
Co. Antrim	Garron Plateau	1886	GAR-1886	54.990	-6.096	5	18	0.600	0.768
		1889	GAR-1889			2	N/A	N/A	N/A
		1914	GAR-1914			3	N/A	N/A	N/A
		1920	GAR-1920			2	N/A	N/A	N/A
		1922	GAR-1922			2	N/A	N/A	N/A
		1955	GAR-1955			8	23	0.417	0.649
		1957	GAR-1957			2	N/A	N/A	N/A
		2011	GAR-2011			13	19	0.459	0.445
	Rasharkin [b]	1837	RAS-1837	54.9	-6.4	10	27	0.533	0.735
		1853	RAS-1853			2	N/A	N/A	N/A
		1857	RAS-1857			6	23	0.556	0.767
		1873	RAS-1873			5	19	0.500	0.750
		1884	RAS-1884			6	20	0.389	0.711
Co. Derry	Coleraine [b]	1800s	COL-18XX	55.1	-6.7	7	20	0.572	0.698
Co. Mayo	Largan Mor	2011	LMA-2011	54.140	-9.694	24	28	0.743	0.581
			LMB-2011	54.154	-9.686	24	27	0.750	0.643
	Sheean	2011	SHA-2011	54.118	-9.653	24	26	0.660	0.630
			SHB-2011	54.119	-9.652	24	30	0.882	0.790
			SHC-2011	54.117	-9.656	24	27	0.812	0.699
			SHD-2011	54.119	-9.651	18	25	0.759	0.770
	Uggoll	2011	UGG-2011	54.108	-9.644	24	25	0.778	0.512
	Barroosky	2011	BAR-2011	54.195	-9.631	24	30	0.785	0.808
	Sheskin	2011	SKA-2011	54.201	-9.562	24	23	0.549	0.673
			SKB-2011	54.198	-9.557	24	23	0.500	0.619
	Croaghaun	2011	CRO-2011	54.182	-9.469	24	26	0.708	0.664
	Formoyle	2011	FOR-2011	54.141	-9.448	24	22	0.840	0.560
	Aghoo	2011	AGH-2011	54.257	-9.408	24	32	0.729	0.764
	Bellacorick	1857	BEL-1857	54.2	-9.5	11	25	0.576	0.703
		1858	BEL-1858			13	27	0.615	0.712
		1965	BEL-1965			3	N/A	N/A	N/A
		1968	BEL-1968			6	19	0.750	0.821
		1970	BEL-1970			3	N/A	N/A	N/A
Co. Westmeath	Lisclogher [b]	1880	LIS-1880	53.6	-7.1	3	N/A	N/A	N/A
		1888	LIS-1888			4	N/A	N/A	N/A

[a] Number of alleles, observed heterozygosity and phenotype diversity only calculated for samples with $N \geq 5$

[b] Extinct population: Latitude / Longitude approximate

Table 2 *S. hirculus* nuclear microsatellite primers used in this study. A – number of alleles.

Locus	Repeat	Primers	A	Allele size range
SH-1-B08	$(AGC)_5$	CCCGCCATTTCTCTATACCA	7	119-137
		GGTTGAGCCAGTCCAAGAAG		
SH-2-D03	$(CTA)_5$	GCTTTTCCATTTTTAGGGCTTT	10	139-169
		AAAAGGAAAGTGAGATACTAATTAGAACAG		
SH-3-A03	$(AT)_6$	TCAAAATATTATTAAGGGAAAAATTCTCA	8	156-188
		CCAAATGTTTGAGTTATGTATAGTTACG		
SH-3-B11	$(TCTT)_7$	TGGCTACTACAATGTAAAGTTGTCTC	8	132-160
		CATAAGTCAAAAGTCAAGGTGTCG		
SH-4-E03	$(AAAT)_4$	TGTCTGTTTGGACATTCCCTTA	11	136-208
		TCAATATATTCTTAAGTTGATTATTAAGTGTG		
SH-4-F10	$(TA)_6$	GGATCCCTCACTTGAAGCTC	16	122-160
		TGTATAGATCAACTCTGCCAAAAA		

Forward tailed with CACGACGTTGTAAAACGAC

Reverse tailed with GTGTCTT

Table 3 Mean levels of genetic diversity calculated over all individuals in the historical vs. extant samples from the Garron Plateau population. A – mean number of alleles; H_O – mean observed heterozygosity; H – mean phenotype diversity (Gini-Simpson index). Significance of differences in mean values was estimated using a t-test.

