

Species Delimitation under the General Lineage Concept: An Empirical Example Using Wild North American Hops (Cannabaceae: *Humulus lupulus*)

PATRICK A. REEVES AND CHRISTOPHER M. RICHARDS*

United States Department of Agriculture, Agricultural Research Service, National Center for Genetic Resources Preservation,
1111 South Mason Street, Fort Collins, CO 80521, USA;

*Correspondence to be sent to: United States Department of Agriculture, Agricultural Research Service, National Center for Genetic Resources
Preservation, 1111 South Mason Street, Fort Collins, CO 80521, USA; E-mail: crichard@lamar.colostate.edu.

Received 13 April 2009; reviews returned 1 October 2009; accepted 4 June 2010
Associate Editor: Mark Fishbein

Abstract.—There is an emerging consensus that the intent of most species concepts is to identify evolutionarily distinct lineages. However, the criteria used to identify lineages differ among concepts depending on the perceived importance of various attributes of evolving populations. We have examined five different species criteria to ask whether the three taxonomic varieties of *Humulus lupulus* (hops) native to North America are distinct lineages. Three criteria (monophyly, absence of genetic intermediates, and diagnosability) focus on evolutionary patterns and two (intrinsic reproductive isolation and niche specialization) consider evolutionary processes. Phylogenetic analysis of amplified fragment length polymorphism (AFLP) data under a relaxed molecular clock, a stochastic Dollo substitution model, and parsimony identified all varieties as monophyletic, thus they satisfy the monophyly criterion for species delimitation. Principal coordinate analysis and a Bayesian assignment procedure revealed deep genetic subdivisions and little admixture between varieties, indicating an absence of genetic intermediates and compliance with the genotypic cluster species criterion. Diagnostic morphological and AFLP characters were found for all varieties, thus they meet the diagnosability criterion. Natural history information suggests that reproductive isolating barriers may have evolved in var. *pubescens*, potentially qualifying it as a species under a criterion of intrinsic reproductive isolation. Environmental niche modeling showed that the preferred habitat of var. *neomexicanus* is climatically unique, suggesting niche specialization and thus compliance with an ecological species criterion. Isolation by distance coupled with imperfect sampling can lead to erroneous lineage identification using some species criteria. Compliance with complementary pattern- and process-oriented criteria provides powerful corroboration for a species hypothesis and mitigates the necessity for comprehensive sampling of the entire species range, a practical impossibility in many systems. We hypothesize that var. *pubescens* maintains its genetic identity, despite substantial niche overlap with var. *lupuloides*, via the evolution of partial reproductive isolating mechanisms. Variety *neomexicanus*, conversely, will likely persist as a distinct lineage, regardless of limited gene flow with vars. *lupuloides* and *pubescens* because of ecological isolation—adaptation to the unique conditions of the Rocky Mountain cordillera. Thus, we support recognition of vars. *neomexicanus* and *pubescens* as species, but delay making a recommendation for var. *lupuloides* until sampling of genetic variation is complete or a stable biological process can be identified to explain its observed genetic divergence. [AFLP; Cannabaceae; hops; niche model; phylogeography; species delimitation.]

“...how entirely vague and arbitrary is the distinction between species and varieties.”
(Darwin 1859, p. 48)

Historically, the two fundamental goals of systematic biology have been the discovery of new species and the reconstruction of organismal relationships. The need for species discovery and description (i.e., species delimitation) has only increased in past decades due to the extinction crisis (Pimm et al. 1995; Thomas et al. 2004). However, until recently, the attention of systematists has been focused on developing and applying new methods for building phylogenies. Important advances, beginning with the conceptualization of species, have fostered the reemergence of species delimitation as a major focus of systematic biology (Wiens 2007).

de Queiroz (1998, 2007) has argued that at the root of all modern species concepts, there is general agreement on the fundamental nature of species: species are separately evolving metapopulation lineages. What differs between concepts are the criteria, or lines of evidence, used to identify such lineages. For example, intrinsic reproductive isolation (Mayr 1942) and the existence of diagnostic characters (Cracraft 1983; Nixon and Wheeler 1990) are both properties of evolutionarily distinct lin-

eages. The observation of either property would suggest that the group in question is a species and that it may warrant taxonomic recognition as such. But, the order in which properties of lineages appear during cladogenesis, or whether they appear, cannot always be predicted, thus the application of several different criteria may be necessary. The perspective that species are lineages, and that multiple criteria may be used to identify them, has been termed the general lineage concept of species (de Queiroz 1998).

Numerous criteria for the identification of lineages have been proposed, some of which are integral to popular species concepts (de Queiroz 1998, 2007), whereas others are ad hoc (e.g., Pons et al. 2006; Knowles and Carstens 2007) (reviewed by Sites and Marshall 2003, 2004). The relative performance of some of these criteria under different evolutionary scenarios has been investigated (Hudson and Coyne 2002; Wiens and Penkrot 2002; Dettman et al. 2003; Marshall et al. 2006; Reeves and Richards 2007). It is reasonable to propose, as has de Queiroz (2007), that the greater the number of species criteria satisfied by a group, the more likely it becomes that the group is a distinct lineage—an identifiable biological entity on an independent evolutionary trajectory. At a minimum, as more criteria are satisfied, the

proposal that a group merits recognition as a species should become less contentious. In spite of offering a resolution to the “species problem” (de Queiroz 2005) and the potential for stable noncontroversial species circumscriptions, the general lineage concept of species has rarely been applied, although the practice is increasing (e.g., Marshall et al. 2006; Alström et al. 2008; Light et al. 2008; Padial and De La Riva 2009).

Under the general lineage concept, the only necessary property of a species is existence as a separately evolving metapopulation lineage (de Queiroz 2005). Thus, the concept itself is egalitarian in its treatment of species delimitation criteria: any criterion that identifies lineages identifies species. However, the tests that are employed to determine whether groups meet criteria are not necessarily equal in their abilities. For example, the commonly implemented test for monophyly, inspection of a reconstructed phylogenetic tree, may perform poorly for identifying lineages in the presence of gene flow (Reeves and Richards 2007) or due to errors associated with randomly sampling few individuals from a complex underlying genealogy (Rosenberg 2007). Additionally, in the case of clinal variation or isolation by distance, imperfect geographical sampling can lead to erroneous conclusions under some species criteria, most notably those reliant on tests of character divergence (see, e.g., Rosenberg et al. 2005; Schwartz and McKelvey 2009). Therefore, although all species criteria identify valid properties of lineages, the particular tests used to evaluate compliance with those criteria may vary in their efficacy.

We note two different categories of species delimitation criteria: those that seek to discern patterns in data consistent with evolution along lineages and those that attempt to identify evolutionary processes capable of maintaining distinct lineages. For example, monophyly (Donoghue 1985; de Queiroz and Donoghue 1988), fixed character state differences (i.e., diagnosability; Cracraft 1983; Nixon and Wheeler 1990), exclusive coalescence of alleles (Baum and Shaw 1995), and formation of a distinct genotypic cluster (i.e., absence of genetic intermediates; Mallet 1995) are pattern-oriented criteria, whereas intrinsic reproductive isolation (Mayr 1942), shared specific mate recognition (Paterson 1985), and occupation of a distinct niche (which implies adaptation; Van Valen 1976) are process-oriented criteria. One set of criteria proposes an evolutionary cause for the persistence of a lineage; the other reflects the effect of having existed as a lineage for some period of time. Without specific knowledge of the relative performance of the various tests for compliance with species criteria, affirmation from tests of complementary pattern and process-oriented species criteria would seem to provide strong corroborating evidence for the existence of a lineage and, hence, the delimitation of a species.

Our study organism, *Humulus lupulus*, or hops, is a perennial, dioecious, wind-pollinated vine in the family Cannabaceae. Wild populations are typically riparian, climbing to heights in excess of 6 m, supported

by deciduous trees. In recognition of differences between wild North American *Humulus* and Old World congeners, Nuttall (1848) described a single North American taxon, *H. americanus*, distributed throughout the United States. The specific epithet *americanus* was never uniformly used. Examination of distinctive specimens from New Mexico led to the proposal of the subspecific taxon *H. lupulus* variety *neomexicanus*, the first recognition of geographical differentiation within New World *Humulus* (Nelson and Cockerell 1903). Subsequently, var. *neomexicanus* was treated as a species, *H. neomexicanus*, by Rydberg (1917). Most recently and authoritatively, *H. lupulus* was divided into five taxonomic varieties based on morphology and geography (Small 1978, 1981). These varieties include 1) var. *lupulus*, which has a native distribution in Europe and Asia but has naturalized in many regions of the world following escape from commercial hop yards (female inflorescences of var. *lupulus* are used to impart characteristic aromatic and bitterness qualities to beer), 2) var. *cordifolius*, restricted to east Asia and Japan, 3) var. *neomexicanus*, a native of the arid western United States, 4) var. *pubescens*, from the Midwestern United States, and 5) var. *lupuloides*, distributed from the northern Great Plains eastward.

The three North American varieties together form a diagnosable monophyletic group distinct from European and Asian varieties (Pillay and Kenny 1996; Murakami et al. 2006a, 2006b). The geographic distributions of the varieties overlap at their peripheries (Fig. 1) suggesting the possibility for gene flow. Although numerous studies have described genetic variation in

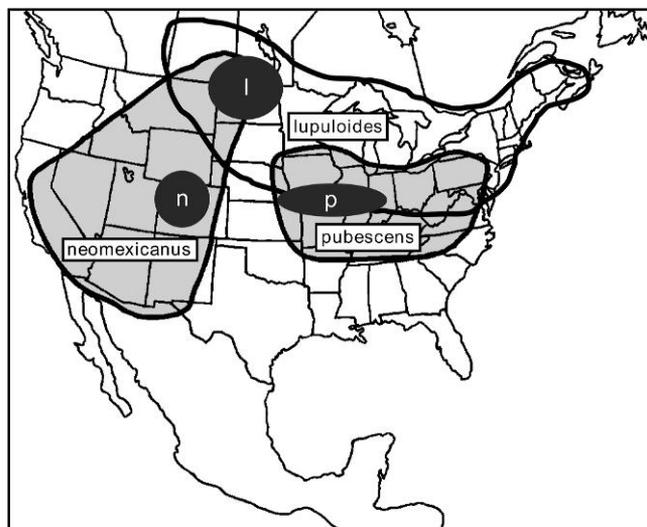


FIGURE 1. Approximate ranges of North American *Humulus lupulus* varieties (based on Small 1978). Shaded ovals indicate regions where DNA samples were taken from individuals conforming to each variety. Variety *lupuloides* (“l”) was sampled from a region where it cooccurs with var. *neomexicanus*, var. *pubescens* (“p”) was sampled from a region where it cooccurs with var. *lupuloides*, and var. *neomexicanus* (“n”) was sampled in allopatry.

cultivated hops (Small 1981; Brady et al. 1996; Hartl and Seefelder 1998; Šuštar-Vozlič and Javornik 1999; Seefelder et al. 2000; Jakše et al. 2001; Patzak 2001; Stajner et al. 2008) and its wild progenitors (Small 1978; Stevens et al. 2000; Henning et al. 2004; Murakami et al. 2006a, 2006b; Townsend and Henning 2009), the extent and mechanisms of evolutionary isolation among *H. lupulus* varieties in North America have not been specifically examined.

