Genetic evidence for predominantly hydrochoric gene flow in the invasive riparian plant *Impatiens glandulifera* (Himalayan balsam)

Heather M. Love[1], Christine A. Maggs[1], Tomás E. Murray[1,†] and Jim Provan[1,2,*]

[1]*School of Biological Sciences and* [2]*Institute for Global Food Security, The Queen's University of Belfast, 97 Lisburn Road, Belfast BT9 7BL, UK*

[†]*Present address: Institute for Biology, Martin-Luther-Universität, Halle-Wittenberg, Hoher Weg 8, D-06099, Halle (Saale), Germany.*

* For correspondence. E-mail J.Provan@qub.ac.uk

• *Background and Aims* Riparian systems are prone to invasion by alien plant species. The spread of invasive riparian plants may be facilitated by hydrochory, the transport of seeds by water, but while ecological studies have highlighted the possible role of upstream source populations in the establishment and persistence of stands of invasive riparian plant species, population genetic studies have as yet not fully addressed the potential role of hydrochoric dispersal in such systems.

• *Methods* A population genetics approach based on a replicated bifurcate sampling design is used to test hypotheses consistent with patterns of unidirectional, linear gene flow expected under hydrochoric dispersal of the invasive riparian plant *Impatiens glandulifera* in two contrasting river systems.

• *Key results* A significant increase in levels of genetic diversity downstream was observed, consistent with the accumulation of propagules from upstream source populations, and strong evidence was found for organization of this diversity between different tributaries, reflecting the dendritic organization of the river systems studied.

• *Conclusions* These findings indicate that hydrochory, rather than anthropogenic dispersal, is primarily responsible for the spread of *I. glandulifera* in these river systems, and this is relevant to potential approaches to the control of invasive riparian plant species.

Key words: Dispersal, gene flow, Himalayan balsam, hydrochory, *Impatiens glandulifera*, invasive species, propagule pressure, riparian.

INTRODUCTION

The introduction of non-native species into new environments represents one of the major threats to ecosystems and biodiversity worldwide and is now ranked second only to habitat loss as a potential cause of ecological catastrophe (Gurevitch and Padilla, 2004). Indeed, a review of recent animal extinctions found that invasive species were responsible for >50 % of cases in which the cause of extinction was known (Clavero and García-Berthou, 2005). Plant invasions can also be extremely detrimental and are usually associated with deliberate anthropogenic introductions, including exotic ornamentals and feral crop species (Pimentel *et al.*, 2001; Weber, 2003). Problems associated with the introduction of exotic plant species into new habitats include hybridization with native species, impacts on species–species interactions and alteration of soil chemistry (Sakai *et al.*, 2001; Tickner *et al.*, 2001; Henderson *et al.*, 2006; Pearson, 2009), but the key detrimental effect is usually the rapid formation of largely monospecific stands that reduce biodiversity in impacted habitats (Cronk and Fuller, 2001; van der Wal *et al.*, 2008; Hejda *et al.*, 2009). Invasive plant species can often outcompete native species for a wide range of resources, including light, nutrients and pollinators, and such 'biotic homogenization' has been implicated as a potential causal factor in future mass extinction events (McKinney and Lockwood, 1999).

Riparian systems are of particular ecological interest as they encompass both terrestrial and freshwater habitats where biodiversity is maximal and important ecological, hydrological and geomorphological processes occur (Tickner *et al.*, 2001). They are especially prone to invasion by exotic plant species as their dynamic physical nature regularly provides suitable habitats for colonizing species due to the periodic deposition of nutrient-rich sediments and the creation of large areas with sufficient space, light and moisture for the successful germination of seedlings (Hupp and Osterkamp, 1996; Myers and Bazely, 2003). In addition to the impacts on native biodiversity mentioned above, invasive plants can also alter the geomorphological and hydrological aspects of riparian ecosystems (Tickner *et al.*, 2001). The establishment of dense populations of non-native plants can significantly lower water-table levels with serious consequences not only for the existing flora but also for entire mesocosms (Loope *et al.*, 1988; DiTomaso, 1998). Newly established populations of invasive plant species may also interfere with sediment distribution, leading to aggradation, which in turn can radically alter the geomorphology of river systems and may increase the risk of periodic flooding (Graf, 1982; Birkeland, 1996).

Propagule pressure (in terms of seeds, fruits or vegetative fragments) has been shown to be a major factor determining the successful establishment and spread of invasive species (Lockwood *et al.*, 2005; Colautti *et al.*, 2006; Martínez-Ghersa and Ghersa, 2006; Eschtruth and Battles, 2009; Simberloff, 2009). This is particularly important in riparian systems, as propagules may be transported longitudinally downstream via hydrochory, the dispersal of seeds via water, potentially facilitating the spread of invasive riparian plants (Burkart, 2001; Nakayama *et al.*, 2007; Tassin *et al.*, 2007). Thus, the unidirectional, linear connectivity between potentially suitable habitats may compound the potential for invasion due to increased propagule pressure. Ecological studies have highlighted the possible role of upstream source populations in the establishment and persistence of stands of invasive plant species in riparian habitats (e.g. Dawson and Holland, 1999; Samuel and Kowarik, 2010). The occurrence of such patterns of dispersal has implications for control strategies: when downstream dispersal is a factor, eradication programmes may be most efficient when they specifically target upstream populations, thus reducing propagule pressure.

