Reconstructing historical ranges with fossil data at continental scales

Jason S. McLachlan\textsuperscript{a,b,*}, James S. Clark\textsuperscript{a,c}

\textsuperscript{a}Department of Biology, Duke University, P.O. Box 90338, Durham, NC 27708, USA
\textsuperscript{b}University Program in Ecology, P.O. Box 90329, Duke University, Durham, NC 27708, USA
\textsuperscript{c}Nicholas School of the Environment, Duke University, Durham, NC 27708, USA

Abstract

Sedimentary records indicating the changing distribution of temperate tree species since the last ice age are widely used to understand the rates and patterns of population response to environmental change and the genetic consequences of such dynamics. It is well known that fossil pollen and plant macrofossil data provide limited information on distribution and abundance of small or diffuse tree populations. Nevertheless, the estimates of potential migration rates that are currently in use rely on assumptions about how fossil pollen and macrofossils record such populations. To better understand how such assumptions may influence estimates of spread rate, we related modern tree distributions to surface pollen abundance, using assumptions that are routinely applied to paleoecological data. We were unable to reconstruct the modern ranges of two species that are well represented in sediments using pollen and macrofossils from sediment surface samples. The ranges of American beech (\textit{Fagus grandifolia}) and yellow birch (\textit{Betula alleghaniensis}) include large areas where these trees are not abundant, and these were impossible to identify using pollen and macrofossil data, respectively. Our analysis supports the view that populations of these and other species existed at low densities over extensive areas in the past, and such areas would not be accurately mapped using fossil data. If so, estimates of past migration rates could be inaccurate.

\#2004 Elsevier B.V. All rights reserved.

Keywords: Fossil pollen; Macrofossil; Post-glacial migration

1. Introduction

One of the most persistent ideas to have emerged from the field of paleoecology is that the ranges of temperate tree species expanded rapidly after the last Ice Age. This consensus view has emerged from syntheses of fossil pollen data at continental scales over the course of the Holocene (Davis, 1983; Huntley and Birks, 1983; Delcourt and Delcourt, 1987; MacDonald, 1993). According to such reconstructions, species with a variety of life histories and dispersal mechanisms migrated northwards at rates of $10^2$ to $10^3$ m per year. Although skepticism about these reconstructions has occasionally arisen (Benett, 1985), the view that trees are capable of rapid migration has become paradigmatic. It has been incorporated in models used to assess the community and ecosystem consequences of contemporary climate change (van Minnen et al., 2000; Iverson et al., 1999; Malanson and Cairns, 1997). It has spurred analysis of long distance seed dispersal, because modern measurements of seed dispersal and life history are difficult to reconcile with such rapid
rates of migration (Clark et al., 2002; Higgins and Richardson, 1999). Fossil-based range reconstructions are also increasingly invoked to help explain population genetic structure within species at broad scales (Brewer et al., 2002; Petit et al., 1997).

Despite the widespread acceptance of continental-scale range reconstructions, paleoecologists have long recognized that species’ range limits are difficult to detect, even for taxa with identifiable pollen or macrofossils that are well represented in sediments (MacDonald, 1993; Davis et al., 1991). The distribution of fossil data from American beech (Fagus grandifolia) dating from the last glacial maximum (LGM) provides a useful illustration. Reconstructions of LGM ranges are important because the LGM was the coldest period in recent climatic history (Jackson et al., 2000). At temperate latitudes, thermophilous species may have consequently been restricted to small southerly ranges with implications for subsequent migration patterns and for population genetic structure. Fossil deposits from the LGM are few and of variable quality. Jackson et al. (2000) reviewed the available LGM fossil pollen data for eastern North America and selected 21 sites of reliable quality dated from 21.5 ± 1.5 ka (eliminating sites with uncertain age-models, taphonomy, etc.). Fig. 1A shows the distribution of sites containing beech pollen in any amount in the 20 sites from this group that we had access to via the North American Pollen Database (NAPD: http://www.ngdc.noaa.gov/paleo/data.html).

Although Fig. 1A suggests that beech was broadly distributed during the LGM, interpretation is frustrated by the fact that pollen can be transported long distances. Trace amounts of pollen can overestimate the range of a species. Accounting for this, past ranges are often delineated by an average threshold percentage (King and Herstrom, 1997; Brewer et al., 2002). For beech this threshold is often set at 0.5% (Webb, 1986; Davis et al., 1991; King and Herstrom, 1997). None of the LGM sediments shown in Fig. 1A meet this threshold (Fig. 1B). Macrofossil data can provide supporting evidence about species’ ranges (Jackson et al., 1997; Willis et al., 2000; Birks, 2003). Unlike pollen, they are typically only dispersed locally. For the same reason, however, the absence of macrofossils does not necessarily mean that populations were absent (Jackson et al., 1997). Beech macrofossils have been identified at a single site.