Wild North American hops cannot be directly used in brewing because of undesirable chemical properties that produce excessive bitterness and objectionable aromas. However, many modern hop cultivars with desirable pathogen resistance and elevated bittering properties trace a portion of their pedigree to crosses between European cultivars (var. *lupulus*) and a wild individual (var. *lupuloides*) from Manitoba (Salmon 1934; Neve 1991). North American *H. lupulus* continues to be explored as a source for agronomically desirable traits such as disease, drought, and insect resistance (Hampton et al. 2001).

We are interested in determining whether the three named taxonomic varieties of hops native to North America form evolutionarily distinct lineages. North American *H. lupulus* shares with many temperate plant species factors that complicate species delimitation such as broad geographic distribution, incomplete reproductive isolation, and distributional overlap among groups. Wild hops thus represent an intriguing case for applying the principles of the general lineage concept to the problem of species delimitation. Moreover, the identification of discrete evolutionary lineages among relatives of crop species should facilitate the use of wild genetic resources in crop improvement programs, a strategy receiving increased attention (Tanksley and McCouch 1997; Gur and Zamir 2004). Genes and gene complexes responsible for adaptation to evolutionary challenges should occur at high frequency in lineages persisting under the environmental conditions presenting the challenge. Thus, in our view, the accurate identification of distinct evolutionary lineages has economic as well as taxonomic value.

In this study, we apply the following five criteria for species delimitation to DNA polymorphism, natural history, and distributional data from wild North American *H. lupulus*:

1. monophyly (Donoghue 1985; de Queiroz and Donoghue 1988),
2. diagnosability, that is, the appearance of fixed differences (Cracraft 1983; Nixon and Wheeler 1990)
3. absence of genetic intermediates (Mallet 1995)
4. intrinsic reproductive isolation (Mayr 1942), and
5. niche specialization (Van Valen 1976).

Whether varieties merit taxonomic recognition as species is discussed in light of the plurality of evidence in favor of their existence as distinct evolutionary lineages.

MATERIALS AND METHODS

DNA Sampling and Amplified Fragment Length Polymorphism Analysis

Leaf tissue was collected from 131 *H. lupulus* var. *lupuloides* individuals from 29 wild populations in the Great Plains of southern Canada and the northern United States (Table 1). Samples were collected from riparian areas and river terraces along the Souris, Qu'Appelle, and Assiniboine rivers of the Red River drainage, and the Missouri, Knife, Little Knife, and White Earth rivers of the Mississippi River drainage. Wild var. *lupuloides* populations from this region of Manitoba, Saskatchewan, and North Dakota may be an important source of genes for crop improvement (Hampton et al. 2001). Seventeen *H. lupulus* var. *pubescens* individuals were sampled from four wild populations along the Missouri River in southeastern Nebraska. Two additional individuals of var. *pubescens* and 9 var. *neomexicanus* individuals were sampled from herbarium specimens (Table 2).

Genomic DNA was extracted from fresh leaves (preserved in vapor phase of liquid nitrogen) or herbarium tissue using the DNeasy 96 Plant Kit (Qiagen). Amplified fragment length polymorphism (AFLP) reactions were performed using the method of Vos et al. (1995) as modified by Marques et al. (1998) and Myburg et al. (2001). Ligation reactions were diluted 10-fold prior to preamplification. Preamplification reactions were diluted 1:40 prior to selective amplification. Eighteen duplexed primer pair combinations, chosen as optimal from a preliminary screen of 32, were used for selective amplification (see online Appendix 1, <http://www.sysbio.oxfordjournals.org>). AFLPs were visualized on a model 4200 LI-COR automated DNA sequencer following Myburg et al. (2001). Polymorphic loci from 50 to 500 bp were manually assigned binary scores (present = 1, absent = 0) using SAGA MX software (version 2.1; LI-COR). AFLP quality and comparability were ensured by confirming that monomorphic bands observed in samples derived from fresh material were also found in herbarium samples, following Lambertini et al. (2008). To establish scoring error rates, DNA from three individuals was subjected to the experimental protocol in duplicate, from DNA extraction through scoring.

Phylogenetic Analyses

To test for monophyly of varieties, two approaches were used. First, the data set containing 159 individuals and 555 AFLP characters was subjected to parsimony analysis using PAUP 4.0b10 (Swofford 1999). The heuristic search option and tree bisection-reconnection (TBR) branch swapping were used with the MULTREES setting "on," and characters were treated as equally weighted following the suggestions of Koopman (2005). Maximum parsimony searches were conducted using an idle-time distributed computing cluster (Reeves et al. 2005). Jobs were run until >3000 random-order taxon addition replicates had been completed (3681 in

TABLE 1. Wild collections of *Humulus lupulus* in North America

Population identifier	State/province	County/census division	Drainage	River	Latitude	Longitude	<i>n</i>	Number of polymorphic loci
A	North Dakota	Ward	Red River	Souris	48.1587°	-101.1603°	5	84
B	North Dakota	Ward	Red River	Souris	48.2046°	-101.2343°	5	101
C	North Dakota	Ward	Red River	Souris	48.2808°	-101.4323°	6	104
D	North Dakota	Ward	Red River	Souris	48.2869°	-101.4520°	6	130
E	North Dakota	Mountrail	Mississippi River	White Earth	48.3296°	-102.7610°	8	124
F	North Dakota	Mountrail	Mississippi River	White Earth	48.3142°	-102.7640°	4	91
G	North Dakota	Mountrail	Mississippi River	Little Knife	48.1495°	-102.4561°	4	21
H	North Dakota	Mountrail	Mississippi River	Little Knife	48.1337°	-102.4393°	5	110
I	North Dakota	Renville	Red River	Souris	48.7640°	-101.7764°	5	78
J	North Dakota	Renville	Red River	Souris	48.9724°	-101.9521°	7	89
K	North Dakota	Renville	Red River	Souris	48.9677°	-101.9497°	2	69
L	Saskatchewan	1, Estevan	Red River	Souris	49.1823°	-102.0259°	5	105
M	Saskatchewan	1, Estevan	Red River	Souris	49.2199°	-102.1807°	8	171
N	Saskatchewan	6, Regina	Red River	Qu'Appelle	50.6563°	-103.5879°	4	106
O	Saskatchewan	6, Regina	Red River	Qu'Appelle	50.6253°	-103.5443°	4	115
P	Saskatchewan	6, Regina	Red River	Qu'Appelle	50.6115°	-103.5561°	1	n/a
Q	Saskatchewan	6, Regina	Red River	Qu'Appelle	50.5955°	-103.4630°	3	55
R	Saskatchewan	6, Regina	Red River	Qu'Appelle	50.5722°	-103.4108°	2	32
S	Saskatchewan	6, Regina	Red River	Qu'Appelle	50.5588°	-103.3273°	5	100
T	Saskatchewan	5, Melville	Red River	Qu'Appelle	50.5266°	-103.2573°	7	135
U	Saskatchewan	5, Melville	Red River	Qu'Appelle	50.6355°	-102.8786°	7	128
V	Saskatchewan	5, Melville	Red River	Qu'Appelle	50.6441°	-102.8229°	5	100
W	Saskatchewan	5, Melville	Red River	Qu'Appelle	50.6294°	-102.7774°	3	104
X	Saskatchewan	5, Melville	Red River	Qu'Appelle	50.5392°	-102.4786°	1	n/a
Y	Manitoba	9, Portage la Prairie	Red River	Assiniboine	49.9462°	-98.0331°	5	57
Z	Manitoba	7, Brandon	Red River	Souris	49.5355°	-100.0549°	1	n/a
AA	Manitoba	7, Brandon	Red River	Souris	49.6130°	-100.2344°	7	146
AB	North Dakota	McHenry	Red River	Souris	48.4150°	-100.3917°	2	49
AC	North Dakota	Mercer	Mississippi River	Knife	47.3281°	-101.4287°	4	75
AD ^a	Missouri	Holt	Mississippi River	Missouri	40.0594°	-95.3993°	5	148
AE ^a	Missouri	Holt	Mississippi River	Missouri	40.0613°	-95.3920°	6	164
AF ^a	Missouri	Holt	Mississippi River	Missouri	40.0601°	-95.3831°	3	158
AG ^a	Nebraska	Nemaha	Mississippi River	Missouri	40.3750°	-95.6557°	3	39

Note: n/a, not applicable.

^aVar. *pubescens*, remainder are var. *lupuloides*.

total). Support values for clades in the strict consensus tree were estimated using 10,000 bootstrap replicates (Felsenstein 1985), the heuristic search option, one random-order taxon addition per replicate, TBR branch swapping, and MULTREES off, a strategy that produces accurate bootstrap proportions while minimizing search time (DeBry and Olmstead 2000). Resulting trees were rooted using midpoint rooting, a method shown to perform well in data sets sampled at low taxonomic levels (Hess and De Moraes Russo 2007). Rooted parsimony trees were then examined to determine the degree of support for monophyletic varieties.