This study aims to test the hypothesis that hydrochory, or water-mediated seed dispersal, is a key factor in the spread of the invasive riparian plant species *Impatiens glandulifera* (Himalayan balsam). The species was introduced into Britain and Ireland as an ornamental in the 19th century and has subsequently become a naturalized invader. It is an annual herb which can outcompete even hardy, well-established, native perennial species in riparian habitats (Beerling and Perrins, 1993). Flowers are self-compatible and protandrous, and are primarily pollinated by Diptera and Hymenoptera (Lopezaraiza-Mikel *et al.*, 2007). Individual plants produce around 500–5000 seeds and short-distance dispersal (<5 m) occurs through explosive dehiscence, whereas long-distance dispersal is believed to occur either anthropogenically or via hydrochory. To date, only a single population genetic study of *I. glandulifera* has been published, and this did not explicitly test hypotheses on hydrochoric dispersal (Walker *et al.*, 2009). The two main hypotheses concerning the role of hydrochory to be tested here are based on the patterns of gene flow expected under a model of linear, unidirectional gene flow. First, there is expected to be an increase in levels of genetic diversity downstream due to the accumulation of propagules from upstream source populations. Secondly, if hydrochory, rather than overland transport (anthropogenic or by animals), is the major factor in the dispersal of seeds, populations on different tributaries not linked by water flow will be more genetically differentiated than populations from different areas of the river which are linked by the flow of the river (i.e. above and below a confluence in a bifurcate or dendritic network), which in turn will be more genetically differentiated than populations from the same tributary. The first hypothesis has been tested previously in several studies (e.g. Ritland, 1989; Gornall *et al.*, 1998; Russell *et al.*, 1999; Lundqvist and Andersson, 2001; Tero *et al.*, 2003; Liu *et al.*, 2006; Markwith and Scanlon, 2007; Pollux *et al.*, 2009; Honnay *et al.*, 2010; Schleuning *et al.*, 2011), but no overall consensus has emerged. The second hypothesis, however, has not been explicitly tested, as studies have not used a bifurcate experimental design, and it has been suggested that the dendritic organization of river systems should be taken into account to gain further insights into the role of hydrochory in the dispersal of riparian plants (Pollux *et al.*, 2009).

MATERIALS AND METHODS

Experimental design, sampling and DNA extraction

To test the predictions of the hydrochory hypothesis, it was necessary to sample *Impatiens grandulifera* over entire river drainage basins (i.e. with river lengths of at least 25 km and with a bifurcate organization). Replication involved studying populations in two contrasting river basins in Britain and Ireland (Fig. 1). The Western Cleddau in Pembrokeshire, Wales, was chosen based on data from the Joint Nature Conservation Committee (www.jncc.gov.uk). The system has been designated as a Special Area of Conservation (SAC), and is characterized by *Callitricho–Batrachion* Type 4 (CB4) vegetation communities, where *I. glandulifera* is a frequent alien (Hatton-Ellis and Grieve, 2003). The second river system selected was the Tempo/ Colebrooke river system in Co. Fermanagh, Northern Ireland, and was chosen based on river vegetation survey records from the Quercus Centre for Biodiversity and Conservation Biology (www.quercus.ac.uk). The system is part of the Upper Lough Erne river basin, a predominantly limestone catchment surrounded by large, semi-natural alluvial forests. The exact age of the study populations is unknown, but the species was not recorded in either area prior to 1940 (Beerling and Perrins, 1993). Species occurrence data for both river systems was obtained from the National Biodiversity Network Gateway (www.nbn.org.uk) and surveys and sampling of both systems were carried out during the summer of 2006. Nine and ten sites representing populations from two upstream tributaries and populations from below the point of the confluence of those tributaries were sampled from the Western Cleddau and Tempo/ Colebrooke systems, respectively (Table 1). Geographical distances between sites were calculated using the ArcView GIS system. Between 15 and 35 individuals per site were sampled and DNA was extracted from individual leaves using the Qiagen DNeasy Plant Mini system.

Microsatellite genotyping

Individuals were genotyped at ten microsatellite loci, representing all the published markers available for *I. glandulifera* (Provan *et al.*, 2007; Walker *et al.*, 2009) with the exception of locus A2 from Walker *et al.* (2009), which regularly failed to amplify in several populations, probably due to the presence of null alleles. Loci IGNSSR103, IGNSSR105, IGNSSR106, IGNSSR 210 and IGNSSR240 were genotyped following the procedures described by Provan *et al.* (2007). For the remaining five loci, forward primers were modified by the addition of a 19-bp M13 tail (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 PET-, 6-FAM- or HEX-labelled M13 primer, 1 pmol M13-tailed forward primer, 10 pmol reverse primer, $1\times$ PCR reaction buffer, 200 μM each dNTP, 2·5 mM $MgCl_2$ and 0·25 U GoTaq Flexi DNA polymerase (Promega, Sunnyvale, CA). PCR cycling conditions followed those given by Provan *et al.* (2007) and Walker *et al.* (2009). 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.
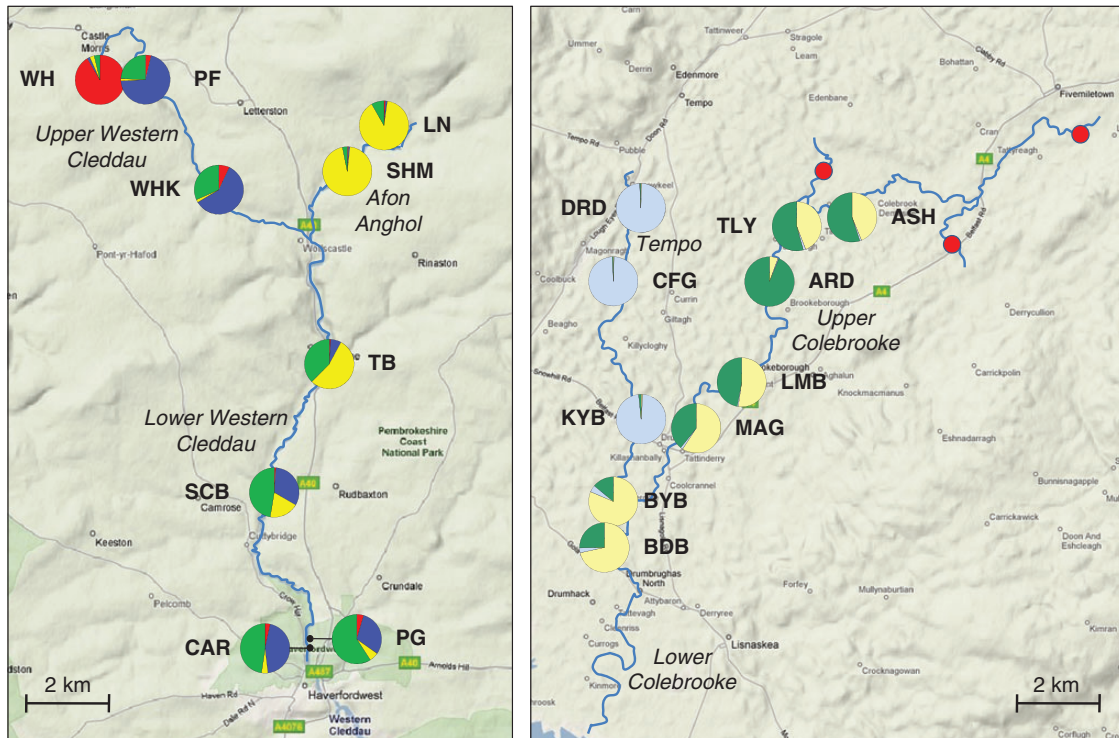
FIG. 1. Maps showing locations of populations sampled in this study. Left, Western Cleddau; right, Tempo/Colebrooke. Population codes are given in Table 1. Pie charts indicate probability of assignment of individuals to one of $K = 4$ (Western Cleddau) or $K = 3$ (Tempo/Colebrooke) genetic clusters based on the STRUCTURE analysis (see Materials and methods). Red circles in the Tempo/Colebrooke system indicate areas surveyed but where no *I. glandulifera* was found. For clarity, only tributaries sampled or surveyed are shown.