Period	Diversity		
	A	*Ho*	*H*
Historical (pre-1958)	1.732	0.507	0.710
Extant	1.722	0.459	0.445
	NS	NS	$P = 0.013$

Figure Legends

Figure 1 Map showing locations of the populations analysed in the present study. Population codes correspond to those in Table 1. Codes in italics represent extinct populations. Locations of the RAS, COL, LIS and BEL populations are approximate.

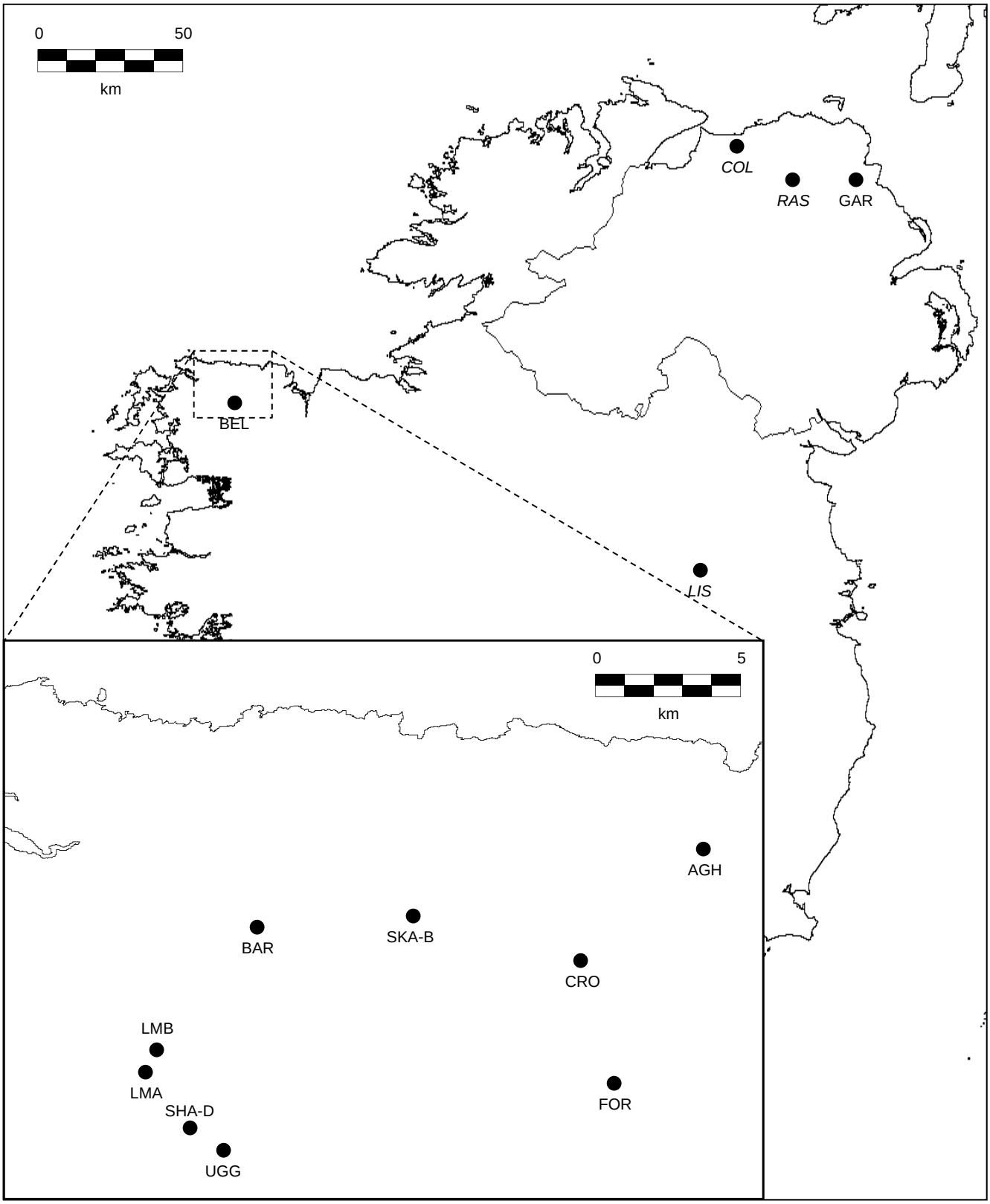
Figure 2 Graph of ΔK values indicating the most likely number(s) of genetic clusters (After Evanno *et al.*, 2005).