Second, the relaxed-clock method of Drummond et al. (2006), which allows the position of the root of a tree to be estimated simultaneous to the phylogeny (Renner et al. 2008), was used to estimate the posterior probability of monophyly for each variety in a Bayesian Markov chain Monte Carlo (MCMC) framework. AFLP data were analyzed using Beast 1.4.8 (Drummond and Rambaut 2007) under a generalized binary reversible substitution model, with branch rates uncorrelated and drawn from a lognormal distribution, and a birth/death prior on tree shape. Two replicate Markov chains were run, each with a total of

TABLE 2. Herbarium specimens of North American *Humulus lupulus* used in AFLP study

Variety	Voucher (Herbarium)	Collection date	Locality ^a
<i>neomexicanus</i>	Phillips BDSR-607 (CSU)	26 July 1972	Larimer, CO
<i>neomexicanus</i>	Mooradian 72-269 (CSU)	14 July 1972	Jefferson, CO
<i>neomexicanus</i>	Harrington 8443 (CSU)	8 August 1956	Moffat, CO
<i>neomexicanus</i>	Harrington 8129 (CSU)	26 July 1955	Routt, CO
<i>neomexicanus</i>	Harrington 4768 (CSU)	11 August 1950	Larimer, CO
<i>neomexicanus</i>	Tabar and Walker 361 (CSU)	16 July 1981	Moffat, CO
<i>neomexicanus</i>	Barrell 185-60 (CSU)	6 August 1960	Gunnison, CO
<i>neomexicanus</i>	Baker and Naumann 82-350 (CSU)	27 August 1982	Rio Blanco, CO
<i>neomexicanus</i>	O'Kane 401 (CSU)	14 August 1981	n/a
<i>pubescens</i>	Friesner 17948 (RM)	11 September 1943	Boone, IN
<i>pubescens</i>	Davis 3784 (RM)	10 February 1916	Ralls, MS

Note: n/a, not available.

^aUS county and state.

5×10^8 steps. Trees were sampled every 5000 steps after the stable posterior distribution of trees was reached (requiring 2.5×10^8 steps). Stationarity and convergence between runs were evaluated by inspection of tree likelihood values using Tracer 1.4 (Rambaut and Drummond 2007). Likewise, mixing was evaluated using effective sample sizes (ESS) of likelihood parameters. In addition to the relaxed-clock model, a nonclock Bayesian likelihood approach that uses a binary stochastic Dollo substitution model (Alekseyenko et al. 2008) was applied (analyzed with Beast 1.5b3). Analogous to Dollo parsimony, under a stochastic Dollo model, the likelihood of the tree is sensitive to the position of the root, hence the model offers an alternative means to find the root position without using an outgroup or a clock. For Dollo runs, two Markov chains were run for 8×10^7 steps. Trees were sampled every 100 steps after a burn-in of 7×10^7 steps, when the tree likelihood had reached apparent stationarity. For both relaxed-clock and Dollo models, the posterior probability of monophyly was estimated as the frequency at which each variety was observed to be monophyletic in the posterior distribution of trees. Beast XML input files containing AFLP genotypes and model details are provided in online Appendix 2, <http://www.sysbio.oxfordjournals.org>.

To test whether named varieties were diagnosable, population aggregation analysis (PAA; Davis and Nixon 1992) was applied. PAA is a procedure for delimiting species via the identification of fixed character differences between sampled populations or groups of populations. Because sampling is seldom exhaustive, in most applications, PAA only identifies fixed character differences between samples. From this, the inference is drawn that the character is fixed within the population or species and is therefore diagnostic (Wiens and Servedio 2000). The estimated scoring error rate per band (0.04685, see online Appendix 3, <http://www.sysbio.oxfordjournals.org>) was used to compensate for AFLP scoring errors that could cause a character to appear polymorphic during PAA, when in fact it was fixed within the sample. Given an error rate of 0.04685 and 131 samples of var. *lupuloides*, the actual minimum number of samples required to define a locus as “fixed” for var. *lupuloides* was adjusted to 125 ($\approx 131 - 131 \times 0.04685$). The correction for the 19 var. *pubescens* and 9 var. *neomexicanus* individuals was less than one individual, so no adjustment was made.

Population Structure Analyses

Humulus is capable of rhizomatous growth, so independently rooted plants could be clones. Here, clones were defined operationally as individuals between which the genetic distance was less than or equal to that between known replicates (i.e., less than or equal to the scoring error rate of 0.04685). When necessary for analysis, the genotypic data from a single randomly selected individual was chosen to represent all members of a set of clones.

To test for genetic subdivision consistent with the genotypic cluster species criterion (Mallet 1995), two analytical procedures were used: statistical inference of clusters in principal coordinate space (PCO-MC; Reeves and Richards 2009), and a Bayesian procedure that minimizes the deviation from Hardy–Weinberg and linkage equilibrium across a data set by the fractional assignment of individual genomes to K populations (STRUCTURE; Pritchard et al. 2000). PCO-MC analysis followed prior recommendations (Reeves and Richards 2009). To maximize sensitivity to subtle population structure while minimizing type I error, the P value cutoff was set to 0.9999, the stability cutoff to 15. A second analysis was performed with a P value cutoff of 0.05 to test for significantly distinct clusters. All principal coordinate axes were used. The data set for STRUCTURE analyses excluded clones. AFLPs were coded as dominant data (Falush et al. 2007). The model assumed admixture of individuals and correlated allele frequencies, which improves sensitivity to subtle population differentiation (Pritchard et al. 2007). The number of populations (K) was estimated using the ΔK method of Evanno et al. (2005). Twenty replicate runs were performed, with a burn-in period of 100,000 generations followed by 100,000 sampled generations. Label switching between replicates was addressed using CLUMPP (Jakobsson and Rosenberg 2007). Multimodality among replicates was addressed by reporting the average membership coefficients calculated from the permuted Q-matrices by CLUMPP. We also calculated the stability coefficient (S_N) of Richards et al. (2009), which indicates the correlation in assignment between runs as a value scaled from 0 (no correspondence) to 1 (perfect correspondence).

Environmental Niche Modeling

Environmental niche modeling was used to estimate the extent to which varieties inhabit distinct habitat. Locality information was obtained from the wild populations sampled for DNA, the USDA Germplasm Resources Information Network (<http://www.ars-grin.gov/npgs/>), and herbarium collections (ARIZ, ASU, BHSC, BUT, CDA, CS, ILLS, ISC, KANU, KSC, MIN, MO, MONT, NEB, OAC, NY, UVSC, WIS, WTU). When specimens had not been identified to variety, determinations were made using Small's (1978) keys, except that geographical information was not used because 1) we did not want to bias our niche predictions and 2) most keys for *Humulus* do not include a geographical component. For specimens without precise latitude and longitude data, coordinates were obtained using locality descriptions and GOOGLE EARTH 4.2 (<http://earth.google.com>). With few exceptions, only those localities that could be georeferenced with error less than 1.14 km (the diagonal length of a quarter section) were included. Care was taken to ensure that localities within regions of sympatry were sampled. A total of 350 unique localities were used for model

estimation (153 var. *lupuloides*, 64 var. *pubescens*, and 133 var. *neomexicanus*).

Niche models were computed using MAXENT 3.2.1 (Phillips et al. 2006). MAXENT calculates a relative index of environmental suitability (interpretable as the probability of species occurrence) across a geographical area using species distribution and environmental data. The maximum-entropy method of model estimation used in MAXENT performs well (Elith et al. 2006; Hernandez et al. 2006) and only requires presence data, but relevant environmental variables must be included. The BioClim data layers, 19 biologically relevant climatic variables obtained from the WorldClim data base (Hijmans et al. 2005), were utilized (see online Appendix 1, <http://www.sysbio.oxfordjournals.org>). Thirty arc-second global grids were cropped to cover the known native range of *H. lupulus* in North America then subjected to analysis using MAXENT.

Our analysis followed that of Raxworthy et al. (2007). Default values were accepted for the “maximum iterations” (500) and “convergence threshold” (1E-5) parameters, which determine the stopping point for the maximization algorithm. The “regularization multiplier,” which controls the degree of over- or underfitting of the model, was set to 1, the default. Duplicate localities were eliminated. Twenty replicate runs were conducted, with 25% of samples randomly set aside as test localities. Validation of models included an examination of extrinsic omission rates (the fraction of test sites not predicted to be suitable habitat) and a binomial test that determines whether the computed model predicts presence at test sites significantly better than a random model (Anderson et al. 2002; Phillips et al. 2006).

To determine whether niche specialization could be inferred among varieties, five taxonomic scenarios were considered:

1. (*lupuloides*, *pubescens*, *neomexicanus*),
2. (*lupuloides*, *pubescens*), (*neomexicanus*),
3. (*lupuloides*, *neomexicanus*), (*pubescens*),
4. (*pubescens*, *neomexicanus*), (*lupuloides*), and
5. (*lupuloides*), (*pubescens*), (*neomexicanus*).

We asked: under which scenario is the geographic distribution of *H. lupulus* in North America best explained by environmental data? This is equivalent to asking: under which scenario can the most probable model be constructed? Or, operationally, is the ability to predict presence at test sites improved by separating individuals by variety? Therefore, for each scenario, the average probability of occurrence across all test sites was calculated. The taxonomic scenario that produced niche models with the highest mean probability of presence at test sites was preferred. Variability in niche predictions caused by random selection of training and test localities was dealt with using the 20 replicate runs. This resampling strategy allowed the error variance around the mean probability of occurrence at test sites to be estimated. The relative predictive power of models created under each taxonomic scenario was then compared using *t* tests.

To estimate niche overlap, we used the statistic *D*, developed by Schoener (1968), and applied to environmental niche models by Warren et al. (2008). *D* describes the difference between two niche models in the predicted probability of presence across a study area, scaled from 0 (no overlap) to 1 (identical models). This statistic is useful because it does not require the conversion of continuously distributed probability of presence values into a binary prediction of cell occupancy via the arbitrary selection of a threshold value defining taxon presence. Measures of niche overlap can be sensitive to choice of threshold (Warren et al. 2008). We calculated a mean estimate of niche overlap (D_m) as the average *D* across replicate runs.