Data analysis

GENEPOP (V3·4; Raymond and Rousset, 1995) was used to test for linkage disequilibrium between nuclear loci. Levels of genetic diversity measured as allelic richness (A_R) and observed and expected heterozygosity (H_O and H_E) were calculated using the FSTAT software package (V2·9.3·2; Goudet, 2001) and the ARLEQUIN software package (V3·01; Excoffier *et al.*, 2005), respectively. Inbreeding coefficients (F_{IS}) were also calculated using the FSTAT software package. To test the hypothesis that genetic diversity increases downstream, a linear regression analysis was carried out for both allelic richness and expected heterozygosity against the distance downstream from the furthest upstream population sampled in each system using the MedCalc software package (V12·2.1; www.medcalc.org, Ostend, Belgium). For both river systems, tributaries plus all sites downstream of their confluence with the main river were considered separately in each case (Upper Western Cleddau/Lower Western Cleddau; Afon Anghol/Lower Western Cleddau; Tempo/Lower Colebrooke; Upper Colebrooke/Lower Colebrooke).

Levels of population differentiation were calculated in the analysis of molecular variance (AMOVA) framework (Excoffier *et al.*, 1992) as pairwise Φ_{ST} values using the ARLEQUIN software package. To test the hypothesis of patterns of differentiation consistent with linear, unidirectional gene flow, populations on different tributaries unlinked by water flow were classified as 'unlinked', whereas those from different areas of the river which are linked by the flow of the river (i.e. above and below

a confluence) were defined as 'linked'. Populations from the same tributary were classed as 'same'. The significance of differences between values for 'unlinked', 'linked' and 'same' groups of populations was determined using the Student–Newman–Keuls *post hoc* test implemented in the MedCalc software package. Genetic structuring was also assessed at the individual level using a Bayesian procedure implemented in the STRUCTURE software package (V2·3.3; Pritchard *et al.*, 2000). The program was run using no prior knowledge and the admixture ancestry model. For each river system, five independent runs were carried out for each value of K, the number of genetic clusters, up to $K = 9$ (Western Cleddau) or $K = 10$ (Tempo/Colebrooke). Each Markov chain Monte Carlo analysis used a burn-in period of 10 000 followed by a further 100 000 iterations. The most likely value for K was estimated using the ΔK statistic of Evanno *et al.* (2005) implemented in the STRUCTURE HARVESTER software package (V0·6.1; Earl *et al.*, 2012).

RESULTS

Two to eight alleles were detected across all populations studied at each of the ten microsatellite loci analysed and no evidence of linkage disequilibrium was detected. All loci were variable in each river system. Mean within-population values for allelic richness (A_R) averaged over all loci ranged from 1·358 (SHM population) to 2·027 (PG population) in the Western Cleddau system, and from 1·500 (CFG population) to 2·525 (BDB