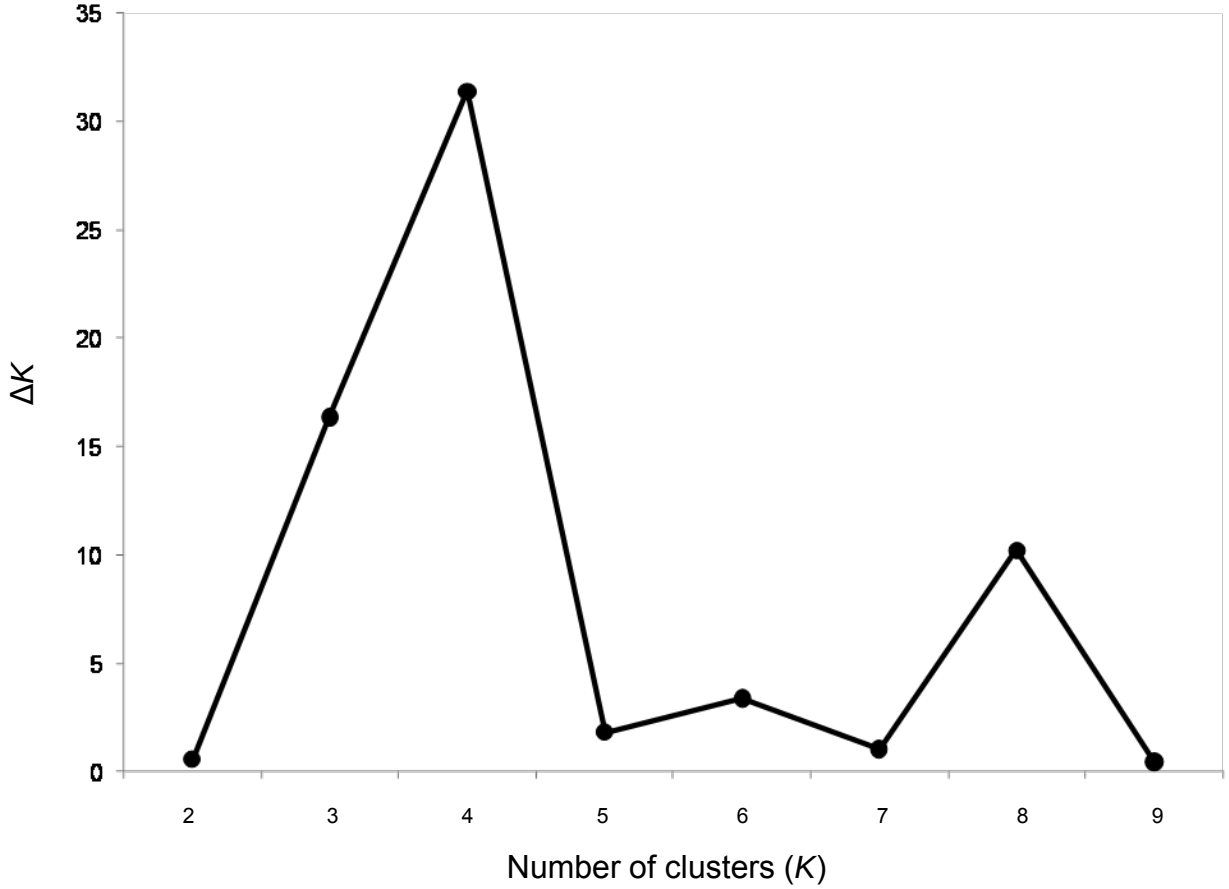
Figure 3 Results of the Bayesian clustering analysis performed using STRUCTURE (V2.3.3). Each column represents an individual, with the height of each coloured segment indicating the probability of membership to each of $K = 4$ (top) or $K = 8$ (middle) genetic clusters. Bottom shows enlarged assignment of Northern Ireland individuals for $K = 8$.