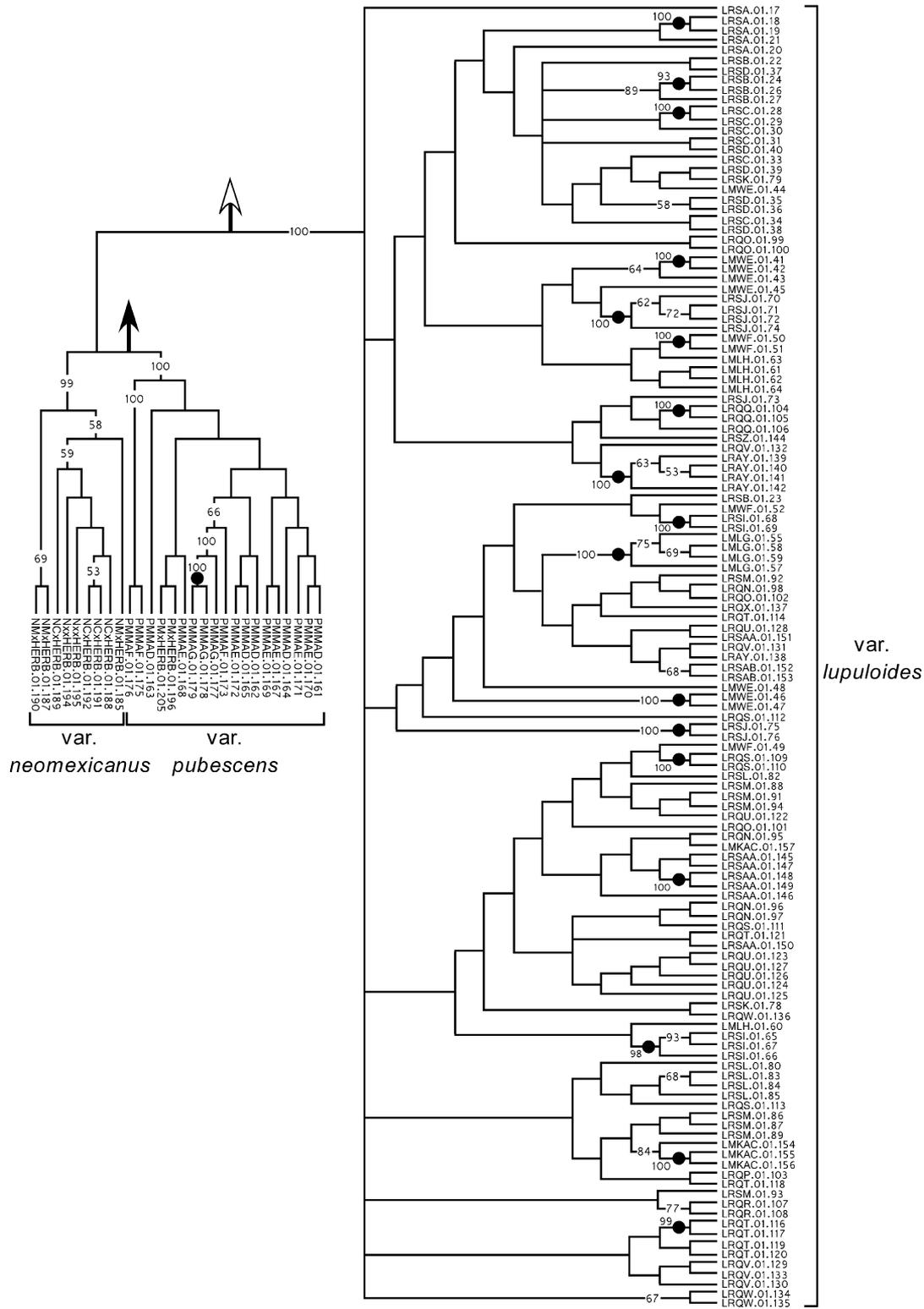
RESULTS

Phylogenetic Analyses

Within the data set containing 555 AFLP loci for 159 individuals, 444 characters were variable and 396 characters were parsimony informative. Parsimony analysis resulted in 43 equally optimal trees of length = 4470 steps. The strict consensus is shown in Figure 2. With a midpoint root position on the branch between the var. *lupuloides* clade and the other two varieties, bootstrap support for a monophyletic var. *lupuloides*, var. *neomexicanus*, and var. *pubescens* was 99% or higher, suggesting that these varieties may be distinct evolutionary lineages.

Although the consensus tree was highly resolved, character congruence was low (consistency index excluding uninformative characters = 0.09), as was bootstrap support for relationships within var. *lupuloides*—less than 50% for most clades. Of 32 clades with bootstrap support values greater than 50%, only 8 were not composed entirely of clones. Therefore, the large majority of the structure found in the tree within var. *lupuloides* was caused by clonality. Only four populations, G, Q, R, and AB, were monophyletic in the consensus tree. Population Q, however, was poorly supported (<50%) and G consisted entirely of clones.

Model-based Bayesian MCMC analyses also supported monophyly of the three varieties. Under a relaxed molecular clock, the probability of monophyly in the posterior distribution of trees for vars. *lupuloides*, *neomexicanus*, and *pubescens* was ≥ 0.999 , ≥ 0.959 , and ≥ 0.999 , respectively. ESS exceeded 250 for all parameters, implying adequate mixing. Under a stochastic Dollo model, the posterior probabilities were similarly high: ≥ 0.999 for all varieties. ESS exceeded 2000 for the tree likelihood parameter for both runs; however, other parameters were below 100, and duplicate chains did not converge to the same mean tree likelihood, suggesting poor mixing. Nevertheless, posterior probabilities of monophyly did not vary from the high values reported, even during early stages of the burn-in phase, implying that monophyly of varieties is an obvious outcome under a nonreversible Dollo model of character evolution. The root position was inferred to be on the branch



Downloaded from sysbio.oxfordjournals.org at DigiTrop USDA's Digital Desktop Library on January 18, 2011

FIGURE 2. Strict consensus of 43 trees resulting from parsimony analysis of AFLP data. Bootstrap support values for clades >50% are shown at nodes. Bullets indicate clades determined to consist entirely of clones. Open arrow indicates position of midpoint root. Closed arrow indicates the root position, relative to the three varieties, found using a relaxed molecular clock or a stochastic Dollo substitution model. Terminal identification: second letter is drainage of origin (C = Colorado River, R = Red River, M = Mississippi River, and x = unknown), third letter is river of origin (A = Assiniboine, L = Little Knife, K = Knife, S = Souris, Q = Qu'Appelle, W = White Earth, M = Missouri, and x = unknown), and fourth or fourth and fifth letters are population identifier from Table 1; "HERB" refers to herbarium specimens from Table 2 and remaining digits identify individuals.

TABLE 3. Population profiles (fixed presence or absence) for 11 diagnostic AFLP characters detected by PAA, plus 2 diagnostic morphological characters, for North American *Humulus lupulus* varieties

Character ^a	<i>lupuloides</i>	<i>pubescens</i>	<i>neomexicanus</i>
mCAC/eACA-212bp	0	0	1 ^b
mCAC/eACA-194bp	1	1	0 ^b
mCAC/eACT-193bp	1	1	0 ^b
mCAG/eACA-339bp	0	1 ^b	0
mCAT/eACA-366bp	1	1	0 ^b
mCAT/eACA-209bp	0	1 ^b	0
mCAT/eACA-167bp	1	1	0 ^b
mCTA/eACA-233bp	1	1	0 ^b
mCTA/eACA-193bp	0	0	1 ^b
mCTC/eACT-197bp	0	0	1 ^b
mCTT/eACA-304bp	0	1 ^b	0
Trichomes present between veins on abaxial leaf surface	0	1 ^b	0
Leaves >10 cm long with 5 or more lobes	0	0	1 ^b

^a"m" and "e" indicate MseI and EcoRI selective nucleotides.

^bDiagnostic character state.

between var. *pubescens* and the others by both Dollo and relaxed clock analyses (Fig. 2).

PAA identified 11 AFLP characters that were seemingly fixed within, but varied among, varieties (Table 3). Three "individually diagnostic" characters (characters that can be used in isolation to distinguish one variety from all others) were found for var. *pubescens*. Eight such characters were found for var. *neomexicanus*, but none were found for var. *lupuloides*. Variety *lupuloides* could, however, be diagnosed using one of several possible combinations of two or more characters. Among populations within var. *lupuloides*, two seemingly fixed differences were detected but they identified undersampled populations ($n = 1$ and $n = 2$), so the finding is likely an artifact (Davis and Nixon 1992). Thus, there is no convincing evidence that any of the 29 var. *lupuloides* populations satisfies the diagnosability criterion for lineage identification.

Population Structure Analyses

PCO-MC yielded four distinct clusters: one for each of the three varieties and a fourth containing var. *lupuloides* and var. *neomexicanus* together (Fig. 3a). Stability was high for all clusters suggesting that they likely represent genetically distinct groups (Reeves and Richards 2009). All but the var. *neomexicanus* cluster (perhaps due to small sample size) were found to be significantly distinct ($P < 0.05$; Fig. 3b).

The number of individuals for the STRUCTURE data set was reduced to 134 by excluding all but one representative from each set of putative clones. K was set to 3, the optimal value found using the method of Evanno et al. (2005). $K = 3$ was corroborated as optimal by PCO-MC as the number of clusters necessary to include all individuals without nesting. The esti-

mated ln probability of the data was similar for most replicate simulations (mean ± 1 standard deviation [SD] = -24512.9 ± 78.7); however, two simulations finished while apparently stuck in suboptimal solutions (ln probability = -25101.1 ± 15.8) and were excluded from further analysis. Assignments were highly stable among the remaining 18 runs ($S_N = 0.94$).

Due to low inferred admixture (Fig. 3b), the most likely clusters corresponded precisely with named taxonomic varieties (Table 4), consistent with PCO-MC analysis. Thus, based on current sampling, the three varieties form genetically distinct units with little evidence for the existence of intermediate genotypes.

Environmental Niche Modeling

A total of 140 niche models were computed using MAXENT (7 configurations of taxa, 20 replicates). MAXENT calculates model validation statistics using 11 different thresholds for defining presence and absence. Regardless of threshold, all models predicted the test sites significantly better than random (P value $\leq 1.7 \times 10^{-8}$). Ninety five percent of comparisons ("comparison" = one model evaluated under one threshold) included 78% or more of the test localities. Thus, for most models under most thresholds, niche models accurately predicted presence at sites that were not used for model construction.