TABLE 1. *Locations of sampled sites and sample numbers*

System	Tributary	Population	Grid ref.	Code	n	H_O	H_E	A_R	F_{IS}
Western Cleddau	Upper Western Cleddau	Withy Hill	SM904297	WH	26	0·366	0·408	1·851	−0·291
		Priskilly Forest	SM919299	PF	21	0·130	0·122	1·461	−0·062
		Welsh Hook	SM932278	WHK	20	0·140	0·164	1·636	0·149*
	Afon Anghol	Little Newcastle	SM979289	LN	19	0·202	0·149	1·547	−0·377
		Sealyham Bridge	SM967279	SHM	24	0·175	0·118	1·358	−0·500
	Lower Western Cleddau	Treffgarne Bridge	SM959230	TB	21	0·195	0·175	1·556	−0·116
		St. Catherine's Bridge	SM945198	SCB	18	0·152	0·206	1·578	0·267*
		Haverfordwest playground	SM953161	PG	17	0·194	0·221	2·027	0·127**
		Haverfordwest car park	SM953164	CAR	15	0·223	0·234	1·993	0·045
Tempo/Colebrooke	Tempo	Drumderg	H341446	DRD	30	0·262	0·217	1·572	−0·211
		Cornafannoge	H335426	CFG	30	0·205	0·194	1·500	−0·053
		Killynure Bridge	H339388	KYB	21	0·222	0·222	1·600	−0·001
		Ballyvelin Bridge	H334372	BYB	30	0·174	0·292	2·170	0·410***
	Upper Colebrooke	Ashbrooke	H392442	ASH	30	0·266	0·282	2·115	0·057**
		Tullyreagh Bridge	H379442	TLY	30	0·273	0·303	2·149	0·100***
		Ardunshin	H372426	ARD	35	0·231	0·231	1·885	0·003
		Littlemount Bridge	H370405	LMB	31	0·245	0·257	2·121	0·045
	Lower Colebrooke	Maguiresbridge	H350387	MAG	30	0·282	0·328	2·327	0·145***
		Ballindarragh Bridge	H331360	BDB	30	0·257	0·326	2·526	0·215***

n, number of individuals sampled; H_O, expected heterozygosity; H_E, expected heterozygosity; A_R, allelic richness; F_{IS}, inbreeding coefficient. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

population) in the Tempo/Colebrooke system. Mean levels of expected heterozygosity averaged over all loci ranged from 0·118 (SHM population) to 0·408 (WH population) in the Western Cleddau system, and from 0·194 (CFG population) to 0·328 (MAG population) in the Tempo/Colebrooke system. Values of F_{IS} ranged from −0·500 (SHM population) to 0·267 (SCB population) in the Western Cleddau system, and from −0·211 (DRD population) to 0·410 (BYB population) in the Tempo/Colebrooke system (Table 1). The regression analyses indicated a significant downstream increase in allelic richness in both systems and a significant downstream increase in expected heterozygosity in all courses except the Upper Colebrooke/Lower Colebrooke (Table 2).

Population-pairwise Φ_{ST} values (see Supplementary Data Table S1) in the Western Cleddau system ranged from zero (PF-WHK) to 0·560 (PF-SHM). In the Tempo/Colebrooke system, values ranged from zero (DRD-CFG and TLY-MAG) to 0·536 (DRD-ARD). In both systems, the Student–Newman–Keuls *post hoc* test indicated significant ($P < 0.05$) differences between population-pairwise Φ_{ST} values for 'unlinked', 'linked' and 'same' groups in the Western Cleddau system. In the Tempo/Colebrooke system, the mean population-pairwise Φ_{ST} between 'unlinked' tributaries was significantly higher ($P < 0.05$) than that for both the 'linked' and the 'same' groups. Although the mean population-pairwise Φ_{ST} between 'linked' tributaries was higher than that between populations on the same tributary, this difference was not significant (Fig. 2). This organization of genetic variation between tributaries was further highlighted by the STRUCTURE analysis, which identified $K = 4$ and $K = 3$ as the most likely numbers of genetic clusters for the Western Cleddau and Tempo/Colebrooke systems, respectively (Fig. 1). In the Western Cleddau, individuals from the WH population, a population isolated from the main river system for >20 years (see Discussion), were almost exclusively assigned to a single cluster, depicted in red, which was present at low frequency elsewhere in the system. The remaining two populations from the Upper Western Cleddau (PF and WHK) were associated with the two clusters indicated in blue and light green. Individuals from both populations on the Afon Anghol tributary (LN and SHM) were generally assigned to a fourth genetic cluster, depicted in yellow. Populations downstream of the confluence of the two tributaries were associated to a greater or lesser degree with combinations of the yellow, blue and light green clusters. In the Tempo/Colebrooke system,

TABLE 2. *Regression analyses of levels of genetic diversity vs. distance downstream from uppermost population*

Tributaries	A_R		H_E	
	R^2	P	R^2	P
Upper Western Cleddau/Lower Western Cleddau	0·657	0·025*	0·945	<0·001***
Afon Anghol/Lower Western Cleddau	0·739	0·014*	0·909	0·002**
Tempo/Lower Colebrooke	0·771	0·025*	0·772	0·025*
Upper Colebrooke/Lower Colebrooke	0·583	0·039*	0·252	n.s.

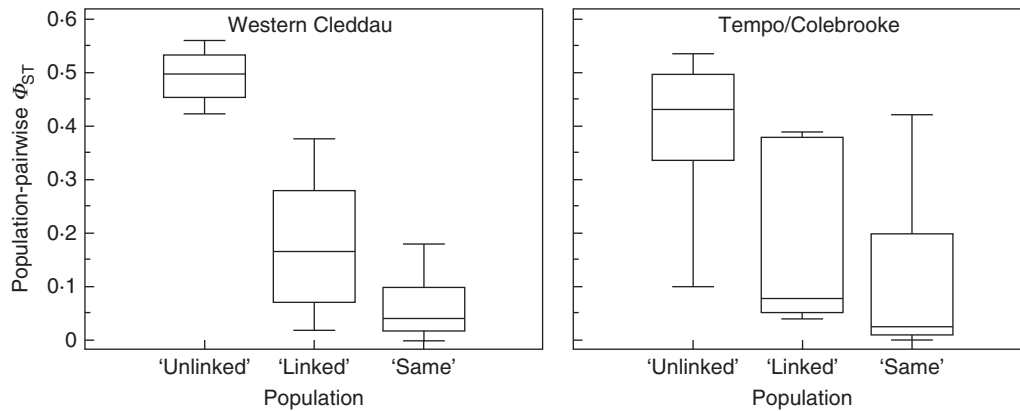A_R, allelic richness; H_E, expected heterozygosity. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

FIG. 2. Box-whisker plots of mean population pairwise Φ_{ST} values between 'unlinked', 'linked' and 'same' populations.

individuals from three of the four Tempo populations (DRD, CFG and KYB) were almost exclusively assigned to a single genetic cluster depicted in light blue, which was almost completely absent elsewhere in the system. Individuals from the remaining population on this tributary (BYB) and those from the populations on the Upper Colebrooke (ASH, TLY, ARD, LMB and MAG) and the sole population below the confluence (BDB) were primarily assigned to one of two clusters, depicted in light yellow and dark green.

