Molecular Phylogenetics and Evolution 52 (2009) 303–311

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Phylogeographic structure of the chloroplast DNA gene pool in North American *Boechera* – A genus and continental-wide perspective

Christiane Kiefer [a], Christoph Dobeš [a,1], Timothy F. Sharbel [b], Marcus A. Koch [a,*]

[a] Heidelberg Institute of Plant Sciences, Department Biodiversity and Plant Systematics, University of Heidelberg, Im Neuenheimer Feld 345, D-69120 Heidelberg, Germany
[b] Apomixis Research Group, Dept. of Cytogenetics and Genome Analysis, Leibnitz Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Gatersleben, Germany

ARTICLE INFO

Article history:
Received 23 October 2008
Revised 10 March 2009
Accepted 11 March 2009
Available online 26 March 2009

Keywords:
Boechera
Brassicaceae
cpDNA
Phylogeography
Hybridisation

ABSTRACT

Continental-wide phylogeographic studies of plants in North America are rare. In our study we examined the phylogeographic history of *Boechera* (Brassicaceae) on a continental-wide scale testing if it is possible to do an analysis for 57 of the currently accepted taxa simultaneously. A large amount of haplotype sharing is explained both by recurrent hybridization and by non-differentiation of haplotypes since speciation. Hence, the chloroplast gene pool in *Boechera* predates speciation and therefore justifies the simultaneous analysis of a large number of taxa. Unrelated from taxon identity we can show that the evolutionary lineages detected have a different phylogeographic history in terms of glacial refugia and recently recolonised areas.

1. Introduction

Phylogeography examines the dispersal of taxa in space and time by relating molecular genetic analyses to geography. The term phylogeography was introduced approximately 20 years ago (Avise et al., 1987), and since then a multitude of studies investigating migration patterns of various plant and animal taxa have been published (for a statistical overview see Avise, 1998). Many of these studies focused on the evolutionary history of European and North American taxa during the Pleistocene (e.g. Brunsfeld and Sullivan, 2005; Koch et al., 2006; Schönswetter et al., 2002), and the field of comparative phylogeography later enabled researchers to reveal common patterns and trends (Hewitt, 2001; Soltis et al., 2006). Present-day distributions of species are not only shaped by habitat preferences but also by glaciation cycles with alternating cool and warm periods during the Quaternary which forced plants and animals to retreat into mostly southern located glacial refugia during colder periods. Warmer periods allowed the taxa to recolonise their original distribution ranges before they had to again retreat into refugia during the next glacial cycle. However, plants adapted to colder habitats could also migrate to and colonise new areas during periods in which temperatures decreased again in the changing climate and landscape which continuously provided new geographically defined corridors. The later fact might have promoted not only migration, but also speciation (Jordon-Thaden and Koch, 2008). These quaternary migrations left their footprints expressed in the geographical distribution of DNA-based polymorphisms, a phenomenon often described as "northern purity verses southern richness" (Hewitt, 2001).

Phylogeographic patterns in Europe were largely influenced by the Alps as an east-west oriented barrier for southwards migration and the Pyrenees as a barrier between the Iberian Peninsula and central Europe. Those two major obstacles lead to clear phylogeographic patterns in several plant and animal taxa e.g. refuge areas close to the eastern margin of the Alps or on the Balkans or differentiation of Iberian from central European populations (Hewitt, 2001; Taberlet et al., 1998).

Phylogeographic patterns in North America are less clear partly due to the north-south orientation of North American mountain chains which meant that plants and animals could migrate along the mountain chains into southern regions without having to cross them. However, a major refuge area for several plant and animal taxa was found in the southern central Rocky Mountains where the canyons were just deep enough to maintain warmer temperatures suitable for survival (Ayoub and Riechert, 2004; Brunsfeld and Sullivan, 2005; Mitton et al., 2000). For other plant and animal taxa glacial refugia were found in the Colorado Rockies and the eastern Great Basin (limber pine [*P. flexilis* James], Mitton et al., 2000) and the eastern Great Lakes (butterflies [*Lycaeides*], Nice et al., 2005).

Most phylogeographic studies focus on the migration of a single species. Often only small study areas are covered due to sampling

---

* Corresponding author. Fax: +49 6221 54 5508.
E-mail address: mkoch@hip.uni-heidelberg.de (M.A. Koch).

[1] Present address: Department of Pharmacognosy, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria.

1055-7903/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2009.03.016

of parts of the distribution ranges only or a narrow distribution of the species. However, in order to investigate large-scale phylogeographic patterns it is necessary to study a system with a wide distribution range (e.g. Jakob and Blattner, 2006; Koch and Matschinger, 2007). One of the few continental-wide phylogeographic studies in North America used three members of the genus *Boechera* (Brassicaceae): the former *Boechera holboellii* sensu Rollins (Rollins, 1993), *Boechera divaricarpa* sensu Rollins (Rollins, 1993) and *Boechera stricta*, respectively (Dobeš et al., 2004a). This study suggested the Rocky Mountains as a primary centre of genetic diversity, and the Sierra Nevada as a secondary centre of genetic diversity (Dobeš et al., 2004a). These three species were taken as representative for the whole genus since they have the widest distribution range of all *Boechera* species.

In the following study we performed a genus-wide phylogeographic study of *Boechera* by including 57 species [57 species according to the taxonomy of introduced by Al-Shehbaz (forthcoming FNA), but 46 species with several subspecies and varieties totaling to 57 taxa as well according to the taxonomy of Rollins (1993)] which represent half of the currently accepted 110 species (Al-Shehbaz, unpublished). If it is possible to do a genus-wide phylogeographic analysis using *Boechera* as a monophyletic group with a continental-wide distribution, this may represent an excellent model system for examining large-scale phylogeographic patterns on the North American continent. Hence we would like to compare patterns revealed by this large-scale study to those found in other studies in order to examine the extent of congruence between the results.