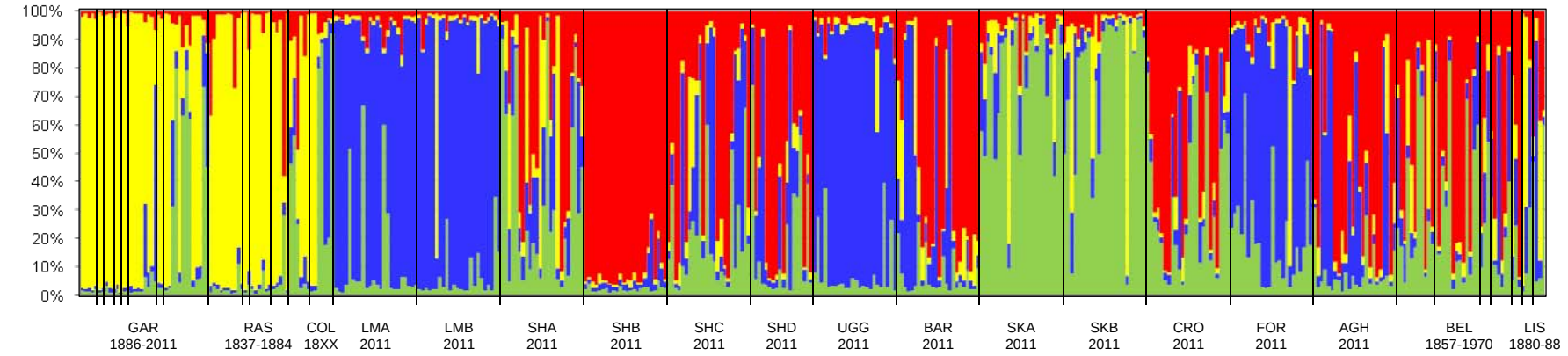
Figure 4 Landscape suitability for the *S. hirculus* throughout Ireland showing areas of high quality (red) in Co. Mayo, and areas of lower quality (turquoise) in Northern Ireland, with areas in dark blue being totally unsuitable.

Figure S1 Jackknife analyses of the importance of environmental variables in maximum entropy modelling of yellow marsh saxifrage distribution. A heuristic estimate of the relative contribution of each variable to the global model is given in parentheses whilst variables are listed in descending order of importance. Grey bars show the performance of the global model (known as % gain) without each variable and black bars show the influence of each variable in isolation (derived from a univariate model only). Percentage contributions that sum to 95% of variance are shown in bold.

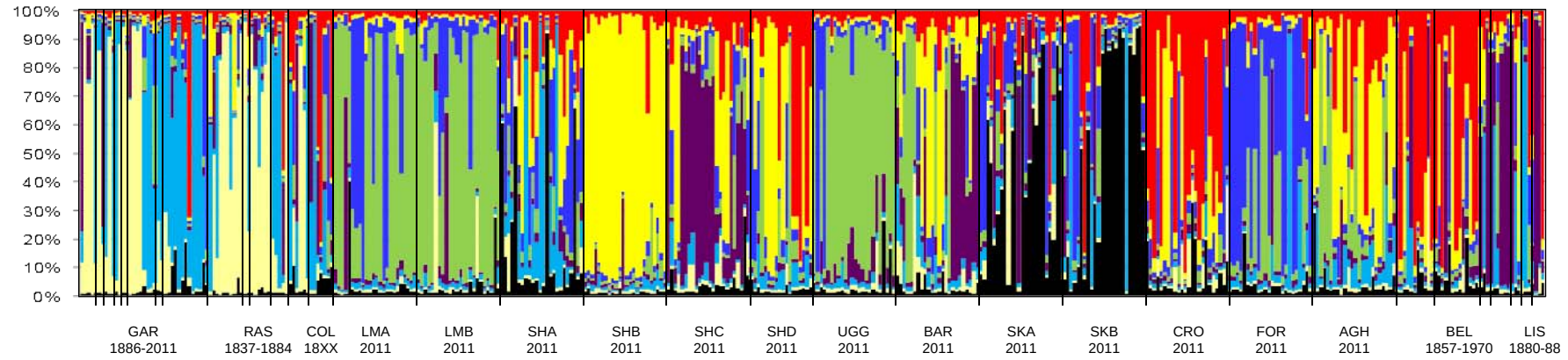
Figure S2 Marginal response curves of the predicted probability of yellow marsh saxifrage occurrence for each explanatory variable that contributed to 95% of the cumulative variance. Curves show logistic predictions when all other environmental variables where maintained at their mean value.