By splitting North American *Humulus* into two or more groups, the predictive power of niche models was improved. The composite model in which vars. *lupuloides* and *pubescens* were treated as occupying a single niche, but var. *neomexicanus* was distinct, was the best model for predicting presence at test sites (Table 5), and was significantly better than a model in which all three taxa were lumped ($P = 0.0004$). When niche models were calculated for the three varieties separately, Schoener's (1968) metric of niche overlap (D) indicated that the predicted niches for var. *lupuloides* and *pubescens* were the most similar ($D_m = 0.182 \pm 0.019$ SD) followed by var. *lupuloides* and *neomexicanus* ($D_m = 0.107 \pm 0.024$). Predicted niches for var. *pubescens* and *neomexicanus* were the least similar ($D_m = 0.050 \pm 0.011$). Thus, predicted niche overlap is substantial between vars. *lupuloides* and *pubescens*, whereas var. *neomexicanus* appears to occupy habitat that is climatically distinct among North American *H. lupulus*. Using binary niche predictions resulting from a single replicate, a sample visualization of niche overlap is displayed in Figure 3c.

DISCUSSION

Monophyly

Parsimony analysis of AFLP data from the three native North American *H. lupulus* varieties revealed strong support for a monophyletic var. *lupuloides*, var. *neomexicanus*, and var. *pubescens* (Fig. 2). Bayesian MCMC analyses using a relaxed molecular clock or a stochastic Dollo

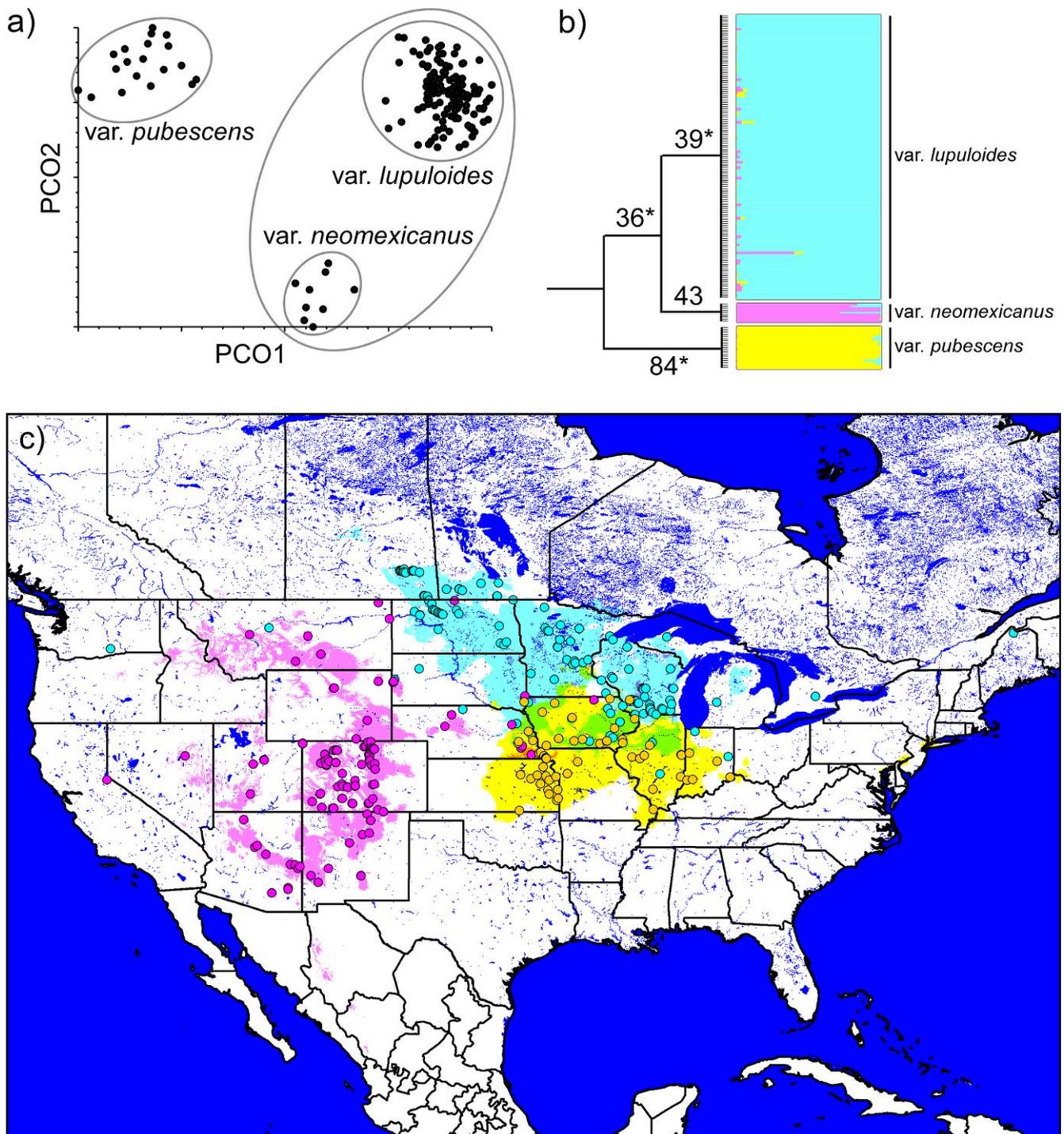


FIGURE 3. Population structure and niche differentiation within native North American hops. a) Clear separation of the three varieties is evident along the first two PCO axes. Ellipses indicate clusters found to be distinct using PCO-MC. b) Tree at left is the hierarchical assignment resulting from PCO-MC analysis. Numbers at nodes are cluster stability values. Asterisks indicate significantly distinct clusters ($P < 0.05$). Admixture within varieties is shown in bar graph at right. The length of colored bars represents the fractional assignment of individuals to each of $K = 3$ genetic clusters inferred by STRUCTURE: cyan = Cluster 1 (*lupuloides*), yellow = Cluster 2 (*pubescens*), and magenta = Cluster 3 (*neomexicanus*). c) Example binary niche prediction using the “10 percentile training presence” threshold: cyan = *lupuloides*, magenta = *neomexicanus*, and yellow = *pubescens*. Predicted regions of sympatry: green = *lupuloides/pubescens*, periwinkle = *lupuloides/neomexicanus* (northeast and south central Wisconsin), and orange = *pubescens/neomexicanus* (southeast Nebraska). Circles indicate localities used for model construction, colored by variety. Unsuitable habitat in the Great Plains forms a plausible barrier to gene exchange between var. *neomexicanus* and the others. A hotspot of *H. lupulus* genetic diversity, with suitable habitat for all varieties, is evident in eastern Nebraska.

TABLE 4. Fractional membership of North American *Humulus lupulus* varieties in each genetic cluster

Taxonomic variety	Cluster 1	Cluster 2	Cluster 3	<i>n</i>
<i>lupuloides</i>	0.9838	0.0052	0.0110	107
<i>pubescens</i>	0.0203	0.9784	0.0013	18
<i>neomexicanus</i>	0.0766	0.0015	0.9219	9

substitution model concurred with this assessment. The posterior probability of monophyly was ≥ 0.959 for all varieties under both models. Thus, all varieties satisfy the most commonly implemented test of monophyly, appearance as a monophyletic group in a reconstructed tree, and accordingly, satisfy the monophyly criterion integral to the monophyletic version of the phylogenetic species concept (e.g., de Queiroz and Donoghue 1988).

We caution, however, that tests of the monophyly criterion that rely on the inspection of phylogenetic trees may be misleading. In simulation studies considering simultaneous analysis of multilocus data sets, tree-based inference of monophyly suffers from a high false-positive rate. Parsimony and distance methods can cause the erroneous inference of monophyly for taxa with an extensive history of gene flow (Allaby and Brown 2003; Reeves and Richards 2007). Whether maximum likelihood analysis is similarly biased is not known. However, in the case of incomplete lineage sorting, when ancestral alleles have been retained through speciation events, trees resulting from likelihood analyses of concatenated data may contain highly supported, but incorrect, clades (Kubatko and Degnan 2007). Inference of monophyly from separate analyses of multilocus data sets (where loci are not concatenated) results in the opposite problem, a high false-negative rate, that is, the failure to identify groups that are, in fact, monophyletic (Knowles and Carstens 2007). Thus, although monophyly is a proper criterion for species delimitation under the general lineage concept—monophyletic groups are distinct evolutionary lineages—the typical test used to discern monophyly, appearance of a clade in a phylogenetic tree, may fail in the face of gene flow and incomplete lineage sorting, factors that are common at low taxonomic levels.

TABLE 5. Relative performance of niche models for North American *Humulus lupulus* varieties under five taxonomic scenarios

Taxonomic scenario	Mean probability of occurrence at test sites	SD	<i>P</i> ^a
1. (<i>lupuloides</i> , <i>pubescens</i> , <i>neomexicanus</i>)	0.5889	0.0231	n/a
2. (<i>lupuloides</i> , <i>pubescens</i>), (<i>neomexicanus</i>)	0.6197	0.0271	0.0004
3. (<i>lupuloides</i> , <i>neomexicanus</i>), (<i>pubescens</i>)	0.5979	0.0254	0.2471
4. (<i>pubescens</i> , <i>neomexicanus</i>), (<i>lupuloides</i>)	0.6018	0.0254	0.0997
5. (<i>lupuloides</i>), (<i>pubescens</i>), (<i>neomexicanus</i>)	0.6076	0.0303	0.0340

Note: n/a, not applicable.

^aStudent's *t* test, versus lumped scenario no. 1.

When divergence between putative species is low, coalescent approaches, which can model the impact of gene flow, incomplete lineage sorting, and demographic factors on reconstructed gene trees, may be better suited for identifying monophyletic groups than phylogenetic trees alone (Rosenberg and Nordborg 2002). Although not yet applicable to binary unlinked markers (e.g., AFLP), emerging hybrid methods that respect the stochastic influence of the coalescent on reconstructed gene trees (e.g., Knowles and Carstens 2007; Liu and Pearl 2007) may form a better test of the monophyly criterion than the current common practice of reconstruction and inspection of phylogenetic trees.

Diagnosability

Under the criterion of diagnosability, species are identified as "the smallest aggregation of populations (sexual) or lineages (asexual) diagnosable by a unique combination of character states in comparable individuals..." (Nixon and Wheeler 1990, p. 211). Using PAA (Davis and Nixon 1992), 3 diagnostic AFLP loci were identified for var. *pubescens* and 8 for var. *neomexicanus*, supporting their status as putative lineages (Table 3). Although var. *lupuloides* lacks any AFLP characters which, taken alone, can be used to distinguish it from the others, several different character combinations can be used. A parallel situation exists for morphological characters. Among native North American *Humulus*, var. *neomexicanus* can be distinguished by conspicuously dissected leaves with five or more lobes. Variety *pubescens* is diagnosable by the presence of trichomes between veins on the abaxial leaf surface. Variety *lupuloides* can be diagnosed by the combined absence of both five-lobed leaves and abaxial leaf trichomes between veins. Thus, all three varieties comply with the diagnosability criterion.