Fig. 1. Fossil data from beech in sediments dating to the last glacial maximum (adapted from Jackson et al., 1997, 2000). The distribution of (A) sites with beech pollen present, (B) fossil pollen sites with beech pollen >0.5%, and (C) beech macrofossils. White circle, absent; black circle, present. The thick black line approximately delineates the southern margin of the Laurentide Ice Sheet at the LGM.
during the LGM (Jackson et al., 2000; Delcourt et al., 1980, Fig. 1C).

The differences between the maps in Fig. 1 raise the question of how well we are able to identify species ranges in the past (and, by extension, how well we can estimate postglacial migration rates). Was beech widely distributed but rare in the late glacial (Fig. 1A)? Was it locally endemic in the lower Mississippi Valley (Fig. 1C)? Do we not have enough information to know (Fig. 1B)? All three scenarios have been suggested (c.f. Jackson et al., 2000; Bennett, 1985; Davis, 1983).

In this paper, we evaluate the evidence used to reconstruct species distributions in the past and to estimate migration rates. We use established paleoecological approaches to reconstruct the modern ranges of species that are well represented in by pollen and macrofossil data (beech and yellow birch (Betula alleghaniensis), respectively). We then compare these results to the modern distribution and abundance of these species. We focus on the reliability of approaches currently used to estimate species range limits with fossil data at broad scales. We address the prospects for future improvements in Section 5.

2. Study organisms

Beech (Fagus) has been widely studied by paleoecologists (Tsukada, 1982; Bennett, 1985; Dexter et al., 1987; Davis et al., 1991; Huntley et al., 1989; King and Herstrom, 1997). It is an important north-temperate genus represented by species common in late-successional forests in North America and Eurasia. Although beech pollen can only be identified to the generic level, a single species dominates in by pollen and macrofossil data (beech and yellow birch (Betula alleghaniensis), respectively). We then compare these results to the modern distribution and abundance of these species. We focus on the reliability of approaches currently used to estimate species range limits with fossil data at broad scales. We address the prospects for future improvements in Section 5.

3. Methods

We used existing data sets to map the modern range and abundance of tree populations and the distribution of pollen and macrofossils in modern sediments. Species ranges are from Thompson et al. (1999: http://pubs.usgs.gov/pp/1999/p1650-b/index.html).

We used the distribution of yellow birch macrofossils in modern sediments to evaluate their ability to identify the range of this species. We reconstructed the modern range of beech from surface pollen data using three criteria commonly used in the fossil pollen literature. The first criterion is a threshold percentage of a pollen type taken with respect to total arboreal pollen (King and Herstrom, 1998). We used this example, because, fossil pollen represents a single species in North America, and pollen is dispersed intermediate distances. Widely dispersed types, such as Pinus, provide little spatial information. Poorly dispersed types, such as Abies, are especially insensitive to low tree abundance. Beech is not well represented in macrofossil records due the absence of well-dispersed identifiable organs (Jackson et al., 1997), though Fagus bud scales have been used to detect local presence (Delcourt et al., 1938).

American beech currently ranges from Florida to the southern boreal forest (Thompson et al., 1999). Its northern distribution may be limited by frost damage at temperatures below −40 °C (Burns and Honkala, 1990). This mesophytic species is limited by soil moisture at its western range limit. Like American beech, yellow birch is most common in the northern hardwood forests of eastern North America. Farther south, it is found primarily along the Appalachian Mountains and has a patchy distribution in the lower Midwest (Burns and Honkala, 1990). Its pollen is indistinguishable from other birch species, but its fruits are small, abundant, well-dispersed, and identifiable in sediments. It has an unusually large macrofossil record (Jackson et al., 1997).

We used the distribution of yellow birch macrofossils in modern sediments to evaluate their ability to identify the range of this species. We reconstructed the modern range of beech from surface pollen data using three criteria commonly used in the fossil pollen literature. The first criterion is a threshold percentage of a pollen type taken with respect to total arboreal pollen in a sediment sample (King and Herstrom, 1997; Brewer et al., 2002). We chose a threshold value of 0.5% to be consistent with paleoecological published pollen data.
estimates of beech’s range (Webb, 1986; Davis et al., 1991; King and Herstrom, 1997). Of course, such thresholds are not independent of the relative abundance of other species (Fagerlind, 1952). Thus, our second criterion is presence–absence: We mapped the distribution of sites with beech pollen present in any abundance. Such trace abundances of pollen can be found well outside of the focal species’ range and the identification of single pollen grains in sediments might be subject to error. Smith and Pilcher (1973) addressed this problem by proposing an “empirical-limit”: taxa that do not produce abundant pollen could be considered locally present when their pollen is first consistently recovered from consecutive samples. Although it is impossible for us to directly apply this approach to modern surface sample data, our third criterion approximates the empirical limit concept. We constructed distribution maps for sites that have had beech pollen continuously present in three or more consecutive fossil pollen samples over the past 2000 radiocarbon years before present.

