BIODIVERSITY RESEARCH

Predicting and explaining plant invasions through analysis of source area floras: some critical considerations

Petr Pyšek,* David M. Richardson† and Mark Williamson‡

*Institute of Botany, Academy of Sciences of the Czech Republic, CZ-252 43 Průhonice, Czech Republic, E-mail: pysek@ibot.cas.cz; †Institute for Plant Conservation, Botany Department, University of Cape Town, Rondebosch 7701, South Africa, E-mail: rich@botzoo.uct.ac.za; and ‡Department of Biology, University of York, York YO10 5 DD, United Kingdom, E-mail: mw1@york.ac.uk

ABSTRACT

Comparing species that become invasive with others from the same regional species pool that do not invade raises several issues about the accuracy of analyses attempting to define the determinants of invasiveness. The delimitation of the source area and deciding which species group(s) to include are especially relevant in analyses focusing on species originating in Europe. Historical patterns of immigration of alien species into Europe must be considered since European floras comprise a complex mix of native species, historical introductions (archaeophytes) and relative newcomers (neophytes). We make three main points: (1) Archaeophytes (species introduced to Europe before the discovery of America) differ from natives in a number of traits and in historical association with people; it is misleading to lump archaeophytes with native taxa. (2) Taxa from climatically and geographically different regions, representing distinct floristic geoelements, need to be treated separately, and not as a homogenous pool of potential invaders. Restricting the source species pool to native taxa with comparable phytogeographical characteristics reduces the variation associated with chance of dispersal by humans from the source area. (3) For prediction, a clear distinction should be made between accuracy (the proportion of those found to be alien that were also predicted to be there) and reliability (or predictive value, the proportion of those predicted to become aliens that do so). Information accumulated over centuries by botanists in Central Europe provides an excellent opportunity to deal with these issues and avoid spurious results. To illustrate these issues, we revisit a recently published study of Central-European plant species as invaders in two Argentinean provinces (Prinzing *et al.*, 2002) to explore and demonstrate the implications of the above points. We hope that future studies will build on these points to achieve more reliable predictions.

Keywords

Archaeophytes, biological invasions, distribution ranges, native species, neophytes, prediction, residence time, source-area approach, species pool.

*Correspondence: Petr Pyšek, Institute of Botany, Academy of Sciences of the Czech Republic, CZ-252 43 Průhonice, Czech Republic. E-mail: pysek@ibot.cas.cz

INTRODUCTION

Biological invasions have received increasing attention worldwide since the 1980s and are now a hot topic in contemporary ecology. This has resulted in the publication of the proceedings of many symposia (e.g. Drake *et al.*, 1989; di Castri *et al.*, 1990; Groves & di Castri, 1991; Pyšek *et al.*, 1995; Carey *et al.*, 1996; Starfinger *et al.*, 1998; Mooney & Hobbs, 2000) and reviews (e.g. Rejmánek *et al.*, 2004). The field is both scientifically interesting and also of considerable practical importance (Luken & Thieret, 1997; Williamson, 1998; Preston *et al.*, 2000; Myers & Bazely, 2003). An understanding of the factors influencing invasiveness and invasibility (which species invade and which factors determine

the susceptibility of an ecosystems to invasion) is a fundamental goal of invasion ecology. Many recent papers have addressed the requirements for the effective management of invasions (e.g. McNeely *et al.*, 2001), and several have provided a critical evaluation of the prospect of predicting the outcome of introductions (Tucker & Richardson, 1995; Williamson, 1996, 1999, 2001ab; Daehler & Carino, 2000; Kolar & Lodge, 2001; Pyšek, 2001).

Much recent research on plant invasions has explored the invasiveness of particular species and the vulnerability of various communities to invasions. This work has yielded some reasonably robust generalizations or rules (Rejmánek & Richardson, 1996; Rejmánek & Richardson, 1996, 2000; Lonsdale, 1999; Tilman, 1999; Grotkopp *et al.*, 2002; Rejmánek *et al.*, 2004). Several

authors have shown that manipulative experiments are essential for gaining a more mechanistic, and hence predictive, understanding of invasions (Williamson, 1999; Kennedy *et al.*, 2002; Klironomos, 2002; see review in Rejmánek *et al.*, 2004). Despite the value of formal experiments, carefully compiled species lists are also useful for understanding invasions, especially for generating hypotheses for testing by experimental and comparative methods (Crawley *et al.*, 1996; Thompson *et al.*, 1995; Williamson & Fitter, 1996b; Weber, 1997; Daehler, 1998, 2001; see also Rejmánek *et al.*, 2004).

Reliable databases of alien species in various countries, summarizing the historical knowledge of generations of botanists, have proved especially useful. In Europe, which has served as a donor of invasive species for many other areas (di Castri *et al.*, 1990; Pyšek, 1998), accurate alien floras (applying objective criteria for the classification of alien taxa) are available for only a few countries, notably Britain and Ireland (Clement & Foster, 1994; Ryves *et al.*, 1996), the Czech Republic (Pyšek *et al.*, 2002), Austria (Essl & Rabitsch, 2002) and Germany (Frank & Klotz, 1990; Klotz *et al.*, 2002; Kühn & Klotz, 2003). Despite the real problems associated with the assessment of species numbers (Williamson, 2002; Pyšek, 2003) and alien/native status (Webb, 1985; Pyšek *et al.*, 2004), such data sets, in association with databases of biological attributes (Frank & Klotz, 1990; Fitter & Peat, 1994) and geographical information, have yielded valuable results that have contributed substantially to the explanation of patterns in alien floras (Crawley *et al.*, 1996; Williamson & Fitter, 1996a,b).

Traditionally, analyses aimed at defining correlates of invasiveness are done on the flora of a target area, either comparing species with different degrees of invasion success in a region (Pyšek *et al.*, 1995; Weber, 1997; Daehler, 1998; Pyšek, 1998), or comparing aliens vs. natives for a given region (Crawley *et al.*, 1996). An alternative approach focuses on the source area by an analysis of species that were introduced to a region from a defined geographical area. Such studies are rather rare and most have been done on subsets of alien species (Scott & Panetta, 1993), taxonomic groups (Rejmánek & Richardson, 1996; Rejmánek, 2000) or congeners (Goodwin *et al.*, 1999). One such study (Prinzing *et al.*, 2002) analyses the flora of a geographical region (East Germany) as a source of invaders for a distant part of the world (Argentina).