DISCUSSION

Although our study revealed low levels of genetic variation, which is characteristic of the founder effects often associated with invasions, they confirm our hypotheses concerning increasing genetic variation downstream and structuring of genetic diversity between tributaries, which are indicative of hydrochoric seed dispersal. Water-mediated dispersal and the subsequent establishment of new individuals and populations of riparian plant species is the end result of a combination of factors including fertilization, seed production, transport and germination. Population genetic approaches can provide an overview of the integrated end result of all these processes and thus offer valuable insights into the spread, persistence and management of invasive riparian plants. The observed patterns of genetic variation within and between the populations analysed in the present study suggest that hydrochory, rather than anthropogenic dispersal, is primarily responsible for the spread of *I. glandulifera* in the Western Cleddau and Tempo/Colebrooke river systems. Hydrogeomorphology of rivers was also found to influence patterns of dispersal. The implications of these findings are relevant to concepts of dispersal in riparian plants in general, and to possible modes of spread and potential approaches to the control of invasive riparian plant species.

Evidence for hydrochoric dispersal

Previous molecular genetic studies of the potential role of hydrochoric dispersal in riparian plants have been somewhat equivocal, with some (Gornall *et al.*, 1998; Lundqvist and Andersson, 2001; Chen *et al.*, 2009; Pollux *et al.*, 2009; Schleuning *et al.*, 2011) reporting the expected increase in

genetic diversity downstream and others (Russell *et al.*, 1999; Markwith and Scanlon, 2007; Prentis and Mather, 2007; Honnay *et al.*, 2010) finding no clear evidence for hydrochory. Here, the significant downstream increase in genetic diversity coupled with the structuring of this diversity between tributaries indicates a major role for hydrochory in the dispersal of *I. glandulifera* seeds. Although the biparentally inherited nuclear markers used to estimate levels of dispersal would also reflect the effects of pollen flow, the effects of downstream seed movement are still evident. The seeds of *I. glandulifera* are negatively buoyant, but they may still be dispersed by currents of fast-flowing water (Beerling and Perrins, 1993). Furthermore, laboratory and field observations have shown that intact seed pods of *I. glandulifera* can float for several days and that 'sprung' pods sometimes retain seeds and thus might also provide a vector for downstream seed dispersal (H. M. Love *et al.*, unpubl. res.). It has been suggested that this dependence on water courses for long-distance dispersal has most likely limited *I. glandulifera* primarily to riparian habitats in its invasive range (Pyšek and Prach, 1993) and a logistic regression analysis at both national and regional levels in England found a correlation between the presence of the species and the abundance of rivers (Collingham *et al.*, 2000).

Further evidence for hydrochoric dispersal can be seen in the organization of genetic variation by tributaries. The pairwise Φ_{ST} values and the results of the STRUCTURE analysis show that gene flow tends to occur primarily between populations within tributaries and, to a lesser degree, between populations on linked tributaries, with little dispersal between unlinked tributaries. Although geographically proximal populations are expected to be genetically similar under a simple, unrestricted two-dimensional model of gene flow, the KYB (Tempo) and MAG (Upper Colebrooke) populations, which are <1 km apart but on two unlinked tributaries, exhibit a pairwise Φ_{ST} value of 0·393 ($P < 0.0001$), whereas the KYB and DRD populations on the same tributary (Tempo) are separated by 8 km but exhibit a pairwise Φ_{ST} value of only 0·002 (n.s.; $P = 0.315$). More random patterns of gene flow such as anthropogenic transport of seeds, i.e. not restricted to unidirectional dispersal between populations linked by watercourses, would be unlikely to give rise to such a pattern of structuring. The only previous study on the population genetics of *I. glandulifera* did not explicitly aim to

test the role of hydrochory in the dispersal of the species but, nevertheless, noted that the average population-pairwise F_{ST} within catchments in north-eastern England was significantly lower than the average value between catchments (Walker *et al.*, 2009). Pollination by bees is important for *I. glandulifera* (Thijs *et al.*, 2012), but because bees are central-place foragers, and must return to their nest after feeding, movement between forage resources is constrained and the resulting pollen-mediated gene flow is normally over short distances (R. J. Paxton, Martin-Luther-Universität, Germany, pers. comm.).

The role of hydrochory in shaping patterns of genetic diversity in *I. glandulifera* is particularly apparent in the Structure analysis of the Western Cleddau system. Populations from the two upstream tributaries are associated with four genetic clusters, which are rarely shared across the two tributaries. Populations below the confluence of the two, however, show evidence of admixture of three of these clusters, reflecting gene flow from the two separate upstream tributaries, and it has been shown that CB4 river systems, like the Western Cleddau (see Methods), are often characterized by recolonization from upstream source populations (Hatton-Ellis and Grieve, 2003). The WH population at the upstream extent of the Upper Western Cleddau consists almost exclusively of individuals associated with the remaining genetic cluster, which is only represented at extremely low frequencies across the remainder of the system. Interviews with local residents and gardeners confirmed that the WH population was planted next to a garden pond which had been dammed off from the main river 22 years earlier. Thus, the genetic distinctness of the WH population reflects its lack of any aquatic connection to the remaining populations, further indicating the importance of hydrochory in gene flow and dispersal in *I. glandulifera*.