Furthermore genetic and species diversity should be compared for revealing the extent of congruence. Congruence of the centres of genetic and species diversity could be interpreted as long term stable populations or as regions subjected to multiple colonisation in which speciation took place over a long time. Incongruence, on the other hand, could be interpreted as either more recent speciation events during which no divergence of cpDNA haplotypes has yet happened or as populations in which speciation had already happened followed by bottlenecks and loss of genetic diversity.

Boechera together with seven other genera belongs to the tribe Boechereae (Bailey et al., 2006). The Boechereae are almost exclusively found in North America (one species in Siberia and Greenland only, respectively). According to the most recent taxonomic concept *Boechera* comprises 110 species. Crossing experiments (Roy, 1995; Schranz et al., 2005) and studies of nrITS variation (Koch et al., 2003) showed the high potential for hybridization within the various members of the genus. Hybrids are often triploid and reproduce asexually by apomixis (Dobeš et al., 2006). Until the complete revision of *Boechera* for the new Flora of North America (Al-Shehbaz, 2009; Windham and Al-Shehbaz, 2006; Windham and Al-Shehbaz, 2007a,b) triploid hybrids were lumped together with the diploid taxa to which they were most similar. Taxonomically they were often treated as a variety. Today the triploid apomicts are separated from the diploid sexuals and are described as distinct species. Hence the 110 recognized species contain 72 sexually reproducing diploid species and 38 triploid apomictic species. These *Boechera* apomicts have been investigated in various studies focusing on genome evolution (Kantama et al., 2007) or addressing the phenomenon of apomixis (Voigt et al., 2007). Here it is still unclear if apomixis is a driving evolutionary agent for speciation or simply the result of hybridization and genomic/genetic stabilization. Apomixis/polyploidy seems to have arisen repeatedly from a sexual background (Dobeš et al., 2004b; Sharbel and Mitchell-Olds, 2001) and has been extensively described for the former "*Boechera holboellii* sensu Rollins" (Rollins, 1993), an almost artificial taxon (Al-Shehbaz, pers. comm.; Windham and Al-Shehbaz, 2006). Hence, we examine both sexual and apomictic lineages within *Boechera* by reconstructing networks and phylogenetic trees based on chloroplast DNA marker sequences *trn*L-F and *rpo*C1 including almost 1300 accessions collected from herbarium specimens representing 57 taxa.

2. Materials and methods

2.1. Plant material

A total of 1286 accessions was examined using leaf material obtained from herbarium specimens from GH, MO and DAO and from collections of Thomas Mitchell-Olds (Duke University, North Carolina, USA). Corresponding accession details are listed in Supplement Table 1, the distribution of the samples is given in Supplement Fig. 7.

The sampling and accession list (Supplement Table 1), shows a large excess of accessions representing *B. holboellii*, *B. divaricarpa* and *B. stricta* [all three taxa sensu Rollins (1993)]. However, this imbalance in taxon sampling does not affect the outcome of the analysis since as we show inhere divergence of chloroplast evolutionary lineages predates speciation.

2.2. DNA extraction and sequencing

Total DNA extraction and PCR reactions were done as described in Dobeš et al. (2004a) with some minor changes. Prior to sequencing PCR products were purified with the NucleoFast Kit (Macherey & Nagel). Cycle sequencing was done in our lab or at the Genome Centre, IPK Gatersleben, using the DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Biosciences) and running a MegaBASE 500 sequencer. Sequencing was done in forward and reverse reaction.

2.3. DNA marker selection

We used various DNA regions from the chloroplast genome to reconstruct maternal phylogenetic lineages which were used to infer past migration history. The *trn*L-F region (*trn*L intron and *trn*L–*trn*F intergenic region separated by three X in the alignment) was analysed to characterise cpDNA haploytypes for further assessment of genetic variation and diversity. The same DNA region was also used to reconstruct haplotype networks and infer phylogenetic hypotheses. In order to increase the significance of networks and phylogenetic trees with respect to their major evolutionary lineages we also added sequence data from the *rpo*C1 intron to a subset of *trn*L-F haplotypes.

2.4. Alignments and haplotype definition

The forward and reverse sequences were aligned and trimmed. The *trn*L intron and the *trn*L-*trn*F intergenic spacer regions were assembled into one sequence (separated by three XXX) and missing bases at the 3′ and 5′ ends were substituted with N. The alignment was done manually using the program GenDoc according to the alignment published by Dobeš et al. (2004a). Haplotypes were defined by running the program TCS 1.21 (Clement et al., 2000). Gaps were set as a 5th state and the connection limit to 95%. In *Boechera* (as in other Brassicaceae genera e.g. *Arabidopsis*, *Cardaminopsis*, refer to (Koch and Matschinger, 2007; Koch et al., 2007), *Trn*F pseudogenes are present in the *trn*L-*trn*F intergenic spacer. In past studies in our group (Dobeš et al., 2004a; 2007; Schmickl et al., 2008) it was shown that pseudogene copies were gained and lost independently several times across the phylogenetic tree. As the mechanism through which such pseudogenes arise, multiply, or are deleted is unknown, we omitted them from the analysis. Neglecting the pseudogenes lead to the collapse of some haplotypes. These haplotypes are hereafter referred to as suprahaplotypes. Twenty suprahaplotypes (S1 to S20) were