Such an approach certainly has great promise for identifying those species with a high probability of becoming invaders. However, as with every new approach, this one has novel theoretical premises. An important issue that, in our view, should be dealt with very carefully in such analyses is the delimitation of the species pool in the source region. Views on which species groups to include may differ among researchers. In this paper we argue that such a decision is crucial. Two aspects are relevant in this regard: (1) whether or not to include long-established, non-native species; and (2) how to take into account the distribution ranges of species that are included.

In this paper, we point out the importance of distinguishing native species in the source area from those that were introduced to that region at different times (which determines their 'residence time', i.e. for how long they have been present in the region). We also stress the importance of a careful consideration of the 'phytogeographical status' of species in the source region, i.e. the character of the primary distribution ranges worldwide of species. In our view, these aspects have not received the attention they deserve in the literature. We will show, using the analysis of Prinzing *et al.* (2002) as an example, how failure in this regard can reduce the value of a potentially useful approach. We argue that in this study (1) archaeophytes, that are not native to the source region, should not have been included as a part of the source species pool (2) natives forming the source species pool might better have been considered by phytogeographical element, and (3) the prediction based on their results uses 'accuracy' (the proportion of those that are aliens and also predicted to be so) rather than 'reliability' — also called 'predictive value' — (the proportion of those predicted to become alien that do so) and is therefore somewhat misleading.

Defining the species pool: the importance of species origin and residence time

An important assumption of studies based on source area is to clarify which of the species constituting its flora should be considered for the pool of potential donor species. In Central Europe, several groups of taxa are recognized depending on their origin: native and alien, the latter divided into archaeophytes and neophytes. Native species (defined as taxa that have arrived there without intentional or unintentional human intervention from an area in which they are native or that originated there naturally, Pyšek *et al.*, 2004) clearly belong to the source species pool but a critical decision is whether to include aliens. Alien plants are those whose presence results from intentional or unintentional human intervention. Many are associated with agriculture, horticulture and forestry.

Plants alien to Central Europe have been traditionally divided into archaeophytes (introduced before the discovery of America in 1492, usually rounded to 1500), and neophytes (introduced after that date; Thellung, 1905; Holub & Jirásek, 1967; Schroeder, 1969). This distinction has been recognized since the start of serious floristic research at the beginning of the 19th century (see Pyšek *et al.*, 2002 for a review of situation in the Czech Republic, Wittig, 2002; Kühn & Klotz, 2003 for Germany, and Pyšek, 1995; Pyšek *et al.*, 2002, 2004 for a comparison of these classification systems with recent usage of the terms). These two groups are also recognized in Great Britain (Preston *et al.*, 2002; Williamson *et al.*, 2003).
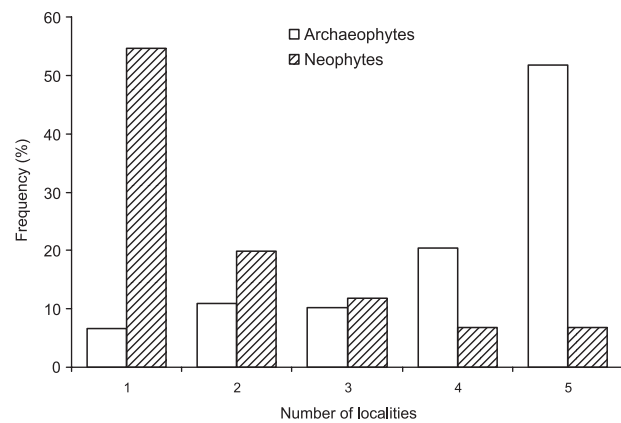
Given that the main rationale behind the source-area approach is to reduce the variation associated with human-mediated dispersal from the source area, it seems logical to exclude neophytes from the source species pool. Neophytes were themselves introduced to the region at various times over the past several centuries, and therefore differ in their residence times. Residence time affects the probability of a species becoming naturalized, as can be inferred from the fact that the longer a neophyte has been present in a region, the more widespread it is. For example, a positive and highly significant dependence of the number of localities (which can be taken as a measure of range size, Williamson, 2001a) on residence time was found for neophytes
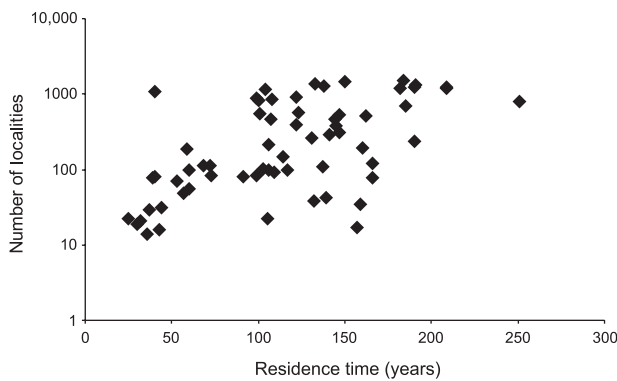
Figure 1 The relationship between the number of reported localities and residence time (time since first record) for 63 neophytes of the Czech Republic. The number of localities has been compiled from regional floristic literature, herbaria and unpublished records up to 2000. Note that the same data sources were used to compile the record for each species. The definition of locality is that in the original records; almost all records are at least 500 m apart. The linear regression, Log Number of localities = 1.41 + 0.0047 Residence time, $n = 63$, $r^2 = 0.38$, $P < 0.001$, is highly significant but the underlying relationship is curvilinear (and would be even more significant) because the number of possible localities is limited.

Figure 2 The distribution of range size of archaeophytes and neophytes of the Czech Republic. Range size is the number of localities of each species, and is plotted using the scale in Clement & Foster (1994): 1 = 1–4 localities; 2 = 5–14; 3 = 15–49; 4 = 50–499; 5 = over 500 localities. Pyšek et al. (2002) for values assigned to particular species. Archaeophytes have significantly greater range sizes than native species: G-test on contingency tables, $\chi^2 = 477.3$, d.f. = 4, $P < 0.001$.

introduced to the Czech Republic (Fig. 1, see also Pyšek et al., 2003). The same relationship holds for naturalized Fabaceae in Taiwan (Wu et al., 2003) and naturalized grasses in Venezuela (M. Rejmánek, unpublished data). As species constituting the pool of neophytes have a large variation in range size, the variation in the chance of each species being moved out of the region by human agency is high and including them into the species pool would increase the bias that the source-area approach attempts to reduce. Moreover, neophytes that were introduced from various parts of the world exhibit considerable variation in life-history characteristics and ecological traits (Pyšek et al., 2002): features that would further increase unwanted variation in the source species pool.