Absence of Intermediates

Mallet's (1995, p. 296) genotypic cluster criterion states that species are "distinguishable groups of individuals that have few or no intermediates when in contact." Although Mallet (1995) promoted the use of genetic data, no particular analytical methods were advocated. To identify genotypic clusters in the data, the principal coordinate-based method PCO-MC was used. PCO-MC has been shown by simulation to outperform tree-based methods for lineage identification in multilocus data sets sampled from taxa with a history of gene flow (Reeves and Richards 2007). The Bayesian MCMC procedure of Pritchard et al. (2000), implemented in the software STRUCTURE, was used to quantify the degree of admixture—in essence, the absence of intermediates—between populations. PCO-MC and STRUCTURE analyses suggest that the named taxonomic varieties are genetically distinguishable from one another, with little evidence for subpopulation structure within varieties or admixture among varieties

(Table 4 and Fig. 3a,b). Thus, both methods used to evaluate genetic subdivision support the recognition of North American *H. lupulus* varieties as distinct lineages under the genotypic cluster species criterion and moreover, predict the existence of biological or ecogeographical barriers to gene flow between varieties.

The admixture model of STRUCTURE allows the fraction of each individual genome drawn from the *K* populations to be estimated when mixed ancestry is suspected. This feature is particularly useful for the analysis of introgression across hybrid zones (Pritchard et al. 2000). STRUCTURE analysis identified one putative hybrid between var. *lupuloides* and var. *neomexicanus* from population W in Saskatchewan (individual LRQW.01.134). The mean fractional ancestry of this individual's genome was estimated to be 53% var. *lupuloides* and 40% var. *neomexicanus*. We consider one putative intermediate genotype among 131 individuals to qualify as "few" under Mallet's (1995) criterion, so this finding does not jeopardize var. *lupuloides*' status as genotypically distinct.

The genotypic cluster criterion functions best when samples are taken from zones where putative lineages are in contact (Mallet 1995). This is an important restriction because 1) allopatric species are common and 2) political boundaries or landscape alteration may preclude collection within contact zones for some species. Hence, the criterion is not universally applicable. In this study, DNA from var. *lupuloides* was sampled from a region where it is known to overlap with var. *neomexicanus*, but *neomexicanus* DNA was sampled in allopatry. Variety *pubescens* DNA was sampled from a region of overlap with var. *lupuloides*, but *lupuloides* DNA was sampled from elsewhere (Fig. 1). Thus, our sampling scheme did not conform precisely to that envisioned by Mallet (1995). In fact, because sampling in this study is geographically localized for all three varieties, we cannot distinguish evolutionary lineages from populations exhibiting isolation by distance (or clinal variation) using the genotypic cluster or any other pattern-oriented species criterion. When sampling is imperfect, compliance with pattern-oriented species criteria is a necessary, but not sufficient condition for lineage identification. For lineages to be identified with certainty, the genetic data must be viewed in light of other factors. Evidence for reproductive isolation or niche divergence between varieties can be used to suggest that mechanisms other than isolation by distance are responsible for the observed genetic divergence.

Reproductive Isolation

Under a criterion of intrinsic reproductive isolation (the "isolation criterion"), species are "groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups" (Mayr 1942, p. 120). Although we provide little new data, we provide here a line of argument based largely on natural history observations to exam-

ine the matter of reproductive isolation among varieties of North American *H. lupulus*.

Numerical taxonomic analyses suggest that gene flow between cultivated hops and nearby wild plants has been common during the history of hop cultivation (Small 1980). Extensive unintentional hybridization between introduced cultivars (derived from European var. *lupulus*) and indigenous plants in Japan (var. *cordifolius*) and the United States (var. *neomexicanus*, var. *lupuloides*) has resulted in the inadvertent production of new geographically distinct cultivars that are genetically similar to native varieties (Small 1980; Šuštar-Vozlič and Javornik 1999; Seefelder et al. 2000). Strong biological barriers to gene exchange do not appear to exist within var. *lupuloides* and var. *neomexicanus*, at least with respect to European var. *lupulus*. Due to their inferred ability to hybridize with cultivated var. *lupulus* (Small 1980), and the observation of intergrading clinal variation in phenotype in areas of sympatry (Small 1978), neither var. *lupuloides* nor var. *neomexicanus* satisfy the isolation criterion.

Variety *pubescens*, on the other hand, may be isolated from the other varieties by reproductive barriers. First, no close relationship was found between any regional cultivars and var. *pubescens* in spite of considerable historical opportunity for hybridization in the hops growing regions of the midwestern United States (Small 1980). Second, var. *pubescens* and var. *lupuloides* differed in flowering phenology in a common garden grown under var. *neomexicanus* field conditions. During two consecutive seasons of observation (2002–2003), var. *lupuloides* individuals of both sexes proceeded to flowering, whereas female var. *pubescens* plants failed to flower prior to the first freeze (data not shown). This suggests the possibility of a premating barrier (temporal isolation) to gene flow that isolates vars. *pubescens* from both *lupuloides* and *neomexicanus*, at least under some environmental conditions. Further experimentation using a common garden, reciprocal transplant approach (see, e.g., Clausen et al. 1940; Angert and Schemske 2005) would be necessary to confirm this preliminary observation. Last, we included two historic herbarium specimens (collected in 1916 and 1943) from localities that were remote (400–800 km) from the contemporary collections. That these samples were nested within the var. *pubescens* clade, and possessed the diagnostic var. *pubescens* AFLP characters, suggests geographic and temporal stability of the character states that support var. *pubescens* as a lineage. Such stability suggests that intrinsic reproductive barriers may isolate vars. *pubescens* and *lupuloides* in regions of sympatry.

Niche Specialization

In studies where imperfect sampling might lead to the erroneous inference of genetic discontinuity, geographic and/or ecological information becomes essential (de Queiroz 2007). With sufficient ecogeographical information, one can determine whether there is substantial

niche overlap between putative lineages or whether they have assumed different ecological roles. In the latter case, the groups would satisfy an ecological species criterion: “a species is a lineage (or a closely related set of lineages) which occupies an adaptive zone minimally different from that of any other lineage in its range and which evolves separately from all lineages outside its range” (Van Valen 1976, p. 233). When combined, geography and ecology form a powerful means to test whether suitable habitat exists between two allopatric groups, or conversely, whether there are substantial ecogeographic barriers to gene flow between them (Ramsey et al. 2003; Wiens 2007). Geographic and ecological information (used for environmental niche modeling) have been combined with conventional methods (using molecular or morphological character divergence) to provide greater strength of inference for species delimitation (Raxworthy et al. 2007; Rissler and Apodaca 2007). In this study, niche models were constructed using 19 biologically relevant climatic variables and 350 localities where *H. lupulus* had been sampled and the taxonomic variety was known.

We present a simple test, following that introduced by Raxworthy et al. (2007), to determine whether varieties occupy significantly distinct niches. The test involves computing niche models for all possible split and lumped arrangements of varieties, then determining which of these taxonomic scenarios produces the best model for predicting presence at test sites (which are not used during model construction). If all individuals share a single niche, organizing them into varieties should not improve our ability to predict presence at test sites using environmental data. But if varieties have different preferred habitats, partitioning should improve test site prediction because a refined model, one that more efficiently explains the distribution of individuals in terms of environmental preferences, can then be constructed. Raxworthy et al.’s (2007) test favored those taxonomic scenarios that minimized omission rates on test data. This approach requires the arbitrary application of one or more threshold values to determine binary presence or absence from continuously distributed probability of presence data. We measured goodness of fit to expectations using the probability of occurrence at test sites under the various taxonomic scenarios, negating the need for threshold selection. The scenario that maximized the probability of occurrence at test sites was deemed best.

The sampled distributions of North American *H. lupulus* varieties were best predicted using climatic data when var. *neomexicanus* was treated as occupying a distinct niche, and vars. *lupuloides* and *pubescens* were treated as occupying the same niche. A more general

model, in which all varieties were treated as one, had significantly less power to predict the occurrence of unknown populations (Table 5). Thus, environmental niche modeling suggests that var. *neomexicanus* persists in habitat with quantifiable climatic differences from that of the other varieties. Niche similarity, as measured using Schoener’s (1968) metric of niche overlap (D), was significantly higher ($P < 0.0001$) between vars. *lupuloides* and *pubescens* than between either of the two and var. *neomexicanus*. Although some overlap in distribution between varieties occurs at the margins of their ranges (Fig. 1), the Great Plains appear to form an important contemporary ecological barrier to gene flow between var. *neomexicanus* and the others (Fig. 3c). The most important single factor forming the barrier would seem to be the scarcity of riparian habitat. Hence, var. *neomexicanus*, which occupies a distinct geographically isolated niche, satisfies the ecological species criterion, but vars. *lupuloides* and *pubescens* do not. Environmental niche modeling suggests that, in spite of the opportunity for (and evidence of) introgression of alleles from other *H. lupulus* varieties, var. *neomexicanus* retains its distinct genetic identity due to adaptation to its core habitat, the Rocky Mountain cordillera.

Summary and Conclusions

Lineage identification is a necessary precursor to species discovery (Mayden 1999; Sites and Marshall 2003; de Queiroz 2007). In the absence of direct observation of speciation, satisfaction of a plurality of species criteria can ensure accurate identification of lineages and stable uncontroversial species delimitations. For wild relatives of crop species, accurate identification of evolutionary lineages may facilitate crop improvement efforts by focusing breeding programs on distinct evolutionary units possessing desirable genes and traits. Accordingly, in order to determine whether subspecific taxa of wild North American hops persist as evolutionarily distinct lineages and thus merit recognition as distinct species, compliance with five species criteria was examined using molecular polymorphism, natural history, and distributional data from the three native varieties (Table 6).

