Zooming in on the lettuce genome:

species relationships in *Lactuca* s.l., inferred from chromosomal and molecular characters

Wim J. M. Koopman

Promotoren: Prof. Dr. Ir. L. J. G. van der Maesen
                      Hoogleraar Plantentaxonomie, Leerstoelgroep Biosystematiek
                      Wageningen Universiteit
                      Prof. Dr. Ir. E. Jacobsen
                      Hoogleraar Plantenveredeling, Laboratorium voor Plantenveredeling
                      Wageningen Universiteit

Co-promotor: Dr. R. G. van den Berg
                      Universitair Hoofddocent, Leerstoelgroep Biosystematiek
                      Wageningen Universiteit

Promotiecommissie: Prof. Dr. P. Baas
                      Universiteit Leiden
                      Dr. J. H. de Jong
                      Wageningen Universiteit
                      Dr. J. C. M. den Nijs
                      Universiteit van Amsterdam
                      Prof. Dr. R. F. Hoekstra
                      Wageningen Universiteit
                      Prof. Dr. A. Lebeda
                      Palacký University

Stellingen

I

L. sativa, L. serriola, L. saligna, en *L. virosa* stammen af van een gemeenschappelijke voorouder: *L. sativa* is het product van selectie door de mens, *L. serriola* ontstond als cultuurvolger, *L. saligna* ontstond door natuurlijke selectie en *L. virosa* ontstond door hybridisatie met een nog onbekende mannelijke ouder.

Dit proefschrift, H. 2, 4.

II

De Zuid-Afrikaanse soort *L. dregeana* is een 17e eeuwse ontsnapper uit de slateelt in de Compagnies Tuin van de VOC bij de Kaap, en conspecifiek met *L. sativa, L. serriola, L. dregeana* en *L. altaica*.

Dit proefschrift, H. 5.

III

Gezien de nauwe verwantschap van *L. tatarica* met *L. sibirica* (zoals blijkt uit hun morfologie, ITS-1 sequenties, kruisbaarheid, AFLP patronen en DNA gehaltes), is de indeling van deze soorten in aparte genera niet gerechtvaardigd.

Dit proefschrift, H. 4, 5, 6.
Contra Shih (1988).

IV

The subsectie *Cyanicae* soorten *L. tenerrima* en *L. perennis* (en mogelijk *L. graeca*) zijn niet nauw verwant met de andere Europese *Lactuca* soorten, en moeten uit het geslacht *Lactuca* worden verwijderd.

Dit proefschrift, H. 4, 5.
Contra Feráková (1977).

V

In afwijking van het algemene beeld in hogere planten (Barow and Meister, 2002) zijn DNA hoeveelheid en AT gehalte in *Lactuca* s.l. significant (en negatief) gecorreleerd.

Dit proefschrift, H. 6.

VI

Omdat op AFLP data gebaseerde fenogrammen en cladogrammen over het algemeen een hoge mate van congruentie vertonen, leidt het interpreteren van UPGMA clustering als phylogenetische analyse (bijv. Sharma et al., 1996; Caicedo et al., 1999; Aggarwal et al., 1999) waarschijnlijk toch tot betrouwbare conclusies over evolutionaire verwantschappen.

Koopman, ungepubliceerde gegevens.

Propositions

I

L. sativa, L. serriola, and L. virosa originate from a common ancestor: L. sativa evolved through selection by man, L. serriola as the related camp-following weed, L. saligna evolved by natural selection, and the evolution of L. virosa involved hybridization with a yet unknown male parent.

This thesis, Ch. 2, 4.

II

The South-African species L. dregeana is an escape from 17th century lettuce cultivation in the Company's Garden of the VOC at the Cape, and conspecific with L. sativa, L. serriola, L. dregeana, and L. altaica.

This thesis, Ch. 5.

III

Considering the close relationship between L. tatarica and L. sibirica (as apparent from their morphology, ITS-1 sequences, crossability, AFLP patterns, and DNA contents), the classification of these species in separate genera is not justified.

This thesis, Ch. 4, 5, 6.
Contra Shih (1988).

IV

The subsection Cyanicae species L. tenerrima and L. perennis (and possibly L. graeca) are not closely related to any of the other European Lactuca species, and should be excluded from the genus Lactuca.

This thesis, Ch. 4, 5.
Contra Feráková (1977).

V

In contrast to the common situation in higher plants (Barow and Meister, 2002), DNA content and AT frequency among Lactuca s.l. species are significantly (and negatively) correlated.

This thesis, Ch. 6.

VI

Considering the general congruence of AFLP based phenograms and AFLP based cladograms, mistaking UPGMA clustering for a phylogenetic analysis (e.g. Sharma et al., 1996; Caicedo et al., 1999; Aggarwal et al., 1999) probably still results in reliable conclusions on evolutionary relationships.

Koopman, unpublished data.

Zooming in on the lettuce genome:

species relationships in *Lactuca* s.l.,
inferred from chromosomal and molecular characters

Wim J. M. Koopman

Contents

Introduction

The genus *Lactuca* L. and its relatives

According to Bremer (1994), the genus *Lactuca* can be classified in the Asteraceae subfamily Cichorioideae, tribe Lactuceae Cass., subtribe Lactucinae Dumort. Tribe Lactuceae is subdivided into 11 subtribes, together comprising more than 1550 species in 96 genera (excluding microspecies of *Taraxacum* Weber in F. H. Wigg., *Hieracium* L., and *Pilosella* Hill). Subtribe Lactucinae comprises ca. 270 species in 17 genera. Tribe Lactuceae is unequivocally characterized by its ligulate capitula and its milky latex. The circumscription of the subtribe Lactucinae is less straightforward, involving terms such as "often", "frequently", "usually", and "sometimes" (Bremer, 1994, p. 183). To make things worse, the generic boundaries within the tribe are even more obscure, causing genera to be repetitively synonymized and resurrected (Cronquist, 1985).

In this thesis, I will refer to *Lactuca* and related genera according to the genus concept of Feráková (1977). Her concept represents a view somewhere between splitting and lumping, including genera such as *Mulgedium* Cass., *Lactucopsis* Schultz-Bip. ex Vis. et Panc., and *Phaenixopus* Cass. in *Lactuca*, while for example the genera *Mycelis* Cass. and *Cicerbita* Wallr. are considered separate. *Lactuca* according to Feráková (1977) is subdivided into four sections: *Phaenixopus* (Cass.) Benth., *Mulgedium* (Cass.) C. B. Clarke, *Lactucopsis* (Schultz-Bip. ex Vis. et Panc.) Rouy, and *Lactuca*. Section *Lactuca* is subdivided into two subsections: *Cyanicae* DC. and *Lactuca* (including *L. sativa* L., the cultivated lettuce). *L. sativa* will be referred to as "lettuce".

In addition to *Lactuca* sensu Feráková (1977), *Lactuca* sensu lato (s.l.) is used for easy reference to *Lactuca* sensu Feráková and a group of closely related genera. I define *Lactuca* s.l. approximately according to Stebbins (1937). In his concept, *Lactuca* not only includes *Mulgedium*, *Lactucopsis*, and *Phaenixopus*, but *Mycelis* and part of *Cicerbita* (excluding *C. alpina* (L.) Wallr. and *C. pancicii* (Vis.) Beauverd.) as well. *Lactuca* s.l. as used in this thesis includes *Lactuca* sensu Stebbins (1937), supplemented with *C. alpina* (*C. pancicii* was not examined), and *Steptorhamphus tuberosus*. The latter was not mentioned by Stebbins (1937), but fits his description of the genus *Lactuca*. The genus concepts of Feráková (1977), Stebbins (1937), and others are discussed in more detail in chapter 4.

In the course of my research, living material from 18 species of *Lactuca* s.l. became available, and all species were included. The species represent all sections and subsections of Feráková (1977). In order to properly evaluate the relationships in *Lactuca* s.l., I also included four outgroup species from genera less closely related to *Lactuca*, viz. *Prenanthes purpurea* L., *Chondrilla juncea* L., *Taraxacum officinale* Weber in F. H. Wigg., and *Sonchus asper* (L.) Hill (Stebbins, 1937; Bremer, 1994). Among these species, *P. purpurea* is considered most closely related to *Lactuca* s.l. Both *Lactuca* and *Prenanthes* are in subtribe Lactucinae (Bremer, 1994), and according to Stebbins (1953) they are connected in the *Prenanthes-Lactuca* line. The remaining species are classified in subtribe Crepidinae Dumort. (*C. juncea* and *T. officinale*), and

in subtribe Sonchinae K. Bremer (*Sonchus asper*). Apart from *P. purpurea* and species from *Lactuca* s.l., I also included *Cichorium intybus* L. The subtribal affinities of *Cichorium* L. are unclear, and I hoped that the inclusion of a *Cichorium* species would contribute to clarifying the subtribal position of the genus. In addition, it enabled a connection of my work to that of Annemieke Kiers (Kiers, 2000) in *Cichorium*.

Objectives

The research for the present thesis had two main objectives.

The first objective was to determine the boundaries and species relationships in the genus *Lactuca*. Because the boundaries of *Lactuca* are a subject of controversy among systematists, species from genera related to *Lactuca* were included as well. Morphology-based descriptions and classifications of all species were available at the start of the project, and therefore I focused on resolving the evolutionary relationships among *Lactuca* s.l. species. I specifically addressed 1) the boundaries and phylogenetic relationships among the closely related species *L. sativa*, *L. serriola*, *L. dregeana*, *L. altaica*, and *L. aculeata* (in this thesis referred to as "*serriola*-like species"); 2) the evolutionary relationships among *L. sativa*, *L. serriola*, *L. saligna*, and *L. virosa*; 3) the evolutionary relationships in *Lactuca* s.l. in relation to the classification of Feráková (1977).

The second objective was to relate the phylogenetic position of *Lactuca* s.l. species to the position of these species in the lettuce gene pool (see Harlan and De Wet (1971) and Discussion). Traditionally, the position of a species in a gene pool is determined by hybridization experiments (e.g. crossing or somatic hybridization), but hybridization data were available for a limited number of *Lactuca* s.l. species only. A preliminary literature search, however, revealed that the position of a species in the lettuce gene pool is related to its evolutionary distance to cultivated lettuce (discussed in chapter 4). I used this relation to predict the position of a species in the lettuce gene pool from its position in *Lactuca* s.l. phylogenies.

In addition to the main objectives, there were two secondary objectives. The first one was to contribute to the development of a theoretical framework for the use of AFLP markers in systematics, the second one was to look for additional practical applications of the research.

Outline and concept of the thesis

The present thesis is divided into four parts, related to different areas of systematic research: herbarium systematics, chromosome morphology and molecular systematics, cladistic theory, and practical applications.

Part I: Herbarium systematics

This first part of the thesis is on 'classical' herbarium study, basic to the chromosome - and molecular work in the remaining parts.

Chapter 1
A format for Asteraceae model voucher collections, developed in Lactuca L.

The basis of all plant systematic research is the voucher collection, that is used for research and future reference. In Lactuceae, plants typically show three growth stages, with an often distinct morphology. During my herbarium studies, I noticed that most collections included only the flowering stage, usually lacking ripe fruits. As a consequence, important diagnostic characters are missing in these collections. I developed a guideline format to ensure the inclusion of all important characters in my own vouchers. Because the format is generally applicable in Asteraceae, I named it Asteraceae Herbarium Collection (AHC) format. The first chapter of this thesis describes the format.

Part II: Zooming in on the lettuce genome

This part includes five chapters, comprising the core of the thesis. The concept of the thesis, "zooming in on the lettuce genome", is developed in this part.

   The usual approach towards molecular systematics is to select a number of sequences showing the proper level of variation for the research questions at hand, perform cladistic analyses using these sequences, and compile the various results into a final conclusion regarding species relationships. For the present thesis, I decided on a different approach. Instead of focusing on various sequences as information sources, I collected my data by examining the nuclear genome (in the following referred to as "genome") at different levels of detail. I named the underlying concept "zooming in on the lettuce genome". There are four "zoom levels", considering the genome in increasing detail: chromosome morphology, total DNA amounts and base composition, genome fragments (AFLP markers), and genome sequences (ITS-1). I first imagined the process to be a "zooming in", but organisatorial factors caused the projects for the three most detailed zoom levels to be carried out and published in reverse order. Because each of the papers includes references to earlier ones, they are best appreciated when read in the same order as they were published. Therefore, this part of the "zooming in" (chapters 4, 5, 6) is presented as a "zooming out".

   The benefit of the "zooming" approach is that it provides the characters needed to determine species boundaries and relationships, as well as information on evolution of the genome itself. In my opinion, the ultimate study of species evolution is a study of the evolution of entire genomes. The evolution of the *Lactuca* s.l. genome is discussed in chapter 6, and in a separate paragraph of the general discussion.

Chapter 2

Chromosome banding patterns in lettuce species (Lactuca sect. Lactuca, Compositae)

Literally, the most superficial level at which one can look at genomes is that of chromosome morphology. The most informative moment to observe chromosome morphology is when its constituent parts, the chromosomes, are at metaphase. In metaphase, individual chromosomes can usually be distinguished. For this chapter, we treated lettuce metaphase chromosomes with various chemicals, in order to reveal banding patterns (indicating differences between DNA classes), and proteins associated with the DNA. We examined the chromosomes of cultivated lettuce (*L. sativa*) and its three most important wild genitors *L. serriola*, *L. saligna*, and *L. virosa*. Phenetic relationships among the four species were determined from the similarities in chromosome morphology.

Chapter 3

A numerical analysis of karyotype and DNA amounts in lettuce cultivars and species (Lactuca subsect. Lactuca, Compositae)

For this chapter, we determined various parameters describing the morphology of the metaphase chromosomes. Additionally, we determined the total DNA amount in the chromosome complement, relative to the DNA amount of tomato. The measurements were conducted in the same accessions of *L. sativa*, *L. serriola*, *L. saligna*, and *L. virosa* that were examined in chapter 2. All parameters were entered in a numerical analysis (Principal Component Analysis or PCA), and we determined the phenetic relationships among the four species from their positions relative to each other. DNA amounts of the four species are examined more extensively in chapters 6 and 8.

Chapter 4

Phylogenetic relationships among Lactuca (Asteraceae) species and related genera based on ITS-1 DNA sequences

The previous chapters were limited to phenetic analyses of *L. sativa* and its most important wild lettuce genitors. In this chapter we included all 23 available species, and performed both phenetic and cladistic analyses. The genomes were sampled at their most basic level: that of individual nucleotides. We determined ITS-1 (Internal Transcribed Spacer-1) sequences, and analyzed them using Fitch parsimony, neighbor-joining, and maximum-likelihood. Evolutionary relationships among the species were inferred from the resulting phylogenies, and related to the gene-pool concept of Harlan and De Wet (1971).

Chapter 5
Species relationships in Lactuca s.l. (Lactuceae, Asteraceae) inferred from AFLP fingerprints

In this chapter, we generated AFLP profiles for *P. purpurea* and all ingroup species that were sequenced for the previous chapter. Using AFLP fragments, we examined the variation in short nucleotide stretches from the genome without identification of individual nucleotides. We performed and compared a variety of phenetic and phylogenetic analyses on the AFLP data. We examined the species relationships in *Lactuca* s.l. in comparison with the sequence information from chapter 4. Phylogenetic analysis of AFLP data is disputed by some researchers, based on theoretical considerations. However, this dispute is not yet very evident in literature. We engaged in the dispute by comparing phenetic and phylogenetic analyses as a method to detect the influence of theoretically expected problems in actual data sets. To our knowledge, this paper is the first to discuss the issue in relation to an actual AFLP phylogeny. The discussion is elaborated upon in chapter 7.

Chapter 6
Evolution of DNA content and base composition in Lactuca (Asteraceae) and related genera

This chapter is the closing piece of part II, connecting all other chapters in this part. The genome is examined at its second least discriminating level, that of total DNA amount and base composition. Absolute DNA content and base composition were estimated for all species from chapters 4 and 5, constituting first estimates for most of them. The DNA contents and base compositions were analyzed both phenetically and cladistically. To our knowledge we are the first to present a genus level cladistic analysis of DNA contents in dicotyledons, and the first to present a cladistic analysis of base composition data in higher plants. We examined the relation between genome size and base composition proposed by Vinogradov (1994) and disputed by Meister and Barow (2001). We related the DNA content and AFLP data in order to examine, for the first time in higher plants, the relation between genome size and number of AFLP bands as assumed by Vos et al. (1995).

Part III: Cladistic theory

AFLP is a DNA fingerprinting technique developed in the mid-1990s, (Vos et al., 1995), and the Biosystematics Group of Wageningen University was among the first to apply AFLP markers to plant systematics (Kardolus, Van Eck, and Van den Berg, 1998). In this theoretical part of the thesis, I take this development a step further, being among the first to address the use of AFLP markers as cladistic characters.

Chapter 7

Phylogenetic signal in AFLP data sets

Elaborating upon the results from chapter 5, I inventoried the features that - in theory - could make AFLP data unsuitable for cladistic analysis. The main problem is that these features may obscure the phylogenetic signal in the data sets. Therefore, I tested the presence of this signal in the AFLP data sets from chapter 5, using various statistical procedures. The phylogenetic nature of the signal was examined by comparison with the ITS-1 results from chapter 4. The general applicability of the results and conclusions on phylogenetic signal in the *Lactuca* s.l. data sets was examined by comparison with results of previous studies on a wide range of taxa.

Part IV: Practical applications

In part II, boundaries and species relationships in *Lactuca* s.l. were examined in relation to the gene-pool concept of Harlan and De Wet (1971). Although the purpose of the research was practical (generating insight in the lettuce gene pool), much consideration was given to theoretical issues as well. Part III is dedicated to theoretical issues entirely. In contrast, Part IV deals with practical applications: chapter 8 addresses the identification of lettuce genitors, while chapter 9 demonstrates the use of plant systematic research for plant breeders.

Chapter 8

Identifying lettuce species (Lactuca subsect. Lactuca, Asteraceae): A practical application of flow cytometry.

L. serriola, L. saligna, and *L. virosa* are the most important wild genitors of cultivated lettuce, but the distinction of these species is sometimes problematic. The identification problem hinders their use as lettuce genitors. In chapter 3, I demonstrated significant differences in relative DNA amounts among *L. serriola, L. saligna,* and *L. virosa.* Here, I examine whether such differences can be used to identify arbitrary samples of the species.

Chapter 9

Plant systematics as a useful tool for plant breeders: examples from lettuce.

According to Stuessy (1990), the field of systematics covers three related areas: taxonomy (i.e. classification with its related activities of identification and nomenclature), the study of phylogeny, and the study of the process of evolution. In this chapter, the use of plant systematic knowledge to plant breeders is demonstrated with examples from *Lactuca* for each area of systematics. The examples are taken from chapters 2, 3, 4, and 8.

Literature

Bremer, K. 1994. Asteraceae: cladistics and classification. Timber Press, Portland, Oregon, USA.

Cronquist, A. 1985. History of generic concepts in the Compositae. *Taxon* 34: 6-10.

Feráková, V. 1977. The genus *Lactuca* L. in Europe. Univerzita Komenského, Bratislava.

Harlan, J. R., and J. M. J. De Wet. 1971. Towards a rational classification of cultivated plants. *Taxon* 20: 509-517.

Kardolus, J. P., H. J. Van Eck, and R. G. Van den Berg. 1998. The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy (Solanaceae). *Plant Systematics and Evolution* 210: 87-103.

Kiers, A. M. 2000. Endive, chicory, and their wild relatives. A systematic and phylogenetic study of *Cichorium* (Asteraceae). *Gorteria Suppl.* 5: 77 pp.

Meister, A., and M. Barow. 2001. Lack of correlation between AT frequency and genome size in higher plants and the effect of non-randomness of base sequences on dye binding. *Cytometry* 47: 1-7.

Stebbins, G. L. 1937. Critical notes on *Lactuca* and related genera. *Journal of Botany*: 12-18.

Stebbins, G. L. 1953. A new classification of the tribe Cichorieae, family Compositae. *Madrono* 12: 33-64.

Stuessy, T. F. 1990. Classification, Taxonomy, and Systematics. P. 5-13 *In* T. F. Stuessy [ed.], Plant Taxonomy: The systematic evaluation of comparitive data. Columbia University Press, New York.

Vinogradov, A. E. 1994. Measurement by Flow Cytometry of Genomic AT/GC Ratio and Genome Size. *Cytometry* 16: 34-40.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407-4414.

1

A format for *Asteraceae* model voucher collections, developed in *Lactuca* L.

W. J. M. Koopman[1]

[1] Biosystematics Group, Nationaal Herbarium Nederland - Wageningen University branch, Generaal Foulkesweg 37, 6703 BL Wageningen, The Netherlands

Summary

Many *Asteraceae* species exhibit a life cycle comprising rosette, bolting, and flowering stages, while fruits are developed only at the end of the flowering stage. Most diagnostic characters are present in the fruits, which are often indispensable for species determination. Additional diagnostic features may be present in other parts of the plants, but not all features are present at all growth stages. To preserve all characters, the ideal *Asteraceae* herbarium collection should comprise plants at all growth stages, including ripe fruits. In reality, most collections presently comprise only one of the stages, often without fruits. To ensure the inclusion of all possible characters in *Lactuca* s.l. voucher collections prepared for the NHN (National Herbarium of the Netherlands), we devised a special format for a comprehensive *Asteraceae* voucher collection. The so-called *Asteraceae* Herbarium Collection (AHC) format includes herbarium specimens at all three growth stages, descriptions of these specimens, photographs of the living plants, a photograph of the flower head in polar and side view, a sample of 25 fruits including pappus and a description, and a permanent microscopic preparation of the pappus in polar and in side view. A sample of total genomic DNA completes the voucher. The AHC format may serve as a blueprint for future *Asteraceae* voucher collections, both within and outside *Lactuceae*.

Key words: *Asteraceae*, herbarium collection, DNA samples, *Lactuca*, *Asteraceae* Herbarium Collection (AHC) format, voucher specimens.

The NHN *Lactuca* projects

The "Nationaal Herbarium Nederland" (National Herbarium of the Netherlands; NHN; http://www.nationaalherbarium.nl/) harbors an estimated total of 5,800,000 specimens; 4,100,000 in Leiden (L), 840,000 in Utrecht (U), and 825,000 in Wageningen (WAG). These specimens include herbarium sheets, wood samples, and liquid-preserved material. *Poaceae* are the most important family in the collection, while *Asteraceae* and *Leguminosae* rank second and third, respectively. At present, *Asteraceae* research within the NHN focuses on *Lactuca* L. and *Senecio* L. (Pelser, 2000; Pelser & al., 2000), while a project on *Cichorium* L. was recently completed (Kiers, 2000; Kiers & al., 1999, 2000).

The research program on *Lactuca* has been ongoing for nearly 20 years, organized in three research projects. The projects primarily dealt with cultivated lettuce (*Lactuca sativa* L.) and its most important wild genitors *L. serriola* L. (prickly lettuce), *L. saligna* L. (least lettuce), and *L. virosa* L. (great lettuce). The first project was carried out in Wageningen by I. M. De Vries, and yielded a total of 265 specimens of *L. sativa*, *L. serriola*, *L. saligna*, and *L. virosa*, comprising 610 herbarium sheets (De Vries, 1990, 1996, 1997; De Vries & Jarvis, 1987; De Vries & Van Raamsdonk, 1994). The collection was deposited in Wageningen. The second project was carried out in Wageningen and Leiden by F. T. Frietema de Vries, and yielded about 200-250 specimens of *L. sativa*, *L. serriola*, and *L. virosa*, deposited in Leiden (De Vries & al., 1992; Frietema de Vries, 1996; Frietema de Vries & al., 1994). The third project was carried out in Wageningen by W. J. M. Koopman, and yielded a total of 757 specimens on 3022 sheets, comprising *L. sativa* (including 33 cultivars), *L. indica* L. (including 3 cultivars), 24 wild species related to *L. sativa* and *L. indica*, and interspecific hybrids from 5 species combinations (Koopman, 1999, 2000; Koopman & al., 1993; Koopman & De Jong, 1996; Koopman & al., 1998, 2001). The collection was deposited in Wageningen.

Lactuceae voucher collections

During the *Lactuca* projects, we realized that *Lactuca* species may have distinctly different habits at various stages of their life cycle. As most tribe *Lactuceae* species, the majority of *Lactuca* genotypes start their life cycle growing a leaf rosette. Following the rosette stage, the plants develop a flowering stem (the process called bolting), and an inflorescence. During bolting, the rosette leaves usually die off. In some cases, the cauline leaves also die off before or during flowering. The rosette leaves, cauline leaves, and the leaves in the inflorescence (if present) usually show a distinct morphology. Although the characters of the various leaf types may be a useful aid in species identification, most diagnostic characters are found in the fruits. The fruits of Asteraceae typically include a fruit body, a beak, and a pappus disc, and each of these structures (or the absence thereof) may contribute characters essential for species determination.

A herbarium study carried out during the third *Lactuca* project revealed that most existing *Lactuceae* collections comprised material from only one of the growth stages. Usually, this was the flowering stage. Since ripe fruits are often developed only at the end of the flowering stage, many accessions representing flowering plants lacked ripe fruits. Consequently, they also lacked the important fruit characters. Obviously, the ideal *Lactuceae* voucher collection should comprise specimens at all three growth stages (representing all leaf types), as well as ripe fruits. To comply to this standard, we devised a format for *Asteraceae* model voucher collections, to be used in the *Lactuca* s.l. study.

The *Asteraceae* Herbarium Collection (AHC) format

The *Asteraceae* Herbarium Collection format aims at preserving as many plant characters as possible, both in and outside the actual herbarium. The format recognizes four elements: the plant material itself, descriptions of the plant material, photographs of the living plants, and a sample of total genomic DNA.

Plants are collected at three growth stages: rosette stage, bolting stage, and flowering stage. Entire plants are collected for all stages whenever possible. Long plant stems are collected in several pieces, but with preservation of leaf shape - and spinulosity patterns along the stem. Plants at flowering stage are collected with open flowers. Additionally, a sample of 25 mature fruits (including pappus) is collected, and a preparation of the pappus is made. The fruit sample is included on the herbarium sheet containing the inflorescence. The pappus preparation consists of a microscope slide with one pappus disc in polar view and one in side view, the latter including the beak of the fruit, if present. Preparations are made permanent by mounting the pappus discs in a solidifying medium under a cover glass, enabling a detailed examination of the pappus characters with hand lens or (dissecting) microscope. Mounting media such as Entellan® Neu (Merck, Darmstadt, Germany) or Lamb DPX (RA Lamb, Apex, NC, USA) are suitable for this purpose, and the preparations can simply be air-dried.

Detailed descriptions of the living plants at all three stages along with locality information are added to the herbarium labels, according to good herbarium practice (see e.g. Bridson & Forman, 1992). A description of the fruits is added to the fruit sample, stressing the most important fruit characters. Information on the position of the involucrum in heads with ripe fruits is added to the description.

Information on growth habit and general plant morphology is preserved by photographs of the plants at rosette stage (polar view), bolting stage (side view), and flowering stage (side view). Flower heads in full bloom are photographed in polar and in side view, including a scale.

The model collection is completed with a sample of total genomic DNA for use in molecular systematic studies. General DNA isolation protocols based on the CTAB method of Doyle & Doyle (1987) will be applicable for most plant species, although specific protocols may be

necessary when the general protocols fail. Alternatively, leaf samples can be collected for future DNA isolation (Miller, 1999). DNA samples should be taken from the vouchered individuals, or from genotypes closely related to them.

Elements of the AHC format

Figures 1-9 show the various elements of a model voucher collection as prepared for *Lactuca aculeata* Boiss. & Kotschy ex Boiss. accession CGN 15692, according to the AHC format.

Figs. 1, 2. Examples of herbarium sheets, including labels with plant description. Scale bar = 50 mm. **1**, Herbarium sheet of plant at rosette stage. **2**, Herbarium sheet of plant at bolting stage.

In species showing a strongly elongated stem at flowering stage, the (lower) cauline leaves frequently die off before or during flowering. In these species, the bolting stage is included in the herbarium collection to preserve the lower cauline leaves (Fig. 2). For species without strongly elongated flowering stems (e.g. *Lactuca perennis* L., *Lactuca tenerrima* Pourr.), bolting stage and flowering stage exhibit a similar leaf morphology. For these species the bolting stage is not included in the herbarium collection.
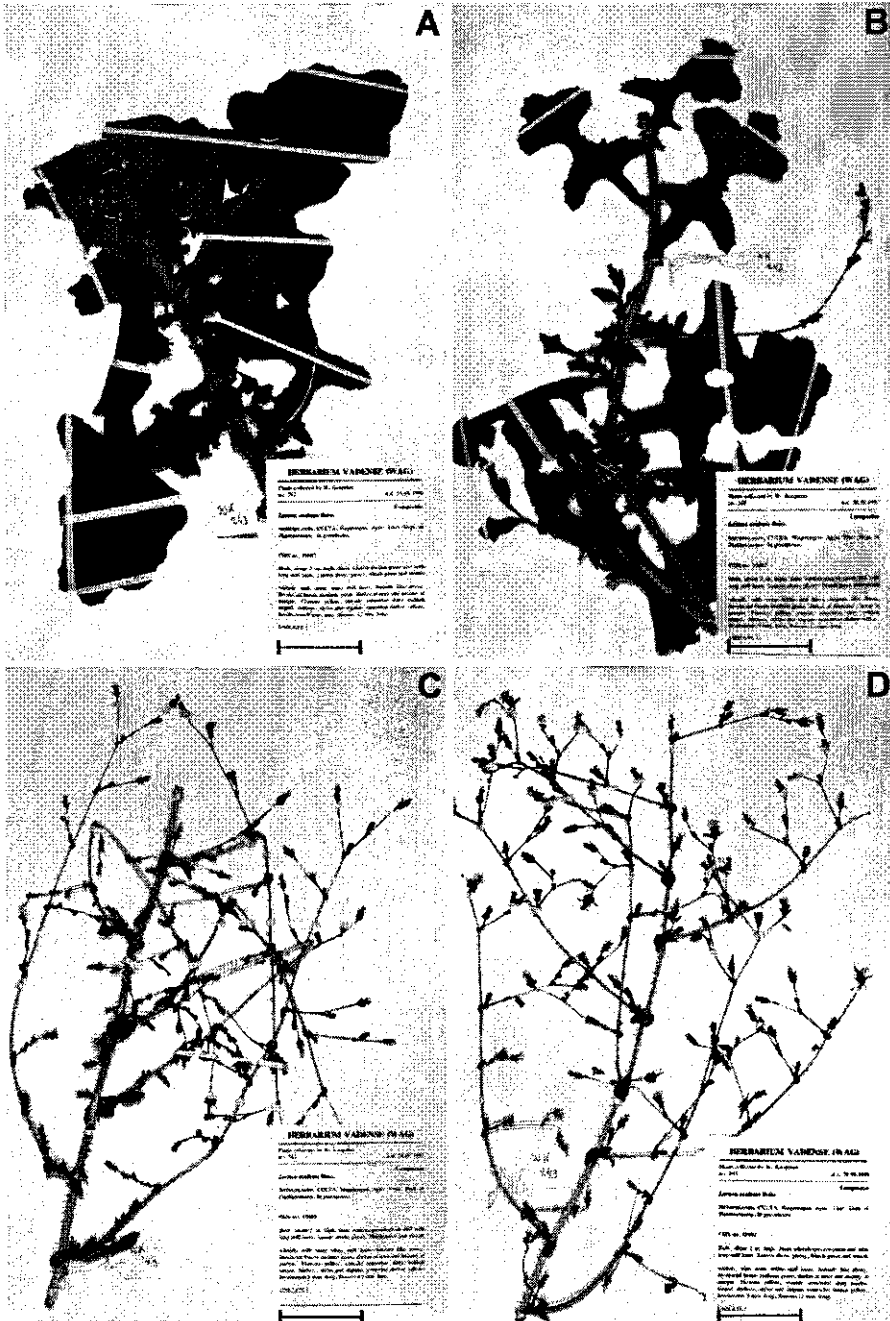
Fig. 3. A-D, herbarium sheets of plant at flowering stage, including labels with plant description. The entire plant is mounted on the four sheets. Scale bars: **A-D** = 50 mm.

Plants at flowering stage frequently show an elongated stem, which may be up to several meters long, especially under greenhouse conditions. Most plant specimens exhibit a gradual transition of leaf shape or spinulosity along the stem, which may be important in identifying the specimen. Therefore, special care should be taken to include all variation along this gradient in the herbarium material. This is done by either preserving the whole plant (Fig. 3), or by preserving several parts of the stem, sampled in such a way that the gradient is adequately reflected in the herbarium accession. The incision of the middle cauline leaves, for example, is diagnostic for distinguishing *L. saligna* var. *saligna* and *L. saligna* var. *runcinata* Gren. & Godr. (Feráková, 1977).
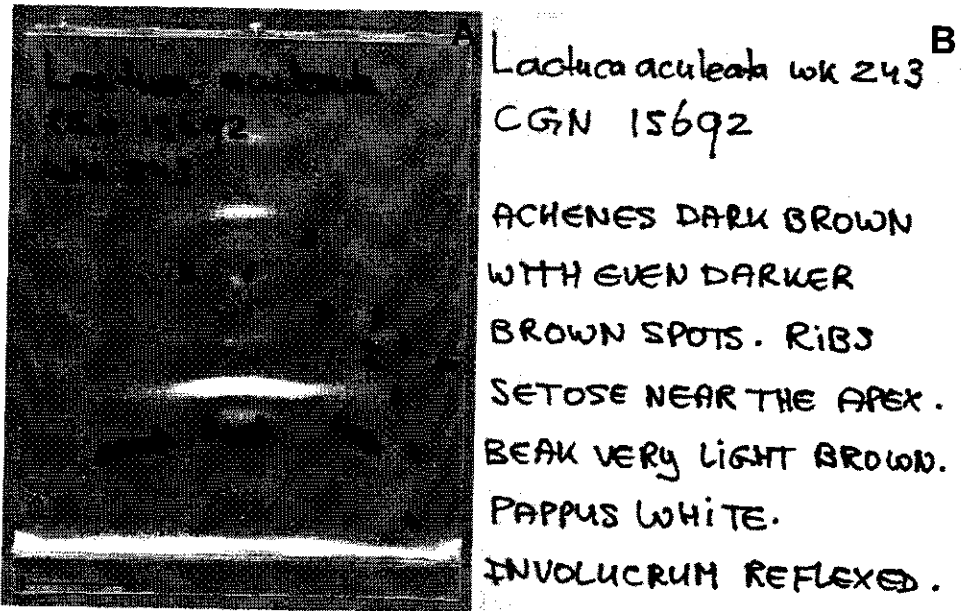
Fig. 4. A, fruit sample. B, description of fruit sample. Scale bar = 15 mm.

A description of the position of the involucrum in flower heads with ripe fruits is included in the description of the fruit sample (Fig. 4), because information on this character is usually not preserved on any of the herbarium sheets. The character is important in distinguishing e.g. *L. sativa* from *L. serriola* (e.g. Feráková, 1977; Frietema de Vries & al., 1994).

The pappus preparation (Fig. 5) enables detailed examination of the finer pappus structures, such as the number of cell rows in the pappus (important in distinguishing e.g. *L. serriola* from *L. saligna*), or the number of basal cells in the pappus hairs (important in distinguishing e.g. *Lactuca* sensu Stebbins from *Prenanthes* L. (Stebbins, 1937)).

Fig. 5. Pappus preparation on microscope slide. Scale bar = 10 mm.

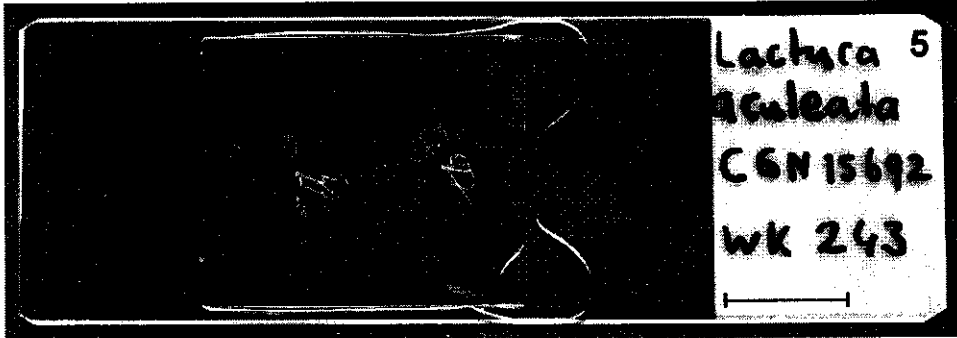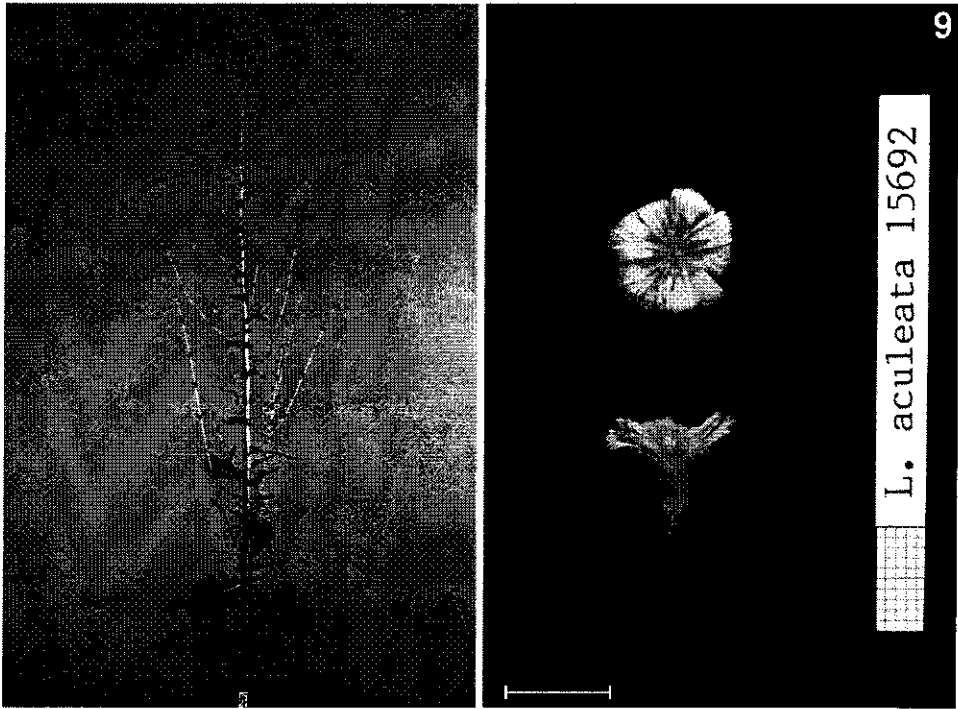Figures 6-9 show photographs of living plants at various stages of the life cycle, exhibiting habitus characteristics that are lost in the preparation of herbarium specimens.

Figs. 6, 7. Photographs of living plants. 6, Plant at rosette stage. Scale bar = 25 mm. 7, Plant at bolting stage. Scale bar = 100 mm.

Figs. 8, 9. Photographs of flowering plant, and of individual flower heads. 8, Plant at flowering stage. Scale bar = 150 mm. 9, Flower heads in polar view and in side view. Scale bar = 10 mm.

Discussion and conclusion

The *Asteraceae* Herbarium Collection format comprises four elements, the more traditional of which are the plant specimens (supplemented with fruits and pappus), and their descriptions. The value of using the AHC format for preparing these elements is that it ensures a coverage of all possible characters. Because at present usually only a part of these characters is preserved in herbarium collections, application of the AHC format is an important improvement on common collecting practice.

Less traditional elements of the AHC format are the photographs and the DNA sample. Including photographs in the AHC format ensures the preservation of two important sets of characters that otherwise would be lost, viz. characters of growth habit and characters of flower head morphology. Growth habit involves characters that can only be observed in intact, living plants, e.g. leaf orientation and three dimensional structure of the inflorescence. Obviously, these characters are lost in the flattened and dried "two dimensional" herbarium specimens. Growth habit can be described, but this usually yields lengthy texts, while it is still difficult to make an adequate reconstruction of the habit based on the description. A photograph of the living plant shows the growth habit information directly, avoiding the twofold loss of

17

information involved in description and subsequent reconstruction of the habit from the description. Flower head morphology also involves characters that are easily lost in the drying process. For example, many species of *Asteraceae* have flower heads with yellow, ligulate flowers. Subtle differences in flower morphology exist among the species, but dried heads containing yellow ligulate flowers usually all look very similar. Again, a description would often be lengthy and inadequate, while a photograph represents the information instantly and completely. By photographing flower heads in both polar and side view, morphological information on the involucrum is also included. A practical point is that the flower heads of *Asteraceae* are especially vulnerable to herbarium insects (see Bridson & Forman (1992) for an overview). The photographs may serve as a source of information on flower head morphology when the actual flower heads are lost due to suboptimal storage conditions (although, of course, the photographs are subject to deterioration, too).

Second of the less traditional elements in the AHC format is the sample of total genomic DNA, although it has to be noticed that suppletion of herbarium accessions with DNA samples is rapidly becoming common practice. It is important that the sample includes nuclear, mitochondrial, and chloroplast DNA, because all three genomes are employed for molecular systematic studies (e.g. Qiu & Palmer, 1999; Qiu & al., 2000; Bowe & al., 2000).

Including all elements, the AHC format is a comprehensive format for collecting herbarium material, ensuring the preservation of all possible characters from the living material. Using the format, one strives to prepare "model collections". In practice it will often not be possible to entirely comply to the format, e.g. because not all elements are represented by the material at hand or because the collecting capacity is limited. This will especially be the case during field expeditions. Returning on a later date to collect additional specimens, or growing additional material to add to the collection, will often be impossible or too expensive. However, in these cases the AHC format still has its value because it provides a clear notion of the elements from which to choose. In cases when ample seeds and greenhouse facilities are available, the full potential of the AHC format can be realized. In accessions prepared using this full potential, the AHC format not only prevents the loss of character information in the preparation of herbarium accessions, but also enables a detailed and comprehensive comparison with other accessions prepared using the format.