Comparisons of modern tree populations and the occurrence of pollen or seeds in lake deposits were used to evaluate the fossil indicators. A ‘false positive’ occurs when pollen is present, but derives from distant populations. A ‘false negative’ occurs when samples from deposits within the species range do not contain pollen or seeds. We evaluated the risk of false positives by calculating the proportion of sites outside the modern range of beech where beech was recorded as present. The risk of false negatives was evaluated by recording the basal area of beech in counties that contained surface sediments with beech pollen abundances below each threshold.

In the absence of sediment dating error, false positives are unlikely for macrofossil data (yellow birch seeds). We evaluated the prevalence of false negatives by calculating the proportion of sites within the range of yellow birch that contain macrofossils and by noting the relationship between false negatives and county basal area.

Estimates of population spread require that spatial information from individual sites are smoothed in some fashion. The way in which this smoothing is done affects estimates. Range maps have been interpolated subjectively (Davis, 1983), using linear models (Bernabo and Webb, 1977), and using moving windows (Huntley et al., 1989; Webb et al., 1993; Williams et al., 2004). We chose the latter approach because it can be compared to other recent studies. We followed Williams et al. (2004) and used a search window measuring 300 km × 300 km (horizontal) × 500 m (vertical) averaging pollen percentages within the window using a tri-cubic distance weighting. Kriging and inverse distance weighting gave similar results to the contours produced by our windowing technique using the 0.5% threshold (not shown).

4. Results

Two independent estimates of the contemporary range of beech and yellow birch are consistent within the US (Figs. 2A and 3A). Beech is most abundant in the Northeast and along the Appalachian Mountains, but is consistently present at lower abundances even at its range limits (Fig. 2A). Yellow birch is most abundant in the north of its range and has a patchy distribution between the southern Appalachians and the Upper Midwest (Fig. 3A).

The 0.5% threshold that is often used to delineate the range of beech in fossil assemblages does not correspond to the modern range limits of American beech (Fig. 2B and C). The 0.5% threshold approximately matches the northern limit of the species, as was found by Davis et al. (1991). However, false negatives predominate in its southern range, where beech is broadly distributed, but less abundant than in the north. Pines produce copious pollen and are abundant in the south, driving down the percentage of beech pollen in sediments. Only 25% of sites in the southern half of beech’s range exceeded the 0.5% threshold. In US counties in the top quartile of beech basal area, beech pollen was usually found in surface sediments, but over 60% of sites from counties in the lowest quartile did not reach the 0.5% threshold (Fig. 4). On the other hand, false positives were common beyond 100 km from the range limit (Fig. 5). Our interpolated contour of beech at the 0.5% threshold approximately identifies the species’ northern range limit but does not include most of its southern range (Fig. 2C). It is similar to the 1% beech contour of Williams et al. (2004) and Webb et al. (1998).

Lowering the threshold to include all surface samples with any beech pollen captures the southern
distribution of beech, but it increases the number of false positives outside of the species’ range (Fig. 2B and C). Sites over 500 km from the current range limit had trace abundances of beech pollen, and over 21% of sites between 100 and 200 km beyond the range limit contained beech pollen (Fig. 5). Within the US range of beech, over 40% of surface sediments in counties in the lower quartile of basal area contained no beech pollen (Fig. 4).

The risk of false negatives is expected to decrease in lakes with larger surface areas because larger lakes collect a higher proportion of regional pollen influx (Sugita, 1993). Surface area data are available for only a subset of lakes in the NAPD. Fig. 6 shows the relationship between Lake surface area and beech pollen abundance in sediment sites in counties where beech is rare. As expected, most false negatives come from smaller deposition basins, but pollen in trace amounts (<0.5%) are found in larger lakes as well.

Fossil pollen sites are an order of magnitude less common that surface samples. Consequently, we found fewer sites with three or more pollen assemblages within the last 2000 years, especially in the Southeastern US where sedimentary sites are less abundant (Fig. 2D). Sites consistently containing at least trace abundances of beech over this time period are found only in the northern two thirds of beech’s modern range.
Fig. 3. (A) The modern range and abundance of yellow birch (see caption Fig. 1). (B) The location of macrofossil surface sediment sites. Open symbols identify sites with no yellow birch macrofossils in surface sediments. Brown symbols indicate sites containing yellow birch macrofossils in surface sediments.