0
50
km
COL
RAS
GAR
BEL
LIS
0
5
km
AGH
BAR
SKA-B
CRO
LMB
LMA
SHA-D
UGG
FOR

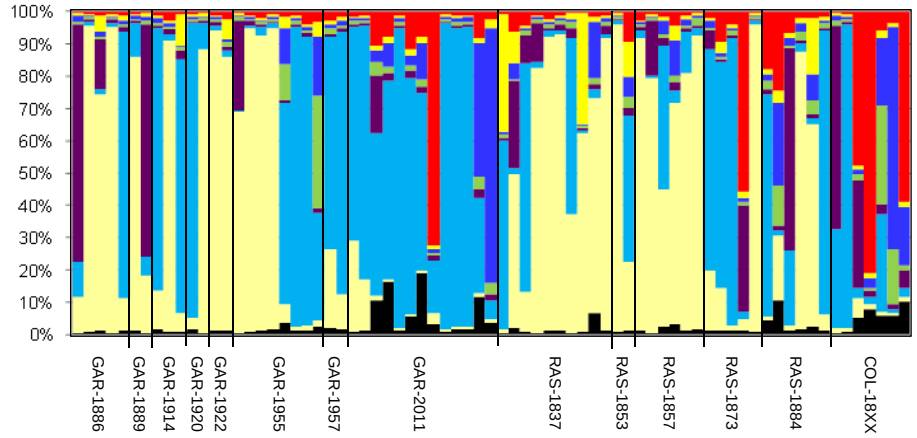

35
30
25
20
15
10
5
0
ΔK
2
3
4
5
6
7
8
9
Number of clusters (K)


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
GAR 1886-2011
RAS 1837-1884
COL 18XX
LMA 2011
LMB 2011
SHA 2011
SHB 2011
SHC 2011
SHD 2011
UGG 2011
BAR 2011
SKA 2011
SKB 2011
CRO 2011
FOR 2011
AGH 2011
BEL 1857-1970
LIS 1880-88

K = 4


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
GAR 1886-2011
RAS 1837-1884
COL 18XX
LMA 2011
LMB 2011
SHA 2011
SHB 2011
SHC 2011
SHD 2011
UGG 2011
BAR 2011
SKA 2011
SKB 2011
CRO 2011
FOR 2011
AGH 2011
BEL 1857-1970
LIS 1880-88

K = 8


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
GAR-1886
GAR-1889
GAR-1914
GAR-1920
GAR-1922
GAR-1955
GAR-1957
GAR-2011
RAS-1837
RAS-1853
RAS-1857
RAS-1873
RAS-1884
COL-18XX