The situation is different with archaeophytes. They have been present in Central and NW Europe for several thousand years. Their long-term introduction dynamics has been estimated in detail for the Czech Republic: 35.2% of presently known species were introduced in the Neolithic/Chalcolithic period (5300–2200 BC) and more than half (52.7%) are thought to have been present by the end of the Bronze Age, 750 BC (Pyšek et al., 2003). Archaeophytes are therefore well adapted to the regional climate of Central Europe, but are not native and were not present there before the start of Neolithic agriculture (Holub & Jirásek, 1967; Zajac, 1979; Williamson, 2002; Pyšek et al., 2004). They are alien species, because their presence in the region is due to intentional or unintentional human involvement, or they have arrived there without the help of people from an area in which they are alien (Pyšek et al., 2004). Archaeophytes differ ecologically from the neophytes introduced later, as is documented in an extensive data set published recently for the Czech Republic (Pyšek et al.,

2002). That country's alien flora consists of 1378 species, of which 332 (24.1%) are considered archaeophytes. Archaeophytes and neophytes have dramatically different proportions of naturalization, and the ratio of casual/naturalized (in the sense of Richardson et al., 2000) is reversed in the two groups: 77.8% of archaeophytes are naturalized, whereas only 21.9% of neophytes are. Consequently, archaeophytes make up 59.7% of naturalized alien Czech species (Pyšek et al., 2002). They also have larger range sizes: 51.8% of archaeophytes in the Czech Republic are known from more than 500 localities and only 6.8% are reported from less than five localities. This pattern is reversed in neophytes with values of 6.8% and 54.6%, respectively (Fig. 2). It has been shown that in regional floras, alien species occur at fewer localities and occupy smaller ranges, on average, than native species. However, in terms of range size, archaeophytes are, unlike neophytes, rather similar to native species.

The resemblance of archaeophytes to native species, in terms of distribution and range size, is probably the main reason for a tendency among botanists to treat long-established taxa as native (as pointed out by Webb, 1985). However, the two groups are clearly different in many respects. Archaeophytes should, like neophytes, also be excluded from the species pool when using the source-area approach. A separate analysis could be made using archaeophytes as a source species pool. Comparing such results with those from using the native species pool might provide interesting insights into the role that residence time plays in shaping the set of potential invaders to other parts of the world.

Defining the source area: the importance of size and placement of distribution ranges

Studies based on the analyses of the alien floras of target areas are usually limited by an important constraint — the source pool of the

invading species is heterogeneous as species come from various regions. Hence, such analyses necessarily confuse the biological characters of the invading species with the circumstances associated with dispersal (i.e. the effect of the vector and the distance they must overcome to reach the target area). Comparing a species arriving from the opposite side of the world with one having its primary distribution range adjacent to the target area introduces a bias of unknown dimensions. In such studies, we are not able to filter out the variation imposed by the species being translocated from various source areas.

In the source-area approach, the null hypothesis should be that each species has (in geographical terms) the same chance of being moved by people from its natural range. Any differences between species in invasive success in the target region would then be attributable to their biological and ecological features. Of course this is not necessarily true, since common species in urban areas have a better chance of being moved than rare species restricted to natural communities. Ideally, this could be controlled for by using information on the habitats occupied and range size in the source region. After filtering out the effect of these two factors (i.e. how common a species is and how often it occurs in habitats where the likelihood of coming into contact with human dispersers is highest), it should be possible to attribute differences in the inherent capability of species to establish and/or become invasive in a target region to their biological and ecological features alone.

The size of the primary (native) geographical ranges of plant species is considered to be one of the most promising predictors of their ability to become invasive (Forcella & Wood, 1984; Williamson, 1996, 2001b; Rejmánek *et al.*, 2002). Rejmánek (1996, 2000) has shown that species (in selected families, Asteraceae, Poaceae, Fabaceae) with large primary distributions in Europe have larger distributions as invaders in North America than related species with smaller natural ranges. Similarly, Goodwin *et al.* (1999), who compared species of European origin occurring as naturalized aliens in Canada with paired European species that had not invaded, concluded that those more widely distributed in Europe had a significantly higher chance of becoming invaders in Canada. In the first example, wider natural range translated to wider adventive range, whereas in the second, the wide primary range results in enhanced probability of becoming naturalized. However, they both demonstrate the importance of the size of the primary range as a predictor of invasion success. Although it is impossible to control completely for the effect of the size of primary geographical range, enormous variation in this could be reduced by working with geographically well defined and restricted areas. Size of the range is closely associated with the probability that a species will be transported from the source region; the larger the area it occupies, the more contacts with humans can be expected. In studies considering complete floras of source areas, sizes of distribution ranges of particular species can be filtered out to reduce the variation in probability that a species comes into contact with people and is translocated elsewhere.

Placement of the primary distribution range is another important characteristic that affects the chance of a species being

moved from the source region. Some species occur across the whole source area while others occupy, for various reasons, only a part of the source area. As we will show below, these characteristics should be taken into account when deciding about the composition of the species pool considered. Through careful delimitation of the source species pool, variation in dispersal-related variables can be further reduced. Central Europe provides a good opportunity for this as the information on the size and placement of primary ranges of species constituting its flora is available in several works (Tutin *et al.*, 1964–1980 but see Meusel *et al.*, 1965; Hultén & Fries, 1986; Pyšek, 2003; Jalas *et al.* 1972–1999) and in national sources (e.g. Slavík, 1986–1998; Haeupler & Schönfelder, 1989; Zajac & Zajac, 2001).

Source area and its species pool: getting the variables right

Prinzing *et al.*'s (2002) paper on Central European plant species as invaders in two Argentinean provinces is the first analysis of the alien flora of a large geographical region using the source-area approach. Instead of comparing the alien and native flora of the target area, these authors (after appropriately controlling for phylogenies) compared those species from the source area (Central Europe) that became aliens (taken to include 'casual', 'naturalized' and 'invasive' taxa as defined by Richardson *et al.*, 2000) in Argentina with those that did not. Only a few species in a species pool become aliens and until now, it was unclear how those that do so differ from non-aliens. They concluded that these alien species are characterized by (a) frequency (i.e. large range sizes) in central Europe and utilization by people (b) preference for warm, dry and nitrogen-rich conditions in central Europe (c) a native range that covers several floristic zones and (d) a ruderal life strategy. They suggested that the traits of aliens indicate that they are in frequent contact with humans as their dispersal vector, which made them preadapted to the abiotic conditions in Argentina, and they are versatile and can colonize disturbed sites. Their results have a high accuracy but a low reliability as is shown below. They explain why most of the species found alien are there, but they do not explain why many more species in Germany with similar characteristics have not been found in Argentina.