The assignment of individuals from the BYB population on the Tempo river to genetic clusters associated with the Colebrooke system highlights the importance of the hydrogeomorphology of individual river systems and tributaries in hydrochoric dispersal. This population is situated near the confluence of the Tempo and the much larger Colebrooke river, but the distance from the point of confluence (approx. 100 m) means that backing up of water from the larger tributary is unlikely to account for the observed pattern of gene flow that contradicts the expected linear, unidirectional movement of propagules. The region is prone to flooding, however, and this is most likely responsible for the observed gene flow between the two tributaries at this point. Previous genetic (Jacquemyn *et al.*, 2006) and ecological (Boedeltje *et al.*, 2004; Gurnell *et al.*, 2008; Moggridge and Gurnell, 2009) studies have indicated the importance of periodic flooding events in the dispersal of riparian plants, including invasive species (Truscott *et al.*, 2006). The low levels of assignment of individuals to the light blue cluster in the downstream BDB population indicates that dispersal is also limited from the Tempo into the much larger Colebrook watercourse, further confirming the role of hydrological factors.

Although the results of this study strongly support hydrochoric downstream dispersal of *I. glandulifera*, it should be borne in mind that there are several other potential, non-exclusive processes that could also shape the observed pattern of genetic diversity in the river systems studied. First, lower upstream genetic diversity could result from founder effects as a result of human-mediated dispersal, as human population densities tend to be higher in the lower reaches of river systems. This, however, would not necessarily lead to the observed structuring of diversity between tributaries, as human-mediated dispersal would be equally likely between tributaries. Secondly, the distribution of genetic diversity might reflect a series of garden escapes, although this would not explain the increase in genetic diversity downstream. Finally, increased habitat heterogeneity downstream might lead to relaxed selection pressures, thus supporting a wider range of genotypes (Ward and Stanford, 1995). Despite these possibilities, hydrochoric dispersal would appear to be by far the most parsimonious explanation for the observed distribution of genetic diversity in populations of *I. glandulifera* in both river systems.

In the majority of angiosperm species, including *I. glandulifera* (van Went, 1984), the plastid genome is transmitted uniparentally via the female parent and thus molecular markers specific to the plastid genome can provide valuable insights into seed-mediated gene flow, including hydrochory (McCauley, 1995). In the present study, we hoped to be able to dissect the effects of pollen and seed flow by comparing patterns of genetic variation in nuclear and plastid markers. Eight chloroplast microsatellite markers were developed and utilized for this study, but failed to display any polymorphism (data not shown). Given that such markers represent the most variable regions of the plastid genome (reviewed by Provan *et al.*, 2001; Ebert and Peakall, 2009), this lack of cytoplasmic variation, coupled with the low levels of diversity found at the nuclear loci analysed in the present study, suggests that *I. glandulifera* populations in Britain and Ireland have a narrow genetic base and is consistent with a limited number of introductions. Furthermore, it suggests that large-scale sequencing of the plastid genome of *I. glandulifera* for single nucleotide polymorphisms (SNPs) may be necessary to develop seed-specific markers for the further study of hydrochory in this species.

Implications for control of invasive riparian plant species

The primary point of control with respect to invasive species is to prevent introduction from the native range, but for already established populations, knowledge of modes and patterns of secondary dispersal will help to prevent the subsequent spread of these species (Kowarik and Samuel, 2008). The unique value of genetic approaches lies in revealing the accumulation over time of the results of multiple complex processes rather intractable to experimental ecological study in natural systems (variation in water flow, interactions between aquatic organisms, etc.). *I. glandulifera* can be treated with glyphosate, but the widespread use of such herbicides in riparian habitats should be viewed as an extreme measure. It has previously been suggested that targeting upstream populations might reduce the impact of propagule pressure where hydrochory plays a major role in dispersal (Dawson and Holland, 1999; Wadsworth *et al.*, 2000; Clements *et al.*, 2008) and the results of the present study give added weight to these recommendations. In addition, the observed genetic differentiation between tributaries would imply that separate tributaries should have their own eradication programmes, again probably focusing on upstream populations, but consideration should also be given to hydrological factors, i.e. targeting larger tributaries first as they typically represent greater sources for downstream areas, as highlighted by the fact

that the predominant genetic cluster in the Tempo is barely represented below its confluence with the much larger Colebrooke. Given that intensity (i.e. the percentage of occupied area) of control is the major determinant of the cost of eradication programmes, and that control intensities of a previous modelling study were far above those typical of real programmes in the UK yet failed to control the spread of *I. glandulifera*, such a multi-faceted, targeted approach might offer the most rational and cost-effective means of control. This concept might also be extended to the control of other invasive riparian species that display evidence of primarily hydrochoric dispersal. The fact that we found similar results in two contrasting systems means that we can infer that our observations are probably widely applicable.

ACKNOWLEDGEMENTS

SUPPLEMENTARY DATA

Supplementary data are available online at www.aob.oxford-journals.org and consist of Table S1: population-pairwise Φ_{ST} values.

LITERATURE CITED

Beerling DJ, Perrins JM. 1993. *Impatiens glandulifera* Royle (*Impatiens roylei* Walp.). *Journal of Ecology* **81**: 367–382.

Birkeland GH. 1996. Riparian vegetation and sandbar morphology along the lower Little Colorado River, Arizona. *Physical Geography* **17**: 534–553.

Boedeltje G, Bakker JP, Ten Brinke A, van Groenendaal JM, Soesbergen M. 2004. Dispersal phenology of hydrochorous plants in relation to discharge, seed release time and buoyancy of seeds: the flood pulse concept supported. *Journal of Ecology* **92**: 786–796.

Burkart M. 2001. River corridor plants (Stromtalpflanzen) in Central European lowland: a review of a poorly understood plant distribution pattern. *Global Ecology and Biogeography* **10**: 449–468.

Chen YY, Yang W, Li W, Li ZZ, Huang HW. 2009. High allozyme diversity and unidirectional linear migration patterns within a population of tetraploid *Isoetes sinensis*, a rare and endangered pteridophyte. *Aquatic Botany* **90**: 52–58.