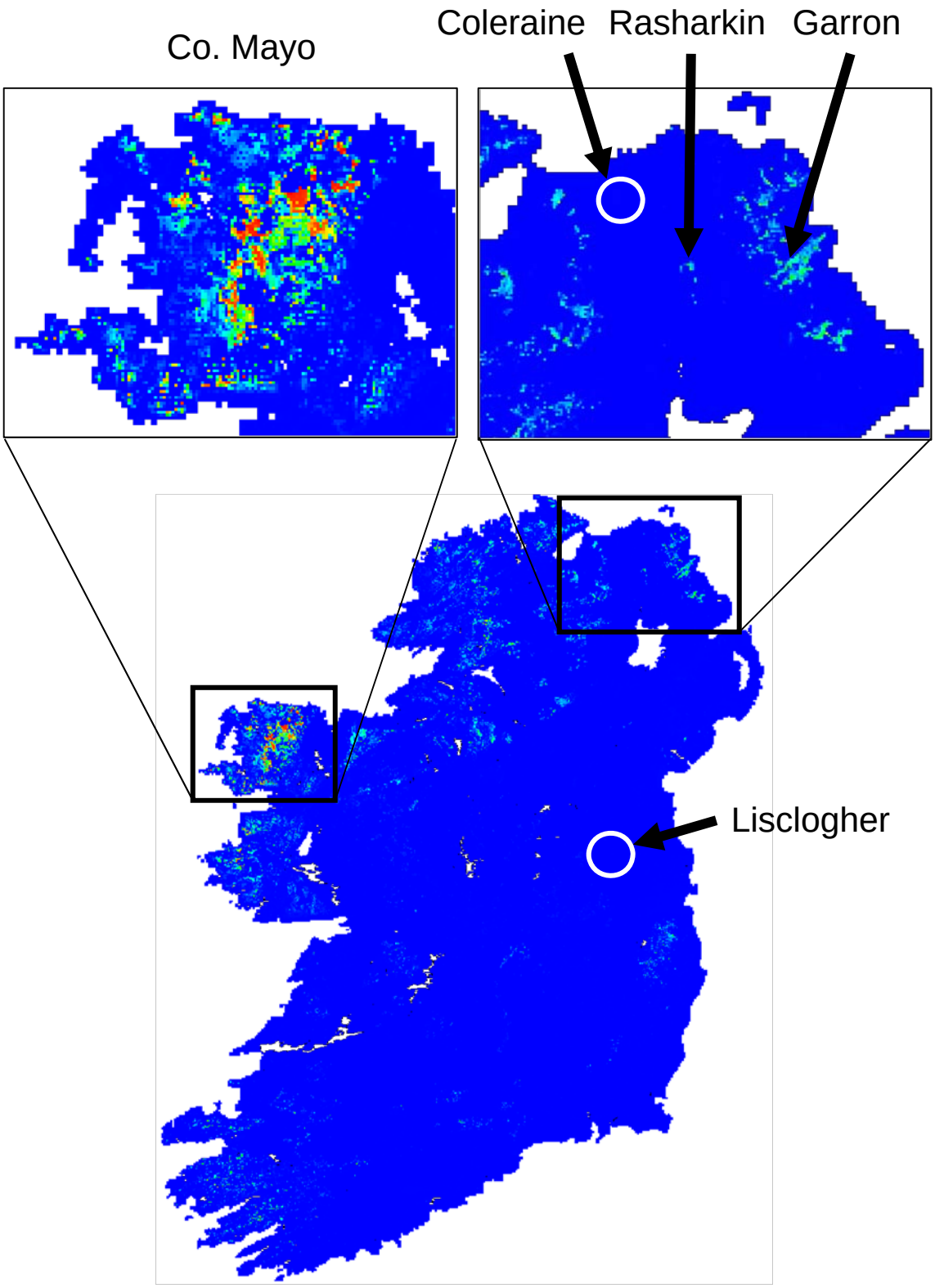
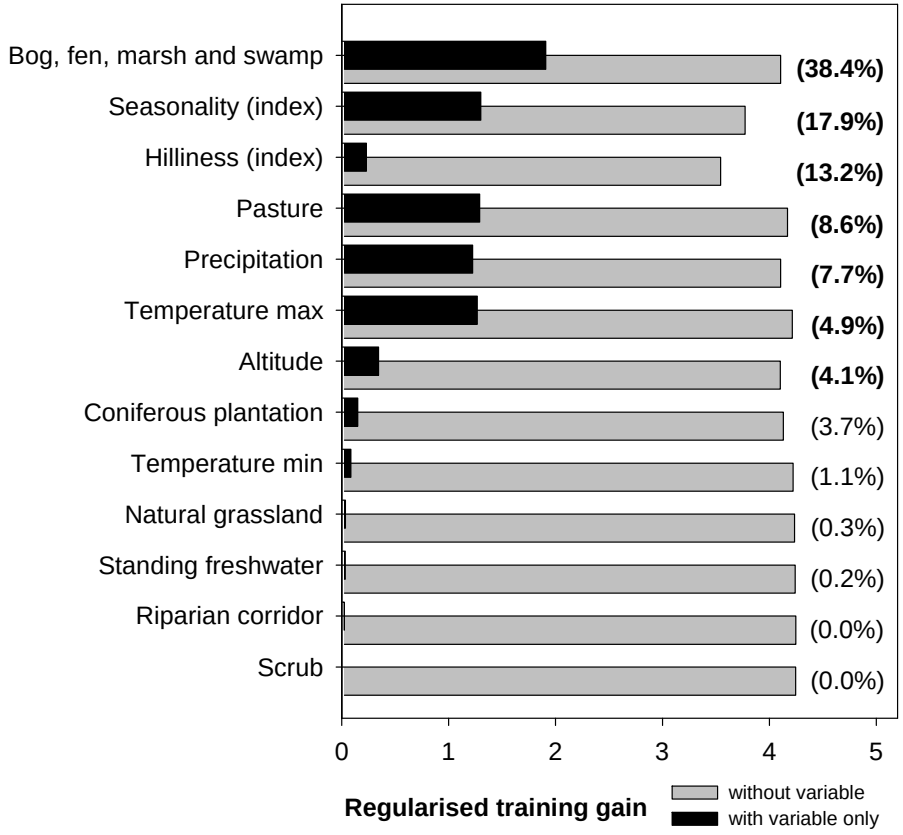