This paper by Prinzing *et al.* (2002) can be used to demonstrate the issues we outlined above. The first issue is how to define the source species pool. By excluding neophytes from the flora of eastern Germany, the data source in their paper, they limited their source pool of species to two groups: (a) native species, and (b) archaeophytes. There are some differences between the classification of alien species used in Germany and the Czech Republic, mostly reflecting the fact that both countries lie on a 'critical point of the SE–NW gradient' (Skalický, 1988) where many species of southern origin reach their north-western distribution. Some species are therefore considered native in the Czech Republic but archaeophytes in Germany as can be seen from the comparison of, e.g. Rothmaler *et al.* (2002) or Klotz *et al.* (2002) with Pyšek *et al.* (2002). Nevertheless, the representation of archaeophytes in the alien floras of Germany and the

Table 1 Summary of the ecological differences between archaeophytes and native species based on data sets from eastern Germany (Frank & Klotz, 1990), the whole of Germany (Klotz *et al.*, 2002), and the Czech Republic (Kubát *et al.*, 2002; Pyšek *et al.*, 2002; this study). The number of species with a given characteristic (taxonomic position, life form, life strategy, month of first flowering, ploidy level, hybrids vs. non-hybrids) was calculated for archaeophytes and native species and the difference in distribution was tested. Grime's system (1979) was used to classify plant strategies; in this system, species are classified according to their response to competition, disturbance and stress. The Raunkiaer scheme (e.g. Ellenberg, 1988) was used for life forms. Ecological demands were expressed by using Ellenberg indicator values of native species and archaeophytes. This system classifies species' response to particular factors on an ordinal scale and is based on extensive field measurements of factors in habitats occupied by species (see Ellenberg *et al.*, 1991 for details). Mean height was obtained from Kubát *et al.* (2002) and Pyšek *et al.* (2002). OR = over-represented

Characteristic	Statistics	d.f.	P	Trends	Source
	G-test on contingency tables				
Taxonomic position	$\chi^2 = 385.41$	77	< 0.001	different families typical for each group	this study
Life form (Raunkiaer)	$\chi^2 = 407.36$	5	< 0.001	therophytes and hemicryptophytes OR in archaeophytes, other life forms in native	Pyšek *et al.* (2002)
Life strategy	$\chi^2 = 505.6$	5	< 0.001	R and CR strategy OR in archaeophytes, C, CS and CSR in native	Klotz *et al.* (2003)
Months of first flowering	$\chi^2 = 7.47$	8	n.s.	no difference	this study
Ploidy level	$\chi^2 = 66.56$	7	< 0.001	diploids OR in archaeophytes, higher ploidy levels in native	Pyšek *et al.* (2002), Kubát *et al.* (2002)
Proportion of hybrids	$\chi^2 = 0.057$	1	n.s.	no difference: 18.7% of all taxa in archaeophytes, 18.1% in native	Pyšek *et al.* (2002), Kubát *et al.* (2002)
	Kruskal–Wallis				
Ecological demands: light	$\chi^2 = 0.61$	1	n.s.	arch = 7.05 ± 1.09 (*n* = 132), native = 6.75 ± 1.53 (*n* = 1314)	Frank & Klotz (1990)
Ecological demands: temperature*	$\chi^2 = 63.30$	1	< 0.001	arch = 6.46 ± 0.96 (*n* = 114), native = 5.52 ± 1.19 (*n* = 916)	Frank & Klotz (1990)
Ecological demands: moisture†	$\chi^2 = 18.72$	1	< 0.001	arch = 4.08 ± 0.98 (*n* = 119), native = 5.91 ± 3.42 (*n* = 1240)	Frank & Klotz (1990)
Ecological demands: soil reaction‡	$\chi^2 = 9.08$	1	< 0.05	arch = 6.84 ± 1.76 (*n* = 95), native = 6.19 ± 2.10 (*n* = 2066)	Frank & Klotz (1990)
Ecological demands: nitrogen§	$\chi^2 = 24.43$	1	< 0.001	arch = 5.44 ± 1.95 (*n* = 110), native = 4.38 ± 2.37 (*n* = 1112)	Frank & Klotz (1990)
	ANOVA:				
Mean height	*F* = 0.033	1, 2493	n.s.	no difference	this study

*A temperature value 5 indicates species confined to mildly warm temperate areas, species assigned value 7 are indicators of warm regions (in Central Europe confined to lowlands), and value 6 is the transition between those (Frank & Klotz, 1990).
†Value 4 corresponds to the transition between indicator species of dry habitats (value 3) and of those with moderately moist soils (value 5). Value 6 indicates the transition between 5 and 7 (species confined to moist soils) (Frank & Klotz, 1990).
‡Value 7 indicates species occurring on slightly acid to slightly basic soils, avoiding strongly acid soils; value 6 represents the transition from 7 to indicators of moderately acid soils (value 5) (Frank & Klotz, 1990).
§Value 5 indicates species confined to moderately nitrogen-rich habitats and that only rarely occur in nitrogen-poor or nitrogen-rich conditions; value 3 indicates species that occur in nitrogen-poor habitats, and value 4 indicates the transition between those (Frank & Klotz, 1990).

Czech Republic is similar. Kühn & Klotz (2003) list 218 archaeophytes and 40 presumed archaeophytes, compared to 655 naturalized and frequent-casual neophytes. So archaeophytes make up 24.9–28.3% of the German alien flora compared with 24.1% in the Czech flora. In eastern Germany, archaeophytes constitute 8.4% of taxa (data based on Frank & Klotz, 1990) present before the start of the invasion of neophytes. The corresponding figure for the Czech flora is 10.8%. This close resemblance indicates that the conclusions drawn above for the Czech flora are probably also valid for German archaeophytes.