Clavero M, García-Berthou E. 2005. Invasive species are a leading cause of animal extinctions. *Trends in Ecology and Evolution* **20**: 110.

Clements DR, Feenstra KR, Jones K, Stainforth R. 2008. The biology of invasive alien plants in Canada. 9. *Impatiens glandulifera* Royle. *Canadian Journal of Plant Sciences* **88**: 403–417.

Cloutier D, Póvoa JSR, Procopio LC, *et al.* 2005. Chloroplast DNA variation of *Carapa guianensis* in the Amazon basin. *Silvae Genetica* **54**: 270–274.

Colautti RI, Grigorovich IA, MacIsaac HJ. 2006. Propagule pressure: a null model for biological invasions. *Biological Invasions* **8**: 1023–1037.

Collingham YC, Wadsworth RA, Huntley B, Hulme PE. 2000. Predicting the spatial distribution of non-indigenous riparian weeds: issues of spatial scale and extent. *Journal of Applied Ecology* **37**(S1): 13–27.

Cronk QCB, Fuller JL. 2001. *Plant invaders: the threat to natural ecosystems*. London: Earthscan Publications.

Dawson FH, Holland D. 1999. The distribution in bankside habitats of three alien invasive plants in the UK in relation to the development of control strategies. *Hydrobiologia* **415**: 193–201.

DiTomaso JM. 1998. Impact, biology and ecology of saltcedar (*Tamarix* spp.) in the southwest United States. *Weed Technology* **12**: 326–336.

Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**: 359–361.

Ebert D, Peakall R. 2009. Chloroplast simple sequence repeats (cpSSRs): technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. *Molecular Ecology Resources* **9**: 673–690.

Eschtruth AK, Battles JJ. 2009. Assessing the relative importance of disturbance, herbivory, diversity and propagule pressure in exotic plant invasions. *Ecological Monographs* **79**: 265–280.

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.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes – application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.

Excoffier L, Laval LG, Schneider S. 2005. ARLEQUIN, Version 3·0: an integrated software package for population genetic data analysis. *Evolutionary Bioinformatics Online* **1**: 47–50.

Gornall RJ, Hollingsworth PM, Preston CD. 1998. Evidence for spatial structure and directional gene flow in a population of an aquatic plant *Potamogeton coloratus*. *Heredity* **80**: 414–421.

Goudet J. 2001. FSTAT, version 2·9.3, A program to estimate and test gene diversities and fixation indices. http://www2.unil.ch/popgen/softwares/fstat.htm.

Graf WL. 1982. Tamarisk and river channel management. *Environmental Management* **6**: 283–296.

Gurevitch J, Padilla DK. 2004. Are invasive species a major cause of extinctions? *Trends in Ecology and Evolution* **19**: 470–474.

Gurnell A, Thompson K, Goodson J, Moggridge H. 2008. Propagule deposition along rivers: linking hydrology and ecology. *Journal of Ecology* **96**: 553–565.

Hatton-Ellis TW, Grieve N. 2003. *Ecology of watercourses characterized by* Ranunculion fluitantis *and* Callitricho-Batrachion *vegetation.* Conserving Natura 2000 Rivers Ecology Series Number 11. Peterborough: English Nature.

Hejda M, Pyšek P, Jarošik V. 2009. Impact of invasive plants on the species richness, diversity and composition of invaded communities. *Journal of Ecology* **97**: 393–403.

Henderson S, Dawson TP, Whittaker RJ. 2006. Progress in invasive plants research. *Progress in Physical Geography* **30**: 25–46.

Honnay O, Jacquemyn H, Nackaerts K, Breyne P, van Looy K. 2010. Patterns of population genetic diversity in riparian and aquatic plant species along rivers. *Journal of Biogeography* **37**: 1730–1739.

Hupp CR, Osterkamp WR. 1996. Riparian vegetation and fluvial geomorphic processes. *Geomorphology* **14**: 277–295.

Jacquemyn H, Honnay O, van Looy K, Breyne P. 2006. Spatiotemporal structure of genetic variation of a spreading plant population on dynamic riverbanks along the Meuse River. *Heredity* **96**: 471–478.

Kowarik I, Samuel I. 2008. Water dispersal as an additional pathway to invasions by the primarily wind-dispersed tree *Ailanthus altissima*. *Plant Ecology* **198**: 241–252.

Liu Y, Wang Y, Huang H. 2006. High interpopulation genetic differentiation and unidirectional linear migration patterns in *Myricaria laxiflora* (Tamaricaceae), an endemic riparian plant in the Three Gorges Valley of the Yangtze River. *American Journal of Botany* **93**: 206–215.

Lockwood JL, Cassey P, Blackburn T. 2005. The role of propagule pressure in explaining species invasions. *Trends in Ecology and Evolution* **20**: 223–228.

Loope LL, Sanchez PG, Tarr PW, Loope WL, Anderson RL. 1988. Biological invasions of arid land nature reserves. *Biological Conservation* **44**: 95–118.

Lopezaraiza-Mikel ME, Hayes RB, Whalley MR, Memmott J. 2007. The impact of an alien plant on a native plant-pollinator network: an experimental approach. *Ecology Letters* **10**: 539–550.

Lundqvist E, Andersson E. 2001. Genetic diversity in populations of plants with different breeding and dispersal strategies in a free-flowing boreal river system. *Hereditas* **135**: 75–83.

McCauley DE. 1995. The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology and Evolution* **10**: 198–202.

McKinney ML, Lockwood JL. 1999. Biotic homogenization: a few winners replacing many losers in the next mass extinction. *Trends in Ecology and Evolution* **14**: 450–453.