In the present study, the AHC format was developed for application in *Lactuceae*. However, many of the limitations for *Lactuceae* collections outlined in this paper also apply to other *Asteraceae* taxa. Although small adjustments may be necessary, the AHC format can be applied in these taxa as well. Therefore, we propose the *Asteraceae* Herbarium Collection (AHC) format as a general model for *Asteraceae* herbarium collections. Adjusted AHC formats may also be applicable outside *Asteraceae*.

Acknowledgements

Literature cited

Bowe, L. M., Coat, G., & dePamphilis, C. W. 2000. Phylogeny of seed plants based on all three genomic compartments: Extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. *Proc. Natl. Acad. Sci. USA* 97: 4092-4097.

Bridson, D. & Forman, L. 1992. The herbarium handbook, Kew.

De Vries, F. T., Van der Meijden, R. & Brandenburg, W. A. 1992. Botanical Files. A study of the real chances for spontaneous gene flow from cultivated plants to the wild flora of the Netherlands. *Gorteria suppl. 1.*

De Vries, I. M. 1990. Crossing experiments of Lettuce cultivars and species (*Lactuca* sect. *Lactuca*, Compositae). *Pl. Syst. Evol.* 171: 233-248.

— 1996. Characterization and identification of *Lactuca sativa* cultivars and wild relatives with SDS electrophoresis. *Genet. Resour. Crop Evol.* 43: 193-202.

— 1997. Origin and domestication of *Lactuca sativa* L. Genet. Resour. Crop Evol. 44: 165-174.

— & Jarvis, C. E. 1987. Typification of seven Linnean names in the genus *Lactuca* L. (Compositae: Lactuceae). *Taxon* 36: 142-154.

— & Van Raamsdonk, L. W. D. 1994. Numerical morphological analysis of Lettuce cultivars and species (*Lactuca* sect. *Lactuca*, Asteraceae). *Pl. Syst. Evol.* 193: 125-141.

Doyle, J. J. & Doyle, J. L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19: 11-15.

Feráková, V. 1977. The genus *Lactuca* L. in Europe, Bratislava.

Frietema de Vries, F. T. 1996. Cultivated plants and the wild flora. Effect analysis by dispersal codes. Leiden.

—, Van der Meijden, R. & Brandenburg, W. A. 1994. Botanical files on lettuce (*Lactuca sativa*). *Gorteria suppl. 2.*

Kiers, A. M. 2000. Endive, chicory, and their wild relatives. A systematic and phylogenetic study of *Cichorium* (Asteraceae). *Gorteria suppl. 5.*

—, Mes, T. H. M., Van der Meijden, R. & Bachmann, K. 1999. Morphologically defined *Cichorium* (Asteraceae) species reflect lineages based on chloroplast and nuclear (ITS) DNA data. *Syst. Bot.* 24: 645-659.

—, Mes, T. H. M., Van der Meijden, R. & Bachmann, K. 2000. A search for diagnostic AFLP markers in *Cichorium* species with emphasis on endive and chicory cultivar groups. *Genome* 43: 470-476.

Koopman, W. J. M. 1999. Plant systematics as a useful tool for plant breeders: examples from lettuce. Pp. 95-105 *in*: Lebeda, A. & Krístková, E. (eds.), Eucarpia leafy vegetables '99, Olomouc.

— 2000. Identifying lettuce species (*Lactuca* subsect. *Lactuca*, Asteraceae): A practical application of flow cytometry. *Euphytica* 116: 151-159.

—, De Jong, J. H. & De Vries, I. M. 1993. Chromosome banding patterns in lettuce species (*Lactuca* sect. *Lactuca*, Compositae). *Pl. Syst. Evol.* 185: 249-257.

— & De Jong, J. H. 1996. A numerical analysis of karyotypes and DNA amounts in lettuce cultivars and species (*Lactuca* subsect. *Lactuca*, Compositae). *Acta Bot. Neerl.* 45: 211-222.

—, Guetta, E., Van de Wiel, C. C. M., Vosman, B. & Van den Berg R. G. 1998. Phylogenetic relationships among *Lactuca* (Asteraceae) species and related genera based on ITS-1 DNA sequences. *Am. J. Bot.* 85: 1517-1530.

—, Zevenbergen, M. J. & Van den Berg, R. G. 2001. Species relationships in *Lactuca* s.l. (Lactuceae, Asteraceae) inferred from AFLP fingerprints. *Am. J. Bot.* 88: 1881-1887.

Miller, J. S. 1999. Banking Desiccated Leaf Material as a Resource for Molecular Phylogenetics. Pp. 331-344 *in*: Metsger, D. A. & Byers, S. C. (eds.), Managing the Modern Herbarium: An Interdisciplinary Approach, Vancouver.

Pelser, P. B. 2000. Moleculaire evolutie in de systematiek. *Gorteria* 26: 249-250.

—, Gravendeel, B. & Van der Meijden, R. 2000. Molecular phylogeny of *Senecio* sect. *Jacobaea* (Asteraceae). P. 42 *in* Bateman, M. & Sanders, L.M. (eds.), Developmental genetics and plant evolution, an international conference of the Systematics Association and the Linnean Society of London.

Qiu, Y-L & Palmer, J. D. 1999. Phylogeny of early land plants: insights from genes and genomes. *Trends Plant Sci.* 4: 26-30.

—, Lee, J., Bernasconi-Quadroni, F., Soltis, D. E., Soltis, P. S., Zanis, M., Zimmer, E. A. Chen, Z., Savolainen, V. & Chase, M. W. 2000. Phylogeny of basal angiosperms: analyses of five genes from three genomes. *Int. J. Plant Sci.* 161: S3-S27.

Stebbins, G. L. 1937. Critical notes on *Lactuca* and related genera. *J. Bot.* 12-18.

2

Chromosome banding patterns in Lettuce species (*Lactuca* sect. *Lactuca*, *Compositae*)

W. J. M. Koopman[1], J. H. De Jong[2], and I. M. De Vries[1]

[1] Biosystematics Group, Nationaal Herbarium Nederland - Wageningen University branch, Generaal Foulkesweg 37, 6703 BL Wageningen, The Netherlands
[2] Laboratory of Genetics, Wageningen University, Arboretumlaan 4, 6703 BD Wageningen, The Netherlands

Summary

Chromosome banding patterns obtained with C- and N- banding, and AgNO$_3$ staining were studied in somatic metaphase complements of four *Lactuca* species. *L. sativa* and *L. serriola* have almost identical chromosome morphology, and *L. saligna* differs only slightly from them, but *L. virosa* is quite distinct from the other species. A gross comparison of the banded karyotypes suggests a closer relationship of *L. saligna* to *L. sativa/serriola* than to *L. virosa*. Our data agree with the results of previous crossing experiments in these species but conflict partly with recent RFLP data which indicate a closer phenetic relationship of *L. saligna* to *L. virosa* than to *L. sativa/serriola*. Such a discrepancy may be explained assuming that domestication of *L. sativa/serriola* resulted in an increased selection pressure on unique DNA sequences as demonstrated by the RFLP data. Differential evolution of specific hererochromatin classes (and presumably of highly repetitive DNA classes), as revealed by chromosome banding techniques was not linked to domestication. Thus the disparity in conclusions about relationship (in terms of genetic similarity) as based on the different experimental approaches reflects a non-parallel evolution of highly repetitive vs. unique DNA classes.

Key words: Compositae, *Lactuca* sect. *Lactuca, Lactuca sativa, L. serriola, L. saligna, L. virosa*, chromosome banding, karyotype evolution, C-banding, N-banding, Ag-NOR staining.

Introduction

At the Department of Plant Taxonomy, Wageningen Agricultural University, the Netherlands, a biosystematic study on cultivated lettuce, *Lactuca sativa* L., and its wild relatives, *L. saligna* L., *L. serriola* L., and *L. virosa* L. (sect. *Lactuca*, Compositae) was initiated to investigate intra- and interspecific variation of these four species. As part of this project, we studied chromosome morphology in these species in order to find differences, if any, between the karyotypes of these taxa. The information thus obtained may complement other taxonomic data on morphological, genetical, and molecular characteristics, and may enlarge our insight into evolutionary relationships between the cultivated lettuce and its related wild *Lactuca* relatives. The taxonomic status of the species is discussed in De Vries & Jarvis (1987).

The most comprehensive study on chromosome morphology of the above mentioned *Lactuca* species was made by Lindqvist (1960a), who also gave an overview on other cytogenetic studies of *Lactuca* species. The basic chromosome number for all *Lactuca* sect. *Lactuca* species was found to be 2n = 18. The karyotypes of *L. sativa*, *L. serriola*, and *L. saligna* showed clear morphological similarities. They all have two satellite chromosome pairs, which are the third and the fifth largest in the complement, but *L. virosa* exhibits differences in chromosome morphology. It has two obvious subtelocentric chromosome pairs with centromere indices of less than 25%, whereas only one chromosome pair possesses microsatellites. Lindqvist (1960a) suggested that the latter chromosome pair in *L. virosa* corresponds to the shorter satellite pair of the other three species. In the present paper we will extend Lindqvist's karyotype analyses with our observations of C-banding, N-banding, and Ag-NOR staining patterns in the chromosomes of these *Lactuca* species.

Materials and methods

Chromosome preparations were made from root tips of young plants of *Lactuca sativa*, *L. saligna*, *L. serriola*, and *L. virosa*. Information on the accessions mentioned in the present study is given by De Vries (1990). Root tips were pretreated in 1.5 mM 8-Hydroxyquinoline, 2¼ hours at 18 °C for metaphase arrest and chromosome shortening. Fixation followed in acetic acid ethanol (1:3) for at least half an hour. Preparations were made according to (i) the standard squash preparation technique and (ii) a cell spreading technique according to Pijnacker & Ferwerda (1984). Briefly, for the squash technique root tips were macerated in 0.2 M HCl for 60 min at room temperature (c. 20 °C) and were rinsed thoroughly in tap water. A single root tip was transferred to a slide and a squash preparation was made in acetic acid 45%. The cover slip was removed after freezing the slide in liquid nitrogen and the preparation was left to dry overnight. For the cell spreading technique root tips were softened in a solution containing a pectolytic enzyme mixture of 0.1% cytohelicase (Institute Biologique Française), 0.1% cellulase RS (Onozuka) and 0.1% pectolyase Y23 in 10 mM citrate buffer, pH 4.3 for 30 - 60 min at 30 °C. Other enzyme compositions may work as well as this solution, but pH,

temperature, and duration of the enzymatic digestion will have to be adapted. After the enzymatic maceration the fragile root meristems were dissected on a clean slide with fine needles in a small drop of acetic acid 60% until a suspension of single cells and small cell clumps was obtained. The cells were spread out on the slide by surrounding the suspension with ice cold, freshly prepared acetic acid ethanol fixative. This process could be helped by tilting the slide and adding some more drops of fixative. Then the preparations were dried overnight.

C-banding

The cell-spread slides were incubated in 0.2 M HCl for 60 min at room temperature. Incubation in 6% Barium hydroxide followed for 15 min at room temperature. The squash preparations were treated in the Barium hydroxide for 5 min at 50 °C. Both types of preparations were rinsed for 20 min and, subsequently, renatured in 2xSSC, pH 7.0 for 30 min at 60 °C. All preparations were finally stained in 1.5% Giemsa in 0.07 M (1/15 M) Sørensen buffer (pH 6.8), air-dried and mounted.

N-banding

One-day-old cell spread slides were incubated in 1M NaH_2PO_4, pH 4.2 for 2 min at 92 ± 2 °C, rinsed in distilled water, stained in 2% Giemsa in 0.1 M Sørensen buffer (pH 6.8), air-dried and mounted.

Ag-NOR staining

Some of the slides were incubated with silver nitrate according to a technique slightly modified after Kodama et al. (1980). A few drops of aqueous 50% (w/v) silver nitrate were put on 2-5 days-old cell spread preparations. The silver nitrate solution was covered with a wet piece of nylon gauze, type Nybolt 3XXX-300 (Swiss, Silk Bolting Cloth Mfg Co. Ltd. Zürich). The slides were transferred to a Petri dish with moistened filter paper and were incubated at 50 °C for about 60 min until the gauze turned brown. Then the preparations were thoroughly rinsed in distilled water, air-dried and mounted.

Microphotographs were taken with a Zeiss Photomikroskop II equipped with high numerical aperture bright field objectives. All chromosome measurements of both banded and unbanded metaphase plates, were taken from the prints, using a digitizing tablet connected to an MS-DOS personal computer. Relative lengths of chromosome arms, and the positions of the chromosome bands were calculated and were used for drawing idiograms.
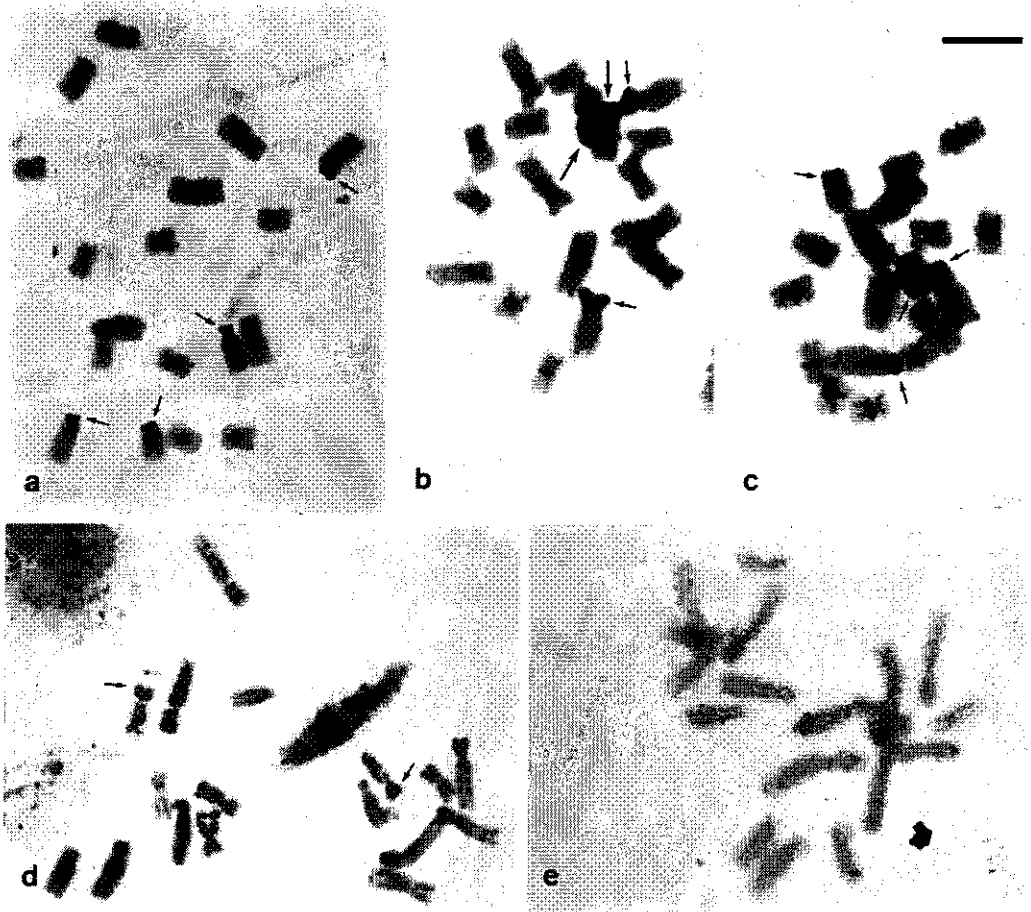
Fig. 1. Giemsa C-banding patterns (**a-d**) in metaphase cells of *Lactuca* species using the cell spreading technique. (**a**) *L. sativa*; (**b**) *L. serriola*; (**c**) *L. saligna*; (**d**) *L. virosa*. (**e**) N-banding in *L. virosa*. The large and small arrows refer to the satellite chromosomes. Bar: 5 μm. In d one chromosome is missing from the complement.

Results

C-banding patterns

The C-banding preparations of *L. sativa* obtained with the squash-technique showed interphase and metaphase nuclei with only large bands in the microsatellites and with no or only very few faint proximal or interstitial chromosome bands. The cell spreading technique, however, revealed metaphase cells with clear proximal bands in addition to the bands in the satellites. The latter bands were the first to appear in insufficiently stained preparations. In darker, well differentiated preparations additional small and variable bands could be detected in the
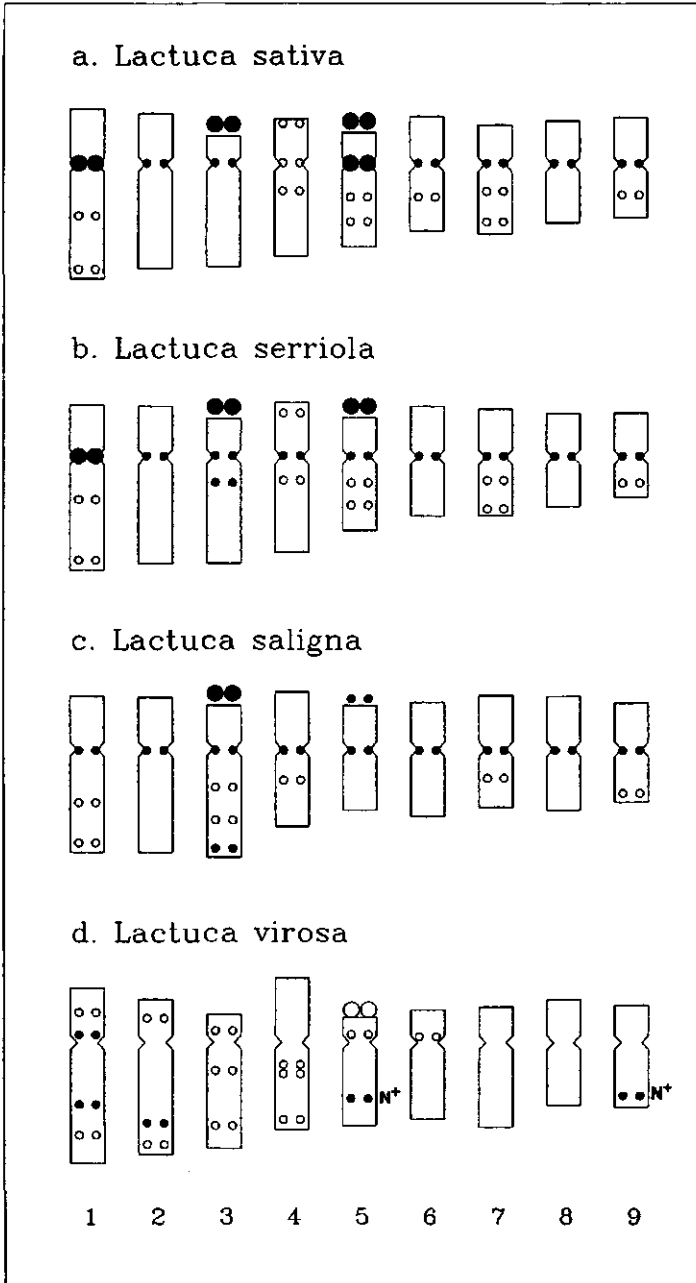
Fig. 2. Chromosome banding idiograms of *Lactuca sativa*, *L. serriola*, *L. saligna*, and *L. virosa*. Filled circles: obvious C-bands; Open circles: variable C-bands. N^+ refer to the C^+N^+ bands in the *L. virosa* chromosomes. For establishing chromosome identification we used relative length, centromere position, presence of satellites and banding pattern as criteria.

interstitial regions of all chromosomes but one (number 8). In most cases such bands serve as diagnostic bands for identifying individual chromosomes. Examples of C-banded metaphase complements are given in Fig. 1a-d. Figure 2 shows the generalized C/N-bands idiograms for the four species.

The overall C-banding profiles in *L. serriola* correspond to that of *L. sativa*. In our material we detected differences only in the occurrence of some bands of chromosomes 3 and 6. In contrast, *L. saligna* has fewer C-bands in common with the former two species. The observed differences mainly concern chromosomes 3 (the large satellite chromosome), 4, 5, and 7. In addition, the satellite bands and the C-bands in the centromere regions of *L. sativa* and *L. serriola*, were generally more obvious than in *L. saligna*. Such a difference may well be explained by differences in staining intensities or by C-banding polymorphism. Diagnostic bands for *L. saligna* were seen in chromosomes 3, 4, 5 and 7, and those in the long arm of chromosomes 1 and 9 were also helpful for identification purposes. In *L. saligna* we found a unique distal band in the long arm of the first satellite chromosome pair (chromosome 3), which was not detected in the corresponding chromosomes of *L. sativa* and *L. serriola*. In most metaphase cells of *L. sativa*, *L. serriola*, and *L. saligna* we found a small interstitial C-band in the smallest chromosome pair of the complement, whereas the same band was detected in the second smallest chromosome pair in some other metaphase cells, which is likely to be the result of small contraction differences between these two chromosome pairs. Such a variation stresses the importance of the diagnostic C-bands in chromosomes 7 and 9 for the identification of the four smallest submetacentric chromosomes.

The C-banding profile in *L. virosa* was found to differ completely from the other three *Lactuca* species. Firstly, we detected no or very faint C-bands at the centromeres, though small C-bands could be discerned in the small arms, close to the centromere of chromosomes 1, 3, 5, and 6. Secondly, the satellite of chromosome 5 was C-positive in some of the cells only. Thirdly, we observed clear interstitial bands in chromosomes 1, 2, 5, and 9, and smaller ones in chromosomes 3 and 4. Though chromosome 9 with its diagnostic band in the distal region of the long arm was chromosome 8 as ordered in sequence of decreasing length, we have classified this chromosome as # 9 according to its putative homeology with the chromosomes 9 of the three other *Lactuca* species.

N-banding patterns

When it became obvious that some of the C-bands, especially those of the satellite regions, differ in the reaction to the BSG-technique, we supposed that probably two or more closely related heterochromatin classes occur in these *Lactuca* species. Surprisingly, the N-banding technique, which we applied for that reason, did not reveal any consistent bands in *L. sativa*, *L serriola* or *L. saligna*, but did show very few minute dots on some of the chromosomes in *L. virosa* (Fig. 1e). Karyotyping such N-banded metaphase plates made it clear that the most striking bands could be assigned to the distal part of the long arm of chromosome 9, and two

other, less noticeable bands were probably located on the long arm of the satellite chromosome (number 5 in the complement). Comparing the N-bands with the C-bands in this species indicated that the two N-bands correspond to the C-bands at similar positions on the chromosomes, a situation comparable for instance, to that in rye (Schlegel & Gill, 1984). These bands on chromosome pairs 5 and 9 are therefore designated C^+N^+ in the idiogram of *L. virosa* in Fig. 2. The remaining faint N-bands were inconsistent and could be shown to correspond to C-bands as well.

Ag-NOR staining

Incubating unstained chromosome preparations in silver nitrate clearly demonstrates locations of metabolically active nucleolar organizer regions in both interphase nuclei and metaphase chromosomes. We applied this technique to *L. virosa* to find out whether chromosomes other than the satellite chromosome pair 5 display any nucleolar activity. As a control we applied Ag-NOR staining in some *L. sativa* preparations. In this species we found the expected four satellite chromosomes with brownish silver deposits (Fig. 3a), though by far the most metaphase plates depict only two or three chromosomes with silver spots at their distal ends. It is assumed that the dark regions represent the NORs of the two satellite chromosomes. Their number varies as a result of physiological suppression. It is not clear as to whether one of the satellite chromosomes acts as a regulator for the total nucleolar activity or that NOR activity is just a process of random regulation. *L. virosa*, the species with only one satellite chromosome pair, always showed metaphase plates with two active NORs (Fig. 3b). Cells with nucleolar activity in more than one chromosome pair, as demonstrated by the silver staining, were not observed.
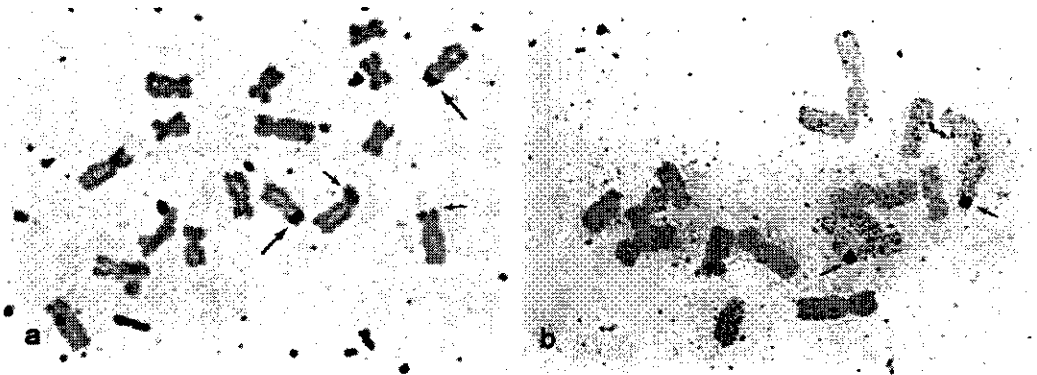

Fig. 3. Ag-NOR staining in metaphase cells of (**a**) *L. sativa* and (**b**) *L. virosa*. The arrows in **a** indicate the silver deposits at the secondary constrictions of the satellite chromosomes.

Discussion

The chromosome banding patterns in preparations obtained with the cell spreading technique are strikingly better than in those made with the squash technique. This improvement of banding differentiation is believed to be the result of both the prolonged pectolytic digestion, which removes most cell wall material that hampers the penetration of chemicals used in the BSG (Barium-SSC-Giemsa staining) procedure, and of the use of acetic acid-ethanol fixative for spreading the softened digested cells. It is also probable that the additional treatment of the fixative on spread cells strongly affects chromatin structure and may so have a positive effect for accomplishing well differentiated banding patterns under standard conditions of the BSG-technique.

The banding patterns of *L. sativa*, *L. serriola*, and *L. saligna* are remarkably similar. All three species have C-bands in all proximal chromosome regions and in the satellites, though their size and staining intensity differ among several putative homeologous chromosomes. The *L. sativa* and *L. serriola* accessions show differences in only three bands on chromosomes 3, 4, and 6; between *L. serriola* and *L. saligna* six bands are lacking in one or the other species, whereas between *L. sativa* and *L. saligna* the C-band karyotypes differ in eight bands. Notwithstanding the C-band polymorphism that causes variation in banding patterns between plants of the same species (see Sumner 1990: Ch.14), comparison of the C-band karyotypes indicates that the (phenetic) relationship between *L. sativa* and *L. serriola* is closer than between those two and *L. saligna*. Lindqvist's (1960a) analysis of unbanded karyotypes and data from crossing experiments in the same species demonstrates a comparable relationship between *L. sativa*, *L. serriola*, and *L. saligna*. He proposed *L. sativa* and *L. serriola*, as the *L. sativa/L. serriola* ecospecies, distinguishing them from *L. saligna*. In the view of De Vries (1990) this ecospecies should be considered as a crop-weed complex. The *L. sativa* types in this complex are the crop component, whereas *L. serriola* is the weed. *L. saligna* is probably also involved. Moreover, Haque & Godward (1985) mentioned the remarkable similarity between three *L. sativa* cultivars and *L. serriola*, both with respect to chromosome length and arm ratio as well as to the characteristics of the satellite chromosomes.

The karyotype of *L. virosa* differs in several aspects from *L. sativa*, *L. serriola*, and *L. saligna* as a group. Not only its unbanded chromosomes in the complement do deviate from those of the other three species, i.e. several chromosomes have more asymmetric centromere positions, but it has only one pair of satellite chromosomes and exhibits an unique C and N banding pattern. Most conspicuous is the absence of centromere bands, the unique C^+N^+ bands and the occurrence of specific intercalary bands not found in the former three species.

The classification of the four species as based on banding patterns supports the view of Zóhary (1991) that *L. serriola* contributes to the primary gene pool of cultivated lettuce, *L. saligna* to the secondary, and *L. virosa* to the tertiary gene pool. Support is given by the crossing experiments of De Vries (1990) which do not contradict previous results of other authors as described in that article.

The above-mentioned results on chromosome morphology and banding patterns partly conflict with recent RFLP data by Kesseli et al. (1991). They constructed a phenogram which shows a division into a wild group including *L. virosa* and *L. saligna* on one hand, and the crop-weed complex on the other. If we consider the degree of similarity in the phenogram as an indication for relationship, this puts *L. saligna* further from *L. sativa/serriola*, but closer to *L. virosa*. Such a discrepancy can be explained by the fact that chromosome banding studies actually elucidate location and variation of chromatin classes containing significant amounts of (highly) repetitive DNA, mainly concentrated in large blocks around centromeres, telomeres, and satellites, perhaps fulfilling some chromosome housekeeping functions, whereas RFLP data are based on variation of cDNA, and thus reflect variation of unique DNA sequences. Studying different types of DNA can thus lead to a disparity in conclusions about relationships (in terms of similarity) between taxa if these types do not show parallel evolution.

The following evolutionary pathway can be postulated: before cultivation began, the evolutionary distance between *L. sativa*, *L. serriola*, and *L. saligna* was close, extremely so for *L. sativa* and *L. serriola*. At that time *L. virosa* had already separated as a distinct species. *L. sativa* and *L. serriola* became influenced by cultivation: *L. sativa* as a cultivated species and *L. serriola* more indirectly, as the related camp follower or weed. *L. saligna* for some reason did not (See also the hybrid pathway scheme for the origin of domesticates and weeds in Small 1984: 198).

Cultivation processes are assumed to cause increased selection pressure on specific genetic properties and so on unique DNA-sequences. As such, *L. sativa* and *L. serriola* developed into a crop-weed complex, leaving *L. saligna* apart as a separate wild species. Selection pressure on those parts of the genome affecting chromosome architecture, banding patterns, and control of nucleolar organizing regions of the satellite chromosomes was evidently different and in the case of these *Lactuca* species not linked to domestication processes.

Lindqvist (1960b) discussed three hypotheses for the origin of cultivated lettuce and concluded that it most probably arose as the product of hybridization between several species. He suggested three possibilities. Firstly, both *L. sativa* and *L. serriola* have developed from heterogeneous hybrid populations, *L. sativa* through selection by man and *L. serriola* through adaptation to man-made waste habitats. Secondly, the progenitors of *L. sativa* were hybrids between *L. serriola* and some other species. Thirdly, *L. serriola* is a product of hybridization between cultivated forms of *L. sativa* and some other species. Our data are in better agreement with his first view.

Acknowledgements

excellent technical assistance. The help of Dr Laas Pijnacker and Mrs Margriet Ferwerda (Department of Genetics, University of Groningen) for the cell spreading technique to root tip meristems is also highly appreciated. Prof. Dr J. van der Maesen, Dr R.G. van den Berg and Dr J.C. Arends of the Department of Plant Taxonomy and Ir. W. Brandenburg of the CPRO-DLO, Wageningen, are acknowledged for their useful comments on the manuscript.

References

Haque, M.Z. & Godward, M.B.E., 1985: Comparison between two genera, species and cultivars in Lactuceae. I. Karyotype analysis. *Cytologia* 50: 725-738.

Kesseli, R., Ochoa, O. & R. Michelmore, R., 1991: Variation at RFLP loci in *Lactuca* ssp. and origin of cultivated lettuce (*L. sativa*). *Genome* 34: 430-436.

Kodama, Y., Yoshida, M.C. & Sasaki, M., 1980: An improved silver staining technique for nuclear organizer regions by using nylon cloth. *Jpn. J. Human Genet.* 25: 229-233.

Lindqvist, K., 1960a: Cytogenetic studies in the *Serriola* group of *Lactuca*. Hereditas 46: 75-151.

Lindqvist, K., 1960b: On the origin of cultivated Lettuce. *Hereditas* 46: 319-350.

Pijnacker, L.P. & Ferwerda, M.A., 1984: Giemsa C-banding chromosomes of potato chromosomes. *Can. J. Genet. Cytol.* 26: 415-419.

Schlegel, R. & Gill, B.S., 1984: N-banding analysis of rye chromosomes and the relationship between N-banded and C-banded heterochromatin. *Can J. Genet. Cytol.* 26: 765-769.

Small, E., 1984: Hybridization in the weed-wild complex. *In* Grant, W.F., (Ed): Plant Biosystematics, pp. 195-210 - Don Mills: Academic Press Canada.

Sumner, A.T., 1990: Chromosome banding. London: Unwin Hyman.

Vries, I.M. DE., 1990: Crossing Experiments of Lettuce cultivars and species (*Lactuca* sect. *Lactuca*, Compositae). *Pl. Syst.Evol.* 171: 233-248.

Vries, I.M. DE & Jarvis, C.E., 1987: Typification of seven Linnaean names in the genus *Lactuca* L. (Compositae: Lactuceae). *Taxon* 36: 142-154.

Zóhary, D., 1991: The wild genetic resources of cultivated lettuce (*Lactuca* sativa L.). *Euphytica* 53: 31-35.

3

A numerical analysis of karyotypes and DNA amounts in lettuce cultivars and species (*Lactuca* subsect. *Lactuca*, Compositae)

W. J. M. Koopman[1] and J. H. De Jong[2]

[1] Biosystematics Group, Nationaal Herbarium Nederland - Wageningen University branch, Generaal Foulkesweg 37, 6703 BL Wageningen, The Netherlands
[2] Laboratory of Genetics, Wageningen University, Arboretumlaan 4, 6703 BD Wageningen, The Netherlands

Summary

Karyotype and relative DNA content were used to characterize *Lactuca sativa*, *L. serriola*, *L. saligna* and *L. virosa* and to determine their evolutionary relationships. In these species karyotype analyses requiring the identification of the homologues are unreliable, because not all chromosomes can be distinguished by their length and centromere position, and no useful additional cytological markers are available. Therefore the karyotypes were established using numerical parameters describing the whole metaphase complement rather than the individual chromosomes, namely: intra and interchromosomal asymmetry index, total chromosome length and area, and number of discernable satellites. The karyotype data were supplemented with data on relative DNA content. No significant differences were found between *L. sativa* and *L. serriola*, whereas *L. saligna* differed significantly from *L. sativa/serriola* only in its relative DNA amount. *L. virosa* differed from *L. saligna* and *L. sativa/serriola* for all parameters. The largest differences were found between *L. saligna* and *L. virosa*, although both have asymmetric karyotypes compared to *L. sativa/serriola*. Since asymmetric karyotypes in Compositae tribe Cichorieae (including *Lactuca*), are considered to be derived it follows that *L. saligna* and *L. virosa* are advanced species that evolved in different directions.

Key words: Asteraceae, Compositae, karyotype analysis, *Lactuca*, lettuce.

Introduction

Ferakova (1977) proposed a subdivision of the west European *Lactuca* L. species into four sections: *Phaenixopus* (Cass.) Benth., *Mulgedium* (Cass.) C.B. Clarke, *Lactucopsis* (Schultz-Bip. ex Vis. et Panc.) Rouy and *Lactuca*. In the section *Lactuca* two subsections were recognized: *Lactuca* and *Cyanicae* DC. The subsection *Lactuca* comprises the species *L. serriola* L., *L. sativa* L. (cultivated lettuce), *L. saligna* L., *L. altaica* Fisch. et Mey., *L. virosa* L. and *L. livida* Boiss. et Reut. All species in the subsection *Lactuca* are diploids with 2n = 18 chromosomes.

Since 1984, the species *L. serriola*, *L. sativa*, *L. saligna* and *L. virosa* have been the subject of a biosystematic study at the Department of Plant Taxonomy, Wageningen Agricultural University. The objectives of the study were to examine the species boundaries and to determine evolutionary relationships among the species. Karyotype studies and analyses of DNA sequences were part of this study.

Karyotype study is a useful tool in taxonomy either to characterize taxa or to reconstruct their phylogeny (see e.g. Stebbins, 1971). Its value for phylogeny reconstruction in Compositae has been amply demonstrated by Babcock (1947) for *Crepis*. Lindqvist (1960) was the first to establish detailed *Lactuca* karyotypes from chromosome measurements. The karyotypes of *L. sativa* (six accessions) and *L. serriola* (eight accessions) were found to be identical for all cases, which was confirmed later by Chatterjee & Sharma (1969) and Haque & Godward (1985). The karyotype of *L. saligna* (two accessions) was found to be slightly different from that of *L. sativa/L. serriola* while distinct differences were observed between *L. virosa* (three accessions) and the other three species. Lindqvist also studied the shape and number of the microsatellites of the nucleolar organizing chromosomes and found one pair for *L. virosa* and two pairs for the other three species.

Lindqvist (1960) and Haque & Godward (1985) described the chromosome pairs on the basis of length, centromere position and presence of microsatellites. The values for chromosome lengths and arm ratios of putative homologues from different complements were averaged, assuming these to be characteristic for a particular chromosome pair in the karyogram. However, in the case of only slight differences among the non-homologues, chromosome length and arm ratio are unreliable parameters to identify chromosomes. Matérn & Simak (1968), Bentzer et al. (1971) and Fukui & Kakeda (1994) demonstrated that analyses based only on these parameters, give rise to considerable numbers of misidentifications. Because misidentified chromosomes will not be properly ranked when ordered by length in a karyogram, Simak (1962) designated these misidentifications as "reversal of order". Misidentification of chromosome arms of metacentric chromosomes was designated as "arm reversal". Since the differences in lengths and arm ratio values of the subsequent chromosome pairs in the diploid chromosome sets of *Lactuca* are small, there is an actual risk of arm

reversal and reversal of order. Consequently, karyograms as constructed by Lindqvist (1960) and Haque & Godward (1985) are unreliable.

The identification problem could be solved by the use of cytological markers such as C- and N-bands. However, the banding patterns of the individual *Lactuca* chromosomes were insufficiently different to enable identification of all chromosomes in the complement (Koopman et al., 1993). As yet, other cytological markers have not been tested for this purpose and therefore an alternative approach was chosen.

In this paper we applied a numerical analysis of the karyotypes of *L. sativa, L. serriola, L. saligna* and *L. virosa* using parameters for the total cell complement rather than for individual chromosomes. Thus, identification of homologues was no longer necessary and the risk of reversals was avoided. The parameters used describe the karyotype in terms of symmetry (intra and interchromosomal asymmetry index) and amount of chromosome material (total chromosome length and total chromosome area). The karyotype data were supplemented with data on relative DNA content. Using these five parameters, ten metaphase plates per accession were compared in a principal component analysis and an analysis of variance followed by a Tukey HSD procedure. Based on the results, species boundaries and phylogenetic relationships of the four species were discussed. Because the differences in visibility of the satellites among *L. sativa, L. serriola* and *L. saligna* were assumed to have no taxonomical significance, data on the satellites were excluded from the analyses.

Materials and Methods

Plant material

Table 1. *Lactuca* accessions of the Centre for Genetic Resources, The Netherlands (CGN), used for karyotyping. The cultivar groups are according to Rodenburg (1960). *L. serriola, L. saligna* and *L. virosa* are wild species.

species	CGN accession nr.	cultivar name	cultivar group
L. sativa	5979	Balady	Cos
L. sativa	4546	Celtuce	Stalk
L. sativa	4600	Great Lakes 65	Crisphead
L. sativa	4707	Oak Leaf	Cutting
L. sativa	5135	Saffier	Butterhead
L. sativa	4869	Tetue de Nimes	Latin
L. sativa	5208	Mataro Tres Ojos	Cos
L. serriola	10881		
L. saligna	5310		
L. virosa	9315		

Ten *Lactuca* accessions of the Centre for Genetic Resources, The Netherlands (CGN) collection, including the species *L. serriola*, *L. saligna*, *L. virosa* and six *L. sativa* cultivar groups (Rodenburg, 1960) were used (Table 1). Voucher specimens of all accessions were deposited at the Herbarium Vadense (WAG), supplemented with photographs of the plants in rosette, bolting and flowering stage, and with pappus preparations and seed samples.

Young plantlets were grown in the greenhouse at 18/22 °C. Actively growing root tips and young leaves were collected for chromosome preparations and DNA measurements, respectively.

Chromosome preparations

Root tips were collected between 0800 and 0900 h and pretreated in 1.5 Mm 8-hydroxyquinoline for 2¼ h at 18 °C for metaphase arrest and chromosome shortening. The material was fixed in acetic acid - ethanol 1:3 and stored at -20 °C until use.

After carefully rinsing in deionized water root tips were hydrolyzed in 1 N HCl at 58 °C for 6¼ min. Subsequently, the root meristems were rinsed again and squashed in a drop of acetic acid 45% on a glass slide. After freezing the slide in liquid nitrogen, the cover slip was removed and the slide was briefly rinsed in acetic acid - ethanol and ethanol 96% steps, respectively. The squash preparations were left to dry overnight, stained in 1% Giemsa in deionized water for three minutes, air-dried and mounted in Entellan-Neu (Merck, Darmstadt, Germany).

For each of the accessions a sample of ten different plants was used for chromosome study. Only one metaphase complement per plant was selected showing well-spread chromosomes with distinctive centromeres, chromatids and satellites.

Chromosome measurements

Chromosomes were measured on enlarged prints at a final magnification of $c. \times 3200$ using a digitizing tablet connected to a PC. To minimize observation inaccuracies, both the short and the long arm lengths of every chromatid were measured three times and their values were averaged.

Three parameters were derived from the arm length data: (i) intrachromosomal and (ii) interchromosomal asymmetry index (Romero Zarco, 1986) and (iii) total chromosome length in μm. The intrachromosomal asymmetry index A_1 equals (1 - complement mean of the ratio of the short and long arm of each chromosome). The interchromosomal asymmetry index A_2 is the ratio of the standard deviation and the mean chromosome length for a complement.

Total chromosome area was estimated by computer imaging. The photo prints were recorded with a CCD camera and digitized by a DT-1451 framegrabber (Data Translation, Marlboro, USA). The final resolution was 0.054 μm/pixel in both directions (image size 512x712 pixels). The images were analyzed using standard routines of the software package Scil-Image (TPD-TNO, Delft, The Netherlands).

DNA measurements

Relative DNA content of four plants of each accession was determined by Plant Cytometry Services (Schijndel, The Netherlands) using a method modified from De Laat & Blaas (1984). The analysis was performed with the ICP 22 (Ortho Diagnostic Systems, Beerse, Belgium) flow cytometer using *Lycopersicon esculentum* 'Tiny Tim' as internal reference. The relative DNA content of each sample was calculated by dividing the median value of the obtained DNA histogram of a *Lactuca* sample by that of the reference.