Among the 197 alien species in two Argentinean provinces, 64 species are archaeophytes in East Germany, making up 32.5% of the total (the difference between the proportion of German archaeophytes in East Germany and in Argentina is statistically significant, *G*-test on contingency tables, $\chi^2 = 78.09$, d.f. 1, *P* < 0.001). Archaeophytes contribute substantially to the pool of species analysed by Prinzing *et al.* (2002). However, archaeophytes are ecologically distinct from native taxa (Pyšek *et al.*, 2002; Kühn & Klotz, 2003); the differences in a number of traits are summarized in Table 1. Different families are typical for each group (Asteraceae are the commonest family in both; Rosaceae, Cyperaceae, Scrophulariaceae and Orchidaceae are over-represented in native species, whereas Poaceae, Brassicaceae, Chenopodiaceae, Lamiaceae, Caryophyllaceae and Apiaceae are over-represented in archaeophytes). Among Raunkiaer life forms (Ellenberg, 1988), therophytes and hemicryptophytes are

over-represented in archaeophytes, while geophytes, chamaephytes and phanerophytes are over-represented in natives. As to the life strategies (*sensu* Grime, 1979; where C means competitive, S stress tolerant and R ruderal), R and CR categories are over-represented among archaeophytes, whereas C, CS and CSR strategies are over-represented in natives. There is a significantly higher proportion of diploids in archaeophytes than in native species (64.2% vs. 44.7%), but triploids (0.0% vs. 5.2%), tetraploids (23.5% vs. 32.1%), pentaploids (0.7 vs. 2.4%) and octoploids (2.8% vs. 4.1%) are less well represented among the former. Both groups also differ in their ecological demands as expressed by Ellenberg indicator values. This system, based on field measurements of factors over the range of habitats occupied by a species, classifies species response to particular factors on an ordinal scale (Ellenberg *et al.*, 1991). Archaeophytes have higher demands for temperature and nitrogen and are confined to less acid and drier soils than native species (Table 1). Indeed, some traits that distinguish archaeophytes from native species also distinguish European species recorded as aliens in Argentina by Prinzing *et al.* (2002) from those European species that do not occur in this country: ruderal strategy, existence in dry, warm and nitrogen-rich conditions. This indicates that the effect of these traits on the results obtained by these authors was not negligible.

It could be argued that it is legitimate to consider archaeophytes among the species pool of the source area because they were present in Central Europe long before travel to Argentina started. However, this probably holds for some neophytes as well, and so the decision to exclude a particular group of species becomes a quantitative one (whereas the Pre-Columbian-Post-Columbian dividing line is based on a qualitatively distinct event). Moreover, we suggest that, for the sake of scientific consistency, archaeophytes and native species should not be treated as a single ecological/phytogeographical group. Having different traits (Table 1), the groups have different capacities to naturalize or invade in the New World. For that reason, they are better treated separately in analyses and, in our view, it therefore seems desirable that only native species should be considered for the source species pool in the first place.

The second issue relates to defining the source area. By including archaeophytes, Prinzing *et al.* (2002) did not reduce the geographical and natural-historical bias associated with the fact that western-Asian and Mediterranean species (most of the archaeophytes come from these regions; Zajac, 1979; Pyšek *et al.*, 2002) have experienced radically different associations with people than species of the Central-European floristic geoelement (Walter & Straka, 1970; di Castri, 1989). In addition, archaeophytes present in Central Europe whose native distribution area covers the Western Mediterranean, including Spain (from were the first colonists departed for Argentina; Prinzing *et al.* 2002), might have a very different probability of reaching Argentina than species also present in Eastern Germany as archaeophytes but having a native range centred on Western Asia and only reaching their NW distribution in the source region, i.e. East Germany.

Furthermore, this issue is not limited to archaeophytes and is relevant when the source species pool is limited to native species.

Prinzing *et al.* (2002) used all species native to Germany, regardless of their phytogeographical classification. Not all native species present in Central Europe belong to the Central-European floristic geoelement; some have their evolutionary centres in adjacent regions and Central Europe is at the edge of their distribution. Such species are still native but are of distinctly different original distribution, ranging across the whole Eurasia. There are enormous differences in the size of primary distribution ranges (Meusel *et al.*, 1965; Hultén & Fries, 1986). Hence, there is much variation in which part of the distribution range a species could have been moved to the target area from.

The major advantage of the known-source approach is that it reduces the extreme variation in the probability of particular species coming into contact with people and, even more importantly, with people that behave socially in a similar way, so that their lifestyle and chance of travelling to other regions is comparable, i.e. it reduces the bias in the chance of being transported to a source region. We believe that failure to consider these underlying circumstances compromises the potential of these generally valuable analyses.

Prediction: towards acceptable levels of accuracy and reliability

Prinzing *et al.* (2002) say that, based on the traits responsible for a species' success as an alien in Argentina, they were 'able to predict correctly 81% of alien species'. Prediction is an important goal of invasion biology and the known-source approach, if performed properly, has great promise because it reduces the bias otherwise associated with species' invasion history. It is necessary to point out some dangers in interpreting prediction models, again using Prinzing *et al.* (2002) as an example, as their calculations have been done impeccably. Logistic regression as used in their paper is undoubtedly a good tool and made better by their cross-validation. Its interpretation is, however, not straightforward. Essentially, it describes the differences between the two classes (successful and unsuccessful). This description may provide an explanation depending on what the characters are that produce a significant regression. A useful explanation is not necessarily a useful prediction (Williamson, 1999). This can be seen in Fig. 3, an adaptation of Fig. 1 of Prinzing *et al.* (2002).

Note first that although the ordinate runs from 0.0 to 1.0, it is not the probability of presence in Argentina. That interpretation of the logistic implies that species scoring more than 0.5 can be expected to be present. On that criterion, something like 90% of those present would be expected to be present, which is satisfactory, but over 75% of those not present would be expected, which is not at all what was intended.