Fig. 4. False negatives. The percentage of sites with beech pollen absent (white) or present at <0.5% (black) found in counties in each quartile of beech basal area.

Fig. 5. False positives. The percentage of sites within distance classes of modern range of beech containing beech pollen >0.5% (black) and <0.5% (white).
Yellow birch macrofossils occur in only 20% of sites within the modern range of the species (Fig. 3B). In the US, the four counties containing yellow birch macrofossils in modern sediments are all in the top quartile of yellow birch abundance.

5. Discussion

Our results suggest that migration rates estimated from paleoecological reconstructions of tree populations could be extremely inaccurate. The limitations of sediment data that we examined here have long been recognized. Yet, the migration rates that derive from them have been widely accepted. We found that we could not systematically identify the range of American beech where it is currently a relatively rare tree, although beech pollen is abundant in sediments where beech trees are common (Fig. 4). It is not unusual for species to be rare across much or all of their range. Recall that during the LGM, many fossil assemblages contain trace quantities of beech pollen, but none reach the 0.5% threshold (Fig. 1).

All reconstructions of postglacial range distributions are based on data similar to those used in this study (Williams et al., 2004; Brewer et al., 2002; Webb et al., 1993; Delcourt and Delcourt, 1987; Huntley and Birks, 1983; Davis, 1983). Map contours in these studies are derived using a variety of methodologies, all of which appear reasonable. Inconsistencies between maps of the same taxa in these studies suggest that the signal in the data is so weak that model assumptions dominate the results (compare reconstructions for American beech in Williams et al., 2004; Delcourt and Delcourt, 1987; Benett, 1985; Davis, 1983).

The amount of uncertainty shown here suggests that the case for rapid postglacial migration is speculative. This does not mean that species did not sometimes migrate rapidly, only that the data provide limited evidence for it. Studies of contemporary invasions show that some species are capable of rapid migration (e.g., Mack, 1981). In order to make the same argument for climate driven range expansions in the early Holocene, paleoecologists will have to show that a standard methodology can accurately estimate the range limits of a species when conditions are unfavorable, such as during the LGM, when temperate species were rare throughout their ranges, fewer sediment sites existed, and pollen spectra were dominated by pine.

The difficulty of identifying the range of species at broad scales does not mean that fossil evidence is uninformative. Macrofossils and high pollen percentages in spatially contiguous sites do not occur fre-
quently outside species ranges (Jackson et al., 1997; Davis et al., 1991, Figs. 4 and 5). The same can probably be said for fossil stomata, which are more abundant than macrofossils in sediments (Parshall, 1999). These indicators are subject to false negatives at broad scales, however, so it cannot be assumed that sites without these indicators are outside of a species’ range. Networks of pollen sites may be able to distinguish local from long-distance pollen dispersal (Davis, 2000; Sugita, 1994), but such networks are not currently available for extensive areas and will never be available where deposition sites are rare, such as southeastern North America.

Population genetic data offer an independent source of information for identifying past range distributions under certain conditions. The small late-glacial populations that are difficult to detect in the fossil record leave a signal in the genetic structure of trees in temperate forests. Genetic drift in isolated glacial refuges is the force that leads to the largest component of molecular variation across the range of many species (Petit et al., 2002; Demesure et al., 1996). As species expanded during the postglacial period, founder effects in outlying populations also had a persistent genetic signal (Petit et al., 1997). An iterative approach, in which molecular data suggest geographic patterns that could be further explored with fossil evidence would take advantage of the complementary nature of the two data sources.

Modern life history evidence and modeling studies provide additional insight. In principal, long-distance dispersal can be invoked to ‘explain’ any estimate of spread rate. Yet, the techniques used to interpret past range limits and population spread has uniformly led to estimates of rapid migration, even for species that lack long-distance dispersal vectors, such as Carya. The weight of the dispersal evidence and modeling studies, combined with limitations of the paleoecological record, suggest that these estimates should not yet be the sole basis for models of contemporary population migration (Clark et al., 2001, 2003).

Acknowledgements

We thank the many contributors to the North American pollen and macrofossil databases of the NOAA paleoclimatology program, the US Forest Service Forest Inventory Analysis and the US Geological Survey for making their data available to the public. S.T. Jackson, T. Webb III, and J.W. Williams provided additional data and ideas. M. Rocca and M. Dietz provided technical assistance. A. Kremer, M. Wolosin, S. Ladeau, K. Brown, I. Ibanez, M. Dietz, and two anonymous reviewers provide helpful comments on the manuscript. We also thank the organizers of the 2002 DYGEN Conference for providing a stimulating intellectual forum. This research was supported by NSF grant DEB-0089769.

References