Statistics

Differences among the accessions regarding asymmetry indices, total chromosome length, total chromosome area and relative DNA content were tested for significance at the 5% level in a one-way analysis of variance followed by a Tukey-HSD procedure using SPSS/PC 4.0 (Norusis, 1990). The NTSYS-PC program version 1.80 (Rohlf, 1993) was used to perform a principal component analysis on A_1, A_2, total chromosome length, total chromosome area and relative DNA content.

Results

Figure 1 gives examples of metaphase complements of the four species. Note the satellite chromosome pairs 7/8 and 9/10 for *L. sativa*, 4/5 and 9/10 for *L. serriola*, 5/6 and 10/11 for *L. saligna* and 9/10 for *L. virosa*. The mean number of visible satellites per complement in *L. sativa* is 3.5 for 'Balady' and 'Tetue de Nimes', 3.8 for 'Celtuce', 'Great Lakes 65' and 'Oak Leaf', 3.9 for 'Saffier' and 4.0 for 'Mataro Tres Ojos'. In *L. serriola* the mean number of visible satellites per complement is 2.5 and in *L. saligna* 3.5. In *L. virosa* two satellites were visible in all cells. Figure 2 shows schematic representations of the metaphase plates of Fig. 1. The chromosomes are ordered in sequence of decreasing length. Reversals of order in the complements of *L. serriola* and *L. saligna* become obvious by the odd number of chromosomes between the satellite chromosome pairs in their karyograms (Fig. 2).

Table 2 presents the data on asymmetry indices, total chromosome area, total chromosome length and relative DNA content of all accessions. The intrachromosomal asymmetry index A_1 of *L. virosa* is significantly higher than that of all other accessions. The differences in A_1 among the other accessions are not significant. *L. saligna* has the highest interchromosomal asymmetry index A_2, followed by *L. sativa* 'Saffier'. The remaining *L. sativa* accessions and *L. serriola* form a group with lower A_2 values than *L. saligna* and *L. sativa* 'Saffier', while only small differences in A_2 within this group occur. *L. virosa* has the smallest A_2 of all of the accessions. Only the differences of *L. saligna* and *L. sativa* 'Saffier' versus *L. virosa* are significant. *L. virosa* has the largest total chromosome area, followed by 'Saffier', 'Great Lakes 65', 'Celtuce', 'Mataro Tres Ojos', 'Balady', *L. serriola*, 'Tetue de Nimes', 'Oak Leaf' and *L.*

saligna, in order of decreasing area. The differences of *L. virosa* versus 'Celtuce', 'Mataro Tres Ojos', 'Balady', *L. serriola*, 'Tetue de Nimes', 'Oak Leaf' and *L. saligna*, and that of *L. saligna* versus *L. sativa* 'Saffier' are significant.
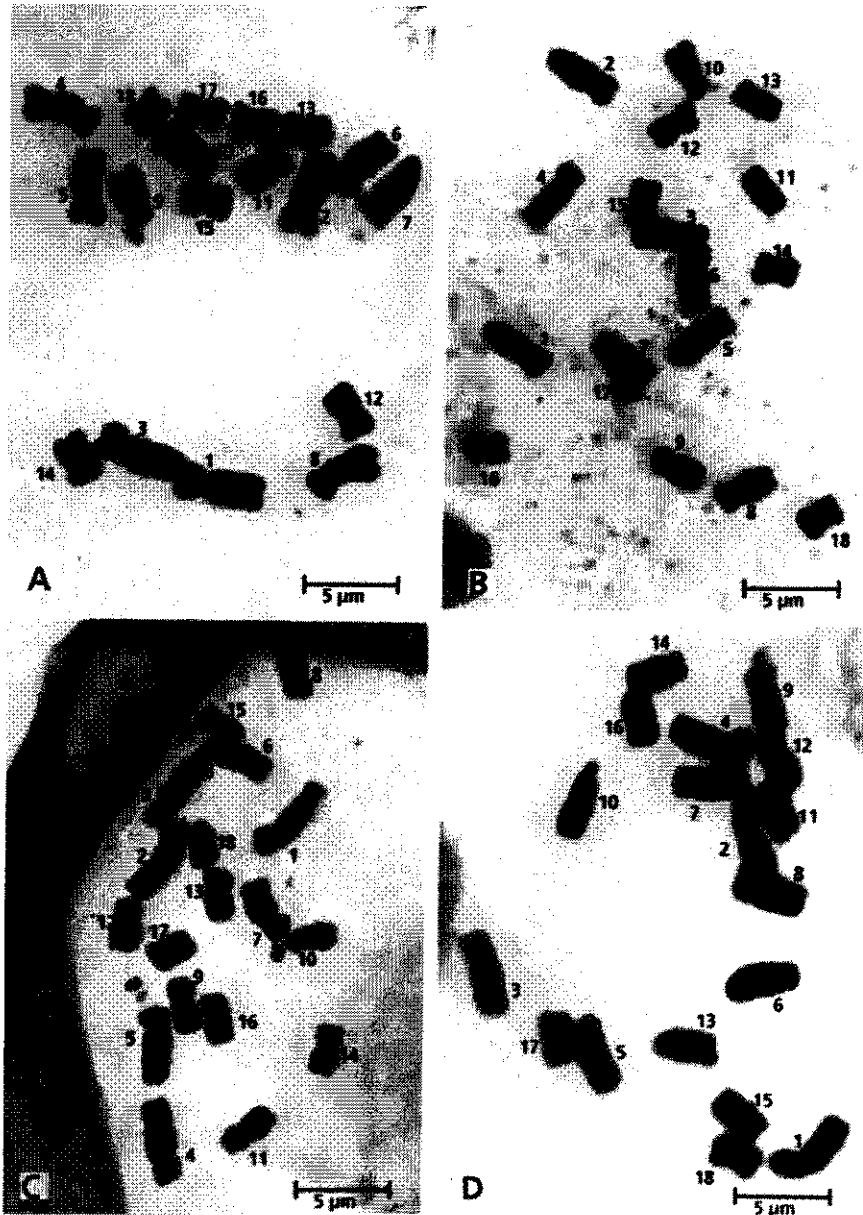

Fig. 1. Metaphase plates of *L. sativa* (**A**), *L. serriola* (**B**), *L. saligna* (**C**) and *L. virosa* (**D**).
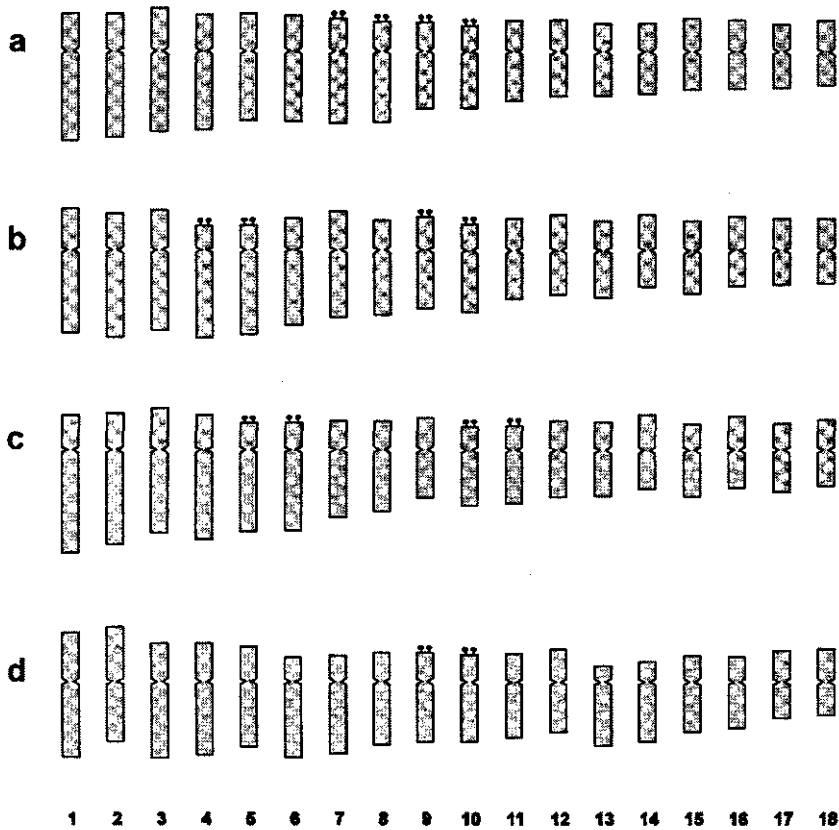
Fig. 2. Idiograms representing the metaphase plates of Fig. 1. *L. sativa* (a), *L. serriola* (b), *L. saligna* (c) and *L. virosa* (d).

L. virosa and *L. sativa* 'Saffier' have the largest total chromosome length. The remaining accessions form a variable group and are in order of decreasing length: *L. serriola*, 'Great Lakes 65', 'Balady', 'Tetue de Nimes', 'Celtuce', 'Oak Leaf', *L. saligna* and 'Mataro Tres Ojos'. *L. virosa* is significantly different from all accessions within this group, apart from 'Saffier'. *L. sativa* 'Saffier' is significantly different from 'Balady', 'Tetue de Nimes', 'Celtuce', 'Oak Leaf', *L. saligna* and 'Mataro Tres Ojos'. None of the other accessions are significantly different from each other.

L. virosa shows the highest and *L. saligna* the lowest relative DNA content of all accessions. The remaining accessions form a group with DNA contents intermediate between *L. saligna* and *L. virosa*. In order of decreasing DNA content these are: 'Tetue de Nimes', 'Balady', *L. serriola*, 'Celtuce', 'Oak Leaf', 'Mataro Tres Ojos', 'Saffier', 'Great Lakes 65'. *L. virosa* and *L. saligna* are significantly different from all of the other accessions. Within the remaining group

of accessions only the differences between 'Tetue de Nimes' and 'Saffier', 'Tetue de Nimes' and 'Great Lakes 65' and between 'Balady' and 'Great Lakes 65' are significant.

Table 2. Means and standard errors of the means of the parameters intrachromosomal asymmetry index (A_1), interchromosomal asymmetry index (A_2), total chromosome area in μm^2, total chromosome length in μm and relative DNA content.

Accession	A_1	A_2	area	length	DNA
L. sativa					
'Balady'	0.377 ± 0.0059	0.203 ± 0.0064	63.87 ± 2.44	1819 ± 27	1.316 ± 0.007
'Celtuce'	0.367 ± 0.0035	0.208 ± 0.0029	65.39 ± 2.12	1764 ± 34	1.308 ± 0.002
'Great Lakes 65'	0.374 ± 0.0059	0.204 ± 0.0043	70.50 ± 1.77	1845 ± 31	1.292 ± 0.002
'Oak Leaf'	0.371 ± 0.0081	0.205 ± 0.0035	62.24 ± 3.64	1758 ± 41	1.306 ± 0.001
'Saffier'	0.383 ± 0.0070	0.219 ± 0.0052	75.18 ± 4.53	2089 ± 99	1.295 ± 0.004
'Tetue de Nimes'	0.376 ± 0.0059	0.207 ± 0.0037	62.97 ± 3.28	1814 ± 56	1.319 ± 0.004
'Mataro Tres Ojos'	0.365 ± 0.0058	0.204 ± 0.0051	64.00 ± 3.05	1669 ± 31	1.302 ± 0.004
L. serriola	0.384 ± 0.0072	0.207 ± 0.0044	63.83 ± 3.56	1867 ± 53	1.314 ± 0.007
L. saligna	0.391 ± 0.0085	0.223 ± 0.0077	58.33 ± 3.35	1727 ± 66	1.120 ± 0.004
L. virosa	0.500 ± 0.0082	0.187 ± 0.0054	81.68 ± 3.78	2131 ± 60	1.669 ± 0.007

PCA results using A_1, A_2, total chromosome length, total chromosome area and relative DNA content are given in Fig. 3. The first axis (PC 1), describing 54% of the variation, is composed of A_1, total chromosome area, total chromosome length and relative DNA content in about equal proportions, and by a smaller proportion of A_2. The second axis (PC2), accounting for 22% of the variation, is mainly determined by A_2 and smaller proportions of total chromosome length and relative DNA content. The third axis (PC3), which describes 16% of the variation is mainly determined by A_1 and total chromosome area and smaller proportions of A_2 and relative DNA content.

All *L. sativa* accessions form a single group, except for three aberrant 'Saffier' complements which are separated along the first and second principal component. The *L. serriola* complements are scattered among those of *L. sativa*. The *L. saligna* complements form a group that is only partially separated from the *L. sativa/serriola* group, along the first and second principal component. *L. virosa* occupies an isolated position mainly due to a separation along the first principal component. *L. virosa* and *L. saligna* are the most dissimilar groups in the PCA.
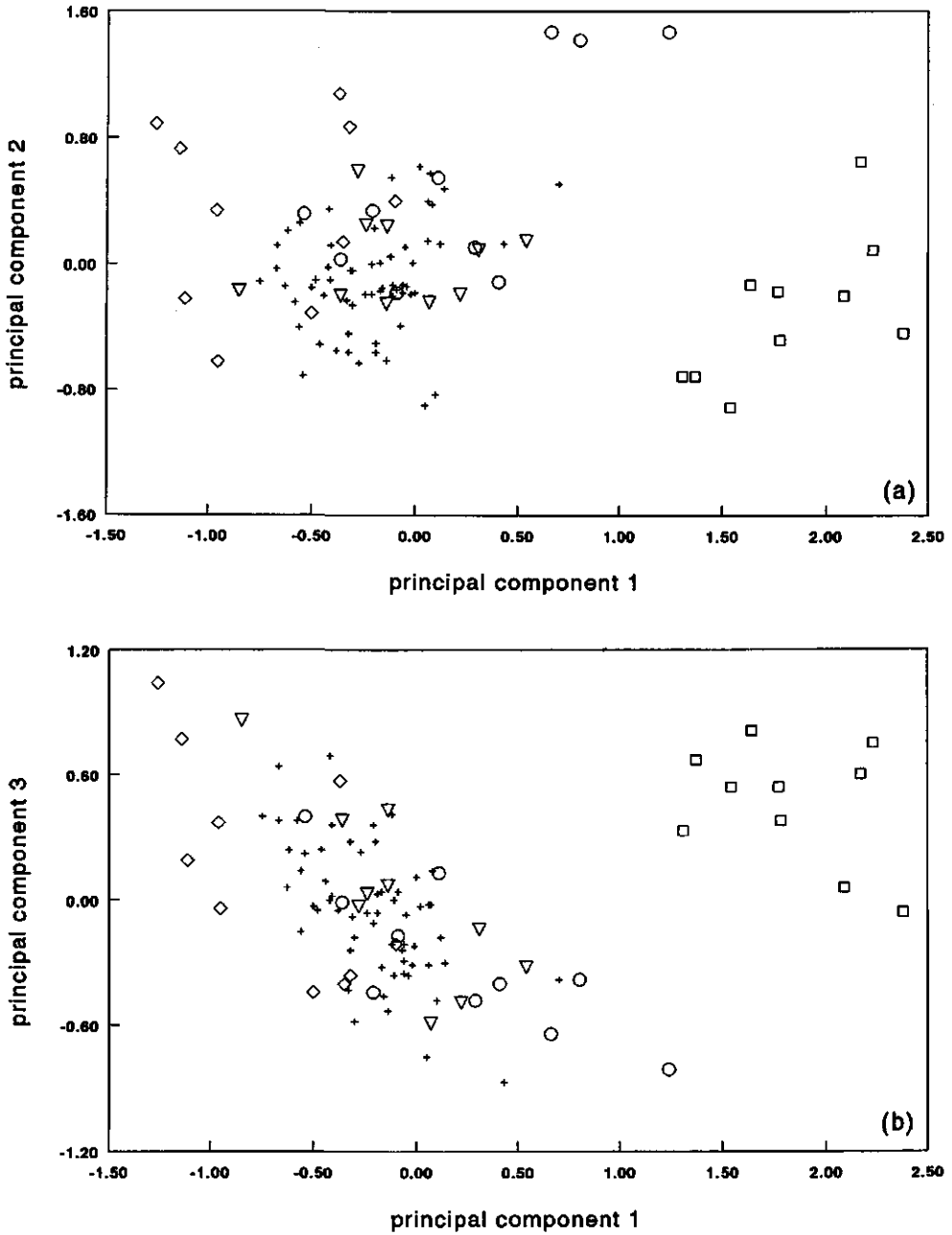
Fig. 3. Principal Component Analysis based on the parameters from Table 2. In both plots each of the ten accessions is represented by 10 metaphases. ○, *L. sativa* 'Saffier'; +, *L. sativa*, remaining accessions; ▽, *L. serriola*; ◇, *L. saligna*; □, *L. virosa*.

Discussion

Our data confirm that arm length differences between the subsequent chromosomes in the chromosome set of *Lactuca* are too small to avoid "reversal of order" and "arm reversal". This makes unequivocal identification of the homologues on the basis of length and centromere position impossible and the karyotypes based on these identifications unreliable. However, the use of numerical parameters for the total cell complement avoids this identification problem.

The karyotype of *L. virosa* as described by these numerical parameters differed from that of the other species in several respects. *L. virosa* had the largest intrachromosomal asymmetry (A_1), the smallest interchromosomal asymmetry (A_2) and the largest genome in terms of total chromosome area and total chromosome length. *L. saligna* had a relatively asymmetric karyotype and a smaller genome. It had the second highest A_1, the highest A_2, the smallest total chromosome area and the second smallest total chromosome length of all accessions. The karyotypes of *L. serriola* and all *L. sativa* accessions except *L. sativa* 'Saffier' were similar to each other and occupied an intermediate position between *L. virosa* and *L. saligna*, but closer to *L. saligna* than to *L. virosa*. For all accessions except *L. sativa* 'Saffier', the differences in relative DNA content showed a similar pattern.

Due to high mean values of A_1, A_2 and total chromosome length and area, the karyotype of *L. sativa* 'Saffier' differed from those of the other accessions in the *sativa/serriola* group (Table 2). The values of total chromosome length and area of the species are positively correlated with those of relative DNA content, suggesting a causality between the parameters. Since this is not the case for *L. sativa* 'Saffier', its large total chromosome length and area likely reflect a lower contraction degree at the time of fixation rather than a large amount of chromosome material. The PCA (Fig 3) showed that this is caused by only three deviating complements. The results of *L. sativa* 'Saffier' are therefore not representative for a regular *L. sativa* karyotype and will be excluded from further discussion.

L. sativa and *L. serriola* cannot be discriminated by their karyotype or relative DNA content, as none of the parameters differed significantly between these species. *L. sativa/serriola* and *L. saligna* can only be discriminated on the basis of their relative DNA content, since they did not show any significant differences for the parameters describing the karyotype. *L. sativa/serriola* and *L. virosa* can be discriminated on the basis of A_1, total chromosome length and relative DNA content. *L. virosa* and *L. saligna* can be discriminated by all five parameters.

These results are in agreement with the conclusions of Lindqvist (1960), Chatterjee & Sharma (1969) and Haque & Godward (1985) that *L. sativa* and *L. serriola* have identical karyotypes. The karyotype of *L. virosa*, described by Lindqvist (1960) as containing more asymmetric chromosomes compared to that of *L. sativa/serriola* was confirmed by our results. The *L. saligna* karyogram established by Lindqvist (1960) shows chromosomes that are shorter and more unequal in length compared to those of *L. sativa/serriola*. Although none of the differences between *L. saligna* and *L. sativa/serriola* were found to be significant in our study,

our data confirm Lindqvist's observations. Therefore his conclusion that the *L. saligna* karyogram is only slightly different from that of *L. sativa/serriola* is supported. In accordance with Lindqvist's observations two pairs of satellites were observed in *L. virosa* and four pairs in *L. sativa*, *L. serriola* and *L. saligna*. The variation in mean number of visible satellites among the accessions can be explained by differences in the state of despiralization of the secondary constrictions in part of the nucleolar organizing chromosomes. If the constrictions are completely condensed, the microsatellites remain tightly attached to the chromosome and are therefore invisible. The extent of despiralization of the secondary constriction reflects metabolic activity of that region rather than polymorphisms for the satellite and so makes the number of visible microsatellites inappropriate for using as a taxonomic parameter. Therefore these differences will not be given any further consideration.

In his discussion on karyotype symmetry in relation to phylogeny and evolutionary processes, Stebbins (1971) assumed a predominant evolutionary trend towards increasing asymmetry in the karyotype. Although opposite trends occur in specific genera (Stebbins, 1971; Jones, 1978), the trends towards increasing asymmetry are particularly obvious within the Compositae, tribe Cichorieae (including *Lactuca*) (Babcock, 1947; Stebbins et al., 1953).

Since all four *Lactuca* species in our study have 18 chromosomes, differences in their karyotypes can be ascribed to processes which do not influence the chromosome number, such as rearrangements within the chromosome arms, pericentric inversions and unequal translocations. Lindqvist (1960) found no multivalents in interspecific hybrids within the subsect. *Lactuca*. Therefore he concluded that the differences in chromosome structure among the species of subsect. *Lactuca* originated in pericentric inversions rather than in translocations. Since this is a process driving a primary trend towards increasing asymmetry, the most asymmetric of the karyotypes in our study, namely that of *L. virosa* and *L. saligna*, can be considered the most derived. Because these are also the most dissimilar karyotypes, they apparently evolved in different directions. Alternatively, gradual deletions and/or duplications of repetitive sequences, and so of some heterochromatin classes may contribute to the shift of centromeres. The differences in banding patterns between *L. sativa/serriola*, *L. saligna* and *L. virosa* as shown by Koopman et al. (1993) are in favor of this assumption.

Using the karyotype parameters and relative DNA contents to characterize the species it can be concluded that:

1: *L. sativa* and *L. serriola* are very closely related or even conspecific, *L. saligna* is a dissimilar but not too distinct species and *L. virosa* is clearly separate from the other three. This is in accordance with data on chromosome banding patterns (Koopman et al., 1993) and crossability (De Vries, 1989). The status of *L. virosa* as a separate species is confirmed by numerical morphological analyses (Frietema de Vries et al., 1994; De Vries & Van Raamsdonk, 1994) as is that of *L. saligna* (De Vries & Van Raamsdonk, 1994). The clusters of *L. sativa* and *L. serriola* showed a slight overlap in these analyses. Frietema de Vries et al. (1994) considered this overlap large enough to lump the species, while De Vries & Van

Raamsdonk (1994) considered it small enough to maintain them. Our results do not support the distinction of *L. sativa* and *L. serriola* as separate species.

2: *L. saligna* and *L. virosa* are the most dissimilar of the four species examined, while the karyotype morphology and relative DNA content of *L. sativa* and *L. serriola* are intermediate between that of *L. saligna* and *L. virosa*. Crossability data (De Vries, 1989) support this position of the species relative to each other. Morphological data only support the intermediate position of *L. serriola* relative to *L. saligna* and *L. virosa*. *L. sativa* occupied a different position in morphological analyses, partly due to the presence of characters caused by domestication (De Vries & Van Raamsdonk, 1994). Apparently, the domestication process is not reflected in the karyotype.

Results on DNA and enzyme analyses are not in accordance with the data on karyotype, relative DNA content, banding patterns, morphology and crossability. RFLP analysis of nuclear DNA shows a closer similarity of *L. saligna* and *L. virosa* to each other than to *L. sativa* and *L. serriola* (Kesseli et al., 1991). Analysis of mitochondrial RFLPs (Vermeulen et al., 1994) makes clear that *L. sativa* and *L. serriola* share more mtDNA fragments with *L. virosa* than with each other, while all three species shared the least fragments with *L. saligna*. Results on isozyme analysis (Kesseli & Michelmore, 1986) showed clusters containing *L. sativa*, *L. saligna*, *L. virosa*, *L. serriola*, *L. virosa* and *L. serriola*, in order of increasing Nei's genetic distance relative to *L. sativa*. Considering the fact that the *L. saligna* and *L. virosa* accessions showing the smallest genetic distance to *L. sativa* were not identified with certainty, in must be concluded that the isozyme results give no unequivocal picture of the species relationships. De Vries (1996) showed that SDS-electrophoresis patterns of achene proteins from *L. sativa* and *L. serriola* were similar, while the patterns of *L. saligna* and *L. virosa* differed from that of *L. sativa/serriola* and from each other. The *L. saligna* pattern was the most dissimilar from that of *L. sativa/serriola*. As becomes clear from the contradictory results mentioned above, further research will be needed to obtain a more obvious view of the relationships among *L. sativa*, *L. serriola*, *L. saligna* and *L. virosa*.

Acknowledgements

References

Babcock, E.B. (1947): The genus *Crepis*, Part one. *Univ. Calif. Publ. Bot.* 21(1): 1-197.

Bentzer, B., Von Bothmer, R., Engstrand L., Gustafsson, M. & Snogerup S. (1971): Some sources of error in the determination of arm ratios of chromosomes. *Bot. Notiser* 124: 65-74.

Chatterjee, T., & Sharma A.K. (1969): Cytotaxonomy of Cichorieae. *Genetica* 40: 577-590.

De Laat, A.M.M., & Blaas, J. (1984): Flow cytometric characterization and sorting of plant chromosomes. *Theor. Appl. Genet.* 67: 463-467.

De Vries, I.M. (1989): Crossing experiments of lettuce cultivars and species. *Pl. Syst. Evol.* 171: 233-248.

De Vries, I.M. & Van Raamsdonk, L.W.D. (1994): Numerical morphological analysis of lettuce cultivars and species (*Lactuca* L. section *Lactuca*). *Pl. Syst. Evol.* 193: 125-141.

De Vries, I.M. (1996): Characterization and identification of *Lactuca sativa* cultivars and wild relatives with SDS-electrophoresis (*Lactuca* sect. *Lactuca*, Compositae). *Genet. Res. Crop Evol.* 43: 193-202.

Ferakova, V. (1977): The genus *Lactuca* L. in Europe. Bratislava: Univerzita Komenskeho.

Frietema de Vries, F.T., Van der Meijden, R. & Brandenburg, W.A. (1994): Botanical files on lettuce. *Gorteria* supplement 2.

Fukui, K. & Kakeda K. (1994): A critical assessment of karyotype analysis by imaging methods. *Jap. J. Genet.* 69: 537-544.

Haque, M.Z. & Godward, M.B.E. (1985): Comparison between two genera, species and cultivars in Lactuceae I: Karyotype analysis. *Cytologia* 5: 725-738.

Jones, K. (1978): Aspects of chromosome evolution in higher plants. *Advances Bot. Res.* 6: 119-194.

Kesseli, R.V. & Michelmore R.W. (1986): Genetic variation and phylogenies detected from isozyme markers in species of *Lactuca. J. Heredity.* 77: 324-331.

Kesseli, R., Ochoa, O. & Michelmore, R. (1991): Variation at RFLP loci in *Lactuca* spp. and origin of cultivated lettuce (*L. sativa*). *Genome* 34: 430-436.

Koopman, W.J.M., De Jong J.H., & De Vries I.M. (1993): Chromosome banding in lettuce species (*Lactuca* sect. *Lactuca*, Compositae). *Pl. Syst. Evol.* 185: 249-257.

Lindqvist, K. (1960): Cytogenetic studies in the *serriola* group of *Lactuca. Hereditas* 46: 75-147.

Matérn, B. & Simak, M. (1968): Statistical problems in karyotype analysis. *Hereditas* 59: 280-288.

Norusis, M.J. (1990): SPSS/PC+ Statistics 4.0. SPSS Inc., Chicago

Romero Zarco, C. (1986): A new method for estimating karyotype asymmetry. *Taxon* 35(3): 526-530.

Rodenburg, C.M. (1960): Varieties of lettuce: An international monograph. Varietal description I.V.T, No 3.

Rohlf, F.J. (1993): NTSYS-PC. Numerical taxonomy and multivariate analysis system version 1.80. Exeter Software, New York.

Simak, M. (1962): Karyotype analysis of *Larix decidua* Mill. from different provenances. *Medd. Statens Skogsforskn. Inst.* 41(1): 1-22

Stebbins, G.L., Jenkins, J.A. & Walters M.S. (1953): Chromosomes and phylogeny in the Compositae, tribe Cichorieae. *Univ. Calif. Publ. Bot.* 26(6): 401-430.

Stebbins, G.L. (1971): Chromosomal evolution in higher plants. Edward Arnold (Publishers) Ltd., London.

Vermeulen, A., Desprez, B., Lancelin, D. & Bannerot, H. (1994): Relationships among Cichorium species and related genera as determined by analysis of mitochondrial RFLPs. *Theor. Appl. Genet.* 88: 159-166.

4

Phylogenetic relationships among *Lactuca* (Asteraceae) species and related genera based on ITS-1 DNA sequences

W. J. M. Koopman[1], E. Guetta[1], C. C. M. van de Wiel[2], B.Vosman[2] and R. G. van den Berg[1]

[1] Biosystematics Group, Nationaal Herbarium Nederland - Wageningen University branch, Wageningen University, Generaal Foulkesweg 37, 6703 BL Wageningen, The Netherlands
[2] Plant Research International B.V., Business Unit Biodiversity and Identity, P.O. Box 16, 6700 AA Wageningen, The Netherlands

Summary

Internal transcribed spacer (ITS-1) sequences from 97 accessions representing 23 species of *Lactuca* and related genera were determined and used to evaluate species relationships of *Lactuca* sensu lato (s.l.). The ITS-1 phylogenies, calculated using PAUP and PHYLIP, correspond better to the classification of Feráková than to other classifications evaluated, although the inclusion of sect. *Lactuca* subsect. *Cyanicae* is not supported. Therefore, exclusion of subsect. *Cyanicae* from *Lactuca* sensu Feráková is proposed. The amended genus contains the entire gene pool (sensu Harlan and de Wet) of cultivated lettuce (*Lactuca sativa*). The position of the species in the amended classification corresponds to their position in the lettuce gene pool. In the ITS-1 phylogenies, a clade with *L. sativa*, *L. serriola*, *L. dregeana*, *L. altaica*, and *L. aculeata* represents the primary gene pool. *L. virosa* and *L. saligna*, branching off closest to this clade, encompass the secondary gene pool. *L. virosa* is possibly of hybrid origin. The primary and secondary gene pool species are classified in sect. *Lactuca* subsect. *Lactuca*. The species *L. quercina*, *L. viminea*, *L. sibirica* and *L. tatarica*, branching off next, represent the tertiary gene pool. They are classified in *Lactuca* sect. *Lactucopsis*, sect. *Phaenixopus*, and sect. *Mulgedium*, respectively. *L. perennis* and *L. tenerrima*, classified in sect. *Lactuca* subsect. *Cyanicae*, form clades with species from related genera and are not part of the lettuce gene pool.

Key words: Asteraceae, *Cichorium*, gene pool, *Lactuca*, Lactuceae, phylogenetic relationships, ribosomal DNA internal transcribed spacer (ITS), *Taraxacum*.

Introduction

Asteraceae (Compositae) systematics at the subtribal level is well characterized by the words of Cronquist (1985): "Generic delimitation in the Compositae is notoriously difficult, because a great many of the recognizable groups are connected by palpable intermediates." This lack of clearly delimited taxa has given rise to many different classifications, as can be seen in the tribe Lactuceae Cass. (or Cichorieae Dumort.). Although the tribe itself is well defined by its milky latex and ligulate florets, the delimitations and phylogenetic relationships of many of its genera and species are still disputed. The numerous, often contradictory classifications for *Lactuca* L. (Lactuceae subtribe Lactucinae Dumort.) and related genera illustrate this dispute.

An economically important member of the genus *Lactuca* is the cultivated lettuce (*Lactuca sativa*), while the closely related species *L. serriola*, *L. saligna*, and *L. virosa* are important genitors for lettuce breeding. In recent years, the use of improved breeding techniques has extended the lettuce gene pool to *L. tatarica* (Chupeau et al., 1994; Maisonneuve et al., 1995). In order to make an accurate choice of the most promising wild species to further broaden the gene pool, more insight is needed into the taxonomic relationships of *Lactuca*. Therefore, a study was undertaken to evaluate the various *Lactuca* classifications and to elucidate the link between species relationships and the possibilities of sexual and somatic hybridization.

Three major generic divisions for *Lactuca* are those of Stebbins (1937), Tuisl (1968), and Feráková (1977). Stebbins (1937) applies a broad genus definition and includes *Mulgedium* Cass., *Lactucopsis* Schultz-Bip. ex Vis. et Panc., *Phaenixopus* Cass., *Mycelis* Cass., and part of *Cicerbita* Wallr. Tuisl (1968) takes the genus in a narrow sense, separating the genera *Mulgedium*, *Scariola* F.W. Schmidt (= *Phaenixopus* Cass.), *Cicerbita*, *Cephalorrhynchus* Boiss., and *Steptorhamphus* Bunge. Feráková (1977) takes an intermediate position including *Mulgedium*, *Lactucopsis*, and *Phaenixopus/Scariola* in *Lactuca* as sections *Mulgedium* (Cass.) C.B. Clarke, *Lactucopsis* (Schultz-Bip. ex Vis. et Panc.) Rouy., and *Phaenixopus* (Cass.) Benth., respectively. Feráková (1977) regards *Mycelis*, *Cicerbita*, *Steptorhamphus*, and *Cephalorrhynchus* as separate genera.

In order to evaluate the relationships of *Lactuca* s.l., a number of less closely related genera were included in our research, viz. *Prenanthes* L., *Chondrilla* L., *Taraxacum* Weber, *Sonchus* L., and *Cichorium* L. Stebbins (1953) considers *Prenanthes* to be related to *Lactuca* s.l., forming part of the so called *Prenanthes-Lactuca* line. *Chondrilla*, *Taraxacum*, and *Sonchus* are generally considered to be more distantly related to *Lactuca*, while the affinities of *Cichorium* are unclear (see, e.g., Bremer 1994). In this paper the subgeneric classification of Feráková (1977) and the generic classification of Bremer (1994) will be used as a starting point (Table 1).

The current subdivisions of *Lactuca* s.l. are mainly based on cytological and morphological characters that often fail to clearly delimit taxa and to recognize unambiguous phylogenetic relationships. During the last decades molecular markers have become available as tools to detect taxonomic units and their relationships. In Asteraceae, ITS sequences proved to be useful for phylogenies at the level of species and closely related genera (Baldwin, 1992, 1993; Kim and Jansen, 1994; Sang et al., 1994; Sang, Crawford, and Stuessy, 1995; Susanna et al.,

51

1995; Bain and Jansen, 1995; Kim et al., 1996). Given these examples, we decided to use ITS sequences to evaluate the classifications of *Lactuca* s.l. In the Lactuceae genera examined so far, ITS-1 was longer and mostly more variable than ITS-2, and thus more suitable for phylogenetic analysis. Moreover, Baldwin (1992, 1993) and Kim and Jansen (1994) showed that in Lactuceae the analysis of ITS-1 resulted in phylogenies that were consistent with and only slightly less resolved than those using a combined data set of both ITS-1 and -2. Therefore we limited our research to ITS-1 sequences.

To facilitate the evaluation of *Lactuca* relationships and classifications, we focused on five research topics: (1) the disputed distinction of the species *L. sativa*, *L. serriola*, *L. dregeana*, and *L. aculeata*; (2) the assumed intermediate position of *L. altaica* between *L. serriola* and *L. saligna* (Feráková, 1977); (3) the position of *L. sativa/L. serriola*, *L. saligna*, and *L. virosa* relative to each other (see Koopman and De Jong, 1996); (4) the boundaries of the genus *Lactuca* and its subgeneric division as proposed by Feráková (1977) in relation to the recognition of the genera *Mycelis*, *Cicerbita*, and *Steptorhamphus*; and (5) the taxonomic position of *Cichorium* (Blackmore, 1981; Bremer, 1994) in relation to the monophyly of the *Prenanthes-Lactuca* line (Stebbins, 1953). These research topics are discussed in relation to various *Lactuca* classifications and morphological, crossability, cytological, isozyme, and molecular data from the literature. In conclusion, an adjusted genus concept for *Lactuca* is proposed based on the ITS-1 results and the subdivision of Feráková (1977). This genus concept is discussed in relation to the gene-pool concept of Harlan and de Wet (1971).

Materials and methods

Plant samples
We used 97 accessions, representing 23 species. The accessions included 11 European *Lactuca* species, one *Lactuca* species from the Middle East (*L. aculeata*), one from South Africa (*L. dregeana*) and one from Asia (*L. indica*), as well as five species from related genera within Lactuceae subtribe Lactucinae and four species outside the subtribe (Table 1). Voucher specimens of the plant material in rosette, bolting, and flowering stage were deposited at the Herbarium Vadense (WAG), supplemented with photographs of the plants in all three stages and with pappus preparations and fruit samples. The plants were grown under standard greenhouse conditions, and fresh young leaf tissue from each plant was collected on liquid nitrogen and kept at -70^0C until use. For DNA extraction, nuclei were isolated from one plant per accession and from the nuclei DNA was purified, using phenol/chloroform extraction as described by Vosman et al. (1992).

Table 1. Lactuceae species used in this study. Subtribal classification according to Bremer (1994), generic and specific classification of European species according to Tutin et al. (1976)/Feráková (1977). Although not treated by Tutin/Feráková, *L. aculeata* from the Middle East and *L. dregeana* from South Africa were included in subsect. *Lactuca* because of their close relationship to *L. serriola* (Zohary, 1991). The Asiatic species *L. indica* is classified in the non-European section *Tuberosae*, according to Iwatsuki et al. (1995).

Subtribe Lactucinae Dumort., genus *Lactuca* L.

Lactuca sect. *Lactuca* subsect. *Lactuca*

Lactuca sativa L.: CGN 5979 (cv. Balady, Bani Swif, Egypt, landrace); CGN 4884 (cv. Verte de Cobham, A.L. Tozer Ltd., Cobham, United Kingdom); CGN 5140 (cv. Capitan, Les Graines Caillard, Angers, France); CGN 5999 (Gradina Botanica a Universitatii din Cluy-Napoca, Cluy-Napoca, Romania); CGN 5045 (Prof. B.M. Kozopolansky Botanical Garden of the Voronezh State University, Voronezh, Rossijskaja, former USSR, landrace); CGN 11387 (cv. Tianjin Big Stem, Institute of vegetables and flowers, Chinese Academy of Agricultural Science, Beijing, China); RKO 93130 (cv. Balisto, Rijk Zwaan B.V., De Lier, The Netherlands); RKO 92296 (cv. Karif, Rijk Zwaan B.V., De Lier, The Netherlands).

Lactuca serriola L.: CGN 10881 (Oudewater, The Netherlands); CGN 5900 (Jerusalem, Israel); CGN 4674 (Botanischer Garten de Universität Goettingen, Germany); CGN 5803 (Botanical Garden, Department of Botany, University of Oulu, Oulu, Finland); CGN 14314 (former USSR); CGN 4667 (Botanische tuinen van de gemeente Rotterdam, Rotterdam, The Netherlands); CGN 15671 (Oshakan, Ashtarakskij, Armenia); CGN 15684 (Antakya, Hatay, Turkey).

Lactuca aculeata Boiss. & Kotschy ex Boiss.: CGN 15692 (Kiziloren, Afyon, Turkey); CGN 9357 (Nov, Israel).

Lactuca dregeana DC.: CGN 4790 (Giardino Botanico e Coloniale dell' Universita di Palermo, Palermo, Italy); CGN 5805 (Jardin Botanique de la Ville, Dijon Cedex, France).

Lactuca saligna L.: CGN 5310 (Raananna, Israel); CGN 5327 (Caca de la Selia, Gerona, Spain); CGN 5301 (Lot, France); CGN 4662 (Botanische tuinen van de gemeente Rotterdam, Rotterdam, The Netherlands); CGN 15705 (Avchala-Tiflis, Tiflis, Georgia); CGN 15697 (Akcay, Balikesir, Turkey).

Lactuca altaïca Fisch. et C.A. Mey.: CGN 4664 (Botanicka Zahrada University Karlovy, Prague, Czechoslovakia); CGN 15711 (Ozero, Dzhandargel, Rustavi, Georgia).

Lactuca virosa L.: CGN 9315 (origin unknown); CGN 4682 (Hortus Botanicus Universitatis Varsaviensis, Warsaw, Poland); CGN 4970 (Jardin Botanique de la Ville, Dijon Cedex, France); CGN 4681 (Hortus Botanicus der Universiteit van Amsterdam, Amsterdam, The Netherlands); CGN 15679 (Rushul, Tabasaran, Daghestan); CGN 13349 (Asturia, Spain); CGN 13350 (Asturia, Spain); CGN 13352 (Asturia, Spain); CGN 5941 (Gadot, Israel); CGN 13339 (Spain); 15680 (Trisandij, Urkarach, Daghestan).

Lactuca sect. *Lactuca* subsect. *Cyanicae* DC.

Lactuca tenerrima Pourr.: CGN 13351 (Asturia, Spain); CGN 9387 (Timhadit, Morocco); CGN 9388 (Collscrola, Barcelona, Spain); CGN 9386 (L' Arralassada, Barcelona, Spain); CGN 14217 (Hortus Universitatis Hauniensis, Botanical Garden of the University of Copenhagen, Copenhagen, Denmark).

Lactuca perennis L.: CGN 13299 (between Vallouise and Puy-Aillaud, Hautes Alpes, France); CGN 9318 (Jardin Botanique de la Ville, Dijon Cedex, France); CGN 9323 (Valais, Switzerland); CGN 10884 (Gradina Botanica a Universitatii din Cluy-Napoca, Cluy-Napoca, Romania); CGN 13294 (Botanischer Garten der Universität (TH), Karlsruhe 1, Germany).

Lactuca sect. *Mulgedium* (Cass.) C. B. Clarke

Lactuca tatarica (L.) C.A. Mey.: CGN 9389 (Central Asia); CGN 9390 (Volgograd, former USSR); CGN 930133 (Delta Danube, Tulcea, Maliuc, Romania); CGN 910430 (Insel Usedom, Bansin, Meckelenburg-Vorpommern, Germany); CGN 930119 (Botanischer Garten der Wilhelm Pieck Universität, Rostock, Germany); W9530 (Yinchuan, Ningxia, China).

