Available online at www.sciencedirect.com
SCIENCE DIRECT®
Phytochemistry 66 (2005) 2836–2843
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

Chemical variation within and among six northern willow species

Tommi Nyman [a,*], Riitta Julkunen-Tiitto [b]

[a] Department of Biology, University of Oulu, Linnanmaa A1, P.O. Box 3000, FI-90014 Oulu, Finland
[b] Department of Biology, University of Joensuu, P.O. Box 111, FI-80101 Joensuu, Finland

Received 7 July 2005; received in revised form 29 September 2005
Available online 15 November 2005

Abstract

Plant tissues typically contain a diverse complement of secondary metabolites that serve as protection against various biotic and abiotic hazards. Chemical similarities are commonly used to infer phylogenetic relationships among plant taxa, but the studies are typically based on the mean concentration of each compound in each study species, thus overlooking within-species variability. In order to investigate patterns of intra- and interspecific chemical variation in plants, we measured the concentrations of condensed tannins and 36 other phenolic compounds in 120 leaf samples representing six northern *Salix* species. Multivariate clustering and ordination analyses of the data show that: (1) Despite considerable within-species variation in chemical profiles, intraspecific variability is on average lower than the variation among species. (2) Interspecific similarities are sensitive to the data analysis methods used, and different chemical classes produce partly contradictory results. (3) Compounds within each biosynthetic class tend to behave in a correlated manner and, consequently, overall chemical similarities are weakly correlated with the phylogeny of the studied species. The conclusion is that chemical data are poorly suited for phylogenetic inference, unless methods for data analysis are improved to take into account the biosynthetic routes by which the compounds are produced.

Keywords: *Salix*, Salicaceae; Willows; HPLC; Multivariate analysis; Chemosystematics; Herbivory; Phenolics; Salicylates; Condensed tannins

1. Introduction

In addition to the multitude of compounds that are needed for primary metabolic functions, plants produce a diverse array of chemicals that are collectively referred to as secondary metabolites (Seigler, 1998; Wink, 2003). Many of these compounds protect plants against abiotic hazards such as UV-radiation (Koes et al., 1993), or act as toxins and deterrents against herbivores and pathogens (Berenbaum, 1995; Hartmann, 1999; Wink, 2003). However, secondary compounds constitute a highly heterogeneous group which includes thousands of chemicals with varying structures that are produced via diverse and in many cases unconnected biosynthetic pathways (Berenbaum, 1995; Seigler, 1998; Wink, 2003; Martens and Mithöfer, 2005).

The taxonomic distribution of individual secondary compounds varies considerably. Some, such as many flavonoids, can be found in many distantly related plant groups (Harborne and Turner, 1984; Wink, 2003; Pelser et al., 2005; Martens and Mithöfer, 2005), whereas others may occur only in a few, closely related taxa or even in a single species (Julkunen-Tiitto, 1989; Jenett-Siems et al., 2005). Chemical similarities are frequently used to infer relationships among plant taxa in so-called chemosystematic studies (e.g., Julkunen-Tiitto, 1989; Cool et al., 1998; Keinänen et al., 1999a; Santos and Salatino, 2000; Nogueira et al., 2001), but the seemingly erratic occurrence of certain compounds in highly divergent plant groups has also led to doubts about the usefulness of chemical data in phylogenetic inference (Becerra, 1997; Grayer et al., 1999; Wink, 2003). A further complication is that chemosystematic analyses are typically based on the average concentration of each compound in each study species, which is potentially dangerous if the overall variation within species is

* Corresponding author. Tel.: +358 8 553 1807; fax: +358 8 553 1061.
E-mail address: Tommi.Nyman@oulu.fi (T. Nyman).

0031-9422/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2005.09.040

large. Many plant populations are known to harbour considerable variation in individual chemical profiles (Cool et al., 1998; Laitinen et al., 2000, 2005; Semmar et al., 2005; Windsor et al., 2005), but the relative level of within-species chemical variation in comparison to the variation among species remains largely unexplored.