Co. Mayo
Coleraine
Rasharkin
Garron
Lisclogher

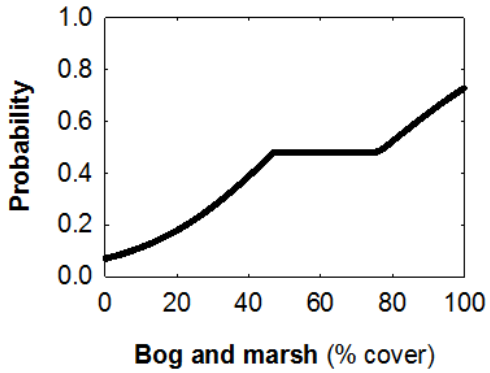
Table S1 Herbarium codes of samples used

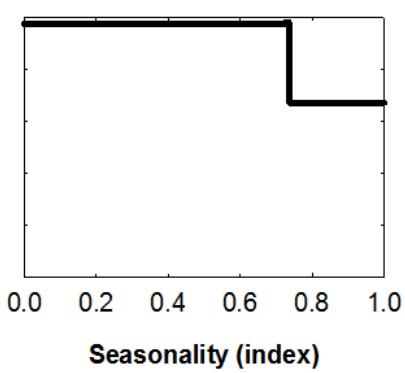
Code	Herbarium	*N*
GAR-1886	BEL-H61348	1
	BEL-H61349	2
	DBN-17-07-1886	2
GAR-1889	DBN-22-1967	2
GAR-1914	DBN-	3
GAR-1920	BEL-H61347	2
GAR-1922	DBN-	2
GAR-1955	BEL-H61345	4
	DBN-4-1958	2
	DBN-32-1980	2
GAR-1957	BEL-H61346	2
RAS-1837	DBN-DM-	10
RAS-1853	DBN-003989	2
RAS-1857	DBN-94-Sh-1857	6
RAS-1873	DBN-Sh-97	5
RAS-1884	BEL-H917	6
BEL-1957	DBN-04001	5
	DBN-3-10-1957	6
BEL-1958	DBN-004002	9
	DBN-003999	4
BEL-1965	DBN-004003	3
BEL-1968	DBN-15-8-68	6
BEL-1970	DBN-003991	3
LIS-1880	DBN-631	3
LIS-1888	DBN-Sh-1888	4

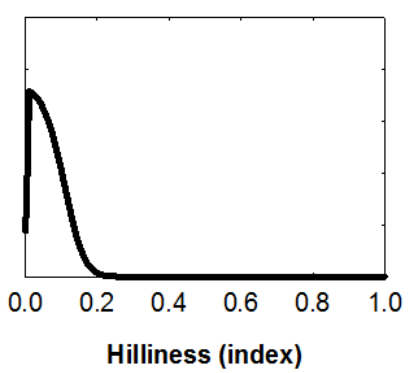
Table S2 Description of variables used to describe landscape suitability for the Yellow marsh saxifrage