Lactuca sibirica (L.) Benth. ex Maxim.: W9513 (River Kulbacksan, Sweden); W9516 (Vanjaurtrask, River Soran, Sweden): W9517 (Knaften, Umeriver, Sweden); W9520 (Lagneset, Oreriver, Sweden); W9523 (Vastana/Mjallan, River Selangersan, Sweden).]

Lactuca sect. *Lactucopsis* (Schultz-Bip ex Vis. et Panc.) Rouy

Lactuca quercina L.: CGN 14220 (Esrergom, Komarom, Hungaria).

Lactuca sect. *Phaenixopus* (Cass.) Benth.

Lactuca viminea (L.) J. & C. Presl: CGN 9326 (Jardim Botanico da Universidade de Coimbra, Coimbra, Portugal); CGN 14301 (Roumodour, France); CGN 16202 (Asktarakskii, Armenia); CGN 926859 (France); CGN 926870 (Italy).

Lactuca sect. *Tuberosae* Boiss.

Lactuca indica L.: CGN 14312 (Cipanas, Cianjur, Indonesia. Landrace, ' Sanelin Lampenas'); CGN 14316 (Anhui, China. Landrace, ' Wo Yang Tai Gan'); CGN 13392 (Hortus Botanicus Pekinensis, Beijing, China); CGN 13393 (Hortus Botanicus Pekinensis, Beijing, China).

Subtribe Lactucinae, other genera.

Mycelis muralis (L.) Dumort.: CGN 9367 (Botanical Garden Floretum Scanicum, Helsingborg, Sweden); CGN 5005 (Botanical Garden of the University of Uppsala, Uppsala, Sweden); CGN 9366 (Passo Muratone, Italy); CGN 9327 (St. Pietro d' Olba, Savona, Italy); CGN 5116 (Botanical Garden of the Armenian Academy of Sciences, Yerevan, Armenia).

Steptorhamphus tuberosus (Jacq.) Grossh.: CGN 9368 (Botanical Garden of the Tel-Aviv University, Tel-Aviv, Israel)

Cicerbita plumieri (L.) Kirschl.: W9501 (Ht. Folin, Nievre, France); W9531 (Vallée de Galbe, Dept. Pyrénées Orientales, France); W9532 (Route Formiguères-Quérigut, Dept. Pyrénées Orientales, France).

Cicerbita alpina (L.) Wallr.: W9507 (Botanischer Garten der Friedrich Schiller Universität, Jena, Germany); W9508 (Botanischer Garten der Martin-Luther Universität, Halle-Saale, Germany); W9527 (Conservatoire et Jardin Botanique de la Ville de Geneva, Chambesy-Genève, Switzerland; wild origin).

Prenanthes purpurea L.: W9534 (KM8 Route Formiguères-Quérigut, Dept. Pyrénées Orientales, France); W9505 (Botanischer Garten der Universität Frankfurk/Main, Germany); W9525 (Conservatoire et Jardin Botanique de la Ville de Genève, Chambesy-Genève, Switzerland; wild origin).

Subtribe Crepidinae Dumort.

Chondrilla juncea L.: CGN 9391 (Instituto Botanico dell' Universita, Siena, Italy); CGN 13308 (Botanischer Garten der Freidrich Schiller Universität, Jena, Germany); CGN 14218 (Bok Ata, Israel).

Taraxacum officinale Weber in F.H. Wigg.: Fr9 (Col de la Croix de Marchampt 670 m, Beaujeu, France); Fr13 (Col de la Croix de Marchampt 670 m, Beaujeu, France); W9606 (Wageningen, The Netherlands).

Subtribe Sonchinae K. Bremer

Sonchus asper (L.) Hill: W9510 (Botanischer Garten der Martin-Luther Universität, Halle-Saale, Germany); W9526 (Conservatoire et Jardin Botanique de la Ville de Genève, Chambesy-Genève, Switzerland; wild origin); W9539 (Coimbra, Portugal).

Unassigned to a subtribe

Cichorium intybus L.: W9601 (Renkum, The Netherlands); W9602 (Botanischer Garten der Universität Dresden, Germany; wild origin); W9603 (cv. Hollandse middel vroeg, Oranjeband zaden/Nunhems, Dronten, The Netherlands).

Sequencing strategy

The ITS-1 was amplified in two steps. In the first step the entire ITS region, including ITS-1 and - 2 plus the interjacent 5.8S rRNA gene, was amplified using the primers "ITS5" and "ITS4" from White et al. (1990). In each PCR (polymerase chain reaction), 10 ng of nuclear DNA were used in a total volume of 25 µL, containing 10 mmol/L Tris-HCl pH 9.0, 50 mmol/L KCl, 1.5 mmol/L $MgCl_2$, 0.01% (w/v) gelatin, 0.1% Triton X-100, 100 µmol/L of each dNTP, 50 ng of each primer, and 0.15 units polymerase (Super Trouper, HT Biotechnology Ltd., UK). Template DNA was denatured for 3 min at 94^0C, followed by 30 cycles of 45 s at 94^0C, 45 s at 55^0C, and 1.45 min at 72^0C on a Hybaid thermal cycler. Final extension consisted of 3 min at 72^0C. In the second step, ITS-1 was amplified with the primers "ITS-5" and "ITS-2" from White et al. (1990) using 40 µL of a 1000 times dilution of the entire ITS-PCR product in a total volume of 100 µL. Again, the template was denatured for 3 min at 94^0C, this time followed by 15 cycles of 45 s at 94^0C, 45 s at 60^0C, and 1.45 min at 72^0C. The final extension consisted of 3 min at 72^0C. The amplified product was purified for sequencing by cutting out the ITS-1 band after electrophoresis on a 2% agarose gel in TBE buffer. DNA was extracted from the agarose by electro-elution, followed by phenol/chloroform extraction and ethanol precipitation according to standard procedures (Sambrook, Fritsch, and Maniatis, 1989). Sequencing of ITS-1 sequences was performed with the Promega Silver Sequence[TM] DNA Sequencing System (Promega, Madison, Wisconsin, USA), according to the manufacturer's instructions, for part of the accessions. The ITS-1 was sequenced in both directions with the above-mentioned primers "ITS5" and "ITS2" using an annealing temperature of 55^0C and 60 cycles on a Perkin-Elmer Thermal Cycler 480. Sequencing reactions were run on a 6% polyacrylamide gel. After silver staining, the gels were recorded on Kodak Duplicating RA 1 film. For another part of the accessions, ITS-1 sequencing was performed with the Applied Biosystems DyeDeoxy Terminator Cycle Sequencing Kit containing the Amplitaq FS DNA polymerase (Perkin-Elmer Applied Biosystems, Foster City, California, USA). Again, ITS-1 was sequenced in both directions with the primers "ITS-5" and "ITS-2" as above. Sequences were run on an Applied Biosystems 373. The few accessions sequenced with both methods gave identical results. All sequences obtained were deposited in the EMBL, GenBank, and DDBJ Nucleotide Sequence Databases under the accession numbers GBANAJ228605 to GBANAJ228661 (Table 3).

Sequence analysis

The single alignments of the ITS-1 sequences were made using the program Micro Genie. Spacer boundaries were determined by comparison with Asteraceae sequences from Baldwin (1993), and tribe Lactuceae species sequences from Kim and Jansen (1994) and Kim et al. (1996). The multiple alignments were done with "PileUp" from the Wisconsin Sequence Analysis Package™ using a Gap Weight of 1.000 and a Gap Length Weight of 0.300.

Data sets and outgroup selection

Two data sets were analyzed. (1) One set contained all different sequences found in this study, supplemented with GenBank sequences GBANL13957 (*L. sativa*), GBANL48143 (*L. perennis*), GBANL48151 (*P. purpurea*), GBANL48301 (*S. asper*), and GBANL48337 (*T. officinale*). Sequences that were found to be identical among different accessions were entered only once in this data set. In the case of identical sequences among accessions, one accession was chosen arbitrarily to represent the sequence. The selected accessions and the corresponding accessions with identical sequences are listed in Table 2.

Table 2. Repeatedly occurring sequences in our study. The accessions selected to represent such sequences in the analyses (first data set, see also Fig. 1) are listed in the left column. Accessions with sequences identical to those of the selected accessions are listed in the right column. Accessions representing sequences unique in our study are not listed.

Accession in analysis		Accessions with identical sequences	
Lactuca sativa	CGN 5999	*L. sativa*	CGN 5979, 4884, 5140, 5045, 11387, RKO 93130, 92296
		L. serriola	CGN 5900, 4674, 5803, 14314, 15671, 15684
		L. dregeana	CGN 4790, 5805
		L. altaica	CGN 4664
L. aculeata	CGN 15692	*L. aculeata*	CGN 9357
L. virosa	CGN 9315	*L. virosa*	CGN 13339
	CGN 4682		CGN 4970, 4681, 5941
	CGN 13349		CGN 13350, 13352
L. saligna	CGN 5310	*L. saligna*	CGN 15705
L. sibirica	W9513	*L. sibirica*	W9517
	W9520	*L. sibirica*	W9523
L. viminea	CGN 9326	*L. viminea*	CGN 16202, 926859, 926870
L. tatarica	CGN 910430	*L. tatarica*	CGN 930119
L. indica	CGN 13393	*L. indica*	CGN 14312, 14316
L. perennis	CGN 9318	*L. perennis*	CGN 13299
Cicerbita plumieri	W9532	*C. plumieri*	W9531
L. tenerrima	CGN 9387	*L. tenerrima*	CGN 9388, 14217
C. alpina	W9508	*C. alpina*	W9527
Cichorium intybus	W9602	*C. intybus*	W9603
Prenanthes purpurea	W9534	*P. purpurea*	W9505, W 9525
Taraxacum officinale	fr 9	*T. officinale*	W9606
Chondrilla juncea	CGN 9391	*C. juncea*	CGN 13308, 14218

The ingroup species were either members of Lactuceae subtribe Lactucinae or, in the case of *Cichorium intybus*, unassigned to a subtribe (Bremer, 1994). Three species from subtribes related to Lactucinae were used as outgroup species: *T. officinale* and *C. juncea* (Lactuceae subtribe Crepidinae) and *S. asper* (Lactuceae subtribe Sonchinae). Since this data set was too large to perform a branch and bound search, it was analyzed using a heuristic search. As a result, not all most parsimonious trees were found. (2) A smaller data set was compiled containing consensus sequences of the subtribe Lactucinae species only. The consensus sequences were obtained by merging all sequences of a species found in our study using the ambiguity codes of the NC-IUB for both ambiguous and variable positions. Additionally, all gaps present in any of the accessions of a species were introduced into its consensus sequence. In this consensus data set, *P. purpurea* was the subtribe Lactucinae species most distantly related to *Lactuca* (Stebbins, 1953) and therefore it was used as outgroup. The resulting data set was small enough to enable a branch and bound search yielding all possible most parsimonious trees.

Phylogenetic analysis

Both data sets were examined with PAUP version 3.1.1 (Swofford, 1993) using Fitch parsimony. Two searches were performed on the first data set. Firstly, to determine the length of the shortest possible tree without the risk of being stuck on an island of suboptimal trees, 1200 replicates of a heuristic search with ACCTRAN (accelerated transformation), multistate taxa interpreted as uncertainty, collapse of zero-length branches, random taxon addition, and TBR (tree bisection-reconnection) without MULPARS were conducted. The MULPARS option requests the saving of all equally most parsimonious trees. Without this option in effect, only one shortest tree was saved in each replicate. Secondly, a heuristic search with ACCTRAN, multistate taxa interpreted as uncertainty, collapse of zero-length branches, simple taxon addition, TBR, and MULPARS was conducted yielding a set of most parsimonious trees from which a strict consensus tree was calculated. The second data set containing the consensus sequences was used to evaluate the topology of the strict consensus trees from the first data set. A branch and bound search was performed in PAUP with ACCTRAN, collapse of zero-length branches, furthest taxon addition with MULPARS, and multistate taxa interpreted as uncertainty (first run) and polymorphisms (second run).

Bootstrap values were calculated in 1000 replications of a heuristic search with ACCTRAN, multistate taxa interpreted as uncertainty (first data set) or polymorphisms (second data set), collapse of zero-length branches, simple taxon addition, and TBR without MULPARS. The amount of phylogenetic signal in the data sets was determined from the tree-length distribution of 100 000 random trees (multistate taxa interpreted as uncertainty) using the g_1 statistic (Hillis and Huelsenbeck, 1992). Sequence divergence values between species were calculated in PHYLIP 3.572 (Felsenstein, 1993) with DNADIST and the Kimura two-parameter method. The transition/transversion ratios were calculated in MacClade 3.04 (Maddison and Maddison,

1993) as the average ratio across 100 most parsimonious trees for the first data set, and as the average ratio across all 17 most parsimonious trees for the second data set. A neighbor-joining tree (based on the sequence divergence values) and a maximum-likelihood tree (using the empirical base frequencies) were calculated in PHYLIP 3.572 for both data sets. A distance matrix containing both absolute and mean pairwise distances between the accessions was generated with PAUP (Table 3).

Results

Both data sets

The ITS-1 sequences ranged in length from 248 to 253 bp for the *Lactuca* species, from 250 to 254 for the remaining subtribe Lactucinae species, and from 252 to 257 for the outgroup species and *C. intybus*, which is well within the range of the Asteraceae ITS-1 lengths published so far (see Baldwin, 1992, 1993; Kim and Jansen, 1994; Sang et al., 1994; Sang, Crawford, and Stuessy, 1995; Susanna et al., 1995; Bain and Jansen, 1995, and Kim et al., 1996). Intraspecific length differences in ITS-1 sequences were found only in *L. serriola*, *L. perennis*, and *L. virosa*, although the sequences found for *L. perennis*, *P. purpurea*, *S. asper*, and *T. officinale* differed slightly in length from those in GenBank. Small differences in individual nucleotides among the accessions were present in almost all species, but the largest intraspecific differences were found relative to the GenBank accessions (Table 3).

First data set

The total aligned length in the data set containing all sequences was 269 bp, including 166 variable sites, 142 of which were phylogenetically informative. The amount of phylogenetic signal was highly significant: the g_1 of -0.75 is considerably lower than the critical value of -0.09 (P = 0.01, Hillis and Huelsenbeck, 1992). In the 1200 replications with random taxon addition, 878 shortest trees of 434 steps, a CI of 0.57 and a RI of 0.85 were found. The search with simple taxon addition and MULPARS yielded 900 trees (the extent of the tree buffer) of 434 steps, a CI of 0.57, a RI of 0.85, and an average transition/transversion ratio across the first 100 trees of 1.26.

In all trees generated, the species *L. sativa*, *L. serriola*, *L. dregeana*, and *L. altaica* form a clade with highly similar sequences. The *L. aculeata* sequence is distinguishable but only slightly different from that of the species in this clade. The sequences of all *L. sativa* accessions in our data set were identical to each other, to that of both *L. dregeana* accessions, to that of six of the eight *L. serriola* accessions and to one of the two *L. altaica* accessions. The *L. serriola* and *L. altaica* sequences that differed from the joint *L. sativa* /serriola /dregeana /altaica sequence deviated in only one or two positions, while the *L. sativa* GenBank accession GBANL 13957 differed in four positions (Table 3).

Table 3. Pairwise distances among species, calculated with PAUP. Absolute distances between the species (in numbers of positions/basepairs) appear in the lower left half of the matrix, the mean distances (calculated according to Swofford, 1993) appear in the upper right half.

Species in our study	Sequence[a]	*S. asper*	*S. asper*[*]	*C. intybus*
S. asper (3,3)[b]	GBANAJ228659-61	0/0	0.004	0.301- 0.305
S. asper[*] (1,1)	GBANL48301	1	--	0.306
C. intybus (3,2)	GBANAJ228653-54	74-75	75	**0.012/3**
T. officinale (3,2)	GBANAJ228656-57	65-66	66	66-68
T. officinale[*] (1,1)	GBANL48337	66-67	67	68-69
L. tenerrima (5,3)	GBANAJ228642-44	69-71	71-72	58-59
L. saligna (6,5)	GBANAJ228618-22	61-64	63-65	49-52
L. virosa (11,5)	GBANAJ228613-17	60-63	62-64	52-57
L. sat/L. dreg (10,1)	GBANAJ228605/09	57-58	59	52-53
L. serriola (8,3)	GBANAJ228606-08	57-59	58-59	52-54
L. aculeata (2,1)	GBANAJ228612	57-58	59	53
L. altaica (2,2)	GBANAJ228610-11	57-60	59-61	52-53
L. sativa[*] (1,1)	GBANL13957	61-62	63	54-55
L. quercina (1,1)	GBANAJ228623	59-60	61	48-49
L. viminea (5,2)	GBANAJ228627-28	67-69	69-70	52-53
L. tatarica (6,5)	GBANAJ228629-33	62-65	63-65	53-56
L. sibirica (5,3)	GBANAJ228624-26	61-64	63-65	54-56
C. alpina (3,3)	GBANAJ228651-52	63-66	65-67	47-50
M. muralis (5,5)	GBANAJ228646-50	68-70	70-71	55-58
L. indica (4,2)	GBANAJ228634/35	63-65	64-65	55-57
C. plumieri (3,2)	GBANAJ228640/41	65-66	67	53-54
L. perennis (5,4)	GBANAJ228636-39	60-63	62-64	52-53
L. perennis[*] (1,1)	GBANL48143	60-61	62	53-54
S. tuberosus (1,1)	GBANAJ228645	63-64	65	55-56
P. purpurea (3,1)	GBANAJ228655	57-58	59	54
P. purpurea[*] (1,1)	GBANL48151	63-64	64	56
C. juncea (3,1)	GBANAJ228658	61-62	63	70

[a] EMBL/GenBank accession number.
[b] Number of accessions/species studied and number of different sequences found among these accessions, respectively.
[*] denote species names corresponding to GenBank entries. Remaining names correspond to EMBL entries.

The various accessions of the species outside the *sativa/serriola/dregeana/altaica* clade form clearly separated basal clades containing accessions of one species only. Apparently, the intraspecific variation within these species is small compared to the interspecific variation, which makes the ITS-1 a good character to identify them.

One of the most parsimonious simple taxon addition trees is shown in Fig. 1. The topology of this tree is identical to that of the strict consensus of all 900 shortest simple taxon addition trees generated, except for its intraspecific branches that often form polytomies in the strict consensus tree (not shown). The topologies of the neighbor joining tree and the maximum likelihood tree were comparable with that of the simple taxon addition tree in the main aspects. All trees show a "*L. sativa*" clade containing *L. sativa*, *L. serriola*, *L. dregeana*, and *L. altaica*, with *L. aculeata* as a sister group of this "*L. sativa*" clade. These five species are part of a

Table 3. Extended.

T. officinale	T. officinale*	L. tenerrima	L. saligna	L. virosa	L. sat/L. dreg
0.263- 0.267	0.267- 0.271	0.286- 0.295	0.251- 0.263	0.247- 0.257	0.235- 0.239
0.268	0.272	0.295- 0.299	0.260- 0.269	0.256- 0.262	0.244
0.266- 0.274	0.273- 0.277	0.236- 0.240	0.199- 0.211	0.211- 0.229	0.211- 0.215
0/0	0.004	0.245- 0.235	0.218- 0.234	0.209- 0.221	0.205- 0.209
1	--	0.235- 0.257	0.226- 0.238	0.217- 0.225	0.213
60-62	62-63	**0/0**	0.163- 0.176	0.167- 0.174	0.163- 0.167
54-58	56-59	40-43	**0.016/4**	0.064- 0.076	0.060- 0.068
52-55	54-56	41-43	16-19	**0.024/6**	0.056- 0.064
51-52	53	40-41	15-17	14-16	--
51-53	53-54	40-43	15-18	14-17	0-1
49-50	51	40-41	15-17	14-16	2
51-52	53	39-41	15-18	14-18	0-2
53-54	55	44-45	19-21	18-20	4
48-49	50	37	16-18	17-18	13
52-54	54-55	42-43	25-28	24-26	20-21
42-45	44-46	41-42	25-29	26-30	26-28
44-46	46-47	37-38	23-28	21-25	21-24
55-57	57-58	33-34	28-31	30-32	28
57-60	59-61	45-47	40-43	39-42	37-39
55-57	57-58	37-38	31-34	36-38	32
48-50	50-51	34	30-34	34-36	32
45-47	47-48	31-34	29-31	30-35	31-33
47-48	49	34	30-31	32-35	33
53-54	55	29-30	31-34	38-39	36
44	45	47	37-39	39-40	40
53-54	55	45	49-50	47-50	44
61-62	63	60-61	54-55	52-53	54

larger clade that also includes *L. saligna* and *L. virosa*, and thus contains all subsection *Lactuca* species examined. This "subsection *Lactuca*" clade is part of a larger "*Lactuca*" clade containing all *Lactuca* species except for the subsection *Cyanicae* species *L. perennis* and *L. tenerrima*, and the Asiatic species *L. indica*. Within this "*Lactuca*" clade, all trees contain the "subsect. *Lactuca*" clade, *L. quercina*, and a clade containing *L. tatarica*, *L. sibirica*, and *L. viminea*. Apart from the "*Lactuca*" clade, all trees contain a clade including *L. perennis*, *C. plumieri*, *L. tenerrima*, and *S. tuberosus*, with *L. perennis* as a sister group of *C. plumieri* and *L. tenerrima* as a sister group of *S. tuberosus*. A separate clade that is present in all trees consists of *M. muralis* and *C. alpina*. The species *C. intybus*, *P. purpurea*, *T. officinale*, *C. juncea*, and *S. asper* branch off in the same order in all the trees. However, in the neighbor-joining and the maximum-likelihood trees *P. purpurea* and *T. officinale* branch off as sister groups and not sequentially. The minor differences between the various trees mostly originate in differences in sister-group relationships. (1) In the neighbor-joining tree *L. saligna* and *L. virosa* are sister groups, while the *L. saligna/L. virosa* clade is a sister group of the clade containing *L. sativa*. In both the simple-taxon addition consensus tree and the maximum-likelihood tree, *L. virosa* and the *L. sativa* clade are sister groups (see Fig. 1). (2) *L. sibirica*

Table 3. Extended.

L. serriola	L. aculeata	L. altaica	L. sativa*	L. quercina	L. viminea
0.235- 0.243	0.235- 0.239	0.235- 0.247	0.251- 0.255	0.242-0.246	0.276- 0.284
0.240- 0.244	0.244	0.244- 0.252	0.260	0.251	0.285- 0.289
0.211- 0.219	0.215	0.211- 0.215	0.219- 0.223	0.194- 0.198	0.211- 0.215
0.205- 0.213	0.197- 0.201	0.205- 0.209	0.213- 0.217	0.193- 0.197	0.209- 0.217
0.213- 0.217	0.205	0.213	0.221	0.201	0.217- 0.221
0.163- 0.174	0.163- 0.167	0.159- 0.167	0.179- 0.183	0.150	0.171- 0.175
0.060- 0.072	0.060- 0.068	0.060- 0.072	0.176- 0.184	0.064- 0.072	0.100- 0.112
0.056- 0.068	0.056- 0.064	0.056- 0.072	0.072- 0.080	0.068- 0.072	0.096- 0.104
0.000- 0.004	0.008	0.000- 0.008	0.016	0.052	0.080- 0.084
0.004/1	0.008- 0.012	0.000- 0.012	0.016- 0.020	0.052- 0.056	0.080- 0.088
2-3	–	0.008- 0.016	0.024	0.052	0.084- 0.088
0-3	2-4	**0.008/2**	0.016-0.024	0.052	0.080- 0.084
4-5	6	4-6	–	0.068	0.096- 0.100
13-14	13	13	17	–	0.072- 0.076
20-22	21-22	20-21	24-25	18-19	**0.004/1**
26-29	26-28	26-28	30-32	16-18	21-24
21-25	21-24	21-24	25-28	14-17	14-19
28-29	28	28	30	26	30-31
37-40	38-39	37-39	39-41	33-35	40-43
32-33	33	32	35	27	35-36
32-33	32	32	36	28	34-36
31-34	31-33	31-33	35-37	24-27	32-36
33-34	33	33	37	26	34-35
36-37	36	35-36	39	33	40-41
40-41	40	40	44	36	41-42
44-45	44	44	47	47	54-55
54-55	54	54	58	49	52-53

and *L. tatarica* are sister groups in both the neighbor-joining tree and the maximum-likelihood tree, while *L. sibirica* and *L. viminea* are sister groups in the simple taxon-addition consensus tree. (3) The clade containing *L. perennis, C. plumieri, L. tenerrima* and *S. tuberosus* is a sister group of the clade containing *M. muralis* and *C. alpina* in both the neighbor-joining and the maximum-likelihood trees, while the larger clade containing these six species is a sister group of a clade containing all other *Lactuca* species. These sister-group relationships are not present in the simple-taxon addition consensus tree. (4) *Prenanthes* and *Taraxacum* are sister groups in both the neighbor-joining tree and the maximum-likelihood trees, but not in the simple taxon-addition consensus tree. (5) The position of *L. indica* is different in the various trees: in the neighbor-joining tree it is a sister group of *L. perennis/C. plumieri*, in the maximum-likelihood tree of *M. muralis/C. alpina*, and in the simple taxon-addition tree of the large *Lactuca* clade.

For all species of which GenBank sequences were available, the variation among the accessions used in our study was smaller than the difference between these accessions and the GenBank sequences. However, the *L. sativa, L. perennis, S. asper,* and *T. officinale* GenBank accessions still form clades with the related sequences from our study. The *P. purpurea*

Table 3. Extended.

L. tatarica	L. sibirica	C. alpina	M. muralis	L. indica	C. plumieri
0.254- 0.266	0.250- 0.262	0.255- 0.267	0.275- 0.283	0.258- 0.266	0.267- 0.272
0.259- 0.267	0.259- 0.267	0.264- 0.272	0.285- 0.289	0.263- 0.267	0.277
0.214- 0.226	0.218- 0.226	0.189- 0.201	0.221- 0.233	0.222- 0.230	0.215- 0.219
0.168- 0.180	0.176- 0.184	0.219- 0.227	0.227- 0.239	0.221- 0.229	0.194- 0.202
0.176- 0.184	0.184- 0.188	0.227- 0.231	0.235- 0.243	0.229- 0.233	0.202- 0.206
0.166- 0.170	0.150- 0.154	0.134- 0.138	0.182- 0.190	0.150- 0.154	0.138
0.100- 0.116	0.092- 0.112	0.112- 0.124	0.160- 0.172	0.124- 0.137	0.121- 0.137
0.104- 0.120	0.084- 0.100	0.120- 0.128	0.152- 0.168	0.144- 0.152	0.137- 0.145
0.104- 0.112	0.084- 0.096	0.112	0.148- 0.156	0.128	0.129
0.104- 0.116	0.084- 0.100	0.112- 0.116	0.148- 0.160	0.128- 0.132	0.129- 0.133
0.104- 0.112	0.084- 0.096	0.112	0.152- 0.156	0.132	0.129
0.104- 0.112	0.084- 0.096	0.112	0.148- 0.156	0.128	0.129
0.120- 0.127	0.100- 0.112	0.120	0.156- 0.164	0.140	0.145
0.064- 0.072	0.056- 0.068	0.104	0.131- 0.139	0.108	0.112
0.084- 0.096	0.056- 0.076	0.120- 0.124	0.160- 0.172	0.140- 0.144	0.137- 0.145
0.008/2	0.040- 0.060	0.124	0.147- 0.155	0.112	0.124
10-15	**0.028/7**	0.108/0.112	0.127- 0.139	0.108- 0.112	0.104- 0.112
31	27- 28	**0.004/1**	0.094- 0.102	0.104	0.092- 0.100
37-39	32-35	24-26	**0.008/2**	0.131- 0.143	0.140- 0.152
28	27-28	26	33-36	**0.000/ 0**	0.096- 0.100
31	26-28	23-25	35-38	24- 25	**0.000/0**
26-30	23-28	22-24	32-36	19- 22	15-16
29	26-27	23-24	32-34	19	15
34	32-33	27-28	35-37	27	24-25
40-42	39-41	36-37	42-43	45-46	42
48-50	46-47	41-42	45-46	45-46	42
50-51	46-49	49	50-51	50-51	52

accession GBANL48151 differed from our own *P. purpurea* accessions by a pairwise distance of 40 positions (Table 3). Because of this large distance, GBANL48151 was a sister group of a *P. purpurea*/*T. officinale* clade in the neighbor-joining analysis and a sister group of *L. perennis* in the maximum-likelihood analysis. However, in the parsimony analysis (Fig. 1), GBANL48151 still formed a clade with the other *P. purpurea* accessions. Apparently the difference with our own sequences is so large that it influences the topologies of the neighbor-joining and maximum-likelihood trees.

Second data set

The total aligned length in the data set with subtribe Lactucinae species was 258 bp, with 118 variable sites, 57 of which were phylogenetically informative. The g_1 statistic of -0.52 was highly significant (< -0.20, P = 0.01), indicating the presence of sufficient phylogenetic signal. With multistate taxa interpreted as uncertainty, 17 trees of 203 steps were obtained with a CI of 0.70 and a RI of 0.64. A set of trees with identical topology, but a length of 255 steps, a CI of 0.76, a RI of 0.64, and an average transition/transversion ratio of 1.43 was obtained with

Table 3. Extended.

L. perennis	L. perennis*	S. tuberosus	P. purpurea	P. purpurea*	C. juncea
0.247- 0.258	0.249- 0.253	0.259- 0.263	0.234- 0.238	0.261- 0.266	0.245- 0.249
0.256- 0.263	0.258	0.269	0.243	0.267	0.254
0.211- 0.215	0.218- 0.222	0.223- 0.227	0.220	0.230	0.280
0.182- 0.190	0.193- 0.197	0.214- 0.218	0.177	0.215- 0.220	0.244- 0.248
0.190- 0.194	0.201	0.222	0.181	0.224	0.252
0.126- 0.138	0.140	0.117- 0.121	0.193	0.186	0.245- 0.249
0.117- 0.124	0.123- 0.127	0.125- 0.137	0.150- 0.159	0.199- 0.203	0.218- 0.222
0.121- 0.141	0.131- 0.143	0.153- 0.157	0.159- 0.163	0.190- 0.202	0.209- 0.214
0.125- 0.133	0.135	0.145	0.163	0.178	0.218
0.125- 0.137	0.135- 0.139	0.145- 0.149	0.163- 0.167	0.178- 0.182	0.218- 0.222
0.125- 0.133	0.135	0.145	0.163	0.178	0.218
0.125- 0.133	0.135	0.141- 0.145	0.163	0.178	0.218
0.141- 0.149	0.152	0.157	0.179	0.190	0.234
0.096- 0.108	0.106	0.133	0.146	0.190	0.197
0.129- 0.145	0.139- 0.143	0.161- 0.165	0.167- 0.171	0.220- 0.224	0.210- 0.214
0.104- 0.120	0.118	0.136	0.162- 0.170	0.194- 0.202	0.201- 0.205
0.092- 0.112	0.106- 0.110	0.128- 0.132	0.158- 0.166	0.186- 0.190	0.185- 0.197
0.088- 0.096	0.093- 0.097	0.108- 0.112	0.144- 0.148	0.166- 0.170	0.194
0.128- 0.143	0.130- 0.138	0.140- 0.148	0.168- 0.172	0.182- 0.186	0.198- 0.202
0.076- 0.088	0.078	0.108	0.181- 0.185	0.182- 0.186	0.201- 0.205
0.060- 0.064	0.061	0.096- .100	0.171	0.171	0.210
0.004/1	0.012	0.092-0.096	0.150	0.135- 0.139	0.173- 0.189
3	--	0.098	0.160	0.146	0.184
23-24	24	--	0.171	0.176	0.202
37	39	42	--	0.163	0.206
33-34	35	43	40	--	0.240
43-47	45	50	51	59	--

multistate taxa interpreted as polytomies. The 50% majority rule consensus of the 17 trees is shown in Fig. 2.

The overall topology of the branch and bound 50% majority rule tree, the neighbor-joining tree, and the maximum-likelihood tree is similar except for the position of *L. indica*, the sister-group relationships within the *tatarica/sibirica/viminea* clade and the sister-group relationship of the *M. muralis/C. alpina* clade. The topologies of the trees based on the second data set corroborate the results obtained using the first data set. All trees show a "*L. sativa*" clade containing *L. sativa*, *L. serriola*, *L. dregeana*, and *L. altaica*, with *L. aculeata* as its sister group. A larger "subsect. *Lactuca*" clade can be identified containing *L. virosa* and *L. saligna* as well, with *L. virosa* as a closer relative to *L. sativa* than *L. saligna*. A still larger "*Lactuca*" clade is formed by the "subsect. *Lactuca*" clade, *L. quercina*, and a clade containing *L. sibirica*, *L. tatarica*, and *L. viminea*. However, the various trees differ with respect to the relationships within the *tatarica/sibirica/viminea* group. Within the branch and bound and the neighbor-joining tree, *L. sibirica* and *L. tatarica* are sister groups, while in the maximum-likelihood tree *L. sibirica* and *L. viminea* are sister groups. In both the branch and bound and the neighbor-joining tree, the "*Lactuca*" clade containing all *Lactuca* species except *L. indica* and the
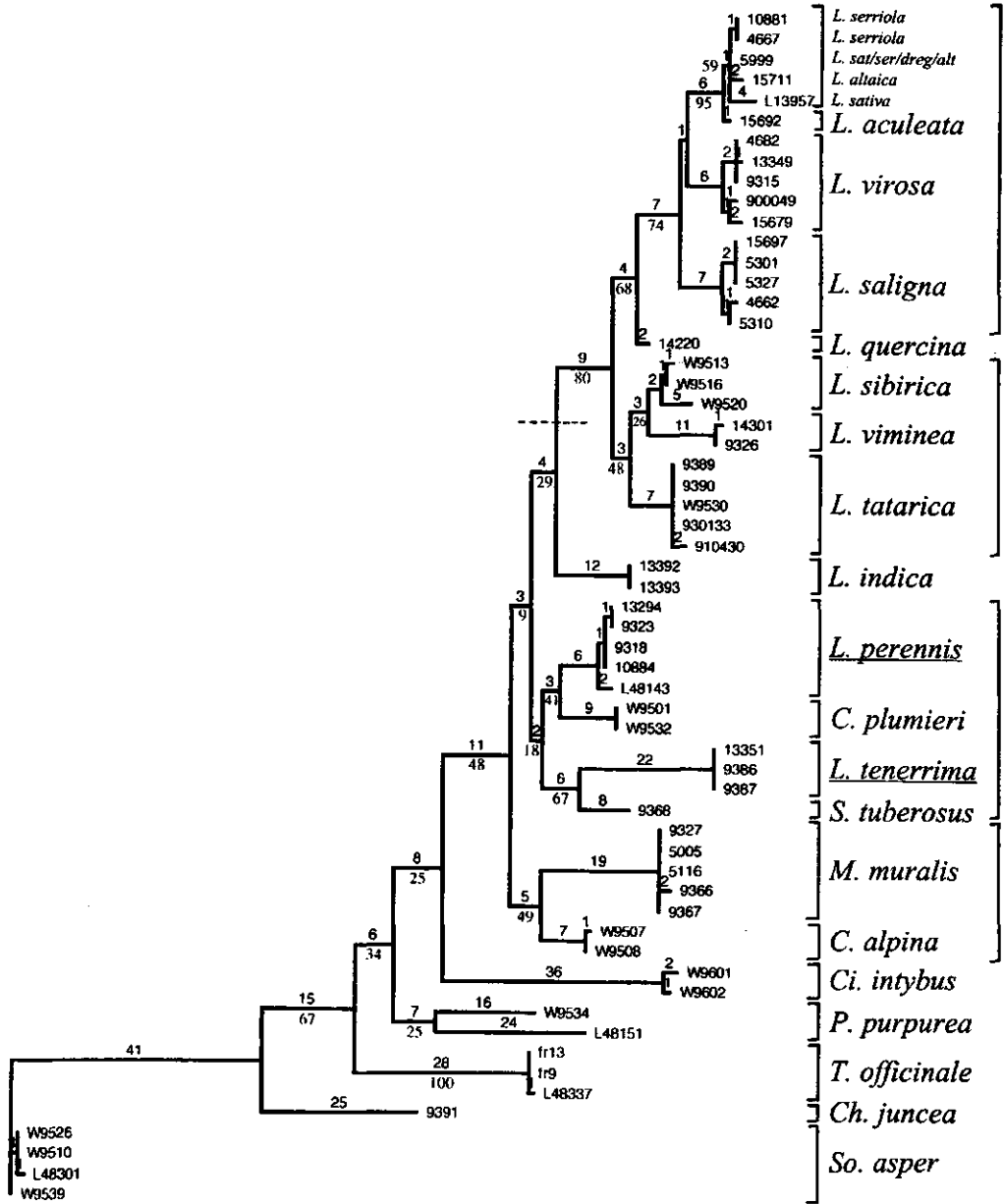
Fig. 1. One of the most parsimonious trees generated with PAUP in a heuristic search using simple taxon addition, TBR, and MULPARS. All sequences from the first data set were used, comprising 23 species of *Lactuca* and related genera. Branch lengths are above branches. The numbers on the internal branches are synapomorphic positions/basepairs, and the numbers on the terminal branches are autapomorphic positions. The numbers below the interspecific branches are the bootstrap values. *L. = Lactuca, C. = Cicerbita, S. = Steptorhamphus, M. = Mycelis, Ci. = Cichorium, P. = Prenanthes, T. = Taraxacum, Ch. = Chondrilla, So. = Sonchus.*
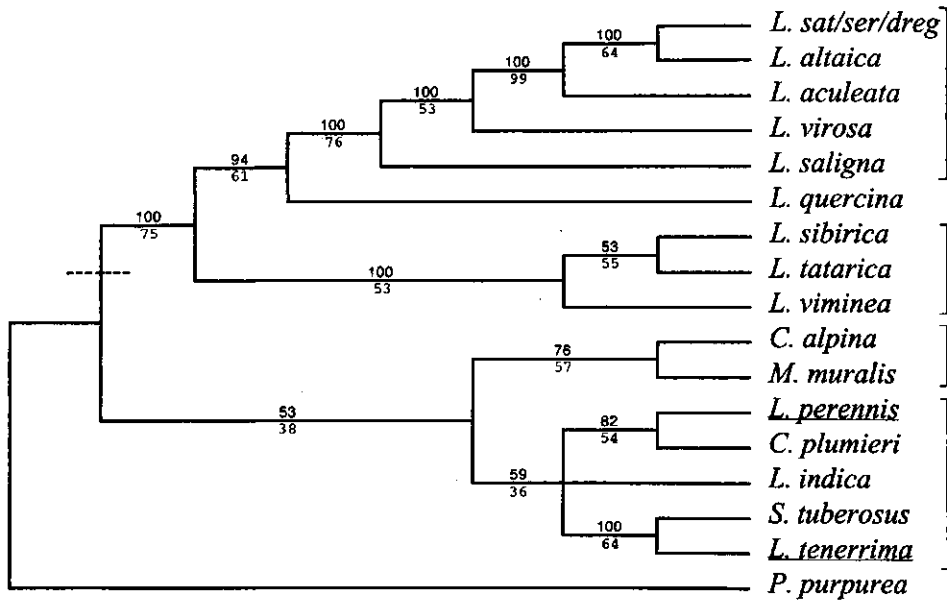
Fig. 2. The 50% majority rule consensus tree calculated from 17 trees generated with PAUP in a branch and bound search using furthest taxon addition and MULPARS. All consensus sequences from the second data set were used, comprising 17 subtribe Lactucinae species. Percentages of original trees showing the clade indicated are above branches. Bootstrap values are below branches. No individual most parsimonious tree is shown since three of these trees can easily be reconstructed by solving the polytomy in the consensus tree. *L.* = *Lactuca, C.* = *Cicerbita, M.* = *Mycelis, S.* = *Steptorhamphus, P.* = *Prenanthes*.

subsect. *Cyanicae* species *L. perennis* and *L. tenerrima* is a sister group of the clade containing all remaining species (except for the outgroup *P. purpurea*). In the maximum-likelihood tree, however, the "*Lactuca*" clade is a sister group of a clade containing *L. perennis, C. plumieri, S. tuberosus,* and *L. tenerrima,* while the *M. muralis/C. alpina* clade is a sister group of *P. purpurea*. The position of *L. indica* is the most variable in the different trees. *L. indica* is a sister group of a *L. perennis/C. plumieri* clade in the neighbor-joining tree and of the '*Lactuca* clade" in the maximum-likelihood tree, while in the branch and bound tree it forms a polytomy with a *perennis/plumieri* and a *tenerrima/tuberosus* clade.

Discussion

The purpose of our study was to gain more insight into the group of species contributing to the lettuce (*Lactuca sativa*) gene pool by evaluating the major classifications of *Lactuca* s.l. using ITS-1 DNA sequences. The research topics formulated in the introduction as guidelines for the

evaluation will now be addressed by discussing literature data on morphology, crossability, cytology, isozyme analyses, and molecular analyses in relation to the major classifications, the ITS-1 data, and the phylogenies generated in our study. In conclusion, an adjustment of the *Lactuca* subdivision of Feráková (1977) is proposed and discussed in relation to the gene-pool concept of Harlan and de Wet (1971).

The work of Kesseli and Michelmore (1986) was excluded from the discussion since the uncertain identities of *L. saligna*, *L. dregeana*, and two of the three *L. virosa* accessions in their data set hamper an unambiguous interpretation of the results in terms of species relationships.