The distribution of chemical variation within and among plant species is directly relevant for the evolutionary relationships between plants and their associated herbivores and pathogens. Especially herbivorous insects tend to utilize only one or a few related plant species, which is generally though to result from chemical differences among plant taxa (Ehrlich and Raven, 1964; Strong et al., 1984; Futuyma and Keese, 1992). However, host use is a dynamic trait which may change during the evolutionary history of an insect lineage, and factors that set the limits for host use will also influence the direction of host shifts. Phylogenetic studies have in most cases found distinct differences between the phylogenies of herbivorous insects and their host plants, but in some cases it has been possible to instead link the evolutionary history of host use to the degree of chemical similarities among host taxa (Becerra, 1997; Wahlberg, 2001).

Willows (*Salix* spp.) constitute a near-ideal model group for investigating intra- and interspecific chemical variation as well as chemistry-mediated evolutionary interactions between plants and their enemies. The genus includes over 400 species, and their ecological diversity ranges from creeping tundra species to large forest trees (Argus, 1997; Skvortsov, 1999). Willow tissues characteristically contain a diverse complement of phenolic compounds such as salicylates, cinnamic acid derivatives, flavonoids, and condensed tannins (Julkunen-Tiitto, 1989; Shao, 1991), many of which have been shown to affect host selection and performance of plant-feeding insects and mammals (e.g., Tahvanainen et al., 1985a,b; Lindroth et al., 1988; Ayres et al., 1997; Roininen et al., 1999; Ikonen et al., 2002; Simmonds, 2003). The herbivore communities of individual willow species often show marked differences, possibly as a result of interspecific variation in phenolic composition (Pasteels and Rowell-Rahier, 1992; Sipura, 1999). Indeed, the mean concentrations of many phenolics differ widely among species, and several studies have attempted to infer relationships among willows based on chemical similarities (Julkunen-Tiitto, 1989; Shao, 1991). Thus far the results have generally been ambiguous or contradictory and, as in most chemosystematic studies, individual species have been represented by the mean concentration in each measured compound, effectively disregarding intraspecific variation in overall chemistry.

The purpose of this study was to gain a more detailed view of intra- and interspecific chemical variability in plants. For this, we used multivariate clustering and ordination methods to analyze a dataset consisting of measurements of the concentrations of condensed tannins and 36 low-molecular weight phenolic compounds in 120 leaf samples representing six northern willow species. Specifically, we wanted to use the dataset as a model system to explore: (1) What is the level of overall chemical variation within species in relation to the variation among species? (2) How do different data analysis methods affect the results and inferences that can be made? (3) Do different chemical classes produce consistent results with respect to among-species chemical similarities? and (4) How do patterns in overall chemical similarities conform to the phylogenetic relationships among the study species? The results are discussed in relation to chemosystematic methods and current hypotheses on the ecology and evolution of plant–herbivore interactions.

2. Results and discussion

2.1. Chemical variation within and among species

The measured compounds mainly include various flavonoids, salicylates, and cinnamic acid derivatives (Table 1), and the leaf samples represent a total of 60 willow individuals; the two leaf samples from each plant were selected so that the upper leaf had a sawfly-induced leaf gall attached to the midrib, whereas the lower leaf was ungalled (see Section 3). Regardless, galled and ungalled leaves from the same individual tend to be clustered in pairs in the Unweighted pair-group method using arithmetic averages (UPGMA) clustering dendrogram (Fig. 1(a)). The pairing is not invariable, but within-individual distances are on average lower than distances among individuals (blocked multi-response permutation procedure tests within species, all $P < 0.001$). More importantly, leaf samples representing the same willow species are grouped together (Fig. 1(a)), so that it is possible to assign each sample to the correct willow species on the basis of its phenolic profile, despite considerable within-species variation. These results are very consistent irrespective of the standardization method (maximum/z-correction) and distance measure (Euclidean/correlation) used, or inclusion or exclusion of condensed tannins (results not shown). The maximum number of samples that are misplaced with regard to the willow species is two (when a z-transformation and Euclidean distances are used).

In contrast, similarities among species are clearly sensitive to changes in data analysis parameters. *Salix phylicifolia* and *S. lapponum* are grouped together in all results, but among-species similarities in the rest of the dendrogram show considerable variability depending on the standardization method and distance measure that are used. For example, *S. glauca* is clustered together with *S. reticulata* (instead of *S. myrsinites*) if Euclidean distances are used instead of correlation distances, and other alternative groupings can also be found (results not shown). In some cases, among-species similarities in the lower part of the dendrogram also change if only ungalled leaves are included, or if condensed tannins are excluded from the dataset. However, both of these factors appear to be of minor importance in comparison to the data analysis methodology.