Using AFLP data, all varieties were found to be monophyletic, were diagnosable, and formed distinct genetic clusters with little admixture. Thus, all varieties satisfy the three pattern-oriented species criteria examined: monophyly (Donoghue 1985; de Queiroz and Donoghue 1988), diagnosability (Nixon and Wheeler 1990), and the formation of distinct genotypic clusters (Mallet 1995). Natural history information and a preliminary examination of temporal isolating barriers

TABLE 6. Conformance of North American *Humulus lupulus* varieties with five distinct criteria for species delimitation

Taxonomic variety	Monophyly	Diagnosability	Genotypic cluster	Reproductive isolation	Distinct adaptive zone
<i>lupuloides</i>	Yes	Yes	Yes	No	No
<i>pubescens</i>	Yes	Yes	Yes	Yes	No
<i>neomexicanus</i>	Yes	Yes	Yes	No	Yes

suggest that var. *pubescens* may possess intrinsic reproductive isolating mechanisms that limit gene flow between it and the other varieties and thus may satisfy the criterion of reproductive isolation (Mayr 1942), a process-based species criterion. Environmental niche modeling suggests that niche specialization has occurred during the evolution of var. *neomexicanus*, but that vars. *lupuloides* and *pubescens* occupy substantially similar niches. Therefore, only var. *neomexicanus* satisfies the criterion of occupying a distinct adaptive zone (Van Valen 1976).

Based upon these findings, vars. *pubescens* and *neomexicanus* are demonstrably distinct lineages, therefore we suggest that they be recognized as species. Variety *pubescens* likely maintains its genetic identity via the evolution of (perhaps incomplete) reproductive isolating mechanisms, a prerequisite for a lineage to persist while sharing substantial niche space with a close relative (Gause 1934). Given the novelty of its predicted niche, adaptation to unique environmental conditions (i.e., ecological isolation) is likely the primary isolating mechanism for var. *neomexicanus*, which retains a distinct genetic identity despite the absence of intrinsic barriers to gene flow from var. *lupuloides*. Given that the current range of var. *lupuloides* corresponds closely with the southernmost extension of the Laurentide ice sheet (Dyke and Prest 1987), it may be a nascent species that has emerged in the 10–18 thousand years since glacial retreat. However, because we were unable to identify any biological process to account for the observed pattern of genetic differentiation, we do not yet recommend it be recognized as a species.

We reiterate that imperfect sampling of molecular polymorphism data, coupled with isolation by distance or clinal variation, can cause the erroneous inference of genetic discontinuity. Under these conditions, the sampled data are likely to mislead, resulting in the false recognition of lineages using species criteria focused on evolutionary pattern. In cases where no mechanistic explanation for the observed genetic divergence is apparent, the only recourse is comprehensive sampling of genetic data from throughout the range. In order to ensure accurate identification of evolutionary lineages for species delimitation, to promote the stability of species circumscriptions, and to minimize taxonomic controversy, satisfaction of complementary pattern- and process-oriented species criteria is important. A coupled hypothesis of cause and effect is a powerful form of corroboration.

SUPPLEMENTARY MATERIAL

Supplementary material can be found at: <http://www.sysbio.oxfordjournals.org/>.

FUNDING

This project was funded by the United States Department of Agriculture, Agricultural Research Service, CRIS Project Number 5402-21000-010-00D. Field collec-

tions were supported, in part, by Busch Agricultural Resources, Inc. AFLP data have been deposited at TreeBASE (www.treebase.org), study number S10570.

ACKNOWLEDGMENTS

We thank John Waddell for collecting most of the native *H. lupulus* tissue used in this study; Scott Dorsch for field collections and common garden observations; Mark Simmons and Jennifer Ackerfield for facilitating the herbarium work; Alex Alekseyenko for assistance applying the stochastic Dollo model; Dale Lockwood for calculating the stability coefficient for the STRUCTURE runs; Amy Angert, Christine Bacon, L. J. Grauke, John R. Paul, Seema Sheth, and Mark Simmons for providing helpful suggestions on the manuscript; and Mark Fishbein, Kevin de Queiroz, and three anonymous reviewers for especially detailed reviews of the paper.

REFERENCES

- Alekseyenko A.V., Lee C.J., Suchard M.A. 2008. Wagner and Dollo: a stochastic duet by composing two parsimonious solos. *Syst. Biol.* 57:772–784.
- Allaby R.G., Brown T.A. 2003. AFLP data and the origins of domesticated crops. *Genome*. 46:448–453.
- Alström P., Rasmussen P.C., Olsson U., Sundberg P. 2008. Species delimitation based on multiple criteria: the spotted bush warbler *Bradypterus thoracicus* complex (Aves: Megaluridae). *Zool. J. Linn. Soc. Lond.* 154:291–307.
- Anderson R.P., Gómez-Laverde M., Peterson A.T. 2002. Geographical distributions of spiny pocket mice in South America: insights from predictive models. *Glob. Ecol. Biogeogr.* 11:131–141.
- Angert A.L., Schemske D.W. 2005. The evolution of species' distributions: reciprocal transplants across the elevation ranges of *Mimulus cardinalis* and *M. lewisii*. *Evolution*. 59:1671–1684.
- Baum D.A., Shaw K.L. 1995. Genealogical perspectives on the species problem. In: Hoch P.C., Stephenson A.G., editors. *Experimental and molecular approaches to plant biosystematics*. St. Louis (MO): Missouri Botanical Garden. p. 289–303.
- Brady J.L., Scott N.S., Thomas M.R. 1996. DNA typing of hops (*Humulus lupulus*) through application of RAPD and microsatellite marker sequences converted to sequence tagged sites (STS). *Euphytica*. 91:277–284.
- Clausen J.C., Keck D.D., Hiesey W.M. 1940. *Experimental studies on the nature of species. I. Effect of varied environments on western North American plants*. Washington (DC): Carnegie Institute of Washington Publication No. 520.
- Cracraft J. 1983. Species concepts and speciation analysis. *Curr. Ornithol.* 1:159–187.
- Darwin C. 1859. *On the origin of species*. London: John Murray.
- Davis J.I., Nixon K.C. 1992. Populations, genetic variation, and the delimitation of phylogenetic species. *Syst. Biol.* 41:421–435.
- DeBry R.W., Olmstead R.G. 2000. A simulation study of reduced tree-search effort in bootstrap resampling analysis. *Syst. Biol.* 49:171–179.
- de Queiroz K. 1998. The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. In: Howard D.J., Berlocher S.H., editors. *Endless forms: species and speciation*. New York: Oxford University Press. p. 57–75.
- de Queiroz K. 2005. Different species problems and their resolution. *BioEssays*. 27:1263–1269.
- de Queiroz K. 2007. Species concepts and species delimitation. *Syst. Biol.* 56:879–886.
- de Queiroz K., Donoghue M.J. 1988. Phylogenetic systematics and the species problem. *Cladistics*. 4:317–338.