The level on the ordinate used by Prinzing *et al.* (2002) of 0.9, shown as a dotted line in Fig. 3, is one possible value giving a good separation of the two distributions. An alternative at a slightly lower level would have the boxes of the box-and-whisker plots entirely on different sides of the line but that would not affect the argument here. About 81% (160/197) of the species recorded in Argentina are above the line and 70% (1408/2011) of those not recorded are below the line. These are the 'predicted'
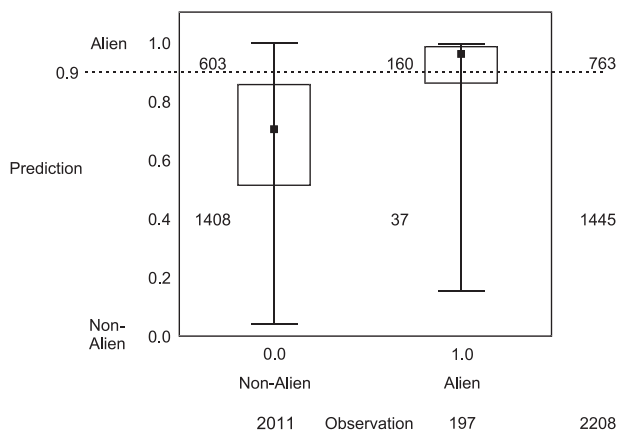
Figure 3 An annotated adaptation of Figure 1 of Prinzing *et al.* (2002). The ordinate (Prediction) values come from their logistic equation. A dotted line has been placed at the 0.9 cut off value they use. The left-hand box and whisker plot is for those species found in East Germany but not Argentina, the right-hand one for those found in Argentina. The four numbers in the boxed area are the total number of species above and below the dotted line in each column. The numbers outside the boxed area are the row and column subtotals and the grand total.

values in Prinzing *et al.* (2002). They are what Smith *et al.* (1999) call 'accuracy' or, in the medical literature (Loong, 2003), 'sensitivity' and 'specificity', respectively. They do not give the probability that a species with a character set found above the dotted line of Fig. 3 will be found as an alien. They give the probability that an alien will have those characters and, conversely, the probability that a German species not found in Argentina (a non-alien) will have characters giving a score below the dotted line.

Accuracy in the sense of Smith *et al.* (1999) is a vertical comparison in Fig. 3 and is a measure of explanation, not of prediction. A more useful comparison for prediction is the horizontal one, called by them 'reliability' and by Lonsdale & Smith (2001) and Loong (2003) 'predictive value'. This form of comparison gives the probability that a species with characters scoring above the line will be found as an alien (positive predictive value) and, conversely, that one with characters scoring below the line will not have been found (negative predictive value). From Fig. 3 these values can be seen to be 21% (160/763) and 97% (1408/1445), respectively; an excellent negative prediction but an unusable positive one. In general it is positive predictions that are needed. Using the criteria from the regression, it is quite easy to say what will not be recorded but distinctly difficult to say what will be.

The reason for the poor positive predictive value in these results is the number of 'false positives', those predicted to occur that do not, 603 of them. Looking at that another way, the 'prevalence' of aliens, of German species found in Argentina, in the total set is too low to allow any useful prediction; it is only 9% (197/2208). With such a low prevalence, such a high frequency of false positives and such a low positive predictive value, there seems little point in refining this particular analysis by distinguishing archaeophytes or species from different phytogeographical regions. Another reason for not doing further

analyses is that the Prinzing *et al.* (2002) rules attempt to predict a mixture of casual, naturalized and invasive species. It is likely that different rules would apply to these different stages (Williamson, 1999; Pyšek *et al.*, 2003; Cassey *et al.*, 2004).

CONCLUSIONS

We have drawn attention to several fundamental issues that affect the value of any analysis that attempts to delimit the determinants of invasiveness from the comparative analysis of species pools. Some of them, such as those related to the decisions about which species should be included in the source area species pool and the primary distribution ranges of species, are especially relevant when considering species originating from Europe. We suggest that these issues are not always understood in other parts of the World. Since so many of the World's major invaders come from Europe, we feel that it is desirable to clarify these issues. There is a wealth of information accumulated over centuries by botanists in Central Europe that is useful when attempting to predict invasiveness. Failing to consider certain underlying circumstances can limit the value of otherwise interesting, methodologically innovative, studies.

ACKNOWLEDGEMENTS

REFERENCES

Carey, J.R., Moyle, P., Rejmánek, M. & Vermeij, G., eds. (1996) Invasion biology. *Biological Conservation*, 78, 1–214.

Cassey, P., Blackburn, T.M., Jones, K.E. & Lockwood, J.L. (2004) Mistakes in the analysis of exotic species establishment: source pool designation and correlates of introduction success among parrots (Aves: Psittaciformes) of the world. *Journal of Biogeography*, **30**, 1–8.

di Castri, F. (1989) History of biological invasions with special emphasis on the Old World. *Biological invasions: a global perspective* (eds J.A. Drake, H.A. Mooney, F. di Castri, R.H. Groves, F.J. Kruger, M. Rejmánek and M. Williamson), pp. 1–30. John Wiley and Sons, Chichester.

di Castri, F., Hansen, A.J. & Debussche, M., eds. (1990) *Biological invasions in Europe and the Mediterranean Basin*. Kluwer Academic Publishers, Dordrecht.

Clement, E.J. & Foster, M.C. (1994) *Alien plants of the British Isles*. Botanical Society of British Isles, London.

Crawley, M.J., Harvey, P.H. & Purvis, A. (1996) Comparative ecology of the native and alien floras of the British Isles. *Philosophical Transactions of the Royal Society B*, **351**, 1251–1259.

Daehler, C.C. (1998) The taxonomic distribution of invasive angiosperm plants: ecological insights and comparison to agricultural weeds. *Biological Conservation*, **84**, 167–180.

Daehler, C.C. (2001) Darwin's naturalization hypothesis revisited. *American Naturalist*, **158**, 324–330.

Daehler, C.C. & Carino, D.A. (2000) Predicting invasive plants: prospects for a general screening system based on current regional models. *Biological Invasions*, **2**, 93–102.

Drake, J.A., Mooney, H.A., di Castri, F., Groves, R.H., Kruger, F.J., Rejmánek, M. & Williamson, M., eds. (1989) *Biological invasions: a global perspective*. John Wiley & Sons, Chichester.

Ellenberg, H. (1988) *Vegetation ecology of Central Europe*. Cambridge University Press, Cambridge.

Ellenberg, H., Weber, H.E., Düll, R., Wirth, V., Werner, W. & Paulissen, D. (1991) Zeigerwerte von Pflanzen in Mitteleuropa. *Scripta Geobotanica*, **18**, 1–284.

Essl, F. & Rabitsch, W., eds. (2002) *Neobiota in Österreich*. Umweltbundesamt GmbH, Wien.

Fitter, A.H. & Peat, H.J. (1994) The ecological flora database. *Journal of Ecology*, **82**, 415–425.