Distinction among L. sativa, L. serriola, L. dregeana, and L. aculeata
The extremely close relationship between *L. sativa* and *L. serriola* is apparent in the good interspecific crossability (e.g., Thompson, Whitaker, and Kosar, 1941; Lindqvist, 1960; De Vries, 1990), the identical karyotype (Lindqvist, 1960; Chatterjee and Sharma, 1969; Globerson, Netzer, and Sacks, 1980; Haque and Godward, 1985; Koopman and De Jong, 1996), chromosome banding pattern (Koopman, De Jong, and De Vries, 1993), and DNA content (Koopman and De Jong, 1996). Furthermore, in phenetic analyses of nuclear RFLP data (Kesseli, Ochoa, and Michelmore, 1991) and nuclear AFLP data (Hill et al., 1996) *L. sativa* and *L. serriola* clustered together. RFLP analysis of mtDNA showed a high proportion of shared fragments (Vermeulen et al., 1994), isozyme analysis of foliar esterases shows patterns common to both species (Roux, Chengjiu, and Roux, 1985), and a phenetic analysis of SDS-electrophoresis patterns of seed proteins (De Vries, 1996) showed *L. sativa* and *L. serriola* as completely interlaced groups. Numerical morphological analyses of plant morphological data showed *L. sativa* and *L. serriola* as separate but partly overlapping groups (De Vries and Van Raamsdonk, 1994; Frietema de Vries, van der Meijden, and Brandenburg, 1994; Frietema de Vries, 1996). While De Vries and van Raamsdonk (1994) maintain *L. sativa* and *L. serriola* as separate species, Frietema de Vries (1996) considers the overlap enough to regard them as conspecific. Using the ITS-1 sequences, *L. sativa* and *L. serriola* cannot be distinguished since six out of eight *L. serriola* sequences were identical to the single sequence that was characteristic to all *L. sativa* accessions in our study. Taking into account this large overlap in ITS-1 sequence and that of characters as shown in the literature cited, it would seem most appropriate to regard *L. sativa* conspecific with *L. serriola*.

The close relationship of *L. dregeana* with *L. serriola* is indicated by its morphology (Zohary, 1991) and the fact that *L. serriola* x *L. dregeana* crosses (Koopman, unpublished data) yielded large numbers of plump seeds. The ITS-1 sequence data from our study corroborate this relationship since the sequences of *L. dregeana*, *L. sativa*, and most of the *L. serriola* accessions were identical. Therefore, *L. dregeana* should be considered conspecific with *L. serriola* as well.

The position of *L. aculeata* close to *L. serriola* is clear from the plant morphology (Zohary, 1991), crossability (Globerson, Netzer, and Sacks, 1980; Koopman, unpublished data),

chromosome morphology (Globerson, Netzer, and Sacks, 1980), and isozyme analyses of foliar esterases (Roux, Chengjiu, and Roux, 1985). The ITS-1 sequence of *L. aculeata* differed only slightly from those of *L. sativa* and *L. serriola*, confirming the close relationship between *L. aculeata* and *L. sativa/serriola*. Further information will be needed to determine the relationship more accurately.

The close relationship of *L. sativa*, *L. serriola*, *L. dregeana*, and *L. aculeata* (and *L. altaica*) as a group is stressed by their pairwise distances. The maximum distance within this group is four positions excluding the *L. sativa* GenBank accession and six positions including it. This is comparable with the largest intraspecific distances within the closely related species *L. saligna* and *L. virosa*, which are four and six, respectively (Table 3).

Position of L. altaica

Based on morphological characters, Feráková (1977) regarded *L. altaica* as a species intermediate between *L. serriola* and *L. saligna*. The close relationship with *L. serriola* is indicated by the results of crossing experiments of Thompson, Whitaker, and Kosar (1941) and Lindqvist (1960). The latter did not even distinguish *L. altaica* as a separate species but considered it a primitive form of *L. sativa*. Furthermore, in (unpublished) crossing experiments by the present first author, *L. altaica* behaved like *L. serriola*. Crosses between *L. serriola* and *L. altaica* yielded many normal seeds of good quality, while crosses between *L. saligna* and *L. altaica* yielded only small seeds of poor quality, comparable to seeds from *saligna* x *serriola* crosses. The ITS-1 results corroborate the close relationships between *L. altaica* and *L. serriola*, since the ITS-1 sequence of one *L. altaica* accession in our study was identical to that of the most common *L. serriola* sequence and that of the other deviated in only two autapomorphic positions. Except for some morphological characters, no evidence was found to support a close relationship to *L. saligna*. Therefore, *L. altaica* should be considered conspecific with *L. serriola* rather than intermediate between *L. serriola* and *L. saligna*.

Position of L. sativa/serriola, L. saligna, and L. virosa relative to each other

Four different possibilities regarding the position of *L. saligna* and *L. virosa* relative to each other and to *L. sativa/serriola* appear from literature. (1) Data on plant morphology (De Vries and Van Raamsdonk, 1994), SDS-electrophoresis patterns of seed proteins (De Vries, 1996), and isozyme analysis of foliar esterases (Roux, Chengjiu, and Roux, 1985) indicate a close relationship between *L. sativa/serriola* and *L. virosa*, while *L. saligna* is the more distinct species. (2) Data on crossability (Thompson, Whitaker, and Kosar, 1941; Lindqvist, 1960; de Vries, 1990), karyotype (Lindqvist, 1960; Koopman and De Jong, 1996), chromosome banding pattern (Koopman, De Jong, and De Vries, 1993), and DNA content (Koopman and De Jong, 1996) indicate that *L. sativa/L. serriola* occupy an intermediate position between *L. saligna* and *L. virosa*, but closer to *L. saligna* than to *L. virosa*. In this case, *L. virosa* is the more distinct species. The position of *L. sativa/serriola* closer to *L. saligna* than to *L. virosa* is corroborated by data on nuclear AFLPs (Hill et al., 1996). (3) Hill et al. (1996) stated that their AFLP results

were similar to previously published nuclear RFLP results of Kesseli, Ochoa, and Michelmore (1991). However, in contrast with the phenetic tree based on AFLPs, the RFLP tree shows that *L. saligna* and *L. virosa* are more related to each other (in terms of similarity) than to *L. sativa/serriola*. (4) A completely different indication of the species relationships within subsect. *Lactuca* was given by Vermeulen et al. (1994). In their analysis of mtDNA RFLPs, the proportion of bands shared by *L. serriola* and *L. virosa* was higher than that shared by *L. serriola* and *L. sativa*. This large proportion of identical bands would indicate that *L. virosa* is at least as related to *L. serriola* as is *L. sativa*, which is in conflict with all previously mentioned results.

In most of our ITS-1 based phylogenies, *L. virosa* is more closely related to *L. sativa/serriola* than is *L. saligna*, which corroborates the results of the isozyme and morphological analyses. However, the order in which *L. virosa* and *L. saligna* branch off in the parsimony tree is determined by only one synapomorphy (Fig. 1). In the neighbor-joining tree *L. virosa* and *L. saligna* are sister groups, which corroborates the RFLP results. Given these ambiguities, the ITS-1 results must be considered inconclusive as to the position of *L. sativa/serriola*, *L. saligna*, and *L. virosa* relative to each other.

The apparently conflicting indications of species relationships reported in literature can be brought in line by postulating (as far as we know for the first time) that *L. virosa* is a hybrid taxon. In that case, the mtDNA RFLPs, which are inherited maternally, can be interpreted as indicating a *serriola*-like species as a female parent to the hybrid. The contribution of a yet unknown male parent is expressed in two *L. virosa* characters that are unique within subsect. *Lactuca*, namely the winged black achene and the presence of one pair of satellite chromosomes instead of two.

In the new hybrid species, different classes of DNA evolved differently, as discussed in Koopman, de Jong, and de Vries (1993). Extensive chromosome rearrangements and dynamic changes of large blocks of repetitive DNA could give rise to the deviating *L. virosa* karyotype, chromosome banding pattern, DNA content, and crossability. The unique DNA sequences reflected in the RFLP and isozyme patterns, as well as in the plant morphology, evolved differently. The ITS-1 results can be explained by assuming that after the hybridization event the ITS-1 sequences derived from both parents were subjected to a homogenization process in the new hybrid (see Elder and Turner [1995] for a discussion on homogenization). The homogenization gave rise to a new ITS-1 sequence characteristic for the hybrid species, while the original parental sequences were lost.

Position of Cichorium
The taxonomic position of *Cichorium* within the Lactuceae is unclear because important characters such as pappus type and pollen morphology are not phylogenetically interpretable. On the one hand, *Cichorium* has a pollen type too widespread within the Lactuceae to be useful for clarifying genus relationships (Blackmore, 1981). On the other hand, *Cichorium* has a

pappus type that is unique within the Lactuceae and thus useless for phylogenetic purposes as well (Bremer, 1994). According to Stebbins (1953) and Jeffrey (1966), *Cichorium* and *Lactuca* are not closely related. Stebbins (1953) recognizes eight subtribes within the Cichorieae and assigns *Cichorium* and *Lactuca* sensu Stebbins (1937) to different subtribes. *Cichorium* is placed in the subtribe Cichorinae, while *Lactuca* is part of the so called *Prenanthes-Lactuca* line of subtribe Crepidinae. Jeffrey (1966) recognizes five groups within the Cichorieae and although both *Cichorium* and *Lactuca* are placed in the same group, *Cichorium* is separated as the monogeneric *Cichorium* subgroup while *Lactuca* is placed in the broadly defined *Crepis* subgroup. Recently, Vermeulen et al. (1994) concluded from mitochondrial RFLP data that *Cichorium* is more closely related to *Lactuca* and *Cicerbita* than to *Chondrilla*, *Taraxacum*, and *Sonchus*. On the other hand, Whitton, Wallace, and Jansen (1995) found that using chloroplast DNA restriction site variation, *Cichorium* appears more closely related to *Prenanthes*, *Chondrilla*, and *Taraxacum* than to *Lactuca* and *Cicerbita*, and least related to *Sonchus*. Our ITS-1 data on *C. intybus* support the view of Vermeulen et al. (1994) that *Cichorium* is closely related to *Lactuca*. According to the ITS-1 data, *Cichorium* is more related to *Lactuca* than is *Prenanthes*, which places *Cichorium* within Stebbins' *Prenanthes-Lactuca* line or Jeffrey's *Crepis*-subgroup.

Evaluation of major generic concepts in Lactuca

According to Feráková (1977) two main generic delimitations of *Lactuca* are possible: taking the genus in a broad sense according to Stebbins (1937), or treating the genus in a narrow sense according to Tuisl (1968). Feráková (1977) herself takes an intermediate position. More recently, Shih (1988) published an extremely narrow genus concept.

The genus *Lactuca* according to Stebbins (1937) is characterized by a corolla tube generally more than half as long as the ligule, a pappus containing at least some bristles that are no more than four-celled in cross section at the base, and a flattened achene with an expanded pappus disc and two lateral ribs or wings more pronounced than the others. The inflorescences are many-headed panicles or racemes. Based on this description, *Lactuca* sensu Stebbins includes the genera *Mulgedium*, *Lactucopsis*, *Phaenixopus*, *Mycelis*, and *Cicerbita*, but *Cicerbita alpina* is excluded because of its coarse pappus and nearly columnar, only slightly compressed achene. Stebbins did not mention *Steptorhamphus*, but it probably should be included as well since it fits Stebbins' genus description of *Lactuca* in general terms, notwithstanding the relatively coarse pappus. Tuisl (1968) takes the genus *Lactuca* in a narrow sense and recognizes *Mulgedium*, *Scariola* (= *Phaenixopus* as recognized by Stebbins), *Cicerbita*, and *Steptorhamphus* as separate genera. *Lactuca* is characterized by two equal rows of pappus hairs, a distinct beak, and many-flowered heads. The genera *Lactuca*, *Mulgedium*, and *Scariola* are separated from *Steptorhamphus* and *Cicerbita* by the presence of two rows of equal pappus hairs. The genus *Mycelis*, which was not mentioned by Tuisl, can also be separated from *Lactuca* by this character. *Lactuca* is separated from *Mulgedium* and *Scariola* because of its

distinct beak, while *Scariola* is separated from *Mulgedium* by its few-flowered heads. According to Tuisl's genus definition, *Lactuca sibirica* should be included in *Mulgedium* as well, although the species was not treated in Tuisl (1968). Tuisl's genus concept also necessitates the separation of *Lactucopsis* from *Lactuca* (Feráková, 1977). Feráková (1977) takes an intermediate position between Stebbins (1937) and Tuisl (1968). She limits the genus *Lactuca* to species with two equal rows of pappus hairs, thus excluding *Steptorhamphus*, *Mycelis*, and *Cicerbita*, but includes Tuisl's genera *Mulgedium* and *Scariola* as sections within *Lactuca*. As a result, Feráková (1977)'s genus *Lactuca* comprises the sections *Phaenixopus* (= *Scariola* as recognized by Tuisl), *Mulgedium*, *Lactucopsis*, and *Lactuca*, the latter of which is divided into subsections *Lactuca* and *Cyanicae* (see also Table 1). More recently, Shih (1988) narrowed down the genus *Lactuca* including only species with numerous yellow florets and pale brown achenes with a clearly distinct, filiform beak. This would limit the genus to species from Feráková's subsection *Lactuca*, but strictly speaking it would exclude *L. virosa* as well as *L. sativa* cultivars with white or blackish achenes.

Our ITS-1 phylogeny is in line with the view of Stebbins (1937) that *Lactuca* in a broad sense should include *Mulgedium*, *Lactucopsis*, *Phaenixopus*, *Mycelis*, and *Cicerbita* and that *Lactuca* can be separated from the closely related genus *Prenanthes*. However, the exclusion of *C. alpina* from *Lactuca* s.l. is not supported since *C. alpina* fell within the group of species from *Lactuca* sensu Stebbins. The ITS-1 data conflict with the subgeneric division of Tuisl (1968) on several points. The separation of *Mulgedium* (represented by *L. tatarica* and *L. sibirica*) from *Scariola* (=*Phaenixopus*, represented by *L. viminea*) is not fully supported since all trees generated showed a *sibirica/tatarica/viminea* clade (Figs. 1, 2) and part of the analyses show a smaller *sibirica/viminea* clade nested within it (see results section and Fig. 1). The distinction between *Lactuca* and *Cicerbita/ Steptorhamphus* or between *Cicerbita* and *Mycelis* is not supported either, since species from the different genera form clades together (Figs. 1, 2). The ITS-1 data partly corroborate the genus concept of Feráková. On the one hand, a clade containing all subsect. *Lactuca* species is present in all ITS-1 analyses, as is a larger clade containing all *Lactuca* species except *L. perennis* and *L. tenerrima*. On the other hand, the distinction between sect. *Mulgedium* and sect. *Phaenixopus* is not confirmed, nor is the position of subsect. *Cyanicae*. The sect. *Mulgedium* species *L. tatarica* and *L. sibirica* form a clade with the sect. *Phaenixopus* species *L. viminea*, while the subsect. *Cyanicae* species *L. perennis* and *L. tenerrima* do not form a clade with *Lactuca* species but with *C. plumieri* and *S. tuberosus*, respectively. The ITS-1 data fit the narrow generic concept of Shih (1988) to a large extent, since this concept would lead to the recognition of separate genera for most of the species used in our study and thus avoids the classification problems. As such, our results would also fit a separation of *L. tatarica* and *L. sibirica* (Shih, 1988) and of *C. alpina* and *C. plumieri*. (Stebbins, 1937; Shih, 1991) However, because the genus description of Shih would exclude *L. virosa* (and even part of *L. sativa*) from *Lactuca* it conflicts with the ITS-1 data showing a *L. sativa/ serriola/ saligna/ virosa* clade.

Delimitation of Lactuca

Regarding the ITS-1 results, there are several options for delimiting the genus *Lactuca*.

1) A large and variable genus could be recognized, approximately according to Stebbins (1937), but with the inclusion of *Cicerbita alpina*. This would "group together species similar in habit, and those whose individual affinities are clearly with each other rather than with species excluded from the genus" however, "occasional, transitional species are found" (Stebbins, 1937). These transitional species show morphological characters grouping them with *Lactuca* as well as morphological characters grouping them with other genera. In our data set, *C. alpina* is such a transitional species, since its general habit and ITS-1 sequence place it well within *Lactuca* sensu Stebbins, but its pappus and achene do not fit his genus description. Including *C. alpina* in *Lactuca* sensu Stebbins would necessitate an expanded genus description including species with a coarse pappus and nearly columnar achenes. This new genus description would probably obscure the boundaries between *Lactuca* s.l. and *Prenanthes* or even less related genera and the genus would become unacceptably variable.

2) A solution to this problem would be to recognize a narrow genus *Lactuca*, identical to Feráková's subsect. *Lactuca*. Depending on the species concept used, the number of species contributing to this genus can be reduced by lumping *L. sativa* and the *L. serriola*-like species *L. serriola*, *L. aculeata*, *L. scarioloides* Boiss., *L. azerbaijanica* Rech., *L. georgica* Grossh., *L. dregeana*, and *L. altaica*, described in Zohary (1991). This solution would also fit an expanded genus description of Shih allowing white or blackish achenes, but necessitates the recognition of many additional genera (see, e.g., Shih, 1988).

3) A third and in our opinion more favorable solution would be to recognize a genus *Lactuca* according to Feráková, but with the exclusion of subsect. *Cyanicae*. This would limit the genus to species with more than three ribs on the achenes. The adjusted genus concept has the benefit of corresponding to the gene pool of cultivated lettuce according to the gene-pool concept of Harlan and de Wet (1971), which facilitates its practical use and acceptance.

Lactuca and the lettuce gene pool

Within Feráková's *Lactuca* subsect. *Lactuca* a group of species can be identified that is closely related to and readily crossable with *L. serriola* and *L. sativa*, containing among others *L. dregeana*, *L. aculeata*, and *L. altaica* (Zohary, 1991). Since all these species have an ITS-1 sequence that is (nearly) identical to that of cultivated lettuce (*L. sativa*), this sequence can be considered characteristic for species contributing to the primary gene pool of lettuce (Harlan and De Wet, 1971). It is expected that the other members of the group, namely *L. scarioloides*, *L. azerbaijanica*, and *L. georgica* (Zohary, 1991) will show similar ITS-1 sequences and crossing behavior. The remaining subsect. *Lactuca* species *L. saligna* and *L. virosa* are characterized by their own distinct ITS-1 sequences and in the analyses they form a clade with the species from the primary genepool. In contrast to what is stated in Zohary (1991), *L.*

saligna as well as *L. virosa* contribute to the secondary gene pool of cultivated lettuce, since both are partly interfertile with *L. sativa* (Lindqvist, 1960; Maisonneuve et al., 1995).

A group of species less related to cultivated lettuce (*Lactuca sativa*) branch off closely to the primary and secondary gene-pool species in ITS-1 based phylogenies. These are *L. viminea* from Feráková's section *Phaenixopus*, *L. tatarica* and *L. sibirica* from section *Mulgedium*, and *L. quercina* from section *Lactucopsis*. Hybridization data on these species are limited to *L. viminea* and *L. tatarica*. *L. viminea* is crossable with *L. virosa* (Groenwold, 1983) yielding a partly fertile hybrid, and *L. tatarica* can be somatically hybridized with *L. sativa* to produce a fertile hybrid (Chupeau et al., 1994; Maisonneuve et al., 1995). Since the genetic diversity of *L. viminea* and *L. tatarica* is not directly accessible for lettuce breeding but requires special techniques such as bridging species or somatic hybridization, they belong to the tertiary gene pool of *L. sativa*. Based on the species relationships proposed by Feráková (1977) and the ITS-1 results it can be expected that *L. quercina* and *L. sibirica* and the species from *Lactuca* sections *Phaenixopus*, *Mulgedium*, and *Lactucopsis* not included in our study, contribute to *L. sativa*'s tertiary gene pool as well.

The species *L. perennis* and *L. tenerrima* from *Lactuca* subsect. *Cyanicae* (Feráková, 1977) do not form a clade with *Lactuca* species in the ITS-1 phylogenies, but with species from related genera outside the lettuce gene pool. This position outside the lettuce gene pool is corroborated by the fact that *L. perennis* is not crossable with subsect. *Lactuca* species (Thompson, Whitaker, and Kosar, 1941) and that the somatic hybrids between *L. perennis* and *L. sativa* reported on in Chupeau et al. (1994) and Maisonneuve et al. (1995), were completely sterile (B. Maisonneuve, INRA, France, personal communication). No literature data were available on the remaining subsect. *Cyanicae* species, but based on the relationships proposed by Feráková (1977), the ITS-1 results on *L. tenerrima* and *L. perennis*, and the literature data on *L. perennis*, it can be expected that the entire subsection falls outside the lettuce gene pool.

The limited data available on the remaining species used in our study do not indicate additional species that could contribute to the gene pool of cultivated lettuce. The tested species *L. indica* (Thompson, Whitaker, and Kosar, 1941) and *S. tuberosus* (=*L. cretica* Desf.; Thompson, 1943) were not crossable with any subsect. *Lactuca* species. Somatic hybridizations of *L. sativa* with *L. indica* (Mizutani et al., 1989) yielded colonies of hybrid callus, but no viable plants could be regenerated from these colonies. Somatic hybridizations between *L. sativa* and *C. plumieri*, *C. juncea*, *C. intybus*, and *T. officinale* completely failed (Chupeau et al., 1994; Maisonneuve et al., 1995). Therefore, the gene pool of cultivated lettuce seems limited to the *Lactuca* species sensu Feráková excluding subsect. *Cyanicae*, as indicated by the dotted line in Figs.1 and 2. The position of *L. indica* relative to the species in the lettuce gene pool is still unclear since its position varied in the different phylogenetic analyses (see Results section and Figs. 1 and 2). However, because *L. indica* can be somatically hybridized with *L. sativa* to produce a viable callus, it is probably closely related to the species in the

lettuce gene pool. Additional research, preferably involving other Asiatic species as well, should elucidate the position of *L. indica*.

Conclusions

Based on our present information, we propose an adjustment of the genus concept of Feráková (1977) implying exclusion of sect. *Lactuca* subsect. *Cyanicae* from *Lactuca*. This limits the genus to species with more than three ribs on the achenes. The adjusted genus coincides with the lettuce gene pool. Section *Lactuca* subsect. *Lactuca* comprises the primary and secondary gene pool, while the sections *Phaenixopus, Mulgedium,* and *Lactucopsis* comprise the tertiary gene pool. Section *Lactuca* subsect. *Cyanicae* is not included in *Lactuca*, nor does it belong to the lettuce gene pool. The position of *L. indica* needs further consideration.

The practical value of our study to plant breeders is that it points out the species contributing to cultivated lettuce's tertiary gene pool. These tertiary gene-pool species will become increasingly important in breeding programs since improved breeding techniques (Chupeau et al., 1994; Maisonneuve et al., 1995) will make them more easily accessible as a gene source.

The usefulness of ITS-1 sequences for phylogenetic analysis and evaluation of existing classifications in *Lactuca* and related genera has been demonstrated in this paper. However, regarding details as, for example, the position of *L. sativa/serriola, L. saligna* and *L. virosa* relative to each other, ITS-1 sequences were inconclusive and supplemental markers are needed to elucidate the relationships. Recently, AFLPs have become available as a tool for systematic studies of closely related species (Sharma, Knox, and Ellis, 1996; Huys et al., 1996; Keim et al., 1997; Kardolus, van Eck, and van den Berg, 1998). Although the relative contributions of restriction site variation and insertions/deletions to the variation sampled with AFLPs is not exactly known, AFLPs proved to be useful molecular markers for phylogenetic purposes. Since Hill et al. (1996) demonstrated that AFLPs can be applied in *Lactuca*, they seem suitable to add information to the ITS-1 phylogeny presented in this paper. Therefore, a phylogenetic analysis using AFLP data of all *Lactuca* accessions from the ITS-1 study will be carried out in the near future.

Acknowledgements

References

BAIN, J. F., AND R. K. JANSEN. 1995. A phylogenetic analysis of the aureoid *Senecio* (Asteraceae) complex based on ITS sequence data. *Plant Systematics and Evolution* 195: 209-219.

BALDWIN, B. G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the Compositae. *Molecular Phylogenetics and Evolution* 1: 3-16.

————. 1993. Molecular phylogenetics of *Calycadenia* (Compositae) based on ITS sequences of nuclear ribosomal DNA: Chromosomal and morphological evolution reexamined. *American Journal of Botany* 80: 222-238.

BREMER, K. 1994. Asteraceae: cladistics & classification. Timber Press, Portland, OR.

BLACKMORE, S. 1981. Palynology and intergeneric relationships in subtribe Hyoseridinae (Compositae: Lactuceae). *Botanical Journal of the Linnean Society* 82: 1-13.

CHATTERJEE, T., and A. K. SHARMA. 1969. Cytotaxonomy of Cichorieae. *Genetica* 40: 577-590.

CHUPEAU, M-C., B. MAISONNEUVE, Y. BELLEC, AND Y. CHUPEAU. 1994. A *Lactuca* universal hybridizer, and its use in creation of fertile interspecific somatic hybrids. *Molecular & General Genetics* 245: 139-145.

CRONQUIST, A. 1985. History of generic concepts in the Compositae. *Taxon* 34(1): 6-10.

DE VRIES, I. M. 1990. Crossing experiments of lettuce cultivars and species (*Lactuca* sect. *Lactuca*, Compositae). *Plant Systematics and Evolution* 171: 233-248.

————. 1996. Characterization and identification of *Lactuca sativa* cultivars and wild relatives with SDS-electrophoresis (*Lactuca* sect. *Lactuca*, Compositae). *Genetic Resources and Crop Evolution* 43: 193-202.

————, AND L. W. D. VAN RAAMSDONK. 1994. Numerical morphological analysis of lettuce cultivars and species (*Lactuca* sect. *Lactuca*, Asteraceae). *Plant Systematics and Evolution* 193:125-141.

ELDER, J. F., AND B. J. TURNER. 1995. Concerted evolution of repetitive DNA sequences in eukaryotes. *Quarterly Review of Biology* 70: 297-320.

FELSENSTEIN, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.572. Computer program distributed by the author, Department of Genetics, University of Washington, Seattle, WA.

FERÁKOVÁ, V. 1977. The genus *Lactuca* L. in Europe. Bratislava, Univerzita Komenského.

FRIETEMA DE VRIES, F.T. 1996. Cultivated plants and the wild flora. Ph.D. dissertation, Rijksherbarium/Hortus Botanicus, Leiden, The Netherlands.

————, R. VAN DER MEIJDEN AND W.A. BRANDENBURG. 1994. Botanical files on lettuce (*Lactuca sativa*). *Gorteria* supplement 2.

GLOBERSON, D., D. NETZER, AND J. SACKS. 1980. Wild lettuce as a source for improving cultivated lettuce. *In* J. W. Maxon Smith, and F. A. Langton [eds.], *Proceedings Eucarpia meeting on leafy vegetables*, 86-96. Glasshouse Crops Research Institute, Littlehampton.

GROENWOLD, R. 1983. Onderzoek van de relatie tussen *Lactuca* en *Bremia Lactucae*. Verslag van een voorlichtingsbijeenkomst voor slaveredelaars (vervolg). *Zaadbelangen* 37: 132.

HARLAN, J. R., AND J. M. J. DE WET. 1971. Towards a rational classification of cultivated plants. *Taxon* 20: 509-517.

HAQUE M. Z., AND M. B. E. GODWARD. 1985. Comparison between two genera, species and cultivars in Lactuceae. I. Karyotype analysis. *Cytologia* 50: 725-738.

HILL, M., H. WITSENBOER, M. ZABEAU, P.VOS, R. KESSELI, AND R. MICHELMORE. 1996. PCR-based fingerprinting using AFLPs as a tool for studying genetic relationships in *Lactuca* spp. *Theoretical and Applied Genetics* 93: 1202-1210.

HILLIS, D. M., AND J. P. HUELSENBECK. 1992. Signal, noise, and reliability in molecular phylogenetic analyses. *Journal of Heredity* 83: 189-195.

HUYS, G., R. COOPMAN, P. JANSSEN, AND K. KERSTERS. 1996. High-resolution genotypic analysis of the genus *Aeromonas* by AFLP fingerprinting. *International Journal of Systematic Bacteriology* 46: 572-580.

IWATSUKI, K., T. YAMAZAKI, D. E. BOUFFORD, AND H. OHBA. 1995. Flora of Japan, vol. IIIb. Kodansha Ltd., Tokyo.

JEFFREY, C. 1966. Notes on Compositae. I. The Cichorieae in east tropical Africa. *Kew Bulletin* 18: 427-486.

KARDOLUS, J. P., H. J. VAN ECK, AND R. G. VAN DEN BERG. 1998. The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy. *Plant Systematics and Evolution*: in press.

KEIM, P., A. KALIF, J. SCHUPP, K. HILL, S.E. TRAVIS, K. RICHMOND, D.M. ADAIR, M. HUGH-JONES, C.R. KUSKE AND P. JACKSON. 1997. Molecular evolution and diversity in *Bacillus anthracis* as detected by amplified fragment length polymorphism markers. *Journal of Bacteriology* 179: 818-824.

KESSELI, R. V., AND R. W. MICHELMORE. 1986. Genetic variation and phylogenies detected from isozyme markers in species of *Lactuca*. *Journal of Heredity* 77: 324-331.

―――, O. OCHOA, AND R. W. MICHELMORE. 1991. Variation at RFLP loci in *Lactuca* spp. and origin of cultivated lettuce (*L. sativa*). *Genome* 34: 430-436.

KIM, K.-J., AND R. K. JANSEN. 1994. Comparisons of phylogenetic hypotheses among different data sets in dwarf dandelions (*Krigia*, Asteraceae): additional information from internal transcribed spacer sequences of nuclear ribosomal DNA. *Plant Systematics and Evolution* 190: 157-185.

KIM, S.-C., D. J. CRAWFORD, J. FRANCISCO-ORTEGA, AND A. SANTOS-GUERRA. 1996. A common origin for woody *Sonchus* and five related genera in the Macaronesian islands: Molecular evidence for extensive radiation. *Proceedings of the National Academy of Sciences, USA* 93: 7743-7748.

KOOPMAN, W. J. M., J. H. DE JONG, AND I. M. DE VRIES. 1993. Chromosome banding patterns in lettuce species (*Lactuca* sect. *Lactuca*, Compositae). *Plant Systematics and Evolution* 185: 249-257.

―――, AND ―――. 1996. A numerical analysis of karyotypes and DNA amounts in lettuce cultivars and species (*Lactuca* subsect. *Lactuca*, Compositae). *Acta Botanica Neerlandica* 45: 211-222.

LINDQVIST, K. 1960. Cytogenetic studies in the *serriola* group of *Lactuca*. *Hereditas* 46: 75-151.

MADDISON, W. P., AND D. R. MADDISON. 1993. MacClade version 3.04. Computer program distributed by Sinauer Associates, Sunderland, MA.

MAISONNEUVE, B., M.-C. CHUPEAU, Y. BELLEC, AND Y. CHUPEAU. 1995. Sexual and somatic hybridization in the genus *Lactuca*. *Euphytica* 85: 281-285.

MIZUTANI, T., X. J. LIU, Y. TASHIRO, S. MIYAZAKI, AND K. SHIMASAKI. 1989. Plant regeneration and cell fusion of protoplasts from lettuce cultivars and related wild species in Japan. *Bulletin of the Faculty of Agriculture, Saga University* 67: 109-118.

ROUX, L., Z. CHENGJIU, AND Y. ROUX. 1985. Caractérisation de *Lactuca sativa* L. et des espèces apparentées par électrofocalisation des estérases. *Agronomie* 5: 915-921.

SAMBROOK , J., E. F. FRITSCH, AND T. MANIATIS. 1989. Molecular cloning, a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

SANG, T., D. J. CRAWFORD, S.-C. KIM, AND T. F. STUESSY. 1994. Radiation of the endemic genus *Dendroseris* (Asteraceae) on the Juan Fernandez Islands: evidence from sequences of the ITS regions of nuclear ribosomal DNA. *American Journal of Botany* 81: 1494-1501.

―――, ―――, AND T.F. STUESSY. 1995. ITS sequences and the phylogeny of the genus *Robinsonia* (Asteraceae). *Systematic Botany* 20: 55-64.

SHARMA, S. K., M. R. KNOX, AND T. H. N. ELLIS. 1996. AFLP analysis of the diversity and phylogeny of *Lens* and its comparison with RAPD analysis. *Theoretical and Applied Genetics* 93:751-758.

SHIH, C. 1988. Revision of *Lactuca* L. and two new genera of tribe Lactuceae (Compositae) on the mainland of Asia. *Acta Phytotaxonomica Sinica* 26: 382-393.

————. 1991. On circumscription of the genus *Cicerbita* Wall., and two new genera of Compositae from Sino-Himalayan region. *Acta Phytotaxonomica Sinica* 29: 394-417.

STEBBINS, G. L. 1937. Critical notes on *Lactuca* and related genera. *Journal of Botany*: 12-18.

————. 1953. A new classification of the tribe Cichorieae, family Compositae. *Madroño* 12: 65-81.

SUSANNA, A., N. GARCIA JACAS, D. E. SOLTIS, AND P. S. SOLTIS. 1995. Phylogenetic relationships in tribe Cardueae (Asteraceae) based on ITS sequences. *American Journal of Botany* 82: 1056-1068.

SWOFFORD, D. L. 1993. PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1.1. Computer program distributed by the Illinois Natural History Survey, Champaign, IL.

THOMPSON, R. C. 1943. Further studies on interspecific genetic relationships in *Lactuca*. *Journal of Agricultural Research* 66: 41-48.

————, T. W. WHITAKER, AND W. F. KOSAR. 1941. Interspecific genetic relationships in *Lactuca*. *Journal of Agricultural Research* 63: 91-107.

TUISL, G. 1968. Der Verwandtschaftskreis der Gattung *Lactuca* L. im Iranischen Hochland und seinen Randgebieten. *Annalen des Naturhistorischen Museums in Wien* 72: 587-638.

TUTIN, T. G., V. H. HEYWOOD, N. A. BURGES, D. M. MOORE, D. H. VALENTINE, S. M. WALTERS, AND D. A. WEBB (eds.). 1976. Flora Europaea, vol. 4, Cambridge University Press, Cambridge.

VERMEULEN, A., B. DESPREZ, D. LANCELIN, AND H. BANNEROT. 1994. Relationships among *Cichorium* species and related genera as determined by analysis of mitochondrial RFLPs. *Theoretical and Applied Genetics* 88: 159-166.

VOSMAN, B., P. ARENS, W. RUS-KORTEKAAS, AND M. J. M. SMULDERS. 1992. Identification of highly polymorphic DNA regions in tomato. *Theoretical and Applied Genetics* 85: 239-244.

WHITE, T. J., T. BRUNS, S. LEE, AND J. TAYLOR. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White [eds.], PCR protocols, a guide to methods and applications, 315-322. Academic Press, San Diego, CA.

WHITTON, J., R. S. WALLACE, AND R. K. JANSEN. 1995. Phylogenetic relationships and patterns of character change in the tribe Lactuceae (Asteraceae) based on chloroplast DNA restriction site variation. *Canadian Journal of Botany* 73: 1058-1073.

ZOHARY, D. 1991. The wild genetic resources of cultivated lettuce (*Lactuca sativa* L.) *Euphytica* 53: 31-35.

5

Species relationships in *Lactuca* s.l. (Lactuceae, Asteraceae) inferred from AFLP fingerprints

W. J. M. Koopman[1], M. J. Zevenbergen[2], and R. G. Van den Berg[1]

[1] Biosystematics Group, Nationaal Herbarium Nederland - Wageningen University branch, Wageningen University, Generaal Foulkesweg 37, 6703 BL Wageningen, The Netherlands
[2] Current address: Keygene N.V., P.O. Box 216, 6700 AE Wageningen, The Netherlands

Summary

An AFLP data set comprising 95 accessions from 20 species of *Lactuca* s.l. (sensu lato) and related genera was generated using the primer combinations E35/M48 and E35/M49. In phenetic analyses of a data subset, clustering with UPGMA based on Jaccard's similarity coefficient resulted in the highest cophenetic correlation, and the results were comparable to those of a principal coordinates analysis. In analyses of the total data set, phenetic and cladistic analyses showed similar tree topologies for the well-supported parts of the trees. The validity of cladistic analysis of AFLP data is discussed. The results do not support a distinction among the *serriola*-like species *L. sativa*, *L. serriola*, *L. dregeana*, and *L. altaica*, which is in line with previous results. Therefore, we postulate that these species are conspecific. The *serriola*-like species *L. aculeata* occupies a clearly separate position, making it an ideal outgroup for studies of the closest relatives of *L. sativa*. The subsect. *Lactuca* as a group is well supported by our data, but the positions of *L. saligna* and *L. virosa* relative to the *serriola*-like species remain unclear. The close relationship between the sect. *Mulgedium* species *L. tatarica* and *L. sibirica* is corroborated by the present AFLP results and by additional crossability data.

Key words: AFLPs, Asteraceae, Compositae, *Lactuca*, Lactuceae, lettuce, molecular phylogeny, phenetic relationships.

Introduction

Cultivated lettuce (*Lactuca sativa* L.) is the world's most important leafy salad vegetable (McGuire et al., 1993). The taxonomic status of this species, the boundaries among *L. sativa* and close relatives, and the boundaries of the genus *Lactuca* L. s.l. (sensu lato) (Lactuceae, Asteraceae) itself have been the subject of controversy among taxonomists for many decades. One of the most widely used classifications today is that of Feráková (1977), comprising the European species of *Lactuca*. She subdivides the genus into four sections: *Lactuca, Mulgedium* (Cass.) C.B. Clarke, *Lactucopsis* (Schultz-Bip. ex Vis. et Panc.) Rouy., and *Phaenixopus* (Cass.) Benth. Section *Lactuca* is subdivided into the subsections *Lactuca* and *Cyanicae* DC. The subsection *Lactuca* comprises *L. sativa, L. serriola* L., *L. altaica* Fisch. et C.A. Mey., *L. saligna* L., *L. virosa* L., and *L. livida* Boiss. et Reut. *Lactuca livida* is closely related to *L. virosa* (Velasco Negueruela, 1981). The species *L. sativa, L. serriola,* and *L. altaica* are closely related and probably conspecific (see Koopman et al., 1998, for a discussion). The lesser known southwest Asian species *L. aculeata* Boiss. & Kotschy ex Boiss., *L. scarioloides* Boiss., *L. azerbaijanica* Rech., *L. georgica* Grossh., and the South-African species *L. dregeana* DC. are also closely related to *L. sativa/serriola/altaica* (Zohary, 1991). These species could all be classified in Feráková's subsection *Lactuca* if her classification were to be extended to include non-European species. The species of subsect. *Lactuca* comprise the readily accessible part of the lettuce gene pool, and form potentially valuable gene sources for lettuce breeding (Zohary, 1991). *Lactuca serriola, L. saligna, L. virosa,* and to a lesser extent *L. altaica* are already commonly used as lettuce genitors. The *Lactuca* species outside subsect. *Lactuca,* as well as species from genera closely related to *Lactuca,* are interesting candidates for broadening the lettuce gene pool (Koopman et al., 1998).

In a previous study, Koopman et al. (1998) used ITS-1 (internal transcribed spacer-1) sequences to examine the relationships of species within or close to the lettuce gene pool. The study enabled straightforward conclusions on the generic and infrageneric boundaries of *Lactuca,* but was inconclusive as to the relationships among closely related species, e.g., within subsect. *Lactuca.* Koopman et al. (1998) concluded that additional information from a more variable marker was needed to resolve these relationships. A study by Hill et al. (1996) demonstrated that AFLPs (Vos et al., 1995) are variable markers useful for studying relationships among closely related species of *Lactuca.* Therefore, in the present study we used AFLP markers to further elucidate the relationships among *Lactuca* species and species from related genera. Our study had four foci: (1) the distinction between *L. sativa* and *L. serriola,* (2) the distinction between *L. serriola* and the *serriola*-like species *L. dregeana, L. altaica,* and *L. aculeata,* (3) the position of *L. saligna* and *L. virosa* relative to these *serriola*-like species, and (4) the detection of clusters/clades of closely related species outside subsect. *Lactuca.*

Data were analyzed both phenetically and cladistically, and the validity of cladistic analysis of AFLP data was discussed. A subset of data was used to compare various combinations of similarity coefficients and clustering methods for phenetic analyses.

Materials and Methods

Plant material

We used 95 accessions from a previous ITS-1 sequence study (Koopman et al, 1998), representing 20 species of *Lactuca* and related genera. The species are listed in Table 1 according to the subtribal classification of Bremer (1994), and the generic and specific classification of Feráková (1977) and Iwatsuki et al. (1995). The choice of species, the major generic concepts in *Lactuca* and related genera, and the delimitation of *Lactuca* were discussed in Koopman et al. (1998). Details on the accessions were given in Koopman et al. (1998) and on the website of the Centre for Genetic Resources, The Netherlands (CGN) at http://www.plant.wageningen-ur.nl/CGN. Each accessions was represented by two plants.

Table 1. Lactuceae species used in this study. The subtribal classification follows Bremer (1994); generic and specific classification of European species follows Feráková (1977). The Asiatic species *L. indica* is classified in the non-European section *Tuberosae*, according to Iwatsuki et al. (1995).

Species	# accessions
Subtribe Lactucinae Dumort., genus *Lactuca* L.	
Lactuca sect. *Lactuca* subsect. *Lactuca*	
Lactuca sativa L.	10
Lactuca serriola L.	10
Lactuca dregeana DC.	2
Lactuca altaica Fisch. et C.A. Mey.	2
Lactuca aculeata Boiss. & Kotschy ex Boiss.	2
Lactuca saligna L.	10
Lactuca virosa L.	11
Lactuca sect. *Lactuca* subsect. *Cyanicae* DC.	
Lactuca tenerrima Pourr.	5
Lactuca perennis L.	5
Lactuca sect. *Mulgedium* (Cass.) C.B. Clarke	
Lactuca tatarica (L.) C.A. Mey.	6
Lactuca sibirica (L.) Benth. ex Maxim.	5
Lactuca sect. *Lactucopsis* (Schultz-Bip. ex Vis. et Panc.) Rouy	
Lactuca quercina L.	1
Lactuca sect. *Phaenixopus* (Cass.) Benth.	
Lactuca viminea (L.) J. & C. Presl	5
Lactuca sect. *Tuberosae* Boiss.	
Lactuca indica L.	5
Subtribe Lactucinae, other genera	
Mycelis muralis (L.) Dumort.	4
Steptorhamphus tuberosus (Jacq.) Grossh.	1
Cicerbita plumieri (L.) Kirschl.	3
Cicerbita alpina (L.) Wallr.	3
Prenanthes purpurea L.	2
Unassigned to a subtribe	
Cichorium intybus L.	3

Voucher specimens of the plant material in rosette, bolting, and flowering stages were deposited at the Herbarium Vadense (WAG), supplemented with photographs of the plants in all three stages and with pappus preparations and fruit samples. All plants were grown under standard greenhouse conditions.