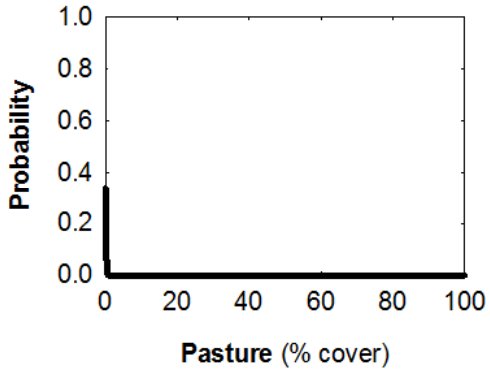
Name	Units	Description
Topography		
Altitude	m	Elevation above sea level in metres
Hilliness	m	Standard deviation in mean elevation above sea level in metres per 500m cell
Habitat composition		
Bog, fen, marsh & swamp	% cover	Coverage representing a composite of Bog, fen, marsh & swamp derived from CORINE 2000
Pasture	% cover	Coverage of pasture derived from CORINE 2000
Coniferous plantations	% cover	Coverage of coniferous woodland derived from CORINE 2000
Natural grass	% cover	Coverage of natural grass derived from CORINE 2000
Scrub	% cover	Coverage of scrub derived from CORINE 2000
Riparian corridor	Km	Total length of river and water body edge including lakes, reservoirs, ponds, rivers, streams and canals in metres
Standing freshwater	% cover	Coverage of lakeland derived from CORINE 2000
Climate		
$Temp_{min}$	^{o}C	Minimum temperature of the coldest month
$Temp_{max}$	^{o}C	Maximum temperature of the warmest month
$Precipitation_{annual}$	mm	Total annual precipitation
Seasonality	Index	Standard deviation of mean monthly temperatures *100

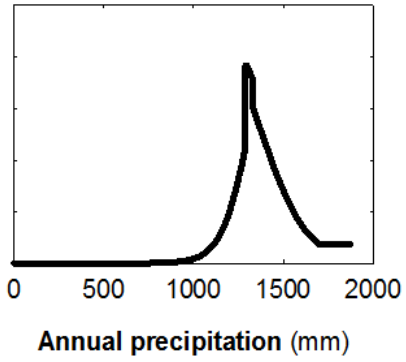

Bog, fen, marsh and swamp
Seasonality (index)
Hilliness (index)
Pasture
Precipitation
Temperature max
Altitude
Coniferous plantation
Temperature min
Natural grassland
Standing freshwater
Riparian corridor
Scrub
(38.4%)
(17.9%)
(13.2%)
(8.6%)
(7.7%)
(4.9%)
(4.1%)
(3.7%)
(1.1%)
(0.3%)
(0.2%)
(0.0%)
(0.0%)
0
1
2
3
4
5
Regularised training gain
without variable
with variable only

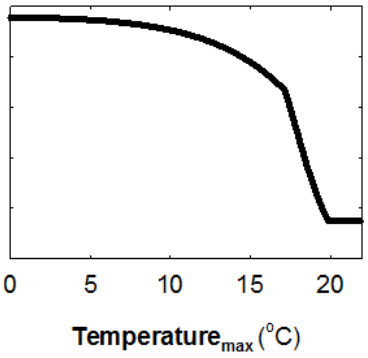

Probability
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
Bog and marsh (% cover)

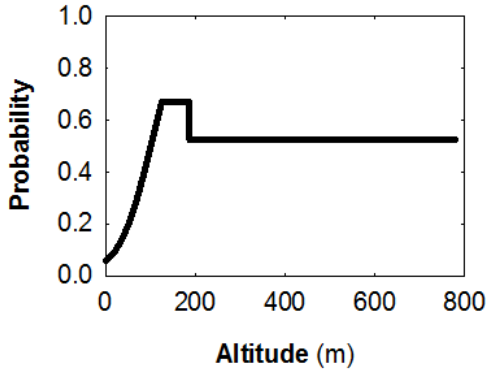

0.0
0.2
0.4
0.6
0.8
1.0
Seasonality (index)


0.0
0.2
0.4
0.6
0.8
1.0
Hilliness (index)


Probability
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
Pasture (% cover)


0
500
1000
1500
2000
Annual precipitation (mm)


0
5
10
15
20
Temperature max (°C)


Probability
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
Altitude (m)