Forcella, F. & Wood, J.T. (1984) Colonization potentials of alien weeds are related to their 'native' distributions: implications for plant quarantine. *Journal of the Australian Institute of Agricultural Science*, **50**, 36–40.

Frank, D. & Klotz, S. (1990) Biologisch-ökologische Daten zur Flora der DDR. *Wissenschaftliche Beiträge Martin-Luther-University Halle-Wittenberg*, **32**, 1–167.

Goodwin, B.J., McAllister, A.J. & Fahrig, J. (1999) Predicting invasiveness of plant species based on biological information. *Conservation Biology*, **13**, 422–426.

Grime, J.P. (1979) *Plant strategies and vegetation processes*. John Wiley and Sons, Chichester.

Grotkopp, E., Rejmánek, M. & Rost, T.L. (2002) Toward a causal explanation of plant invasiveness: seedling growth and life-history strategies of 29 pine (*Pinus*) species. *American Naturalist*, **159**, 396–419.

Groves, R.H. & di Castri. F., eds. (1991) *Biogeography of Mediterranean invasions*. Cambridge University Press, Cambridge.

Haeupler, H. & Schönfelder, P. (1989) *Atlas der Farn- und Blütenpflanzen der Bundesrepublik Deutschland*. Eugen Ulmer Verlag, Stuttgart.

Holub, J. & Jirásek, V. (1967) Zur Vereinheitlichung der Terminologie in der Phytogeographie. *Folia Geobotanica and Phytotaxonomica*, **2**, 69–113.

Hultén, E. & Fries, M. (1986) *Atlas of North European vascular plants north of the topic of cancer*, Vol. 1–3 Koeltz Scientific Books, Königstein.

Jalas, J., Suominen, J. *et al.*, eds. (1972–1999) *Atlas Florae Europaeae distribution of vascular plants in Europe*, Vol. 1–12. The Committee for Mapping the Flora of Europe Societas Biologica Fennica, Vanamo, Helsinki.

Kennedy, T.A., Naeem, S., Howe, K.M., Knops, J.M.H., Tilman, D. & Reich, P. (2002) Biodiversity as a barrier to ecological invasion. *Nature*, **417**, 636–638.

Klironomos, J.N. (2002) Feedback with soil biota contributes to plant rarity and invasiveness in communities. *Nature*, **417**, 67–70.

Klotz, S., Kühn, I. & Durka, W. (2002) *BIOLFLOR — Eine Datenbank zu biologisch-ökologischen Merkmalen der Gefäßpflanzen in Deutschland. Schriftenreihe für Vegetationskunde*, **38**, 1–334.

Kolar, C.S. & Lodge, D.M. (2001) Progress in invasion biology: predicting invaders. *Trends in Ecology and Evolution*, **16**, 199–204.

Kubát, K., Hrouda, L., Chrtek, J. jun., Kaplan, Z., Kirschner, J., Štepánek, J. & Zázvorka, J., eds. (2002) *Key to the flora of the Czech Republic*. Academia, Praha. (in Czech).

Kühn, I. & Klotz, S. (2003) The alien flora of Germany — basics from a new German database. *Plant invasions: ecological threats and management solutions* (eds L.E. Child, J.H. Brock, G. Brundu, K. Prach, P. Pyšek, P.M. Wade and M. Williamson), pp. 89–100, Backhuys Publishers, Leiden.

Lonsdale, W.M. (1999) Global patterns of plant invasions and the concept of invasibility. *Ecology*, **80**, 1522–1536.

Lonsdale, W.M. & Smith, C.S. (2001) Evaluating pest-screening systems — insights from epidemiology and ecology. *Weed risk assessment* (eds R.H. Groves, F.D. Panetta and J.G. Virtue), pp. 52–60. CSIRO Publishing, Collingwood, Victoria.

Loong, T.W. (2003) Understanding sensitivity and specificity with the right side of the brain. *British Medical Journal*, **327**, 716–719.

Luken, J.O. & Thieret, J.W. (1997) *Assessment and management of plant invasions*. Springer Verlag, New York.

McNeely, J.A., Mooney, H.A., Neville, L.E., Schei, P.J. & Waage, J.K. (eds) (2001) *Global strategy on invasive alien species*. IUCN, Gland.

Meusel, H., Jäger, E. & Weinert, E. (1965) *Vergleichende Chorologie der zentraleuropäischen Flora*. Gusta Fischer Verlag, Jena.

Mooney, H.A. & Hobbs, R.J., eds. (2000) *Invasive species in a changing world*. Island Press, Washington DC.

Myers, J.H. & Bazely, D.R. (2003) *Ecology and control of introduced plants*. Cambridge University Press, Cambridge.

Preston, G., Brown, G. & van Wyk, E., eds. (2000) *Best management practices for preventing and controlling invasive alien species*. The Working for Water Programme, Cape Town.

Preston, C.D., Pearman, D.A. & Dines, T.D. (2002) *New atlas of the British and Irish flora*. Oxford University Press, Oxford.

Prinzing, A., Durka, W., Klotz, S. & Brandl, R. (2002) Which species become aliens? *Evolutionary Ecology Research*, **4**, 385–405.

Pyšek, P. (1995) Recent trends in studies on plant invasions (1974–1993). *Plant invasions: general aspects and special problems* (eds P. Pyšek, K. Prach, M. Rejmánek and P. Wade), pp. 223–236. SPB Academic Publishing, Amsterdam.

Pyšek, P. (1998) Is there a taxonomic pattern to plant invasions? *Oikos*, **82**, 282–294.

Pyšek, P. (2001) Past and future of predictions in plant invasions: a field test by time. *Diversity and Distributions*, **7**, 145–151.

Pyšek, P. (2003) How reliable are data on alien species in Flora Europaea? *Flora*, **198**, 499–507.

Pyšek, P., Prach, K., Rejmánek, M. & Wade, M., eds. (1995) *Plant invasions: general aspects and special problems*. SPB Academic Publishing, Amsterdam.

Pyšek, P., Richardson, D.M., Rejmánek, M., Webster, G., Williamson, M. & Kirschner, J. (2004) Alien plants in checklists and floras: towards better communication between taxonomists and ecologists. *Taxon*, **53**, 131–143.

Pyšek, P., Sádlo, J. & Mandák, B. (2002) Catalogue of alien plants of the Czech Republic. *Preslia*, **74**, 97–186.