DNA extraction
Fresh young leaf tissue was collected from each plant, frozen in liquid nitrogen and kept at – 70°C. Nuclei were isolated (one plant per accession), and DNA was further purified using phenol/chloroform extraction as described by Vosman et al. (1992).

AFLP analysis
The AFLP procedure was performed according to Van Eck et al. (1995) with minor modifications. In the restriction/ligation reaction ~ 250 ng of genomic DNA was digested for 1 h at 37°C using 2.5 U (units) EcoRI, 2.5 U MseI, and 8 μL 5× restriction-ligation buffer (5× RL buffer) in a total volume of 40 μL. Restriction/ligation was continued for another 3 h after addition of 10 μL of ligation mixture (containing 5 pmol EcoRI adapter, 50 pmol MseI adapter, 1.0 μL 10 mmol/L ATP, 2.0 μL 5× RL buffer, and 1.0 U T4 DNA ligase). The subsequent selection of biotinylated restriction fragments with streptavidin-coated Dynabeads was replaced by a tenfold dilution of the restriction/ligation mixture with distilled water.

Preselective amplification was performed using the primers E01 (EcoRI + A) and M02 (MseI + C). The resulting product was diluted 50-fold with T0.1E buffer (10 mmol/L Tris pH 8.0, 0.1 mmol/L EDTA) . The final restriction fragment amplification was performed using primers with three selective nucleotides. The EcoRI primer in this reaction was labeled with [33]P. A pilot study was conducted to test ten primer combinations: E33/M59, E35/M48, E35/M49, E35/M59, E35/M60, E38/M54, E44/M48, E44/M49, E45/M48, and E45/M49. The test data set contained four plants of *L. sativa*, two plants of *L. saligna*, and one plant from each of the other species in Table 1. The AFLP fragments for this experiment were separated on a 0.35-mm sequence system (Gibco BRL/Life Technologies, Rockville, Maryland, USA) and visualized on Kodak X-OMAT LS Scientific Imaging Film (Eastman Kodak, Rochester, New York, USA). Selection of primer combinations was based on the number of bands per lane, the number of bands that were constant among the species, and the absence of very fat bands or smears. Primer combinations E35/M48 (EcoRI + ACA/MseI + CAC) and E35/M49 (EcoRI + ACA/MseI + CAG) were selected to generate the final data set. The AFLP procedure for E35/M48 was performed as above. For E35/M49, final restriction fragment amplification and separation and visualization of the AFLP fragments was performed according to Arens et al. (1998). *Lactuca sativa* 'Norden' served as size standard on each gel. A reference gel with fragment lengths of 'Norden' was kindly provided by Keygene N.V. (Wageningen, The Netherlands).

Data analysis

AFLP fragments were scored as present/absent. Fragment scoring and lane matching were performed automatically on digital images of the autoradiograms, using Phoretix 1D advanced Version 4.00 (Phoretix International, Newcastle upon Tyne, UK). All but the faintest bands were scored, where necessary scores and matches were corrected manually. Fragments scored ranged from 112 to 453 nucleotides for E35/M48 and from 111 to 502 nucleotides for E35/M49. Data from both primer combinations were combined in one data set.

The data set was analyzed in two steps:

Firstly, a data subset was constructed comprising *L. sativa* and its closest relatives, *L. serriola, L. dregeana, L. altaica,* and *L. aculeata*. In the following these will be referred to as the "*serriola*-like species." The subset was used to compare various similarity coefficients and clustering methods and to study the relationships among the "*serriola*-like species" in detail. Clustering methods and similarity coefficients were tested using the procedures SIMQUAL, SAHN, and TREE from the program NTSYSpc version 2.02k (Applied Biostatistics, Setauket, New York, USA). The "TM" option was set to "FIND" to enable detection of all possible trees. The clustering methods UPGMA, WPGMA, Complete-link, and Single-link were applied in all possible combinations with the similarity coefficients Dice, Jaccard's, and Simple matching. Clustering methods and similarity coefficients are described in Rohlf (1993). Cophenetic correlation coefficients (*r*) were calculated and compared for each of the combinations using the procedures COPH and MXCOMP from NTSYSpc 2.02k. These coefficients indicate the correlation between a similarity matrix and the phenetic tree resulting from it after a cluster analysis, and thus are a measure for the goodness of fit of the cluster analysis to the similarity matrix.

Species relationships among the *serriola*-like species were studied using a principal coordinates analysis (PCO). Jaccard's similarity coefficient and the procedures DCENTER, EIGEN, and MXPLOT from NTSYSpc 2.02k were used to perform the PCO.

Secondly, analyses were performed on the entire dataset, containing all accessions from Table 1. This data set was used to compare phenetic and cladistic analysis of the AFLP data, and to detect well-supported species clusters/clades within *Lactuca* s.l. The cluster analysis was performed with TREECON 1.2 (Van de Peer and De Wachter, 1994), which enabled bootstrapping of the resulting phenogram. Nei and Li's (1979) dissimilarity coefficient and UPGMA clustering were used; bootstrap values were calculated in 1000 replications. Cladistic analyses and determination of phylogenetic signal in the data set were performed using PAUP version 4.0a (Swofford, 1999). Parsimony settings were: ACCTRAN and "collapse of zero length branches" (max). Phylogenetic signal was determined from the tree-length distribution of 100 000 trees, using the g_1-statistic (Hillis and Huelsenbeck, 1992). The lettuce data set contained >25 taxa and >500 variable characters, and therefore the critical value of −0.08 was used. A g_1-statistic lower than this critical value indicates the presence of significant phylogenetic signal in the corresponding data set (Hillis and Huelsenbeck, 1992).

The cladistic analyses started as a jackknife analysis using 10 000 replicates of a fast heuristic search, nominal deletion of 37% of the characters, and "Jac" resampling. A 50% majority rule consensus tree was calculated based on the jackknife analysis and used as a constraint tree for a heuristic search. The heuristic search comprised 10 000 random-addition sequences and tree bisection-reconnection (TBR) branch swapping with "multrees" switched off. A second search was performed using four cycles of successive weighting. The strict consensus of the heuristic search above was used as a starting point. Characters were reweighted by the maximum value of the rescaled consistency indices, and the searches were conducted with 100 random-addition sequences, TBR branch swapping, and "multrees" on. Jackknife values for the resulting tree were calculated as above.

Results

Total number of bands scored was 544 for E35/M48, and 521 for E35/M49, all of which were polymorphic. Band numbers for the individual accessions ranged from 16 to 109 (average 59.0 bands/lane) for E35/M48 and from 28 to 103 for E35/M49 (average 54.6 bands/lane).

Table 2. Cophenetic correlation coefficients for a data subset containing *Lactuca sativa, L. serriola, L. dregeana, L. altaica,* and *L. aculeata* accessions; total number of trees found in the analysis are in brackets. When multiple trees were found, only the highest cophenetic value is shown. For the Single-link method, all trees are equivalent and thus have the same cophenetic values.

Clustering / Similarity	Dice	Jaccard's	Simple matching
UPGMA	0.974 (1)	0.979 (1)	0.955 (1)
WPGMA	0.963 (1)	0.968 (1)	0.953 (1)
Complete-link	0.969 (2)	0.973 (2)	0.948 (2)
Single-link	0.951 (1)	0.957 (1)	0.941 (8)

Table 2 shows the cophenetic correlation coefficients from analyses of the data subset containing *L. sativa* and its closest relatives, *L. serriola, L. dregeana, L. altaica,* and *L. aculeata* (the *serriola*-like species). UPGMA clustering yielded the highest cophenetic correlation in all cases, Single-link the lowest. The ranking of WPGMA and Complete-link was less consistent (see Table 2). Among the similarity coefficients, Jaccard's consistently yielded the highest cophenetic correlations, followed by Dice and Simple matching. Since a similar ranking of similarity coefficients and clustering methods was found by Mace et al. (1999a, b) for *Solanum* and *Datura/Brugmansia*, it possibly applies to all AFLP data sets. The combination of Jaccard's similarity with UPGMA clustering yielded the highest cophenetic correlation and is therefore considered most suitable for determining phenetic species relationships in *Lactuca* s.l. The combination of UPGMA with the Dice or Nei and Li (equaling 1-Dice) coefficient is also suitable for our data set, since trees based on these coefficients were identical to those based on Jaccard's coefficient. Identical topologies for the

Jaccard's and Dice coefficients were also found by Milbourne et al. (1997) for AFLP data from cultivated potato.
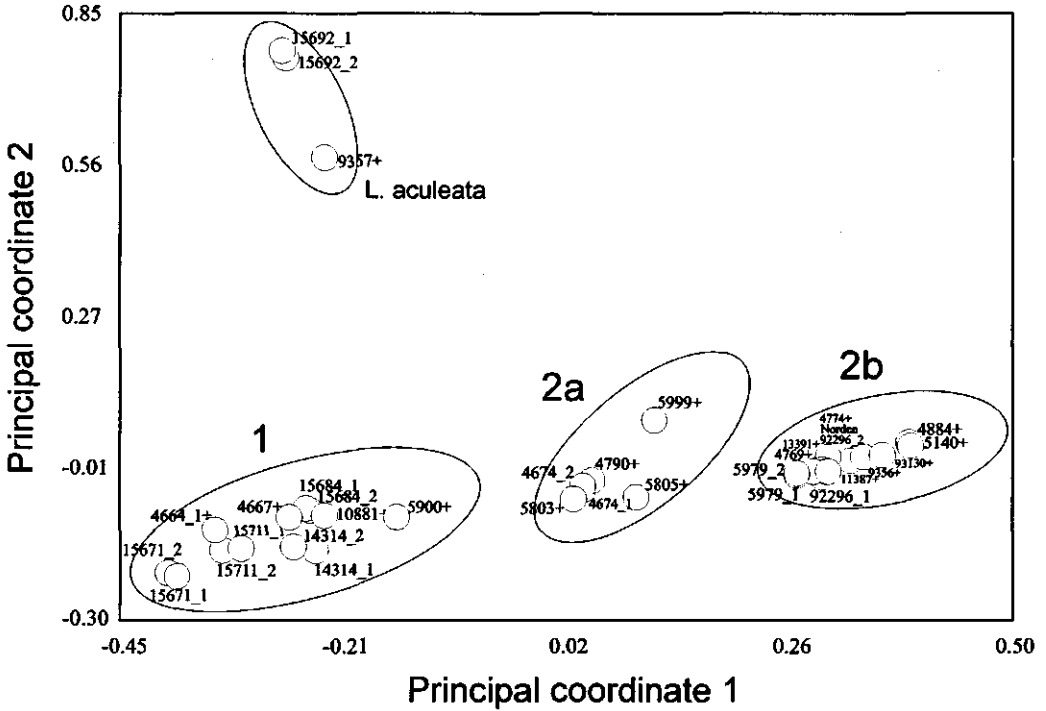

Fig. 1. Principal coordinates analysis of a data subset. Numbers 1, 2a, and 2b indicate the different groups of species referred to in the text. 1 = *L. serriola/L. altaica*, 2a = *L. sativa/L. serriola/L. dregeana*, 2b = *L. sativa/L. serriola*.

Species relationships among the *serriola*-like species were studied in detail with a PCO (Fig. 1). The first principal coordinate describes 18% of the total variation and separates three groups. Group 1 contains the *L. altaica* accessions and some of the *L. serriola* accessions. The *L. serriola* accessions CGN 15684 and CGN 5900 also fall in this group, although they cluster in group 2a in the cluster analysis (see below). The *L. altaica* accessions fall among the *L. serriola* accessions. Group 2a contains *L. sativa*, *L. serriola*, and *L. dregeana*. The *L. dregeana* accessions fall among the *L. serriola* accessions. Group 2b contains most *L. sativa* accessions and the *L. serriola* oilseed accessions. Note that the *L. sativa* accessions also include an oilseed accession, CGN 9356. The second principal coordinate describes 12% of the total variation and clearly sets apart *L. aculeata* from *L. sativa*, *L. serriola*, *L. dregeana*, and *L. altaica*.

In the cluster analysis comprising all accessions, all species except *L. sativa* and *L. serriola* have their own distinct branches (Fig. 2a). However, *L. altaica* (30% support) and *L. dregeana* (98%) cluster within *L. sativa/L. serriola*. Subsection *Lactuca* (the *serriola*-like species

together with *L. virosa* and *L. saligna*) is well supported (99%). *L. virosa* clusters more closely to the *serriola*-like species than does *L. saligna*, but the branch determining this order is poorly supported (52%). The cluster including only the *serriola*-like species is well supported (100%) and consists of four groups. These groups are identical to those in the PCO, except for the position of CGN 15684 and CGN 5900 (see above). The cluster with *L. aculeata* is strongly supported (100%), but the *L. serriola/L. altaica* cluster (group 1) and the *L. sativa/L. serriola/L. dregeana* cluster (group 2a), are not (26% and 18%, respectively). The cluster with only *L. sativa* accessions (group 2b) is strongly supported (96%), but falls entirely within the *L. sativa/L. serriola/L. dregeana* cluster. The only well-supported species cluster outside subsect. *Lactuca* is that of *L. sibirica*, *L. tatarica*, and *L. quercina* (99%).

In the cladistic analyses, the g_1-statistic for the combined data set was -0.39, indicating significant phylogenetic signal. The heuristic search with random-addition sequences yielded 40 shortest trees of 4628 steps (RI = 0.76, CI = 0.23, RC = 0.18). The search with successive weighting yielded a single tree of 752 steps (RI = 0.84, CI = 0.45, RC = 0.38). This single tree was compatible to the strict consensus tree of the search with random-addition sequences, but slightly more resolved (Fig. 2b). Topology and bootstrap/jackknife supports for clusters/clades with a support >70% are similar in the cladogram (Fig. 2b) and the phenogram (Fig. 2a).

Discussion

Distinction between L. sativa and L. serriola

Lactuca sativa and *L. serriola* group in three clusters/clades, but the distinction between these clusters/clades is weakly supported. Two of the clusters/clades contain both *L. sativa* and *L. serriola* accessions. This is consistent with the AFLP results of Hill et al. (1996), showing a large *L. sativa/L. serriola* cluster with *L. serriola* accessions branching off basally to a large subcluster containing all *L. sativa* accessions and one *L. serriola* accession. This *L. serriola* is a "landrace type", intermediate between *L. sativa* and *L. serriola*. The *L. sativa/L. serriola* cluster is clearly separated from *L. saligna*, *L. virosa*, *L. indica*, and *L. perennis*.

Next pages: Fig. 2. (left page: a) UPGMA phenogram based on Nei and Li's distance. Numbers on branches are bootstrap values. (right page: b) Cladogram resulting from successive weighting of the strict consensus tree from 10 000 random-addition sequences with TBR branch swapping and "multrees" switched off. Numbers on each branch are jackknife value (left of slash) and total number of AFLP bands supporting the branch (unweighted)(right of slash). Dotted branches collapse in the strict consensus of the trees from the random-addition sequence searches. Two plants were used for each accession, indicated by _1 and _2. A + sign indicates that the AFLP patterns were identical for the two plants, and only one of the plants is depicted in the phenogram (*sat/ser/dreg/alt* = species cluster containing *L. sativa, L. serriola, L. dregeana,* and *L. altaica*). The boxes between Fig. 2a and 2b indicate well-supported clusters present in both trees. Numbers 1, 2a, and 2b indicate different groups of species referred to in the text. 1 = *L. serriola/L. altaica*, 2a = *L. sativa/L. serriola/L. dregeana*, 2b = *L. sativa/L. serriola*.

a

Nei and Li's distance
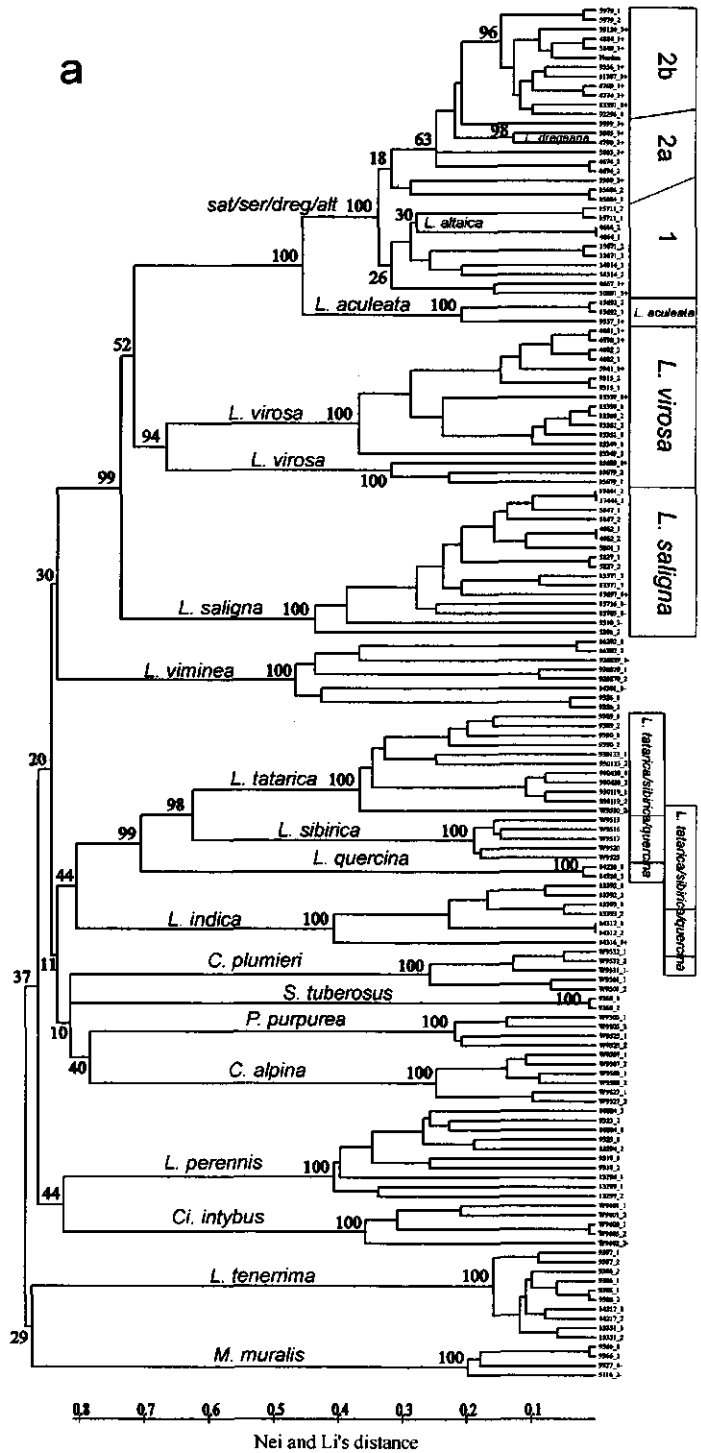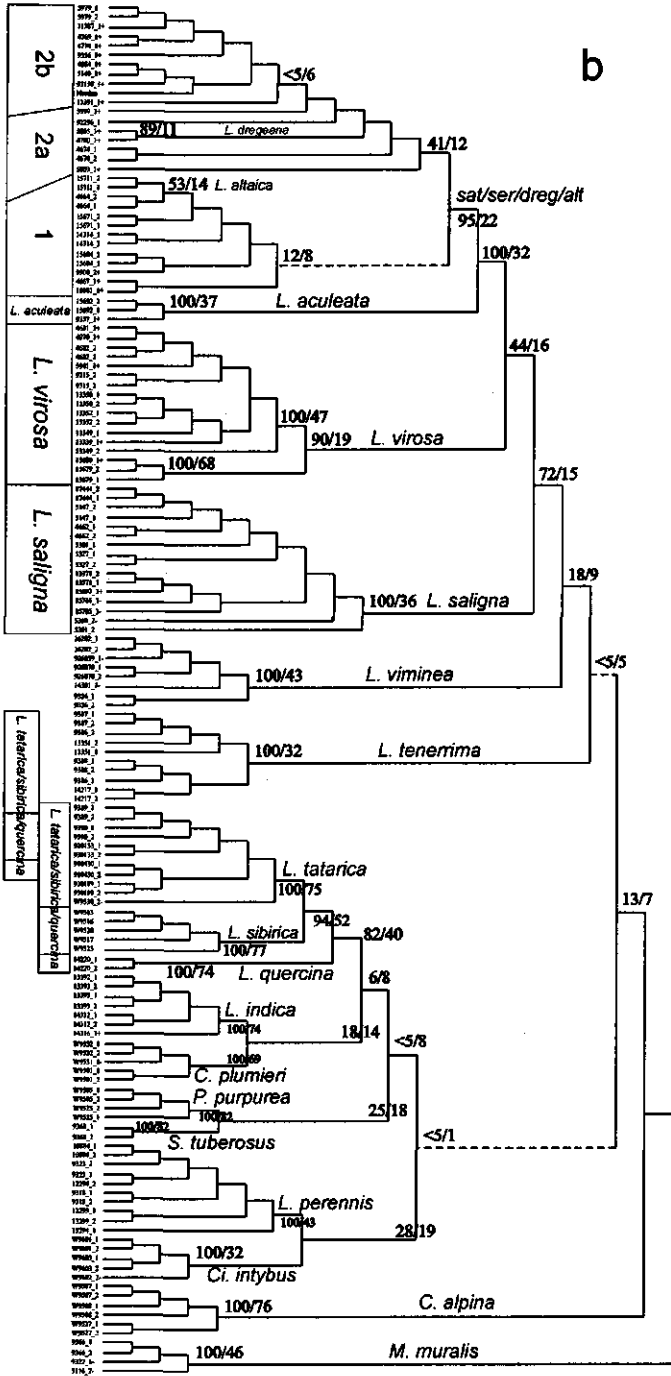
The *L. serriola* and *L. sativa* oilseed accessions in the present study fell within *L. sativa*. This is consistent with the results of Frietema de Vries et al. (1994) and Frietema de Vries (1996). In their principal components analysis of morphological data, the *L. sativa* and *L. serriola* accessions fall in two partly overlapping groups. According to the text, the oilseed accessions are included in the *L. sativa* cluster, although an accompanying figure depicts them as intermediate between *L. sativa* and *L. serriola* (Frietema de Vries et al., 1994; Frietema de Vries, 1996).

Given the lack of distinction between *L. sativa* and *L. serriola* in the present study, the position of the *L. serriola* oilseeds within *L. sativa* in the study of Frietema de Vries et al. (1994) and Frietema de Vries (1996), the presence of a *L. serriola* "landrace type" within *L. sativa* in the study of Hill et al. (1996), and the close similarity of *L. sativa* and *L. serriola* in other characters (discussed in Koopman et al., 1998), we support the conclusion of Frietema de Vries et al. (1994) and Frietema de Vries (1996) that *L. sativa* and *L. serriola* are conspecific. However, we do not support the distinction of *L. sativa* subsp. *sativa* and *L. sativa* subsp. *serriola*, as proposed by Frietema de Vries (1996). In our opinion, the species are too similar even to maintain them as subspecies. Therefore, we consider the earliest name, *L. sativa*, the correct name for both *L. serriola* and *L. sativa*.

Position of L. dregeana, L. altaica, and L. aculeata relative to L. serriola

Lactuca dregeana accessions fell within the mixed *L. sativa*/*L. serriola* cluster/clade in all our analyses. Most accessions within this cluster/clade show a mixture of *L. serriola* and *L. sativa* characteristics. For example, *L. sativa* accession CGN 5999 has an especially rigid, nearly woody stem, and spines on the midribs beneath, characteristics usually associated with *L. serriola*. On the other hand, *L. serriola* accessions CGN 5803 and CGN 4674 show spineless lower midribs, somewhat fleshy leaves, and involucres that are not completely reflexed when the fruits are ripe. These characteristics are usually associated with *L. sativa*. The *L. dregeana* accessions show a similar combination of characteristics. They resemble *L. sativa* in their somewhat fleshy leaves and involucres that are not completely reflexed when the achenes are ripe. On the other hand, they show *L. serriola* characteristics such as a rigid, spiny stem, spiny lower midribs, and dark brown, spotted achenes.

The combination of morphological characteristics and the position of *L. dregeana* in the mixed *sativa*/*serriola* cluster/clade in our AFLP analyses suggest that *L. dregeana* escaped from cultivation. The fact that *L. dregeana* is endemic to South Africa could mean that it originated from the primitive lettuce cultivars introduced there by European settlers in the 17th century. Lettuce seed production in the Cape was reported as early as 1652-1654 (Karsten, 1951) and could easily have led to escapes to the wild by wind dispersal of achenes from cultivars with loose involucres. After taking into account the morphology of *L. dregeana*, its position in the AFLP analyses, and its possible origin in cultivated lettuce, *L. dregeana*

probably does not deserve a species status, but it should be regarded conspecific with *L. sativa/ L. serriola*.

The *L. altaica* accessions in the present study fell within a group of *L. serriola* accessions, corroborating previous ITS-1 results and the conclusion that *L. altaica* is probably conspecific with *L. serriola* (Koopman et al., 1998). However, this conclusion is based on only two *L. altaica* accessions. Recently, additional wild material of *L. altaica* and its relatives *L. serriola* and *L. saligna* was collected in Uzbekistan (Van Soest, 1997). A study on this material is currently being carried out to further elucidate the relationships and taxonomic status of *L. altaica*.

The accessions of *L. aculeata* form a clearly distinct group among the *serriola*-like species, with a 100% jackknife and bootstrap support. The position of *L. aculeata* separate from, yet closely related to, the other *serriola*-like species is well supported by our earlier ITS-1 study (Koopman et al., 1998). This distinct position of *L. aculeata* within the *serriola*-like species makes it an ideal outgroup for studies into *L. sativa*, *L. serriola*, and their closest relatives.

All *serriola*-like species together, i.e., including *L. aculeata*, form a homogeneous group of closely related species within subsect. *Lactuca*. This is indicated by the 100% jackknife and bootstrap support for this group in the present AFLP analysis (Fig. 2ab), the 95% bootstrap support in a previous ITS-1 analysis, and the fact that all *serriola*-like species are fully interfertile (Koopman et al., 1998).

Position of L. saligna and L. virosa within subsect. Lactuca

The results of previous studies on plant morphology (De Vries and Van Raamsdonk, 1994), crossability (Thompson, Whitaker, and Kosar, 1941; Lindqvist, 1960; de Vries, 1990), SDS (Sodium Dodecyl Sulphate) electrophoresis patterns of seed proteins (De Vries, 1996), isozyme analysis of foliar esterases (Roux, Chengjiu, and Roux, 1985), karyotype (Lindqvist, 1960; Koopman and De Jong, 1996), chromosome banding pattern (Koopman, De Jong, and De Vries, 1993), DNA content (Koopman and De Jong, 1996; Koopman, 2000), nuclear RFLPs (restriction fragment length polymorphisms)(Kesseli, Ochoa, and Michelmore, 1991), mtDNA RFLPs (Vermeulen et al., 1994), nuclear AFLPs (Hill et al., 1996), and ITS-1 sequences (Koopman et al., 1998) showed different possibilities for the position of *L. saligna* and *L. virosa* relative to the *serriola*-like species, as was discussed in Koopman et al. (1998).

The present results indicate that *L. virosa* is more closely related to the *serriola*-like species than is *L. saligna*. However, the position of *L. saligna* and *L. virosa* relative to the *serriola*-like species is not very reliable, as is indicated by the low bootstrap and jackknife supports on the branches separating *L. saligna* and *L. virosa*. The results are not consistent with the AFLP analysis of Hill et al. (1996), indicating that *L. saligna* is the closest relative of the *serriola*-like species, and that *L. perennis* is even more closely related to the *serriola*-like species than is *L. virosa*. However, Hill et al. (1996) do not indicate support values for the relationships. Given the different positions of *L. saligna* and *L. virosa* in the study of Hill et al. (1996) and in the

present study, and the lack of branch support for this positions, we conclude that the available AFLP data are inconclusive as to the position of *L. virosa* and *L. saligna* relative to the *serriola*-like species. This is also true for the position of *L. perennis*.

In the present study, *L. virosa* accessions CGN 15679 and 15680 form a separate clade with a 100% bootstrap support. The anomalous position of these accessions may indicate that they are a distinct infraspecific taxon within *L. virosa*. The distinct position of CGN 15679 and CGN 15680 is also reflected by the fact that their DNA content is 1.16 times that of the other *L. virosa* accessions (Koopman, 2000).

Species clusters/clades outside subsect. Lactuca

In a previous study using ITS-1 sequences, four clades of species were detected outside subsect. *Lactuca*: (1) *L. tatarica*/*L. sibirica*/*L. viminea*, (2) *L. perennis*/*C. plumieri*, (3) *L. tenerrima*/*S. tuberosus*, (4) *M. muralis*/*C. alpina*. Only one of these clades could be partially confirmed by our AFLP results. In all phenetic and cladistic analyses, the only well-supported cluster/clade outside subsect. *Lactuca* was one with *L. tatarica*, *L. sibirica*, and *L. quercina*. The relationship between *L. tatarica* and *L. sibirica* is consistent with the ITS-1 results and with the classification of Feráková (1977). The close relationship of *L. tatarica* and *L. sibirica* with *L. quercina* is not. In the ITS-1 phylogeny, *L. quercina* has its own distinct branch, while *L. tatarica* and *L. sibirica* occupy the same clade. In the classification of Feráková (1977), *L. quercina* is classified in section *Lactucopsis*, while *L. tatarica* and *L. sibirica* together make up section *Mulgedium*. The close relationship between *L. tatarica* and *L. sibirica* was confirmed by our recent crossing experiments, reported here for the first time. We conducted reciprocal crosses between four *L. tatarica* accessions and four *L. sibirica* accessions, and these crosses yielded viable seeds for six out of eight combinations of accessions. The F1 plants were vigorous and fully fertile, indicating a close genetic relationship between *L. tatarica* and *L. sibirica*. The possible close relationship of *L. tatarica*/*L. sibirica* with *L. quercina* needs more verification. Species relationships involving *L. viminea, L. indica, L. perennis, L. tenerrima, M. muralis, C. plumieri, C. alpina, S. tuberosus, P. purpurea,* and *C. intybus* could not be assessed in the present study, because the AFLPs were too variable to determine reliable relationships of these species.

Methodological considerations

In the present study, the data were analyzed both phenetically and cladistically. The validity of such analyses is sometimes disputed, although this dispute was not reflected in literature until now. Critics recognize two main sources of error in the cladistic analysis of AFLP data. Firstly, the fact that AFLPs are anonymous markers is a source of error. Because AFLP fragments are identified by their length and not by their base composition, nonidentical fragments of equal length will mistakenly be scored as identical. Secondly, the fact that AFLPs are scored dominantly is a source of error. AFLPs are usually scored as dominant characters, i.e., with only the character states present (1), and absent (0). In reality, at least part of the bands may

represent codominant markers that have three character states, namely 0/0, 1/0, and 1/1. Both sources of error introduce homoplasies in the data set, possibly leading to erroneous tree topologies in cladistic analyses. In our opinion, the impact of these homoplasies on the conclusions regarding species relationships will be minor.

When we compare phenetic and cladistic analysis of AFLP data, there are two possible situations. Firstly, the topologies of the phenogram and the cladogram may be identical. In this case, the homoplasies were too minor to influence the topology of the cladogram. Consequently, they will not affect conclusions on species relationships. Secondly, the topologies of the phenogram and the cladogram may be different. In this case, the homoplasies significantly affected the topology of branches in the cladogram. However, because the differences are caused by homoplasies, there will be internal conflict in the data defining these branches. In branch support analyses such as bootstrapping or jackknifing, the presence of such conflicting data gives rise to low support values. These poorly supported branches will be discarded as uninformative when conclusions on species relationships are drawn. Therefore, in this case, too, the homoplasies in the AFLP data will not affect the conclusions on species relationships. In both cases, cladistic analysis of AFLP data will give rise to reliable phylogenetic conclusions, notwithstanding the validity of the theoretical objections.

The first case is illustrated by a study of Kardolus et al. (1998) in which a cladogram and a phenogram of 16 wild *Solanum* species show highly similar topologies, even for moderately supported groups. In our lettuce study both cases are present: the well-supported parts of the phenogram and the cladogram show similar topologies, while the differences in the remaining parts of the trees are poorly supported.

Conclusions

AFLPs proved to be suitable molecular markers to study the relationships among closely related species of *Lactuca* s.l. In phenetic analyses of a data subset, the combination of Jaccard's similarity coefficient with UPGMA clustering resulted in the highest cophenetic value. The results of a principal coordinates analysis of the subset were comparable to those of the UPGMA analysis. A data set comprising all accessions was analyzed phenetically as well as cladistically, and the well-supported parts of the trees were comparable for both types of analyses. The AFLP results corroborated the conclusions from a previous ITS-1 sequence study (Koopman et al., 1998) that the *serriola*-like species *L. sativa, L. serriola, L. dregeana,* and *L. altaica* cannot be reliably distinguished and are probably conspecific. *Lactuca dregeana* possibly escaped from cultivation. *Lactuca aculeata* is closely related to the other *serriola*-like species, but clearly different. The AFLP results were inconclusive as to the position of *L. saligna* and *L. virosa* relative to the *serriola*-like species, but the status of subsect. *Lactuca* (the *serriola*-like species together with *L. saligna* and *L. virosa*) as a recognizable group within *Lactuca* s.l. was supported in all analyses. In the previous ITS-1 study, a number of species clades outside subsect. *Lactuca* were identified. Among the relationships indicated by these clades, only the close

relationship between *L. tatarica* and *L. sibirica* (together constituting *Lactuca* subsect. Mulgedium) was corroborated by the present AFLP results. The close relationship between these species was also corroborated by our crossability data.

Acknowledgements

References

ARENS, P., H. COOPS, J. JANSEN, AND B. VOSMAN. 1998. Molecular genetic analysis of black poplar (*Populus nigra* L.) along Dutch rivers. *Molecular Ecology* 7: 11-18.

BREMER, K. 1994. Asteraceae: cladistics and classification. Timber Press, Portland, Oregon, USA.

DE VRIES, I. M. 1990. Crossing experiments of lettuce cultivars and species (*Lactuca* sect. *Lactuca*, Compositae). *Plant Systematics and Evolution* 171: 233-248.

DE VRIES, I. M. 1996. Characterization and identification of *Lactuca sativa* cultivars and wild relatives with SDS-electrophoresis (*Lactuca* sect. *Lactuca, Compositae*). *Genetic Resources and Crop Evolution* 43: 193-202.

DE VRIES, I. M., AND L. W. D. VAN RAAMSDONK. 1994. Numerical morphological analysis of lettuce cultivars and species (*Lactuca* sect. *Lactuca*, Asteraceae). *Plant Systematics and Evolution* 193: 125-141.

FERÁKOVÁ, V. 1977. The genus *Lactuca* L. in Europe. Univerzita Komenského Press, Bratislava, Czechoslovakia.

FRIETEMA DE VRIES, F. T. 1996. Cultivated plants and the wild flora. Ph.D. dissertation, Rijksherbarium/Hortus Botanicus, Leiden, The Netherlands.

FRIETEMA DE VRIES, F. T., R. VAN DER MEIJDEN, AND W. A. BRANDENBURG. 1994. Botanical files on lettuce (*Lactuca sativa*). *Gorteria* supplement 2.

HILL, M., H. WITSENBOER, M. ZABEAU, P. VOS, R. KESSELI, AND R. MICHELMORE. 1996. PCR-Based fingerprinting using AFLPs as a tool for studying genetic relationships in *Lactuca* spp. *Theoretical and Applied Genetics* 93: 1202-1210.

HILLIS, D. M., AND J. P. HUELSENBECK. 1992. Signal, noise, and reliability in molecular phylogenetic analyses. *Journal of Heredity* 83: 189-195.

IWATSUKI, K., T. YAMAZAKI, D. E. BOUFFORD, AND H. OHBA. 1995. Flora of Japan, vol. IIIb. Kodansha Ltd., Tokyo, Japan.

KARDOLUS, J. P., H. J. VAN ECK, AND R. G. VAN DEN BERG. 1998. The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy (*Solanaceae*). *Plant Systematics and Evolution* 210: 87-103.

KARSTEN, M. C. 1951. The Old Company's Garden at the cape and its superintendents. Maskew Miller Limited, Cape Town, South Africa.

KESSELI, R. V., O. OCHOA, AND R. W. MICHELMORE. 1991. Variation at RFLP loci in *Lactuca* spp. and origin of cultivated lettuce (*L. sativa*). *Genome* 34: 430-436.

KOOPMAN, W. J. M. 2000. Identifying lettuce species (*Lactuca* subsect. *Lactuca*, Asteraceae): A practical application of flow cytometry. *Euphytica* 116: 151-159.

KOOPMAN, W. J. M., J. H. DE JONG, AND I. M. DE VRIES. 1993. Chromosome banding patterns in lettuce species (*Lactuca* sect. *Lactuca*, *Compositae*). *Plant Systematics and Evolution* 185: 249-257.

KOOPMAN, W. J. M., AND J. H. DE JONG. 1996. A numerical analysis of karyotypes and DNA amounts in lettuce cultivars and species (*Lactuca* subsect. *Lactuca*, Compositae). *Acta Botanica Neerlandica* 45: 211-222.

KOOPMAN, W. J. M., E. GUETTA, C. C. M. VAN DE WIEL, B. VOSMAN, AND R. G. VAN DEN BERG. 1998. Phylogenetic relationships among *Lactuca* (Asteraceae) species and related genera based on ITS-1 DNA sequences. *American Journal of Botany* 85: 1517-1530.

LINDQVIST, K. 1960. Cytogenetic studies in the *serriola* group of *Lactuca*. *Hereditas* 46: 75-151.

MACE, E. S., R. N. LESTER, AND C. G. GEBHARDT. 1999a. AFLP analysis of genetic relationships among the cultivated eggplant, *Solanum melongena* L., and wild relatives (Solanaceae). *Theoretical and Applied Genetics* 99: 626-633.

MACE, E. S., C. G. GEBHARDT, AND R. N. LESTER. 1999b. AFLP analysis of genetic relationships in the tribe Datureae (Solanaceae). *Theoretical and Applied Genetics* 99: 634-641.

MCGUIRE, P. E., E. J. RYDER, R. W. MICHELMORE, R. L. CLARK, R. ANTLE, G. EMERY, R. M. HANNAN, R. V. KESSELI, E. A. KURTZ, O. OCHOA, V. E. RUBATZKY, AND W. WAYCOTT. 1993. Genetic resources of lettuce and *Lactuca* species in California: an assessment of the USDA and UC collections and recommendations for long-term security. Report No. 12, Genetic Resources Conservation Program, Division of Agriculture and Natural Resources, University of California, Davis, California, USA.

MILBOURNE, D., R. MEYER, J. E. BRADSHAW, E. BAIRD, N. BONAR, J. PROVAN, W. POWELL, AND R. WAUGH. 1997. Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. *Molecular Breeding* 3: 127-136.

NEI, M., AND W.-H. LI. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA* 76: 5269-5273.

ROHLF, F. J. 1993. NTSYS-pc: numerical taxonomy and multivariate analysis system, version 1.80. Applied Biostatistics Inc., Setauket, New York, USA.

ROUX, L., Z. CHENGJIU, AND Y. ROUX. 1985. Caractérisation de *Lactuca sativa* L. et des espèces apparentées par électrofocalisation des estérases. *Agronomie* 5: 915-921.

SWOFFORD, D. L. 1999. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer, Sunderland, Massachusetts, USA.

THOMPSON, R. C., T. W. WHITAKER, AND W. F. KOSAR. 1941. Interspecific genetic relationships in *Lactuca. Journal of Agricultural Research* 63: 91-107.

VAN DE PEER, Y., AND R. DE WACHTER. 1994. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Computer Applications in the Biosciences* 10: 569-570.

VAN ECK, H. J., J. ROUPPE VAN DER VOORT, J. DRAAISTRA, P. VAN ZANDVOORT, E. VAN ENCKEVORT, B. SEGERS, J. PELEMAN, E. JACOBSEN, J. HELDER, AND J. BAKKER. 1995. The inheritance and chromosomal localization of AFLP markers in a non-inbred potato offspring. *Molecular Breeding* 1: 397-410.

VAN SOEST, L. J. M. 1997. Report of expedition to Uzbekistan in 1997: itinerary, collected materials and data. Agricultural Research Department, Centre for Plant Breeding and Reproduction Research, Centre for Genetic Resources, The Netherlands.

VELASCO NEGUERUELA, A. 1981. Sobre *Lactuca livida* Boiss. & Reuter. *Anales del Jardín Botánico de Madrid* 38: 316.

VERMEULEN, A., B. DESPREZ, D. LANCELIN, AND H. BANNEROT. 1994. Relationships among *Cichorium* species and related genera as determined by analysis of mitochondrial RFLPs. *Theoretical and Applied Genetics* 88: 159-166.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAN DE LEE, M. HORNES, A. FRIJTERS, J. POT, J. PELEMAN, M. KUIPER, AND M. ZABEAU. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407-4414.

VOSMAN, B., P. ARENS, W. RUS-KORTEKAAS, AND M. J. M. SMULDERS. 1992. Identification of highly polymorphic DNA regions in tomato. *Theoretical and Applied Genetics* 85: 239-244.

ZOHARY, D. 1991. The wild genetic resources of cultivated lettuce (*Lactuca sativa* L.). *Euphytica* 53: 31-35.