- Dettman J.R., Jacobson D.J., Turner E., Pringle A., Taylor J.W. 2003. Reproductive isolation and phylogenetic divergence in *Neurospora*: comparing methods of species recognition in a model eukaryote. *Evolution*. 57:2721–2741.
- Donoghue M.J. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. *Bryologist*. 88:172–181.
- Drummond A.J., Ho S.Y.W., Phillips M.J., Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol*. 4:e88.
- Drummond A.J., Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol*. 7:214.
- Dyke A.S., Prest V.K. 1987. Late wisconsinan and holocene history of the laurentide ice sheet. *Geogr. Phys. Quatern*. 41:237–263.
- Elith J., Graham C.H., Anderson R.P., Dudík M., Ferrier S., Guisan A., Hijmans R.J., Huettmann F., Leathwick J.R., Lehmann A., Li J., Lohmann L.G., Loiselle B.A., Manion G., Moritz C., Nakamura M., Nakazawa Y., Overton J.M., Peterson A.T., Phillips S.J., Richardson K., Scachetti-Pereira R., Schapire R.E., Soberón J., Williams S., Wisz M.S., Zimmermann N.E. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*. 29:129–151.
- Evanno G., Regnaut S., Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol*. 14:2611–2620.
- Falush D., Stephens M., Pritchard J.K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes*. 7:574–578.
- Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. 39:783–791.
- Gause G.F. 1934. *The struggle for existence*. Baltimore (MD): Williams & Wilkins.
- Gur A., Zamir D. 2004. Unused natural variation can lift yield barriers in plant breeding. *PLoS Biol*. 2:e245.
- Hampton R., Small E., Haunold A. 2001. Habitat and variability of *Humulus lupulus* var. *lupuloides* in upper midwestern North America: a critical source of American hop germplasm. *J. Torrey Bot. Soc.* 128:35–36.
- Hartl L., Seefelder S. 1998. Diversity of selected hop cultivars detected by fluorescent AFLPs. *Theor. Appl. Genet*. 96:112–116.
- Henning J.A., Steiner J.J., Hummer K.E. 2004. Genetic diversity among world hop accessions grown in the USA. *Crop. Sci.* 44:411–417.
- Hernandez P.A., Graham C.H., Master L.L., Albert D.L. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*. 29:773–785.
- Hess P.N., De Moraes Russo C.A. 2007. An empirical test of the midpoint rooting method. *Biol. J. Linn. Soc.* 92:669–674.
- Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G., Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol*. 25:1965–1978.
- Hudson R.R., Coyne J.A. 2002. Mathematical consequences of the genealogical species concept. *Evolution*. 56:1557–1565.
- Jakobsson M., Rosenberg N.A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*. 23:1801–1806.
- Jakše J., Kindlhofer K., Javornik B. 2001. Assessment of genetic variation and differentiation of hop genotypes by microsatellite and AFLP markers. *Genome*. 44:773–782.
- Knowles L.L., Carstens B.C. 2007. Delimiting species without monophyletic gene trees. *Syst. Biol*. 56:887–895.
- Koopman W.J.M. 2005. Phylogenetic signal in AFLP data sets. *Syst. Biol*. 54:197–217.
- Kubatko L.S., Degnan J.H. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst. Biol*. 56:17–24.
- Lambertini C., Frydenberg J., Gustafsson M.H.G., Brix H. 2008. Herbarium specimens as a source of DNA for AFLP fingerprinting of *Phragmites* (Poaceae): possibilities and limitations. *Plant Syst. Evol.* 272:223–231.
- Light J.E., Toups M.A., Reed D.L. 2008. What's in a name: the taxonomic status of human head and body lice. *Mol. Phylogenet. Evol.* 47:1203–1216.
- Liu L., Pearl D.K. 2007. Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Syst. Biol*. 56:504–514.
- Mallet J. 1995. A species definition for the modern synthesis. *Trends Ecol. Evol.* 10:294–299.
- Marques C.M., Araujo J.A., Ferreira J.G., Whetten R., O'Malley D.M., Liu B.H., Sederoff R. 1998. AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. *Theor. Appl. Genet*. 96:727–737.
- Marshall J.C., Arévalo E., Benavides E., Sites J.L., Sites J.W. Jr. 2006. Delimiting species: comparing methods for mendelian characters using lizards of the *Sceloporus grammicus* (Squamata: Phrynosomatidae) complex. *Evolution*. 60:1050–1065.
- Mayden R.L. 1999. Consilience and a hierarchy of species concepts: advances toward closure on the species puzzle. *J. Nematol.* 31:95–116.
- Mayr E. 1942. *Systematics and the origin of species*. New York: Columbia University Press.
- Murakami A., Darby P., Javornik B., Pais M.S.S., Seigner E., Lutz A., Svoboda P. 2006a. Microsatellite DNA analysis of wild hops, *Humulus lupulus* L. *Genet. Resour. Crop. Evol.* 53:1553–1562.
- Murakami A., Darby P., Javornik B., Pais M.S.S., Seigner E., Lutz A., Svoboda P. 2006b. Molecular phylogeny of wild hops, *Humulus lupulus* L. *Hereditas*. 97:66–74.
- Myburg A.A., Remington D.L., O'Malley D.M., Sederoff R.R., Whetten R.W. 2001. High-throughput AFLP analysis using infrared dye-labeled primers and an automated DNA sequencer. *BioTechniques*. 30:348–357.
- Nelson A., Cockerell T.D.A. 1903. Three new plants from New Mexico. *P. Biol. Soc. Wash.* 16:45–46.
- Neve R.A. 1991. *Hops*. London: Chapman and Hall.
- Nixon K.C., Wheeler Q.D. 1990. An amplification of the phylogenetic species concept. *Cladistics*. 6:211–223.
- Nuttall T. 1848. Descriptions of plants collected by Mr. William Gambel in the rocky mountains and upper California. *Proc. Acad. Nat. Sci. Philadelphia*. 4:7–26.
- Padial J.M., De La Riva I. 2009. Integrative taxonomy reveals cryptic Amazonian species of *Pristimantis* (Anura: Strabomantidae). *Zool. J. Linn. Soc. Lond.* 155:97–122.
- Paterson H.E.H. 1985. The recognition concept of species. In: Vrba E.S., editor. *Species and speciation*. Pretoria: Transvaal Museum.
- Patzak J. 2001. Comparison of RAPD, STS, ISSR and AFLP molecular methods used for assessment of genetic diversity in hop (*Humulus lupulus* L.). *Euphytica*. 121:9–18.
- Phillips S.J., Anderson R.P., Schapire R.E. 2006. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190:231–259.
- Pillay M., Kenny S.T. 1996. Structure and inheritance of ribosomal DNA variants in cultivated and wild hop, *Humulus lupulus* L. *Theor. Appl. Genet*. 93:333–340.
- Pimm S.L., Russell G.J., Gittleman J.L., Brooks T.M. 1995. The future of biodiversity. *Science*. 269:347–350.
- Pons J., Barraclough T.G., Gomez-Zurita J., Cardoso A., Duran D.P., Hazell S., Kamoun S., Sumlin W.D., Vogler A.P. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol*. 55:595–609.
- Pritchard J.K., Stephens M., Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics*. 155:945–959.
- Pritchard J.K., Wen X., Falush D. 2007. Documentation for STRUCTURE software. Version 2.2 [Internet]. Available from <http://pritch.bsd.uchicago.edu/software/structure22/readme.pdf>.
- Rambaut A., Drummond A.J. 2007. Tracer v1.4 [Internet]. Available from <http://beast.bio.ed.ac.uk/Tracer>.
- Ramsey J., Bradshaw H.D. Jr., Schemske D.W. 2003. Components of reproductive isolation between the monkey flowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). *Evolution*. 57:1520–1534.
- Raxworthy C.J., Ingram C.M., Rabibisoa N., Pearson R.G. 2007. Applications of ecological niche modeling for species delimitation: a review and empirical evaluation using day geckos (*Phelsuma*) from Madagascar. *Syst. Biol*. 56:907–923.
- Reeves P.A., Friedman P.H., Richards C.M. 2005. wolfPAC: building a high-performance distributed computing network for phylogenetic analysis using "obsolete" computational resources. *Appl. Bioinformatics*. 4:61–64.

- Reeves P.A., Richards C.M. 2007. Distinguishing terminal monophyletic groups from reticulate taxa: performance of phenetic, tree-based, and network procedures. *Syst. Biol.* 56:302–320.
- Reeves P.A., Richards C.M. 2009. Accurate inference of subtle population structure (and other genetic discontinuities) using principal coordinates. *PLoS ONE*. 4:e4269.
- Renner S.S., Grimm G.W., Schneeweiss G.M., Stuessy T.F., Ricklefs R.E. 2008. Rooting and dating maples (*Acer*) with an uncorrelated-rates molecular clock: implications for North American/Asian disjunctions. *Syst. Biol.* 57:795–808.
- Richards C.M., Volk G.M., Reilley A.A., Henk A.D., Lockwood D.R., Reeves P.A., Forsline P.L. 2009. Genetic diversity and population structure in *Malus sieversii*, a wild progenitor species of domesticated apple. *Tree Genet. Genomes*. 5:339–347.
- Rissler L.J., Apodaca J.J. 2007. Adding more ecology into species delimitation: ecological niche models and phylogeography help define cryptic species in the black salamander (*Aneides flavipunctatus*). *Syst. Biol.* 56: 924–942.
- Rosenberg N.A. 2007. Statistical tests for taxonomic distinctiveness from observations of monophyly. *Evolution*. 61: 317–323.
- Rosenberg N.A., Mahajan S., Ramachandran S., Zhao C., Pritchard J.K., Feldman M.W. 2005. Clines, clusters, and the effect of study design on the inference of human population structure. *PLoS Genet.* 1:660–671.
- Rosenberg N.A., Nordborg M. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nat. Rev. Genet.* 3:380–390.
- Rydberg P.A. 1917. *Flora of the Rocky Mountains and adjacent plains, Colorado, Utah, Wyoming, Idaho, Montana, Saskatchewan, Alberta, and neighboring parts of Nebraska, South Dakota, and British Columbia*. Published by the Author. New York.
- Salmon E.S. 1934. Two new hops: 'Brewer's Favourite' and 'Brewer's Gold.' *J. Southeast. Agric. Coll. Wye, Kent*. 34:93–106.
- Schoener T.W. 1968. *Anolis* lizards of Bimini: resource partitioning in a complex fauna. *Ecology*. 49:704–726.
- Schwartz M.K., McKelvey K.S. 2009. Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conserv. Genet.* 10:441–452.
- Seefeldt S., Ehrmaier H., Schweizer G., Seigner E. 2000. Genetic diversity and phylogenetic relationships among accessions of hop, *Humulus lupulus*, as determined by amplified fragment length polymorphism fingerprinting compared with pedigree data. *Plant Breed.* 119:257–263.
- Sites J.W. Jr., Marshall J.C. 2003. Delimiting species: a renaissance issue in systematic biology. *Trends Ecol. Evol.* 18:462–470.
- Sites J.W. Jr., Marshall J.C. 2004. Operational criteria for delimiting species. *Annu. Rev. Ecol. Syst.* 35:199–227.
- Small E. 1978. A numerical and nomenclatural analysis of morpho-geographic taxa of *Humulus*. *Syst. Bot.* 3:37–76.
- Small E. 1980. The relationship of hop cultivars and wild variants of *Humulus lupulus*. *Can. J. Bot.* 58:676–686.
- Small E. 1981. A numerical analysis of morpho-geographic groups of cultivars of *Humulus lupulus* based on samples of cones. *Can. J. Bot.* 59:311–324.
- Stajner N., Satovic Z., Cerenak A., Jabornik B. 2008. Genetic structure and differentiation in hop (*Humulus lupulus* L.) as inferred from microsatellites. *Euphytica*. 161:301–311.
- Stevens J.F., Taylor A.W., Nickerson G.B., Ivancic M.I., Henning J.A., Haunold A., Deinzer M.L. 2000. Prenylflavonoid variation in *Humulus lupulus*: distribution and taxonomic significance of xanthogalenol and 4'-O-methylxanthohumol. *Phytochemistry*. 53: 759–775.
- Šuštar-Vozlič J., Javornik B. 1999. Genetic relationships in cultivars of hop, *Humulus lupulus* L., determined by RAPD analysis. *Plant Breed.* 118:175–181.
- Swofford D.L. 1999. PAUP*: phylogenetic analysis using parsimony (and other methods). Version 4.0b10. Sunderland (MA): Sinauer Associates.
- Tanksley S.D., McCouch S.R. 1997. Seed banks and molecular maps: unlocking genetic potential from the wild. *Science*. 277:1063–1066.
- Thomas C.D., Cameron A., Green R.E., Bakkenes M., Beaumont L.J., Collingham Y.C., Erasmus B.F.N., Ferreira de Siqueira M., Grainger A., Hannah L., Hughes L., Huntley B., van Jaarsveld A.S., Midgley G.F., Miles L., Ortega-Huerta M.A., Peterson A.T., Phillips O.L., Williams S.E. 2004. Extinction risk from climate change. *Nature*. 427:145–148.
- Townsend M.S., Henning J.A. 2009. AFLP discrimination of native North American and cultivated hop. *Crop Sci.* 49:600–607.
- Van Valen, L. 1976. Ecological species, multispecies, and oaks. *Taxon*. 25: 233–239.
- Vos P., Hogers R., Bleeker M., Reijans M., van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23:4407–4414.
- Warren D.L., Glor R.E., Turelli M. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*. 62:2868–2883.
- Wiens J.J. 2007. Species delimitation: new approaches for discovering diversity. *Syst. Biol.* 56:875–878.
- Wiens J.J., Penkrot T.A. 2002. Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*). *Syst. Biol.* 51:69–91.
- Wiens J.J., Servedio M.R. 2000. Species delimitation in systematic: inferring diagnostic differences between species. *P. Roy. Soc. Lond. B. Bio.* 267:631–636.