Pyšek, P., Sádlo, J., Mandák, B. & Jarošík, V. (2003) Czech alien flora and a historical pattern of its formation: what came first to Central Europe? *Oecologia*, **135**, 122–130.

Rejmánek, M. (1996) A theory of seed plant invasiveness: the first sketch. *Biological Conservation*, **78**, 171–181.

Rejmánek, M. (2000) Invasive plants: approaches and predictions. *Austral Ecology*, **25**, 497–506.

Rejmánek, M. & Richardson, D.M. (1996) What attributes make some plant species more invasive? *Ecology*, **77**, 1655–1661.

Rejmánek, M., Richardson, D.M., Higgins, S.I., Pitcairn, M. & Grotkopp, E. (2004) Plant invasion ecology: state of the art. *Invasive alien species: a new synthesis* (eds H.A. Mooney, R.N. Mack, J.A. McNeely, L. Neville, P.J. Schei and J. Waage). Island Press, Washington, DC. in press.

Richardson, D.M., Pyšek, P., Rejmánek, M., Barbour, M.G., Panetta, F.D. & West, C.J. (2000) Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions*, **6**, 93–107.

Rothmaler, W., Jäger, E.J. & Werner, K. (2002) *Exkursionsflora von Deutschland*. Spektrum Akademischer Verlag, Heidelberg & Berlin.

Ryves, T.B., Clement, E.J. & Foster, M.C. (1996) *Alien grasses of the British Isles*. Botanical Society of British Isles, London.

Schroeder, F.G. (1969) Zur Klassifizierung der Anthropochoren. *Vegetatio*, **16**, 225–238.

Scott, J.K. & Panetta, F.D. (1993) Predicting the Australian weed status of southern African plants. *Journal of Biogeography*, **20**, 87–93.

Skalický, V. (1988) Phytogeographical characteristic. *Flora of the Czech Socialistic Republic*. (eds S. Hejný and B. Slavík), pp. 103–121. Academia, Praha. (in Czech).

Slavík, B. (1986–1998) *Phytocartographical syntheses of the Czech Republic*, Vol., pp. 1–3. Academia, Prague.

Smith, C.S., Lonsdale, W.M. & Fortune, J. (1999) When to ignore advice: invasion predictions and decision theory. *Biological Invasions*, **1**, 89–96.

Starfinger, U., Edwards, K., Kowarik, I. & Williamson, M. (eds.) (1998) *Plant invasions: Ecological mechanisms and human responses*. Backhuys Publishers, Leiden.

Thellung, A. (1905) Einteilung der Ruderal- und Adventivflora in genetische Gruppen. *Die Flora des Kanton Zürich, 1 Teil. Die Ruderal- und Adventivflora des Kanton Zürich* (eds O. Naegeli and A. Thellung). *Vierteljahrsschrift der Naturforschenden Gesellschaft Kanton Zürich*, **50**, 232–236.

Thompson, K., Hodgson, J.G. & Rich, T.C.G. (1995) Native and alien invasive plants: more of the same? *Ecography*, **18**, 390–402.

Tilman, D. (1999) The ecological consequences of changes in biodiversity: a search for general principles. *Ecology*, **80**, 1455–1474.

Tucker, K.C. & Richardson, D.M. (1995) An expert system for screening potentially invasive alien plants in South African fynbos. *Journal of Environmental Management*, **44**, 309–338.

Tutin, T.G., Heywood, V.H., Burges, N.A., Moore, D.M., Valentine, D.H., Walters, S.M. & Webb, D.A., eds. (1964–1980) *Flora Europaea* Vol. 1–5. Cambridge University Press, Cambridge.

Walter, H. & Straka, H. (1970) *Arealkunde*. Eugen Ulmer, Stuttgart.

Webb, D.A. (1985) What are the criteria for presuming native status? *Watsonia*, **15**, 231–236.

Weber, E.F. (1997) The alien flora of Europe: a taxonomic and biogeographic overview. *Journal of Vegetation Science*, **8**, 565–572.

Williamson, M. (1996) *Biological invasions*. Chapman & Hall, London.

Williamson, M. (1998) Measuring the impact of plant invaders in Britain. *Plant invasions: ecological mechanisms and human responses* (eds U. Starfinger, K. Edwards, I. Kowarik and M. Williamson), pp. 57–68. Backhuys Publishers, Leiden.

Williamson, M. (1999) Invasions. *Ecography*, **22**, 5–12.

Williamson, M. (2001a) Can the impacts of invasive species be predicted? *Weed risk assessment* (eds R.H. Groves, F.D. Panetta and J.G. Virtue), pp. 20–33. CSIRO Publishing, Collingwood, Victoria.

Williamson, M. (2001b) Can the impacts of invasive plants be predicted? *Plant invasions: species ecology and ecosystem management* (eds G. Brundu, J.H. Brock, I. Camarda, L.E. Child and M. Wade), pp. 11–19. Backhuys Publishers, Leiden.

Williamson, M. (2002) Alien plants in the British Isles. *Biological invasions: economic and environmental costs of alien plant, animal and microbe species* (ed. D. Pimentel), pp. 92–112. CRC Press, Boca Raton.

Williamson, M. & Fitter, A. (1996a) The characters of successful invaders. *Biological Conservation*, **78**, 163–170.

Williamson, M. & Fitter, A. (1996b) The varying success of invaders. *Ecology*, **77**, 1661–1666.

Williamson, M., Preston, C. & Telfer, M. (2003) On the rates of spread of British alien plants. *Plant invasions: ecological threats and management solutions* (eds L.E. Child, J.H. Brock, G. Brundu, K. Prach, P. Pyšek, P.M. Wade and M. Williamson), pp. 63–74. Backhuys Publishers, Leiden.

Wittig, R. (2002) *Siedlungsvegetation*. Eugen Ulmer, Stuttgart.

Wu, S.-H., Chaw, S. & Rejmánek, M. (2003) Naturalized Fabaceae (Leguminosae) species in Taiwan: the first approximation. *Botanical Bulletin of Academia Sinica*, **44**, 59–66.

Zajac, A. (1979) Pochodzenie archeofitów wystepujacych w Polsce. *Rozprawy Habilitaczne Uniwersiteta Jagiellonskogo Krakow*, **29**, 1–213.

Zajac, A. & Zajac, M., eds. (2001) *Distribution atlas of vascular plants in Poland*. Institute of Botany, Jagiellonian University, Cracow.