6

Evolution of DNA content and base composition in *Lactuca* (Asteraceae) and related genera

W. J. M. Koopman[1], J. Hadam[2], and J. Doležel[2]

[1] Biosystematics Group, Nationaal Herbarium Nederland - Wageningen University branch, Wageningen University, Generaal Foulkesweg 37, 6703 BL Wageningen, The Netherlands
[2] Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany,
Sokolovská 6, CZ-77200 Olomouc, Czech Republic

Manuscript

Summary

We determined 2C DNA contents (109 accessions) and base compositions (AT%, 24 accessions) for 23 species of *Lactuca* and related genera, as well as relative DNA contents for 21 modern *L. sativa* cultivars. Using these data, we examined the: 1) interspecific variation in DNA content and base composition; 2) intraspecific variation in DNA content; 3) evolution of DNA content and base composition; 4) correlation of DNA content and AT%; 5) correlation of genome size and number of AFLP bands generated from these genomes. Relative to other angiosperms, *Lactuca* s.l. species have low 2C DNA contents (ranging from 1.913 pg in *L. tenerrima* to 13.068 pg in *L. indica*), and high AT contents (ranging from 61.4% in *L. virosa* to 64.2% in *L. perennis*). Interspecific variation in DNA content is not generally applicable for species identification. Six species showed significant intraspecific variation in DNA content (*L. viminea* (25%), *L. virosa* (21%), *L. serriola* (7%), *L. sativa* (6%), *L. sibirica* (4%), and *P. purpurea* (4%)). Only the variation within *L. virosa* seems to have evolutionary significance, indicating two distinct evolutionary lines. *Lactuca* s.l. showed a general increase in DNA content during evolution, accompanied by decrease in AT% in one group of species and increase in another. The general trend is that increase in DNA amount is significantly associated with decrease in AT%. For species with 2C DNA contents of 8.5 pg or less, genome size is positively associated with number of AFLP bands, according to a curvilinear relationship. For species with 2C DNA contents above 8.5 pg, the relationship is more complex.

Key words: AFLP, Asteraceae, AT content, base composition, DNA content, *Lactuca*, Lactuceae, molecular phylogeny.

Introduction

DNA contents vary widely among plant species, with a tendency toward more similar DNA contents among more closely related taxa. Within angiosperms, DNA contents vary over 1000-fold (Bennett and Smith, 1976; Bennett and Smith, 1991; Bennett, Bhandol, and Leitch, 2000; Bennett and Leitch, 2001), amounting up to 100-fold within families (Cox et al., 1998), and typically varying 2 to 10 fold within genera (Price, 1988). Evolutionary patterns in DNA content of angiosperms have been studied since the early 1970's (Price, 1976). Evolution towards increase as well as towards decrease of DNA amount was detected, sometimes within a single genus (Price, 1976; Narayan, 1987; Kiehn, 1995; Bharathan, 1996; Ohri, 1998). Increase in DNA amount seems to be the general rule, at least at higher taxonomic levels (Bharathan, 1996; Bennetzen and Kellogg, 1997; Leitch, Chase, and Bennett, 1998).

In comparison to DNA content, base composition shows little variation among plant groups. Data on base composition are far more limited than those on DNA content, but an extensive literature survey showed that proportions of A+T nucleotides range from 55% to 65% in dicotyledons and non-Poaceae monocotyledons (average 60%), from 52% to 59% in Poaceae (average 56%), and from 60% to 65% in gymnosperms (average 62%). Base composition in relation to phylogeny was studied in e.g. *Hypochaeris* (Cerbah et al., 1999; Cerbah et al., 2001) and *Lathyrus* (Mohammed Ali, Meister, and Schubert, 2000), but did not seem to have much evolutionary significance in these genera.

DNA content and base composition in Asteraceae have been studied in a range of genera, for example in Microseridinae (Price and Bachmann, 1975), in *Helianthus* (Sims and Price, 1985), *Hypochaeris* (Cerbah et al. 1995, 1999), *Crepis* (Dimitrova and Greilhuber, 2000), and *Artemisia* (Torrell and Vallès, 2001). Similar to the situation in other angiosperm families, genera of Asteraceae show evolution towards increase in DNA content (Price and Bachmann, 1975; Cerbah et al., 1995; Cerbah et al., 1999; Torrell and Vallès, 2001) as well as towards decrease in DNA content (Price and Bachmann, 1975; Cerbah et al., 1999; Torrell and Vallès, 2001). Evolutionary advancement in for example *Crepis* was not significantly correlated to DNA content (Dimitrova and Greilhuber, 2000). The only elaborate study on evolution of base composition in Asteraceae showed no clear evolutionary pattern (Cerbah et al., 1999).

Among the numerous studies on evolution of DNA content and base composition, only a few use actual species phylogenies as a reference. In most studies, DNA contents and base compositions are evaluated with reference to a more general framework of knowledge about species relationships, which may not always be reliable. The evolution of DNA content and base composition can be studied more reliably and in more detail using, for example, molecular phylogenies. However, the number of such studies is still very limited.

In Asteraceae, Cerbah et al. (1999) examined the evolution of DNA content and base composition in *Hypochaeris,* using and ITS (Internal Transcribed Spacer) phylogeny for comparison. However, the DNA content and base composition data were not analyzed cladistically. In the present paper, we are the first to perform cladistic analyses of DNA content

and base composition in Asteraceae, using a phylogeny based on combined ITS-1 and AFLP data as a reference. The analyses are performed using DNA amount- and base composition data of 23 species from *Lactuca* and related genera, comprising first estimates for most species.

Our study focuses on five topics: 1) the evolution of DNA content and base composition, testing the hypothesis that DNA amounts increased during the evolution of *Lactuca* and related genera; 2) the correlation of DNA content and base composition as proposed by Vinogradov (1994) and disputed by Meister and Barow (2001); 3) the intraspecific variation in DNA content, testing the hypothesis of DNA constancy (discussed in for example in Bennett and Leitch (1995) and Ohri (1998)) for *Lactuca* and related genera; 4) the interspecific variation in DNA content and base composition, evaluating their use as taxonomic characters; 5) the "almost linear" correlation between genome size and number of AFLP bands assumed by Vos et al. (1995).

Materials and Methods

Plant material

We used 130 accessions from 23 species, including 11 European *Lactuca* species, one *Lactuca* species from the Middle-East (*L. aculeata*), one from South Africa (*L. dregeana*), one from Asia (*L. indica*), five species from related genera within Lactuceae subtribe Lactucinae and four species outside the subtribe. All species are diploids with 2n = 18, except *C. juncea* (2n = 15), *L. tenerrima*, *Cicerbita plumieri*, and *Steptorhamphus tuberosus* (2n = 16), and *Taraxacum officinale* (2n = 24; triploids were removed from the analysis) (Feráková, 1977; Doleжalová et al., In press; Index to plant chromosome numbers at http://mobot.mobot.org/W3T/Search/ipcn.html). Details on the accessions are given in Tables 1 and 2 (all accessions), and on the website of the Centre for Genetic Resources, The Netherlands (CGN) at http://www.plant.wageningen-ur.nl/CGN (CGN accessions only). Voucher specimens of all plant material at rosette, bolting, and flowering stage were deposited at the Herbarium Vadense (WAG), supplemented with pappus preparations and fruit samples. Voucher specimens for accessions in Table 1 were also supplemented with photographs of the plants at all three growth stages. In the present paper, we refer to the genera *Lactuca*, *Mycelis*, *Steptorhamphus*, and *Cicerbita* as *Lactuca* s.l. (sensu lato)(see Koopman et al., 1998).

Opposite page: Table 1. Accessions used to study DNA content and base composition in 23 species of *Lactuca* and related genera. Relative DNA contents were determined for all accessions listed, absolute DNA contents and base composition were determined only for accessions in bold type. Subtribal classification according to Bremer (1994), generic and specific classification of European species according to Tutin et al. (1976)/Feráková (1977). *L. aculeata* from the MiddleEast and *L. dregeana* from South Africa were included in subsect. *Lactuca* because of their close relationship to *L. serriola* (Zohary, 1991), although they were not treated by Tutin/Feráková. The Asiatic species *L. indica* is classified in the non-European section *Tuberosae*, according to Iwatsuki et al. (1995).

Subtribe Lactucinae Dumort., genus *Lactuca* L.

Lactuca sect. *Lactuca* subsect. *Lactuca*

   Lactuca sativa L.: **CGN 5979** ('Balady', landrace, Bani Swif, Egypt); CGN 4884 ('Verte de Cobham', A.L. Tozer Ltd., Cobham, United Kingdom); CGN 5140 ('Capitan', Les Graines Caillard, Angers, France); CGN 5999 (Gradina Botanica a Universitatii din Cluy-Napoca, Cluy-Napoca, Romania); CGN 5045 (Landrace, Prof. B.M. Kozopolansky Botanical Garden of the Voronezh State University, Voronezh, Rossijskaja, former USSR); CGN 11387 ('Tianjin Big Stem', Institute of vegetables and flowers, Chinese Academy of Agricultural Science, Beijing, China); CGN 13391 ('Sudia', Vilmorin-Andrieux, Beaufort-En-Vallee, France); CGN 9356 (Oilseed landrace, Qena, 1km South of Isna, Egypt); RKO 93130 ('Balisto', Rijk Zwaan B.V., De Lier, The Netherlands); RKO 92296 ('Karif', Rijk Zwaan B.V., De Lier, The Netherlands).

   Lactuca serriola L.: **CGN 10881** (Oudewater, The Netherlands); CGN 5900 (Jerusalem, Israel); CGN 4674 (Botanischer Garten de Universität Goettingen, Germany); CGN 5803 (Botanical Garden, Department of Botany, University of Oulu, Oulu, Finland); CGN 14314 (former USSR); CGN 4667 (Botanische tuinen van de gemeente Rotterdam, Rotterdam, The Netherlands); CGN 15671 (Oshakan, Ashtarakskij, Armenia); CGN 15684 (Antakya, Hatay, Turkey); CGN 4774 (Kena, Egypt; oilseed landrace); CGN 4769 (Upper Egypt, Egypt; oilseed landrace).

   Lactuca dregeana DC.: **CGN 4790** (Giardino Botanico e Coloniale dell' Universita di Palermo, Palermo, Italy); CGN 5805 (Jardin Botanique de la Ville, Dijon Cedex, France).

   Lactuca altaïca Fisch. et C.A. Mey.: **CGN 15711** (Ozero, Dzhandargel, Rustavi, Georgia); CGN 4664 (Botanicka Zahrada University Karlovy, Prague, Czechoslovakia).

   Lactuca aculeata Boiss. & Kotschy ex Boiss.: **CGN 9357** (Nov, Israel); CGN 15692 (Kiziloren, Afyon, Turkey).

   Lactuca saligna L.: **CGN 5310** (Raananna, Israel); CGN 5327 (Caca de la Selia, Gerona, Spain); CGN 5301 (Lot, France); CGN 4662 (Botanische tuinen van de gemeente Rotterdam, Rotterdam, The Netherlands); CGN 15705 (Avchala-Tiflis, Tiflis, Georgia); CGN 15697 (Akcay, Balikesir, Turkey); CGN 13371(Sofia, 1 km East of Rozen, Bulgaria); CGN 5147 (Marche, Pesaro, Italy); CGN 17444 (Fobbing, South Essex, United Kingdom); CGN 15716 (Daghestan, Tsahlinskij, Gudermes,10 km West, Union of Soviet Soc. Rep.).

   Lactuca virosa L.: **CGN 9315** (origin unknown); CGN 4682 (Hortus Botanicus Universitatis Varsaviensis, Warsaw, Poland); CGN 4970 (Jardin Botanique de la Ville, Dijon Cedex, France); CGN 4681 (Hortus Botanicus der Universiteit van Amsterdam, Amsterdam, The Netherlands); CGN 13349 (Asturia, Spain); CGN 13350 (Asturia, Spain); CGN 13352 (Asturia, Spain); CGN 5941 (Gadot, Israel); CGN 13339 (Spain); CGN 15679 (Rushul, Tabasaran, Daghestan); CGN 15680 (Trisandij, Urkarach, Daghestan).

Lactuca sect. *Lactuca* subsect. *Cyanicae* DC.

   Lactuca tenerrima Pourr.: **CGN 13351** (Asturia, Spain); CGN 9387 (Timhadit, Morocco); CGN 9388 (Collscrola, Barcelona, Spain); CGN 9386 (L'Arralassada, Barcelona, Spain); CGN 14217 (Hortus Universitatis Hauniensis, Botanical Garden of the University of Copenhagen, Copenhagen, Denmark).

   Lactuca perennis L.: **CGN 9318** (Jardin Botanique de la Ville, Dijon Cedex, France); CGN 13299 (between Vallouise and Puy-Aillaud, Hautes Alpes, France); CGN 9323 (Valais, Switzerland); CGN 10884 (Gradina Botanica a Universitatii din Cluy-Napoca, Cluy-Napoca, Romania); CGN 13294 (Botanischer Garten der Universität (TH), Karlsruhe 1, Germany).

Lactuca sect. *Mulgedium* (Cass.) C. B. Clarke

Lactuca tatarica (L.) C.A. Mey.: **CGN 9389** (Central Asia); CGN 9390 (Volgograd, former USSR); CGN 930133 (Delta Danube, Tulcea, Maliuc, Romania); CGN 910430 (Insel Usedom, Bansin, Meckelenburg-Vorpommern, Germany); CGN 930119 (Botanischer Garten der Wilhelm Pieck Universität, Rostock, Germany); W9530 (Finchuan, Ningxia, China).

Lactuca sibirica (L.) Benth. ex Maxim.: W9513 (River Kulbacksan, Sweden); W9516 (Vanjaurtrask, River Soran, Sweden): W9517 (Knaften, Umeriver, Sweden); **W9520** (Lagneset, Oreriver, Sweden); W9523 (Vastana/Mjallan, River Selangersan, Sweden).

Lactuca sect. *Lactucopsis* (Schultz-Bip ex Vis. et Panc.) Rouy

Lactuca quercina L.: **CGN 14220** (Esrergom, Komarom, Hungaria).

Lactuca sect. *Phaenixopus* (Cass.) Benth.

Lactuca viminea (L.) J. & C. Presl: **CGN 9326** (Jardim Botanico da Universidade de Coimbra, Coimbra, Portugal); CGN 14301 (Roumodour, France); CGN 16202 (Asktarakskii, Armenia); CGN 926859 (France); CGN 926870 (Italy).

Lactuca sect. *Tuberosae* Boiss.

Lactuca indica L.: **CGN 13392** (Hortus Botanicus Pekinensis, Beijing, China); CGN 13393 (Landrace, Hortus Botanicus Pekinensis, Beijing, China); CGN 14312 ('Sanelin Lampenas', landrace, Cipanas, Cianjur, Indonesia); CGN 14316 ('Wo Yang Tai Gan', landrace, Anhui, China); CGN 20713/W9542 (near Deng Ling He, South of Simao, Yunnan, China).

Subtribe Lactucinae, other genera

Mycelis muralis (L.) Dumort.: **CGN 9366** (Passo Muratone, Italy); CGN 9367 (Botanical Garden Floretum Scanicum, Helsingborg, Sweden); CGN 5005 (Botanical Garden of the University of Uppsala, Uppsala, Sweden); CGN 9327 (St. Pietro d'Olba, Savona, Italy); CGN 5116 (Botanical Garden of the Armenian Academy of Sciences, Yerevan, Armenia).

Steptorhamphus tuberosus (Jacq.) Grossh.: **CGN 9368** (Botanical Garden of the Tel-Aviv University, Tel-Aviv, Israel)

Cicerbita plumieri (L.) Kirschl.: **CGN 19090/W9501** (Ht. Folin, Nievre, France); W9531 (Vallée de Galbe, Dept. Pyrénées Orientales, France); W9532 (Route Formiguères-Quérigut, Dept. Pyrénées Orientales, France).

Cicerbita alpina (L.) Wallr.: **W9507** (Botanischer Garten der Friedrich Schiller Universität, Jena, Germany); W9508 (Botanischer Garten der Martin-Luther Universität, Halle-Saale, Germany); W9541 (Institut für Allgemeine Botanik und Botanischer Garten, Universität Hamburg, Germany; collected in Klöntal, Kanton Glarus, Switzerland).

Prenanthes purpurea L.: W9504 (Botanischer Garten der Universität Tübingen, Germany); W9505 (Botanischer Garten der Universität Frankfurk/Main, Germany); W9524 (Conservatoire et Jardin Botanique de la Ville de Genève, Chambesy-Genève, Switzerland; collected in Ain); W9525 (Conservatoire et Jardin Botanique de la Ville de Genève, Chambesy-Genève, Switzerland; collected in Valais); **W9534** (KM8 Route Formiguères-Quérigut, Dept. Pyrénées Orientales, France).

Subtribe Crepidinae Dumort.

Chondrilla juncea L.: **CGN 9391** (Instituto Botanico dell' Universita, Siena, Italy); CGN 13308 (Botanischer Garten der Freidrich Schiller Universität, Jena, Germany); CGN 14218 (Bok Ata, Israel).

Taraxacum officinale Weber in F.H. Wigg.: **W9606** (Wageningen, The Netherlands); W9707 (Gerendal, Limburg, The Netherlands); Fr9 (Col de la Croix de Marchampt 670 m, Beaujeu, France).

Subtribe Sonchinae K. Bremer

Sonchus asper (L.) Hill: **W9510** (Botanischer Garten der Martin-Luther Universität, Halle-Saale, Germany); W9511 (Wageningen, The Netherlands); W9526 (Conservatoire et Jardin Botanique de la Ville de Genève, Chambesy-Genève, Switzerland; wild origin); W9539 (Coimbra, Portugal).

Unassigned to a subtribe

Cichorium intybus L.: **W9601** (Renkum, The Netherlands); W9602 (Botanischer Garten der Universität Dresden, Germany; wild origin); W9603 ('Hollandse middel vroeg', Oranjeband zaden/Nunhems, Dronten, The Netherlands).

Table 2. *Lactuca sativa* cultivars used for study of intraspecific variation. Cultivar groups Butterhead, Crisphead, Cos, and Cutting are according to Rodenburg (1960). We subdivided the Butterhead group into categories indoor and outdoor to reflect the distinct breeding backgrounds of these categories. Breeding companies in brackets. ** indicate cultivars that were developed by Leen de Mos B.V., prior to its merger with Nunhems Zaden B.V.

Butterhead indoor
'Cortina' (Nunhems Zaden B.V.**); 'Flandria' (Rijk Zwaan B.V.); 'Rachel' (Enza Zaden B.V.); 'Susy' (Novartis Seeds B.V.)
Butterhead outdoor
'Milan' (Nunhems Zaden B.V.); 'Rexado' (Nunhems Zaden B.V.**); 'Sander' (Seminis Vegetable Seeds); 'Sunny' (Nickerson-Zwaan B.V.); 'Titan' (Novartis Seeds B.V.)
Cos
'Fredo' (Nunhems Zaden B.V.**); 'Hector' (Enza Zaden B.V.); 'Odessa' (Novartis Seeds B.V.); 'Odra' (Nunhems Zaden B.V.); 'Toledo' (Seminis Vegetable Seeds)
Crisphead
'Calgary' (Seminis Vegetable Seeds); 'Robinson' (Nickerson-Zwaan B.V.); 'Roxette' (Rijk Zwaan B.V.); 'Saladin' (Nunhems Zaden B.V.)
Cutting
'Cordoba' (Enza Zaden B.V.); 'Frisby' (Nickerson-Zwaan B.V.); 'Krizet' (Rijk Zwaan B.V.)

Determination of absolute DNA content and base composition

Seeds were sown in petri dishes on tap water, kept at 7°C for three days to break dormancy, and subsequently put at room temperature until germinated. For *C. alpina* and *P. purpurea*, cold treatment at 7°C was replaced by soaking the seeds in a 500 ppm GA solution (Gibberellic Acid; Sigma, St. Louis, Missouri, USA) for three days at room temperature. Germinated seeds were transferred to pots with peaty soil, grown under standard greenhouse conditions for several weeks, and subsequently reared in open-air.

Samples were prepared according to Galbraith et al. (1998). Small amounts of leaf tissue were chopped with a razor blade in 1 mL of Otto I buffer in a glass Petri dish. Crude suspension was filtered through 50-μm mesh nylon filters and pelleted at 300g for 10 min. The supernatant was discarded, and the nuclei were resuspended in 150μl Otto I buffer. The samples were kept at room temperature for at least 20 min. Prior to analysis, 0.6 ml Otto II

buffer containing either 2μg/ml 4',6-diamidino-phenylindole (DAPI) or 50μg/ml propidium iodide (PI) and 50μg/ml RNase were added to the sample. The nuclei were analyzed with a Partec PAS II flow cytometer (Partec GmbH, Münster, Germany), equipped with a HBO100 mercury arc lamp and a 40x/1.25 glycerine immersion objective. In most cases we analyzed four plants per accession. To avoid bias due to random instrument drift, each plant was measured four times on different days. 2C nuclear DNA content (in pg) was calculated according to the formula:

2C DNA content of accession = (2C DNA content of standard) x ((G_0/G_1 peak mean of specimen) / (G_0/G_1 peak mean of referencestandard)). The AT content of a specimen (in %) was calculated according to the formula proposed by Godelle et al. (1993): AT% of specimen = AT% of reference standard x $(R_{DAPI}/R_{PI})^{1/n}$, where R = peak mean of specimen / peak mean of reference standard, and n is the number of contiguous AT pairs needed as binding site for DAPI. n was assumed equal to 4 (Portugal and Waring, 1988; Wilson et al., 1989).

Because of the large variation in DNA content among individual accessions, three internal reference standards were used: 1) *Glycine max* 'Polanka' was used for *L. viminea*, *L. tenerrima*, *M. muralis*, *C. intybus*, *C. juncea*, *T. officinale*, and *S. asper*; 2) *L. sativa* CGN 5979 was used for *L. quercina*, *L. virosa* CGN 15679, *L. sibirica*, *L. tatarica*, and *P. purpurea*; and 3) *Pisum sativum* 'Ctirad' was used for the remaining accessions (see Table 1). 2C DNA content and base composition of *Pisum sativum* 'Ctirad' was determined previously (2C = 9.07 pg, AT = 61.42%, Dolezel, Sgorbati, and Lucretti, 1992), that of *L. sativa* CGN 5979 (2C = 6.044 pg, AT = 61.722%) and *Glycine max* 'Polanka' (2C = 2.349 pg, AT = 64.049%) was determined in the present study using 'Ctirad' as a primary reference.

Determination of relative DNA content

Plants were grown as above, but kept in the greenhouse during the entire experiment. Young rosette leafs (four plants per accession) were collected for DNA analysis. Relative DNA amounts were determined by Plant Cytometry Services (Schijndel, The Netherlands) on a PAS II flow cytometer (Partec GmbH, Münster, Germany) as described in Koopman (2000). *Lycopersicon esculentum* Mill. 'Tiny Tim' was used as internal reference.

Statistics

All ANOVAs and pairwise comparisons were performed in JMP version 3.1.4 (SAS Institute Inc., Cary, North Carolina, USA). For the pairwise comparisons, we used Tukey-Kramer HSD multiple comparison procedures with α = 0.05. All correlations and non-linear associations were calculated and tested on significance (α = 0.05) in Excel 97 (Microsoft Corporation, Redmond, Washington, USA).

Evolution of absolute DNA content and base composition

To examine the evolution of DNA content and base composition in *Lactuca* s.l., we determined absolute DNA content and base composition for one accession per species (boldface accessions in Table 1). As an exception, *L. virosa* was represented by two accessions (CGN 9315 and CGN 15679), because a previous study indicated two distinct levels of DNA content within this species (Koopman, 2000). Differences among accessions were tested using an ANOVA with fixed effect of species, and random effect of plants within species.

As a basis for tracing evolution of DNA content and base composition, we calculated most parsimonious trees (MPTs) including all boldface accessions of Table 1, using combined ITS-1 and AFLP data. ITS data (267 bp of aligned sequence) and AFLP data (1030 markers) were taken from previous studies (Koopman et al., 1998; Koopman, Zevenbergen, and Van den Berg, 2001). *S. asper*, *C. juncea*, and *T. officinale* were used as outgroup. For the outgroup species and for *P. purpurea* W 9534, no AFLP data were available. Instead, we used missing values for the outgroup species, and data from *P. purpurea* W9525 for *P. purpurea* W9534. W9525 is a valid replacement for W9534, because both accessions have identical ITS-1 sequences (Koopman et al., 1998), and their recalculated absolute DNA contents (see below) are not significantly different. MPTs were calculated in PAUP* 4.0b8 (PPC/Altivec) (Swofford, 1999) using a heuristic search with 10 000 random addition sequences, TBR (tree bisection-reconnection) branch swapping, and "multrees" switched on. Parsimony settings were: acctran and "collapse of zero-length branches". Jackknife support was calculated in a fast-heuristic search with 10000 replicates, nominal deletion of 37% of the characters in each replicate, and "Jac" resampling. Starting trees were obtained using random addition sequences without branch swapping.

The evolution of DNA content and base composition was traced on the MPT using MacClade 4 (Maddison and Maddison, 2000). We reconstructed ancestral character states using linear (Wagner) parsimony (Farris, 1970; Swofford and Maddison, 1987) with the MINSTATE option in effect. Using MINSTATE, only lowest values are reported in case of multiple equally parsimonious reconstructions of ancestral character states.

Association of DNA content and base composition

The association of DNA content and AT% was examined using a series of regression analyses. We calculated and tested linear fits (i.e. correlations) and polynomial fits (i.e. associations) up to the third degree for three sets of distances 1: for all species in the present study; 2: for ingroup species only (i.e. excluding *S. asper*, *C. juncea*, and *T. officinale*); 3: for *Lactuca* species only. The associations were calculated using the original data on individual plants (349 data pairs). For some accessions, DNA content and AT% were not measured on the same plant, and data for these plants (all 4 plants of *L. aculeata*, and 1 plant of *L. tatarica*, *L. quercina*, and *L. indica*, respectively) were excluded from the calculations. Additionally, we determined the associations for a data set comprising only species averages.

Intra- and interspecific DNA content variation in Lactuca s.l.

To study intra- and interspecific DNA content variation in more detail, we determined relative DNA contents for 109 accessions from 23 species of *Lactuca* and related genera (all accessions in Table 1). Subsequently, the relative DNA contents were recalculated into absolute DNA contents. Absolute DNA contents determined in the previous experiment (one accession for each species, two for *L. virosa*) were used as absolute DNA content references. Relative DNA contents were recalculated into absolute DNA contents as: $Abs_{spec} = (Rel_{spec}/Rel_{specref})*Abs_{specref}$; with Abs_{spec} = absolute DNA content of a given accession of a species, Rel_{spec} = relative DNA content of that accession, $Rel_{specref}$ = relative DNA content of the reference accession for that species, and $Abs_{specref}$ = absolute DNA content of the reference accession for the species. A cluster analysis was performed on the mean recalculated DNA content per accession, using NTSYSpc 2.02k (Applied Biostatistics, Setauket, New York, USA). We employed UPGMA clustering based on Manhattan distances. Differences among all accessions were tested using an ANOVA with fixed effect of species, and random effect of accessions within species. The ANOVA was followed by pairwise comparisons among the accessions.

Association of DNA content and number of AFLP bands

In the present study, we determined recalculated absolute DNA amounts for 109 accessions of 23 species. In a previous study (Koopman, Zevenbergen, and Van den Berg, 2001), AFLP patterns were determined for 91 of these accessions. We used the AFLP data to examine the association between the DNA amounts and the numbers of AFLP bands generated for each accession. Eighteen accessions were excluded from the present analysis, because we had no AFLP data available for these accessions (*L. sativa* CGN 5045, *L. indica* W9542, *M. muralis* CGN 9367 and CGN 5005, *C. alpina* W9541, *P. purpurea* W. 9504, W9524, and W9534, and all accessions of the outgroup (*S. Asper, C. juncea*, and *T. officinale*)). The original data sets were generated with primer combination E35/M48 and E35/M49, and contained two plants per accession. To calculate the average number of AFLP bands per accession, we summed the numbers of bands for both primer combinations, and averaged the data of both plants. The association between DNA amounts and numbers of AFLP bands per accession was examined and tested for significance using linear fits and polynomial fits up to the fourth degree. We tested two sets of accessions: one set including all species, and one set excluding species with DNA amounts above 8.5 pg (see discussion).

Intraspecific variation in L. sativa

To examine intraspecific variation in cultivated lettuce, we determined relative DNA contents of 21 modern *L. sativa* cultivars listed in Table 2. The cultivars were subdivided into cultivar groups Butterhead, Crisphead, Cos, and Cutting, according to Rodenburg (1960). The Butterhead group was subdivided into Butterhead indoor and Butterhead outdoor to take account of the distinct breeding background of these two categories. Differences between

cultivars and cultivar groups were tested using an ANOVA with fixed effect of cultivar groups, and fixed effect of cultivars within cultivar groups. The ANOVA was followed by pairwise comparisons among the accessions.
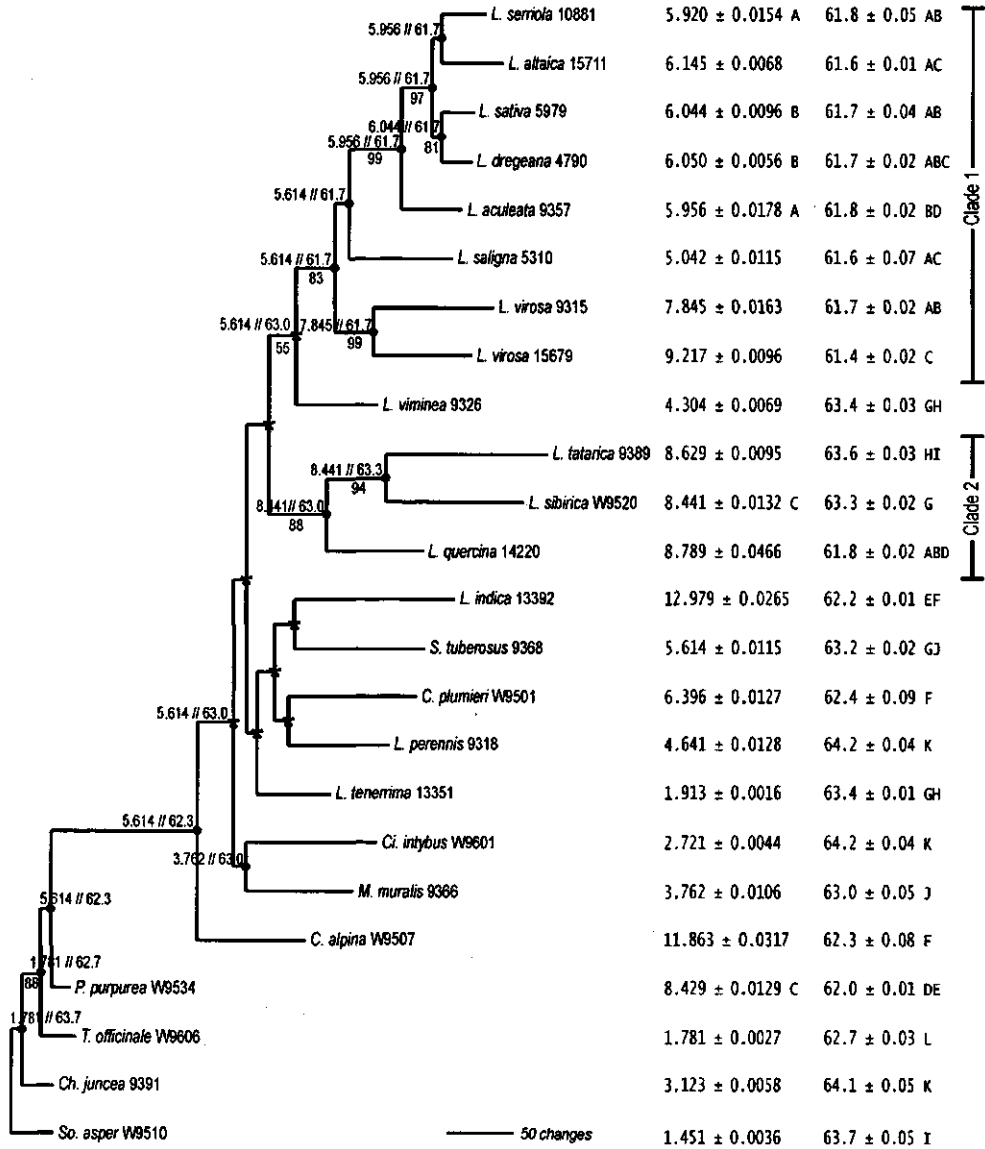
Results

Absolute DNA content and base composition

Absolute 2C DNA contents and AT% as measured for all species are depicted in Fig. 1. 2C DNA contents among *Lactuca* s.l. ranged from 1.913 pg in *L. tenerrima* to 12.979 pg in *L. indica*. Contents in the outgroup ranged from 1.451 pg in *S. asper* to 3.123 pg in *C. juncea*. The ANOVA indicated significant differences among the species ($p < 0.00005$), but no significant differences among plants within species ($p = 0.2594$). Significance of species differences was confirmed by the results of the Tukey HSD procedure, showing only three pairs of species that were not significantly different in their absolute 2C DNA content: *L. dregeana*/*L. sativa*, *L. serriola*/*L. aculeata*, and *L. sibirica*/*P. purpurea* (Fig 1, identical letters indicate species that are not significantly different).

Base compositions among *Lactuca* s.l. ranged from 61.4% AT in *L. virosa* to 64.2% AT in *L. perennis*. Base compositions in the outgroup ranged from 62.7% AT in *T. officinale* to 64.1% AT in *C. juncea*. The ANOVA indicated significant differences for species ($p < 0.00005$), and plants within species ($p = 0.0013$). The Tukey HSD procedure showed four significantly different groups of species, while within the groups many of the species did not differ significantly in base composition. The first and largest group (A-F in Fig.1) shows the lowest AT contents, and comprises all subsect. *Lactuca* species (*L. sativa*, *L. serriola*, *L. dregeana*, *L. altaica*, *L. aculeata*, *L. saligna*, *L. virosa*), supplemented with *L. quercina*, *L. indica*, *C. plumieri*, *C. alpina*, and *P. purpurea*. Within this group, the subsect. *Lactuca* species generally have lower AT contents than the species outside the subsection. The second group (G-J) comprises *Lactuca* species not included in the first group (*L. tatarica*, *L. sibirica*, *L. viminea*, *L. tenerrima*), two other species from subtribe Lactucinae (*M. muralis*, *S. tuberosus*), and *S. asper* (subtribe Sonchinae). The third group (K) has the highest AT content, and comprises *L. perennis* and two species outside subtribe Lactucinae, viz. *C. juncea* and *C. intybus*. The AT content of *T. officinale* (L) differs significantly from that of all other species in the present study, although it is well within the AT content range of *Lactuca* s.l.

Evolution of absolute DNA content and base composition

Cladistic analysis of a combined AFLP/ITS data set yielded one MPT of 2198 steps, a consistency index (CI) of 0.50, a retention index (RI) of 0.38, and a rescaled consistency index (RC) of 0.19. The MPT topology was used as a reference to reconstruct the evolution of DNA content and base composition. Fig. 1 depicts the MPT, with ancestral character states for DNA content and AT% reconstructed on the internal nodes. DNA content and AT% of the individual

L. serriola 10881	5.920 ± 0.0154 A	61.8 ± 0.05 AB
L. altaica 15711	6.145 ± 0.0068	61.6 ± 0.01 AC
L. sativa 5979	6.044 ± 0.0096 B	61.7 ± 0.04 AB
L. dregeana 4790	6.050 ± 0.0056 B	61.7 ± 0.02 ABC
L. aculeata 9357	5.956 ± 0.0178 A	61.8 ± 0.02 BD
L. saligna 5310	5.042 ± 0.0115	61.6 ± 0.07 AC
L. virosa 9315	7.845 ± 0.0163	61.7 ± 0.02 AB
L. virosa 15679	9.217 ± 0.0096	61.4 ± 0.02 C
L. viminea 9326	4.304 ± 0.0069	63.4 ± 0.03 GH
L. tatarica 9389	8.629 ± 0.0095	63.6 ± 0.03 HI
L. sibirica W9520	8.441 ± 0.0132 C	63.3 ± 0.02 G
L. quercina 14220	8.789 ± 0.0466	61.8 ± 0.02 ABD
L. indica 13392	12.979 ± 0.0265	62.2 ± 0.01 EF
S. tuberosus 9368	5.614 ± 0.0115	63.2 ± 0.02 GJ
C. plumieri W9501	6.396 ± 0.0127	62.4 ± 0.09 F
L. perennis 9318	4.641 ± 0.0128	64.2 ± 0.04 K
L. tenerrima 13351	1.913 ± 0.0016	63.4 ± 0.01 GH
Ci. intybus W9601	2.721 ± 0.0044	64.2 ± 0.04 K
M. muralis 9366	3.762 ± 0.0106	63.0 ± 0.05 J
C. alpina W9507	11.863 ± 0.0317	62.3 ± 0.08 F
P. purpurea W9534	8.429 ± 0.0129 C	62.0 ± 0.01 DE
T. officinale W9606	1.781 ± 0.0027	62.7 ± 0.03 L
Ch. juncea 9391	3.123 ± 0.0058	64.1 ± 0.05 K
So. asper W9510	1.451 ± 0.0036	63.7 ± 0.05 I

50 changes

Fig. 1. Most parsimonious tree (MPT) of 23 species from *Lactuca* and related genera, based on combined ITS-1 sequence data and AFLP markers. Above nodes: reconstructed ancestral character states for 2C DNA content (left of slashes) and AT% (right of slashes). Below nodes: jackknife values. Nodes indicated with * all have identical reconstructed character states, viz. 5.614 pg 2C DNA content, and 63.0% AT. Genus abbreviations: *Ch.* = *Chondrilla*, *C* = *Cicerbita*, *Ci.* = *Cichorium*, *L* = *Lactuca*, *M* = *Muralis*, *P* = *Prenanthes*, *So* = *Sonchus*, *S* = *Steptorhamphus*, *T* = *Taraxacum*. DNA content and base composition values for the individual species are given behind the species names: first column: 2C DNA content (pg); second column: base composition (AT%). Identical letters indicate values that are not significantly different. Clade 1 and 2 indicate well supported groups in the MPT (jackknife values > 80%).

species are on the terminal branches. Branch supports are indicated by Jackknife values (only values above 50% are included). The MPT shows two well supported clades (jackknife values > 80%): one with all subsect. *Lactuca* species (Clade 1), and one with *L. tatarica, L. sibirica*, and *L. quercina* (Clade 2). The position of *P. purpurea* outside *Lactuca* s.l. is in agreement with morphology-based classifications (e.g. Stebbins, 1937).

Considering the evolution of DNA content, the outgroup (on average 2C = 2.118 pg DNA) shows low values relative to the ingroup. In contrast, *P. purpurea* shows a relatively high DNA content (2C = 8.429 pg). The high DNA content of *P. purpurea* is reflected in the reconstructed ancestral 2C DNA amount of 5.614 pg for the nodes basal to *Lactuca* clades 1 and 2. In both clades, the 2C DNA amount shows an increase relative to the 5.614 pg in the basal nodes, but the increase in clade 2 (*L. tatarica, L. sibirica, L. quercina*, average 2C = 8.620 pg) is considerably larger than in clade 1 (subsect. *Lactuca*, average 2C = 6.527 pg). Notwithstanding the general trend towards increasing DNA amounts, the terminal branches of the MPT show both increases and decreases. The variation in DNA amounts is apparent among species within subsect. *Lactuca*, but most striking among the *Lactuca* s.l. species outside clades 1 and 2. For example, *L. indica* and *L. tenerrima* show a 6.8 fold difference in DNA content. Unfortunately, the unreliable topology in this part of the MPT (jackknife values < 50%) precluded a reliable reconstruction of ancestral character states.

Considering the evolution of base composition, it is apparent that the outgroup species have relatively high (*T. officinale*, 62.7%; *S. asper*, 63.7%; *C. juncea*, 64.1%) AT contents. These high values for the outgroup give rise to a relatively high reconstructed ancestral state (62.3% AT) for the basal node of the *Lactuca* s.l. clade. Within *Lactuca* s.l., base compositions evolve in different directions: clade 2 shows an increase in AT content for *L. tatarica* and *L. sibirica* (average 63.5%), while *L. quercina* and subsect. *Lactuca* (clade 1) show a decrease (average 61.7% AT in clade 1). Subsect. *Lactuca* shows a remarkable constancy in base composition: among the eight species, only two (*L. aculeata* and *L. virosa*) show significant differences in AT content, but only with part of the other species.

Association of DNA content and base composition

All subdivisions of the data set showed a highly significant negative association of DNA content with AT%. In all cases, a second degree polynomial showed a significant increase in fit relative to a linear association, but the fit did not increase significantly for higher degree polynomials. Therefore, we consider the association between DNA content and AT% best described by a second degree polynomial. For the data set including all species, the R squared value was $R^2 = 0.25$ for the linear association (correlation coefficient -0.50, p = $4.10*10^{-23}$), and $R^2 = 0.33$ for the second degree polynomial fit (p = $3.97*10^{-10}$). The second degree polynomial is depicted in Fig. 2. For the subset of data including only ingroup species, the R squared value was $R^2 = 0.18$ for the linear association (correlation coefficient -0.42, p = $2.57*10^{-14}$), and $R^2 = 0.31$ for the second degree polynomial fit (p = $5.75*10^{-13}$). For the subset
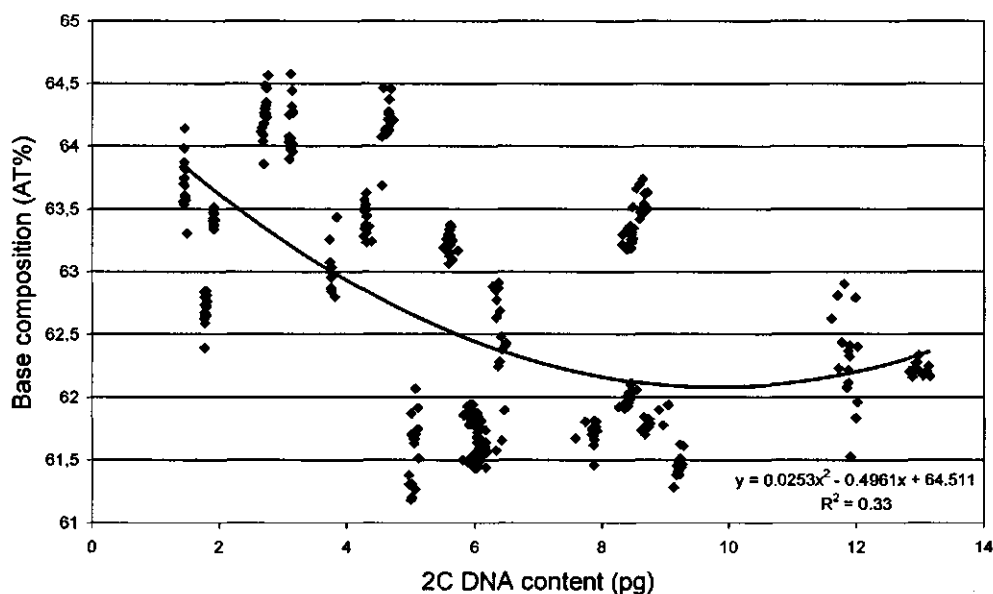
Fig. 2. Association of 2C DNA content (pg) and base composition (AT%). ◆ indicate data points for individual plants.

of data including only *Lactuca* species, the R squared value was $R^2 = 0.09$ for the linear association (correlation coefficient -0.30, $p = 8.92*10^{-6}$), and $R^2 = 0.19$ for the second degree polynomial fit ($p = 1.69*10^{-6}$). For the data set with species averages (values from Fig. 1), only the linear association was significant ($R^2 = 0.23$, correlation coefficient -0.48, $p = 0.02$). The inverse relation between DNA content and AT% is also apparent in *L. virosa*, the only species for which we determined intraspecific variation in AT content. Relative to *L. virosa* CGN 9315 (2C = 7.845 pg DNA, 61.7% AT), *L. virosa* 15679 (2C = 9.217pg DNA, 61.4% AT) showed a 17.5% increase in DNA amount, associated with a small (0.5%) but statistically significant decrease in AT content (Fig.1).