Markwith SH, Scanlon MJ. 2007. Multiscale analysis of *Hymenocallis coronaria* (Amaryllidaceae) genetic diversity, genetic structure and gene movement under the influence of unidirectional stream flow. *American Journal of Botany* **94**: 151–160.

Martínez-Ghersa MA, Ghersa CM. 2006. The relationship of propagule pressure to invasion potential in plants. *Euphytica* **148**: 87–96.

Moggridge HL, Gurnell AM. 2009. Hydrological controls on the transport and deposition of plant propagules within riparian zones. *River Research and Applications* **26**: 512–527.

Myers JH, Bazely D. 2003. *Ecology and control of introduced plants*. Cambridge: Cambridge University Press.

Nakayama N, Nishihiro J, Kayaba Y, Muranaka T, Washitani I. 2007. Seed deposition of *Eragrostis curvula*, an invasive alien plant on a river floodplain. *Ecological Research* **22**: 696–701.

Pearson DE. 2009. Invasive plant architecture alters trophic interactions by changing predator abundance and behavior. *Oecologia* **159**: 549–558.

Pimentel D, McNair S, Janecka J, *et al*. 2001. Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems and Environment* **84**: 1–20.

Pollux BJA, Luteijn A, van Groenendael JM, Ouborg NJ. 2009. Gene flow and genetic structure of the aquatic macrophyte *Sparganium emersum* in a linear unidirectional river. *Freshwater Biology* **54**: 64–76.

Prentis PJ, Mather PB. 2007. Micro-geographic landscape features demarcate seedling genetic structure in the stream lily, *Helmholtzia glaberrima*. *Aquatic Botany*, **87**: 111–115.

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

Provan J, Powell W, Hollingsworth PM. 2001. Chloroplast microsatellites: new tools for studies in plant ecology and systematics. *Trends in Ecology and Evolution* **16**: 142–147.

Provan J, Love HM, Maggs CA. 2007. Development of microsatellites for the invasive riparian plant *Impatiens glandulifera* (Himalayan balsam) using intersimple sequence repeat cloning. *Molecular Ecology Notes* **7**: 451–453.

Pyšek P, Prach K. 1993. Plant invasions and the role of the riparian habitats: a comparison of four species. *Journal of Biogeography* **20**: 413–420.

Raymond M, Rousset F. 1995. GenePop (version 1·2): population genetic software for exact tests and ecumenicism. *Journal of Heredity* **86**: 248–249.

Ritland K. 1989. Genetic differentiation, diversity and inbreeding in the mountain monkeyflower (*Mimulus caespitosus*) of the Washington cascades. *Canadian Journal of Botany* **67**: 2017–2024.

Russell JR, Weber JC, Booth A, Powell W, Sotelo-Montes C, Dawson IK. 1999. Genetic variation of *Calycophyllum spruceanum* in the Peruvian Amazon Basin, revealed by amplified fragment length polymorphism (AFLP) analysis. *Molecular Ecology* **8**: 199–204.

Sakai AK, Allendorf FW, Holt JS, *et al*. 2001. The population biology of invasive species. *Annual Review of Ecology and Systematics* **32**: 305–332.

Samuel I, Kowarik I. 2010. Urban rivers as dispersal corridors for primarily wind-dispersed invasive tree species. *Landscape and Urban Planning* **94**: 244–249.

Schleuning M, Becker T, Vadillo GP, Hahn T, Matthies D, Durka W. 2011. River dynamics shape clonal diversity and genetic structure of an Amazonian understorey herb. *Journal of Ecology*, **99**: 373–382.

Simberloff D. 2009. The role of propagule pressure in biological invasions. *Annual Review of Ecology, Evolution and Systematics* **40**: 81–102.

Tassin J, Riviere JN, Clergeau P. 2007. Reproductive versus vegetative recruitment of the invasive tree *Schinus terebenthifolius*: implications for restoration on Reunion Island. *Restoration Ecology* **15**: 412–419.

Tero N, Aspi J, Siikamäki P, Jäkäläniemi A, Tuomi J. 2003. Genetic structure and gene flow in a metapopulation of an endangered plant species *Silene tatarica*. *Molecular Ecology* **12**: 2073–2085.

Thijs KW, Brys R, Verboven HAF, Hermy M. 2012. The influence of an invasive plant species on the pollination success and reproductive output of three riparian plant species. *Biological Invasions* **14**: 355–365.

Tickner DP, Angold PG, Gurnell AM, Mountford JO. 2001. Riparian plant invasions: hydrogeomorphological control and ecological impacts. *Progress in Physical Geography* **25**: 22–52.

Truscott A-M, Soulsby C, Palmer SCF, Newell L, Hulme PE. 2006. The dispersal characteristics of the invasive plant *Mimulus guttatus* and the ecological significance of increased occurrence of high-flow events. *Journal of Ecology* **94**: 1080–1091.

van der Wal R, Truscott AM, Pearce ISK, Cole L, Harris MP, Wanless S. 2008. Multiple anthropogenic changes cause biodiversity loss through plant invasion. *Global Change Biology* **14**: 1428–1436.

van Went JL. 1984. Unequal distribution of plastids during vegetative cell formation in *Impatiens*. *Theoretical and Applied Genetics* **68**: 305–310.

Wadsworth RA, Collingham YC, Willis SG, Huntley B, Hulme PE. 2000. Simulating the spread and management of alien riparian weeds: are they out of control? *Journal of Applied Ecology* **37**(S1): 28–38.

Walker NF, Hulme PE, Hoelzel AR. 2009. Population genetics of an invasive riparian species, *Impatiens glandulifera*. *Plant Ecology* **203**: 243–252.

Ward JV, Stanford JA. 1995. The serial discontinuity concept: extending the model to floodplain rivers. *Regulated Rivers: Research & Management* **10**: 159–168.

Weber E. 2003. *Invasive plant species of the World: a reference guide to environmental weeds*. Wallingford, UK: CABI Publishing.