Intra- and interspecific DNA content variation in Lactuca s.l.

2C DNA contents among *Lactuca* s.l. and *P. purpurea* ranged from 1.913 pg in *L. tenerrima* to 13.068 pg in *L. indica* CGN 14316, while the Lactuceae outgroup species showed 2C DNA contents between 1.434 pg (*S. asper* W9539) and 3.153 pg (*C. juncea* CGN 14218). The ANOVA for all accessions from Table 1 showed a significant effect of species ($p < 0.00005$), and accessions within species ($p < 0.00005$). The phenogram with all accessions is depicted in Fig. 3. Recalculated absolute DNA contents, and results of pairwise comparisons among the accessions are given behind the accession numbers. Identical letters indicate DNA contents that are not significantly different.
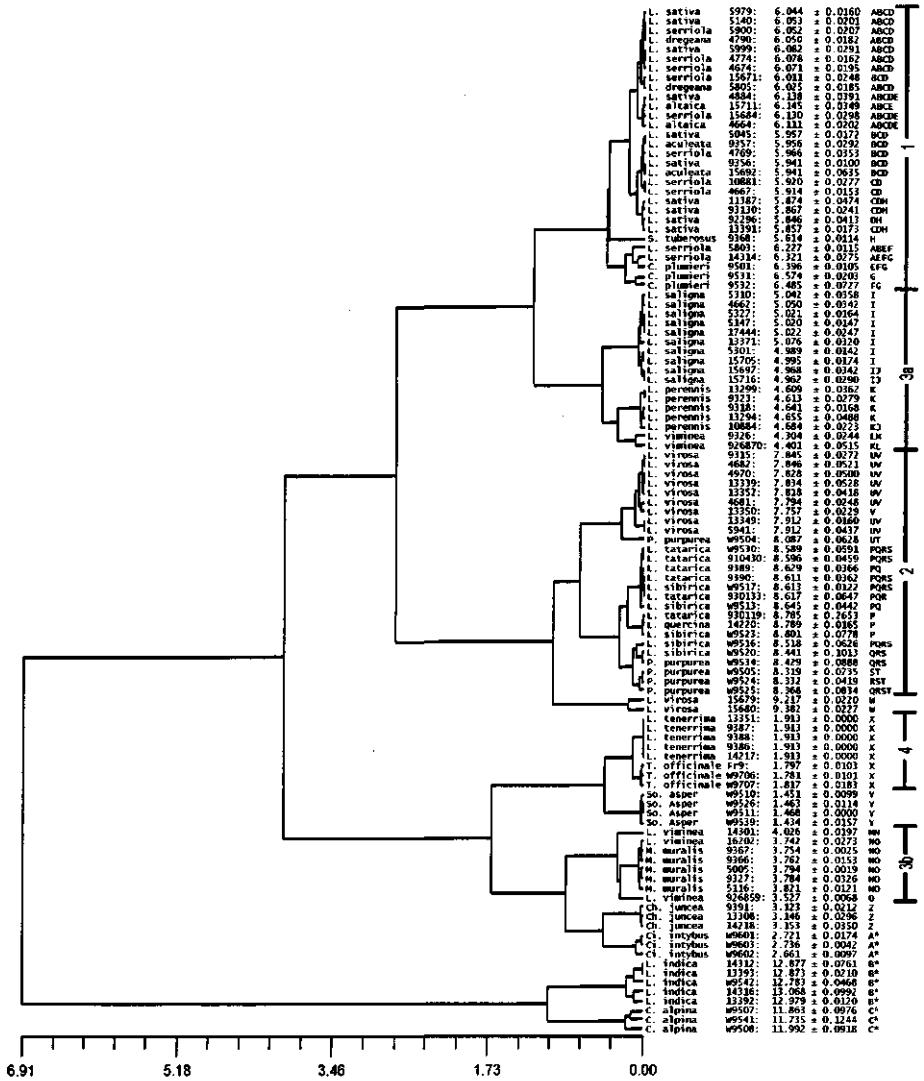
Fig. 3. UPGMA phenogram, based on Manhattan distances among the mean recalculated absolute DNA amounts, for 109 accessions of *Lactuca* and related genera. Genus abbreviations: *Ch.* = *Chondrilla, C = Cicerbita, Ci.* = *Cichorium, L = Lactuca, M = Muralis, P = Prenanthes, So = Sonchus, S = Steptorhamphus, T = Taraxacum*. Recalculated absolute 2C DNA contents are given behind accession names. Identical letters indicate DNA contents that are not significantly different.

111

Among the 23 species studied, only five showed significant intraspecific variation. Intraspecific variation is largest within *L. viminea*, where six out of 10 pairwise comparisons among the accessions were significantly different. DNA contents within *L. viminea* differ up to 25% between accessions CGN 926870 (2C = 4.401 pg) and CGN 926859 (2C = 3.527 pg). *L. virosa* showed two groups of accessions with significantly different DNA contents. The DNA contents of a small group with CGN 15679 and CGN 15680 (average 2C = 9.300 pg) are significantly larger than the DNA contents of the remaining *L. virosa* accessions (average 2C = 7.838 pg). The maximum difference in DNA content within *L. virosa* was 21% between CGN 13350 (2C = 7.757 pg) and CGN 15680 (2C = 9.382 pg). The accessions within both *L. virosa* groups are not significantly different. For *L. serriola*, 6 out of 45 pairwise comparisons were significantly different. All of these differences involve two accessions with relatively high DNA contents: CGN 14314 and CGN 5803. In Fig. 3, *L. serriola* CGN 14314 and CGN5803 (average 2C = 6.274 pg DNA) are separated from the remaining *L. serriola* accessions (average 2C = 6.018 pg DNA) in the phenogram, suggesting two distinct DNA groups within *L. serriola*. However, the majority of pairwise differences between CGN 14314/CGN 5803 and the remaining accessions were not significant. Therefore, our results do not support a subdivision of *L. serriola* accessions, notwithstanding the separation in Fig. 3. The largest difference in DNA content within *L. serriola* was 7% between CGN 4667 (2C = 5.914 pg) and CGN 14314 (2C = 6.321 pg). Within *L. sibirica*, among 10 pairwise differences only that between W9520 (2C = 8.441 pg) and W9523 (2C = 8.801 pg) was significant (4% difference). Within *P. purpurea*, among 10 pairwise differences only that between W9504 (2C = 8.087 pg DNA) and W9534 (2C = 8.429 pg) was significant (4% difference).

Considering interspecific variation, four groups of species can be recognized with overlapping (i.e. not significantly different) DNA contents for at least one accession (Fig. 3). Group 1: *L. sativa, L. serriola, L. dregeana, L. altaica, L. aculeata, S. tuberosus, C. plumieri*. *S. tuberosus* only overlaps with *L. sativa, C. plumieri* only with *L. sativa, L. serriola*, and *L. altaica*. The species from subsect. *Lactuca* (*L. sativa, L. serriola, L. dregeana, L. altaica, L. aculeata*) show overlap for most of their accessions. Group 2: *L. tatarica, L. sibirica, L. quercina, P. purpurea*, and *L. virosa* (excluding CGN 15679/15680). *L. virosa* overlaps only with *P. purpurea* W9504, but differs from all other accessions within group 2. The remaining species show extensive overlap. Group 3: *L. saligna, L. viminea, L. perennis*, and *M. muralis*. *L. saligna* slightly overlaps with *L. perennis, L. perennis* slightly overlaps with *L. viminea*, and *L. viminea* shows extensive overlap with *M. muralis*. In the cluster analysis (Fig. 3), the DNA range in group 3 is divided in two subgroups (3a: *L. saligna, L. viminea, L. perennis*; 3b: *M. muralis, L. viminea*). The subgroups are not entirely separated in the pairwise comparison tests because *L. viminea* CGN 9326 (3a) and CGN 14301 (3b) are not significantly different. The last group (group 4) comprises *L. tenerrima* and *T. officinale*, that show overlap for all accessions. *L. virosa* CGN 15679/CGN 15680, and species not included in the above groups

showed unique DNA contents, i.e. DNA contents that were significantly different from those of all other species.

Association of DNA content and number of AFLP bands

For the data set including all species, the results showed a highly significant positive correlation of DNA content and number of AFLP bands ($R^2 = 0.56$, correlation coefficient 0.75, p = $1.02*10^{-17}$). However, the association between DNA content and number of AFLP bands significantly increased up to the third degree ($R^2 = 0.80$, p = $2.16*10^{-5}$), indicating that the relation is best described by a third degree polynomial. Fig. 4 shows that the association of DNA content and number of AFLP bands is more or less linear for accessions with relatively low DNA contents, but clearly deviates from linear for some of the accessions with a 2C DNA content above 8.5 pg. The clearly deviating accessions belong to three species: *L. tatarica* (average 2C = 8.638 pg DNA and 187.8 bands), *C. alpina* (2C = 11.928 pg DNA, 111.8 bands), and *L. indica* (2C = 12.949 pg DNA, 135.9 bands).
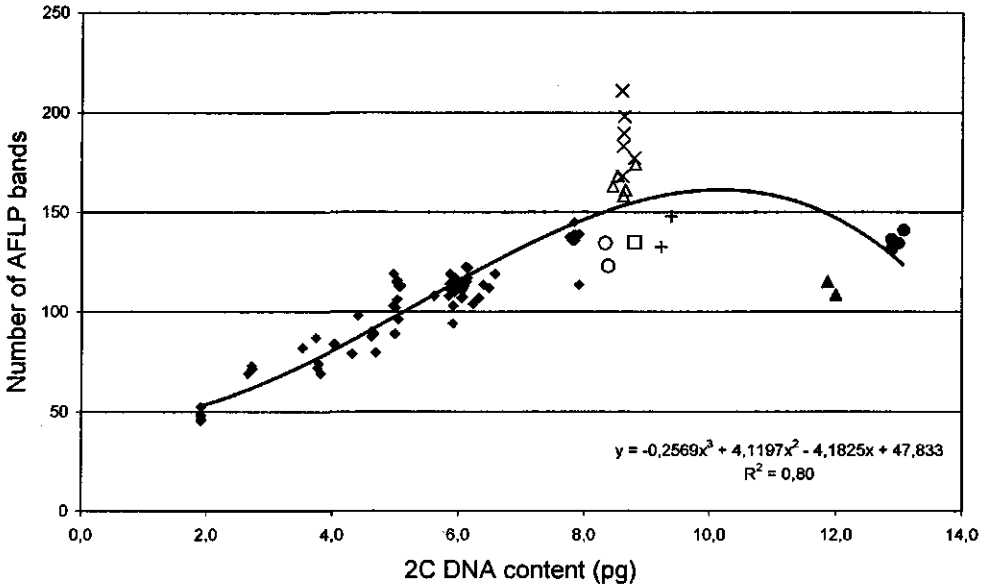
Fig. 4. Association of 2C DNA content (pg) and number of AFLP bands per accession. Symbols: ×, *L. tatarica*; ▲, *C. alpina*; ●, *L. indica*; △, *L. sibirica*; +, *L. virosa* 15679/15680; ☐, *L. quercina*; ○, *P. purpurea*, ♦, remaining species (see Table 1).

Apart from *L. tatarica*, *C. alpina*, and *L. indica*, three more species showed DNA contents above 8.5 pg, viz. *L. quercina* (average 2C = 8.789 pg DNA, 135.0 bands), *L. sibirica* (2C = 8.604 pg DNA, 164.8 bands), and *L virosa* 15679/15680 (2C = 9.300 pg DNA, 140.3 bands).

Table 3. Means and standard errors of the means for relative DNA amounts of cultivars from various cultivargroups of *L. sativa* (groups on first row). The DNA amounts values are relative to *L. esculentum* 'Tiny Tim', and thus in arbitrary units.

Butterhead indoor		Butterhead outdoor		Cos	
'Cortina'	2.776 ± 0.0086	'Milan'	2.646 ± 0.0359	'Fredo'	2.817 ± 0.0090
'Flandria'	2.767 ± 0.0286	'Rexado'	2.682 ± 0.0243	'Hector'	2.691 ± 0.0341
'Rachel'	2.741 ± 0.0360	'Sander'	2.734 ± 0.0086	'Odessa'	2.714 ± 0.0066
'Susy'	2.745 ± 0.0123	'Sunny'	2.705 ± 0.0102	'Odra'	2.757 ± 0.0172
		'Titan'	2.691 ± 0.0245	'Toledo'	2.698 ± 0.0222
Average	2.757 ± 0.0121	Average	2.692 ± 0.0103	Average	2.735 ± 0.0091

For a data set excluding these species, the linear correlation between DNA content and number of AFLP bands was highly significant ($R^2 = 0.88$, correlation coefficient 0.94, $p = 4.57 * 10^{-33}$), as was the association for the second degree polynomial ($R^2 = 0.89$, $p = 0.0053$). The association did not significantly increase for the third degree polynomial.

Intraspecific variation in L. sativa

Table 3 shows the relative DNA amounts for plants from the various cultivar groups. The ANOVA showed significant effect of cultivar groups ($p < 0.00005$) and cultivars within groups ($p = 0.0001$). Pairwise comparisons of the groups showed that the Butterhead Indoor, Crisphead, and Cos groups have significantly more DNA than the Cutting group. The Butterhead Indoor group has significantly more DNA than the Butterhead Outdoor group. Pairwise comparisons among the individual cultivars showed various significant differences. Within the cultivar groups: 'Fredo' has significantly more DNA that 'Toledo' and 'Hector' (Cos); and 'Saladin' has significantly more DNA than 'Roxette' (Crisphead). Among the cultivar groups: the Cos cultivar 'Fredo' has significantly more DNA than 'Titan', 'Rexado', 'Milan' (Butterhead Outdoor), 'Cordoba', 'Frisby' (Cutting), and Roxette (Crisphead); the Crisphead cultivar 'Saladin' has significantly more DNA than 'Rexado', 'Milan' (Butterhead Outdoor), 'Cordoba', and 'Frisby' (Cutting); and the Butterhead Indoor cultivars 'Cortina' and 'Flandria' have significantly more DNA than 'Cordoba', 'Frisby' (Cutting), and Milan (Butterhead Outdoor). The largest difference in relative DNA content was 6% between 'Fredo' (2.817) and 'Frisby'/'Milan' (both 2.646).

Discussion

Calculation of absolute DNA amounts

We measured absolute DNA amounts and AT contents of 24 accessions from 23 species of *Lactuca* and related genera (boldface accessions in Table 1), and DNA amounts relative to *L. esculentum* 'Tiny Tim' for all accessions. The relative DNA amounts of the accessions listed in Table 1 were recalculated into absolute DNA amounts, using the previously measured absolute DNA amounts as references. The calculations involved a comparison of the relative

Table 3. Extended.

Crisphead		Cutting	
'Calgary'	2.731 ± 0.0275	'Cordoba'	2.650 ± 0.0201
'Robinson'	2.759 ± 0.0241	'Frisby'	2.646 ± 0.0206
'Roxette'	2.680 ± 0.0220	'Krizet'	2.739 ± 0.0120
'Saladin'	2.798 ± 0.0012		
Average	2.742 ± 0.0107	Average	2.678 ± 0.0104

DNA amount of a given accession to the relative DNA amount of an absolute DNA content reference. Because the relative DNA amounts were determined using the AT preferent dye DAPI, they can only be recalculated into absolute DNA amounts reliably, when accessions and references have identical base compositions. We ensured this condition by using species specific absolute DNA content references, assuming that the base composition is constant within species. The use of species specific absolute DNA content references effectively excluded the influence of interspecific differences in base composition and fluorochrome binding properties (see Dolezel, Sgorbati, and Lucretti, 1992; Dolezel et al., 1998), and only leaves the intraspecific variation to be dealt with. The intraspecific variation in base composition was estimated using the data on *L. sativa*, *L. serriola*, *L. dregeana*, and *L. altaica*. These species are extremely closely related (probably conspecific, see Koopman et al., 1998; Koopman, Zevenbergen, and Van den Berg, 2001), and therefore their AT percentages can be regarded to reflect the expected intraspecific variation in a *Lactuca* s.l. species. The fact that the AT contents among these species are not significantly different (Fig. 1) indicates that it is reasonable to assume the base composition of a species to be constant. As an exception, we used two reference accessions for *L. virosa*. Previous studies demonstrated the presence of two distinct evolutionary lines in this species, and therefore we used different references for each line.

The relative DNA amounts for the 21 modern *L. sativa* cultivars listed in Table 2 were not recalculated into absolute DNA amounts, because we had no absolute DNA content reference available for any of the cultivars. Consequently, only relative DNA amounts are reported for the modern cultivars.

Absolute DNA content and base composition

2C DNA contents among angiosperms range from about 0.2 to about 250 pg DNA (e.g. Bennett, Bhandol, and Leitch, 2000). In the present study, absolute 2C DNA contents for *Lactuca* s.l. varied between 1.451 pg and 12.979 pg, and therefore are in the lower part of the range for angiosperms. The 6.8 fold variation represented by the 2C DNA content range for *Lactuca* s.l. is well within the 2 to 10 fold variation reported as typical within genera (Price, 1988).

An extensive literature survey showed that the proportion of AT nucleotides in dicotyledonous genomes ranges from 55% to 65% (e.g. Marie and Brown, 1993; Meister and Barow, 2001). The AT contents of the *Lactuca* s.l. species in the present study ranged from 61.4% to 64.2%, and therefore are in the upper part of the range for angiosperm AT contents.

Considering the DNA contents and base compositions established in the present study, *Lactuca* s.l. can be characterized as a genus with a relatively low (but not anomalous) DNA content, and a relatively high (but not anomalous) AT content.

Evolution of DNA content and base composition

The present results indicated an evolutionary trend towards higher DNA amounts in Lactucinae. This trend is apparent in the higher average DNA amount of the Lactucinae species relative to the outgroup, but also in the two supported clades within *Lactuca* s.l.. The increase in DNA amount is in accordance with the general evolutionary trend toward higher DNA amounts in angiosperms, proposed by Bennetzen and Kellogg (1997). The trend is apparent when all angiosperms are considered (Leitch, Chase, and Bennett, 1998), but also within monocotyledons (Bharathan, 1996; Leitch, Chase, and Bennett, 1998) and dicotyledons (Leitch, Chase, and Bennett, 1998). At the family level the trend is less unequivocal, and both increases and decreases in DNA amount occur, sometimes even within the same genus (e.g. Price, 1976; Narayan, 1987; Kiehn, 1995; Bharathan, 1996; Ohri, 1998). In *Lactuca* s.l., we found a general evolutionary trend towards higher DNA amounts, but the low DNA amounts of e.g. *L. viminea*, *L. perennis*, and *M. muralis* indicate that opposite trends may also occur. The co-occurrence of such opposite trends in DNA content evolution of Asteraceae has already been demonstrated in the subtribe Microseridinae (Price and Bachmann, 1975), and in the genus *Hypochaeris* (Cerbah et al., 1999). In *Hypochaeris* however, the evolutionary trend towards decrease in DNA amount was less obvious in an earlier study on a smaller number of species (Cerbah et al., 1995). The difference between both *Hypochaeris* studies may indicate that increase in DNA content is the general rule, and that decrease is a more limited phenomenon, requiring more extensive sampling to be detected. This assumption is corroborated by our *Lactuca* study, showing a general evolutionary trend towards increase in DNA amount, and only a limited number of species with a DNA content below that of their reconstructed ancestral nodes. More extensive sampling in Asteraceae is needed to reveal whether the trend towards increase in DNA amount in *Lactuca* s.l. is also present in Asteraceae as a whole.

Considering base composition, significant differences exist among the closely related subsect. *Lactuca* species (e.g. *L. aculeata* and *L. altaica*), while on the other hand less closely related species are not significantly different (e.g. *L. aculeata* and *L. quercina*, see Fig. 1). In Poaceae, King and Ingrouille (1987a) and King and Ingrouille (1987b) established that base composition can be used to detect taxonomically related groups at the subtribal level, but given the differences and overlaps in our data set, base composition is no reliable indicator of

taxonomic groups in *Lactuca*. However, tracing base compositions on a molecular phylogeny did reveal clear evolutionary patterns, with base compositions evolving towards higher (*L. tatarica* and *L. sibirica*) as well as towards lower (*L. quercina* and subsect. *Lactuca*) AT contents. We are the first to demonstrate such clear evolutionary pattern in base composition traced on a molecular phylogeny.

Association of DNA content and base composition

Notwithstanding the bi-directional evolution of base composition in *Lactuca* s.l., the tests of association showed that in general the increase in DNA amount was accompanied by a decrease in AT content. This decrease was apparent at all taxonomic levels examined, from among subtribes (Lactucinae, Crepidinae, Sonchinae) to within species (*L. virosa*). A·negative correlation of genome size with AT contents in angiosperms was first suggested by Vinogradov (1994). The correlation was recently tested by Meister and Barow (2001), using a larger sample of 54 species from 17 higher plant families. However, they did not find a significant correlation, neither between nor within families. Vinogradov (1994) based his conclusions on only six species, and assembled part of his data from previous studies. Meister and Barow (2001) suggest that these data may have been insufficient to arrive at a reliable conclusion. Moreover, they determined that a correlation of DNA content and base composition was not significant for the six species when the latest available genome size values were used. In contrast, our results on 24 accessions of 23 species unequivocally corroborate the correlation suggested by Vinogradov (1994). A possible explanation of the discrepancy between our results and those of Meister and Barow (2001) may be that there is no general association between DNA amount and AT content in higher plants, but that such a relation does exist for specific plant groups. Indeed, literature data suggest that even within Asteraceae the inverse relation demonstrated in *Lactuca* s.l. might be exceptional. For example, absence of a correlation between DNA content and base composition is reported in *Hypochaeris* (Cerbah et al., 1999). Reports on e.g. Conifers (Miksche and Hotta, 1973), *Pennisetum* (Martel et al., 1997), *Hydrangea* (Cerbah et al., 2001), and *Lathyrus* (Mohammed Ali, Meister, and Schubert, 2000) indicate a lack of correlation in a variety of other plant groups as well.

Intra- and interspecific DNA content variation in Lactuca s.l.

The existence of intraspecific variation in genome size has long been debated (see Bennett and Leitch, 1995, and Ohri, 1998, for an overview). Until the late 1970s, DNA content within a species was considered constant. However, when DNA content data became available for an increasing number of species, considerable intraspecific variation was detected in some of them. Much of the variation reported in early studies was later discovered to results from artifacts such as aneuploidy or methodological errors (notably self-tanning (Greilhuber, 1988)). In the early 1980s however, more reliable reports were published, and the presence of intraspecific variation in many taxa became more generally accepted (though still disputed, see

e.g. Greilhuber (1998)). Recently, Ohri (1998) suggested a reappraisal of intraspecific variation in DNA content as additional taxonomic character, hypothesizing that intraspecific genome size may be fairly stable when a narrow species concept is adopted.

In our *Lactuca* data, we detected six species with intraspecific variation in DNA content. The largest variation was found in *L. viminea* (25%), with the accessions grouping in two distinct clusters. However, the grouping is not conform to previous groupings based on ITS sequences (Koopman et al., 1998) and AFLP markers (Koopman, Zevenbergen, and Van den Berg, 2001), nor to the subspecific classification of Feráková (1977). Therefore, the distinct DNA clusters probably indicate the variability of the species rather than subspecific taxa. Large variation was also detected between two clusters of *L. virosa* accessions (21%). These clusters are in accordance with the ITS and AFLP results, and therefore very likely represent two distinct evolutionary groups within *L. virosa*. These groups were detected with both the relative DNA measurements and the absolute DNA measurements. Both types of measurements were conducted in different years, with different plant material, in different labs, and by different people. Therefore, the differences between the *L. virosa* groups must be considered highly reproducible. However, the groups seem not to be characterized by clear morphological autapomorphies, and more research is needed to establish their formal intraspecific status. Within *L. serriola*, the maximum intraspecific variation was 7%, which probably reflects the variability of the species. The vegetatively propagating *L. sibirica* showed 4% intraspecific variation. Chromosome counts (Koopman and Zevenbergen, unpublished data) indicated that aneuploidy is the most probable reason for this variation. *P. purpurea* was not examined in enough detail to explain the 4% intraspecific variation. Significant variation in *L. sativa* was detected among modern cultivars only (discussed later).

Apart from intraspecific variation, we also detected extensive interspecific variation, suggesting that species could be characterized by their DNA content. The suitability of DNA contents as taxonomic markers in *Lactuca* was already demonstrated for *L. serriola*, *L. saligna*, and *L. virosa* Koopman (2000), but required the a priori knowledge that the "unknown" sample at least belonged to one of the three species. The present study demonstrated that DNA contents in *Lactuca* have only limited value as identification tools, because closely related species may have DNA amounts that are significantly different, whereas the ranges in DNA amount of more distantly related species may overlap (Fig. 3). However, similarities in DNA amount corroborated a number of previously established groups of closely related species, viz. *L. tatarica/L. sibirica/L. quercina* (belonging to the tertiary gene pool of cultivated lettuce, see Koopman et al. (1998), and the group of "*serriola*-like" species *L. sativa*, *L. serriola*, *L. dregeana*, *L. altaica*, and *L. aculeata* (the primary gene-pool species). The clusters in Fig. 3 largely correspond to the species clusters determined in Doležalová et al. (In press). However, a comparison of the present results with those of Doležalová et al. (In press) is hampered by the fact that: 1) they determined relative DNA amounts instead of absolute DNA amounts, using

the AT-binding fluorochrome DAPI (although their results were reported in pg DNA); and 2) many of their species identifications were uncertain or incorrect.

Association of DNA content and number of AFLP bands

In the present study we confirmed, for the first time, the "almost linear" relationship between genome size and number of AFLP bands in higher plants predicted by Vos et al. (1995). We took absolute DNA content as a measure of genome size, which is warranted because absolute DNA content and genome size are directly related according to 1pg DNA = 965 Mbp, or, alternatively, 1pg DNA = 980 Mbp (see Bennett, Bhandol, and Leitch (2000) for a discussion). Vos et al. (1995) indicated that the "almost linear" relationship is lost in the complex genomes of higher plants because of the high numbers of repeated sequences in those genomes. Our data show that the relationship between 2C DNA content and number of AFLP bands is almost linear below 8.5 pg DNA, but deviates from linear above 8.5 pg. Apparently, genome complexity is similar in *Lactuca* s.l. genomes smaller than approximately 8.5 pg (i.e. 4165 Mbp/1C), and the numbers of repeated sequences are limited. Three species show numbers of AFLP bands that clearly deviate from what is expected based on a linear relationship, but there is no general pattern to the deviation: *L. tatarica* showed an excess of bands, while *C. alpina* and *L. indica* showed a shortage. The deviating band numbers could reflect methodological artifacts such as incomplete restriction or template competition, but all AFLP samples were treated in a similar fashion and visual inspection of the gels showed no irregularities. The base compositions of *L. tatarica*, *C. alpina*, and *L. indica* are well within the range established for *Lactuca* s.l., so the base compositions of these species cannot explain their deviating AFLP band numbers. Therefore, the deviating band numbers are most probably due to features related to the larger genome size. Because increase in genome size is generally correlated with increase in repetitive sequences, these sequences are most probably responsible for the deviating band numbers. Thus, our results confirm the suggestion of Vos et al. (1995) that repeated sequences are responsible for a non-linear relation between genome size and numbers of AFLP bands in the complex genomes of higher plants. The large differences in numbers of bands in *L. tatarica*, *C. alpina*, and *L. indica* may originate in differences in the composition and organization of the repeat sequences. For example, the high numbers of bands in *L. tatarica* may result from repetitive sequences containing relatively high numbers of Eco/Mse sites and matching selective bases. Conversely, the low band numbers in *C. alpina* and *L. indica* may result from repetitive sequences with relatively low numbers of such sites.

The fact that *L. tatarica*, *C. alpina*, and *L. indica* all have genome sizes above 8.5 pg DNA indicates that 8.5 pg DNA is the lower limit for "complex genomes" in *Lactuca* s.l. To examine the relation between genome size and number of AFLP bands for species with less complex genomes only, we excluded all species with DNA contents above 8.5 pg, and reanalyzed the data set. According to Vos et al. (1995), the results for this limited data set should demonstrate an "almost linear" relationship. Theoretically, however, a curvilinear relationship is expected

because the chances of two AFLP fragments to coincide in the same AFLP band increase with increasing numbers of bands, and thus with increasing genome size. The presence of such comigrating non-identical fragments has already been demonstrated in *Beta* (Hansen et al., 1999) and *Glycine max* (Meksem et al., 2001). The increasing number of comigrating non-identical fragments in larger genomes result in a curvilinear relationship between genome size and number of AFLP bands, with genome sizes increasing faster than the numbers of AFLP bands. Indeed, our results demonstrated that for genome sizes below 8.5 pg DNA, a curvilinear relationship (the second degree polynomial fit) describes the relation between genome size and number of AFLP bands better than a linear relationship does.

Intraspecific variation in L. sativa

Within *L. sativa* , significant variation in DNA content (6%) was detected only among the modern cultivars listed in Table 2. No significant variation was detected among the accessions from Table 1, mainly comprising landraces and older cultivars. The significant variation among the modern cultivars probably reflects an increased selection pressure and an increased use of wild genitors in these cultivars. For example, traits such as bolting behavior and duration of the growth cycle may be associated with DNA content. Such associations have been demonstrated for various traits, and are known as nucleotypic effects (Bennett, 1985). Selection on extremes in these traits (e.g. fast growth, slow bolting) may result in differences in DNA contents among cultivars. Cultivars with *L. saligna* or *L. virosa* in their progeny may show a decrease or increase in their DNA amount relative to the *L. sativa* parent, reflecting the DNA contents of the wild ancestor.

Conclusions

2C DNA amounts among 19 Lactucinae species (18 species from *Lactuca* s.l., and *P. purpurea*) and *C. intybus* ranged from 1.913 pg in *L. tenerrima* to 13.068 pg in *L. indica*, while three Lactuceae outgroup species outside subtribe Lactucinae showed DNA contents between 1.434 (*S. asper*) and 3.153 pg (*C. juncea*). These values are in the lower part of the DNA content range for angiosperms. Six species showed significant intraspecific differences in DNA amount: *L. viminea* (25%), *L. virosa* (21%), *L. serriola* (7%), *L. sativa* (6%), *L. sibirica* (4%), and *P. purpurea* (4%). Only the variation within *L. virosa* seems to have evolutionary significance, indicating two distinct evolutionary clades. We detected four groups of species with (partially) overlapping DNA content ranges: 1) *L. sativa, L. serriola, L. dregeana, L. altaica, L. aculeata* (i.e. the "*serriola*-like species)", *S. tuberosus, C. plumieri*; 2) *L. tatarica, L. sibirica* (i.e. subsect. *Mulgedium*), *L. quercina, P. purpurea*, and *L. virosa* (excluding CGN 15679/15680); 3) *L. saligna, L. viminea, L. perennis*, and *M. muralis*; 4) *L. tenerrima* and *T. officinale*. The groups only partly conform to the generally accepted classifications, because they include closely related as well as more distinctly related species. Therefore, DNA content in *Lactuca* s.l. has only limited value as a taxonomic character. The evolution in *Lactuca* s.l. was generally directed towards increasing genome size. The genome size of Lactuceae species

with 2C DNA amounts below 8.5 pg (4165 Mbp/1C) was positively associated with the number of AFLP bands generated from it, showing a curvilinear (second degree polynomial) relationship. This curvilinear relationship indicates a similar genome complexity and proportion of repeat sequences for these species. For species with 2C DNA amounts above 8.5 pg, the association was lost, indicating differences in genome complexity and repeat sequences among these species.

AT contents among the 19 Lactucinae species ranged from 61.4% in *L. virosa* to 64.2% in *L. perennis*. The outgroup species showed AT contents between 62.7% (*T. officinale*) and 64.1% (*C. juncea*). These values are in the upper part of the AT content range for Angiosperms. Four groups with partially overlapping AT contents were detected: 1) *L. sativa*, *L. serriola*, *L. dregeana*, *L. altaica*, *L. aculeata*, *L. saligna*, *L. virosa* (i.e. subsect *Lactuca*), *L. quercina*, *L. indica*, *C. plumieri*, *C. alpina*, *P. purpurea*; 2) *L. tatarica*, *L. sibirica* (i.e. subsect. *Mulgedium*), *L. viminea*, *L. tenerrima*, *M. muralis*, *S. tuberosus*, S. aper; 3) *L. perennis*, *C. juncea*, *C. intybus*; 4) *T. officinale*. Again, the groups only partly conform to the generally accepted classifications. The evolution of base composition was bi-directional: towards increase of AT content for *L. tatarica/L. sibirica*, and towards decrease of AT content for *L. quercina* and subsect. *Lactuca*. Notwithstanding the bi-directional evolution of base compositions, DNA content and AT% showed a significant negative association for all subsets of Lactuceae species that were tested.

Acknowledgements

References

BENNETT, M. D. 1985. Intraspecific variation in DNA amount and the nucleotypic dimension in plant genetics. Plant genetics: proceedings of the third annual ARCO Plant Cell Research Institute-UCLA symposium on plant biology., Keystone, Colorado, USA, 35.

BENNETT, M. D., AND J. B. SMITH. 1976. Nuclear DNA amounts in angiosperms. *Philosophical Transactions of the Royal Society of London B* 274: 227-274.

_____. 1991. Nuclear DNA amounts in angiosperms. *Philosophical Transactions of the Royal Society of London B* 334: 309-345.

BENNETT, M. D., AND I. J. LEITCH. 1995. Nuclear DNA Amounts in Angiosperms. *Annals of Botany* 76: 113-176.

_____. 2001. Angiosperm DNA C-values database (release 3.1, Sept. 2001). Royal Botanic Gardens, Kew.

BENNETT, M. D., P. BHANDOL, AND I. J. LEITCH. 2000. Nuclear DNA Amounts in Angiosperms and their Modern Uses-807 New Estimates. *Annals of Botany* 86: 859-909.

BENNETZEN, J. L., AND E. A. KELLOGG. 1997. Do Plants Have a One-Way Ticket to Genomic Obesity? *The Plant Cell* 9: 1509-1514.

BHARATHAN, G. 1996. Reproductive development and nuclear DNA content in angiosperms. *American Journal of Botany* 83: 440-451.

CERBAH, M., J. COULAUD, B. GODELLE, AND S. SILJAK-YAKOVLEV. 1995. Genome size, fluorochrome banding, and karyotype evolution in some *Hypochoeris* species. *Genome* 38: 689-695.

CERBAH, M., J. COULAUD, S. C. BROWN, AND S. SILJAK-YAKOVLEV. 1999. Evolutionary DNA variation in the genus *Hypochaeris. Heredity* 82: 261-266.

CERBAH, M., E. MORTREAU, S. BROWN, S. SILJAK-YAKOVLEV, H. BERTRAND, AND C. LAMBERT. 2001. Genome size variation and species relationships in the genus Hydrangea. *Theoretical and Applied Genetics* 103: 45-51.

COX, A. V., G. J. ABDELNOUR, M. D. BENNETT, AND I. J. LEITCH. 1998. Genome size and karyotype evolution in the slipper orchids (Cypripedioideae: Orchidaceae). *American Journal of Botany* 85: 681-687.

DIMITROVA, D., AND J. GREILHUBER. 2000. Karyotype and DNA-content evolution in ten species of Crepis (Asteraceae) distributed in Bulgaria. *Botanical Journal of the Linnean Society* 132: 281-297.

DOLEŽALOVÁ, I., A. LEBEDA, J. JANEČEK, J. ČÍHALÍKOVÁ, E. KŘÍSTKOVÁ, AND O. VRÁNOVÁ. In press. Variation in chromosome numbers and nuclear DNA contents in genetic resources of *Lactuca* L. species (Asteraceae). *Genetic Resources and Crop Evolution.*

DOLEŽEL, J., S. SGORBATI, AND S. LUCRETTI. 1992. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiologia Plantarum* 85: 625-631.

DOLEŽEL, J., J. GREILHUBER, S. LUCRETTI, A. MEISTER, M. A. LYSÁK, L. NARDI, R. OBERMAYER. 1998. PLANT GENOME SIZE ESTIMATION BY FLOW CYTOMETRY: INTER-LABORATORY COMPARISON. *Annals of Botany* 82 (SUPPLEMENT A): 17-26.

FARRIS, J. S. 1970. Methods for computing Wagner trees. *Systematic zoology* 19: 83-92.

FERÁKOVÁ, V. 1977. The genus *Lactuca* L. in Europe. Univerzita Komenského, Bratislava.

GALBRAITH, D. W., G. M. LAMBERT, J. MACAS, AND J. DOLEŽEL. 1998. Analysis of nuclear DNA content and ploidy in higher plants. *In* J. P. Robinson, Z. Darzynkiewicz, P. N. Dean,

L. G. Dressler, A. Orfao, P. S. Rabinovitch, C. C. Stewart, H. J. Tanke, and L. L. Wheeless [eds.], Current Protocols in Cytometry, 7.6.1 - 7.6.22. John Wiley & Sons, Inc., New York.

GODELLE, B., D. CARTIER, D. MARIE, S. C. BROWN, AND S. SILJAK-YAKOVLEV. 1993. Heterochromatin Study Demonstrating the Non-Linearity of Fluorometry Useful for Calculating Genomic Base Composition. *Cytometry* 14: 618-626.

GREILHUBER, J. 1988. "Self-tanning" - a new and important source of stoichiometric error in cytophotometric determination of nuclear DNA content in plants. *Plant Systematics and Evolution* 158: 87-96.

_____. 1998. Intraspecific Variation in Genome Size: A Critical Reassessment. *Annals of Botany* 82: 27-35.

HANSEN, M., T. KRAFT, M. CHRISTIANSSON, AND N.-O. NILSSON. 1999. Evaluation of AFLP in *Beta. Theoretical and Applied Genetics* 98: 845-852.

KIEHN, M. 1995. Chromosome survey of the Rubiaceae. *Annals of the Missouri Botanical Garden* 82: 398-408.

KING, G. J., AND M. J. INGROUILLE. 1987a. DNA base composition heterogeneity in the grass genus *Briza* L. *Genome* 29: 621-626.

_____. 1987b. Genome heterogeneity and classification of the Poaceae. *New Phytologist* 107: 633-644.

KOOPMAN, W. J. M. 2000. Identifying lettuce species (*Lactuca* subsect. *Lactuca*, Asteraceae): A practical application of flow cytometry. *Euphytica* 116: 151-159.

KOOPMAN, W. J. M., M. J. ZEVENBERGEN, AND R. G. VAN DEN BERG. 2001. Species relationships in *Lactuca* s.l. (Lactuceae, Asteraceae) inferred from AFLP fingerprints. *American Journal of Botany* 88: 1881-1887.

KOOPMAN, W. J. M., E. GUETTA, C. C. M. VAN DE WIEL, B. VOSMAN, AND R. G. VAN DEN BERG. 1998. Phylogenetic relationships among *Lactuca* (Asteraceae) species and related genera based on ITS-1 DNA sequences. *American Journal of Botany* 85: 1517-1530.

LEITCH, I. J., M. W. CHASE, AND M. D. BENNETT. 1998. Phylogenetic Analysis of DNA C-values Provides Evidence for a Small Ancestral Genome Size in Flowering Plants. *Annals of Botany* 82: 85-94.

MADDISON, D. R., AND W. P. MADDISON. 2000. MacClade 4. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.

MARIE, D., AND S. C. BROWN. 1993. A cytometric exercise in plant DNA histograms, with 2C values for 70 species. *Biology of the Cell* 78: 41-51.

MARTEL, E., D. DE NAY, S. SILJAK-YAKOVLEV, S. BROWN, AND A. SARR. 1997. Genome Size Variation and Basic Chromosome Number in Pearl Millet and Fourteen Related Pennisetum Species. *Journal of Heredity* 88: 139-143.

MEISTER, A., AND M. BAROW. 2001. Lack of Correlation Between AT Frequency and Genome Size in Higher Plants and the Effect of Nonrandomness of Base Sequences on Dye Binding. *Cytometry* 47: 1-7.

MEKSEM, K., E. RUBEN, D. HYTEN, K. TRIWITAYAKORN, AND D. A. LIGHTFOOT. 2001. Conversion of AFLP bands into high-throughput DNA markers. *Molecular Genetics and Genomics* 265: 207-214.

MIKSCHE, J. P., AND Y. HOTTA. 1973. DNA Base Composition and Repetitious DNA in Several Conifers. *Chromosoma* 41: 29-36.

MOHAMMED ALI, H. B., A. MEISTER, AND I. SCHUBERT. 2000. DNA content, rDNA loci, and DAPI bands reflect the phylogenetic distance between *Lathyrus* species. *Genome* 43: 1027-1032.

NARAYAN, R. K. J. 1987. Nuclear DNA changes, genome differentiation and evolution in *Nicotiana* (*Solanaceae*). *Plant Systematics and Evolution* 157: 161-180.

OHRI, D. 1998. Genome Size Variation and Plant Systematics. *Annals of Botany* 82: 75-83.

PORTUGAL, J., AND M. J. WARING. 1988. Assignment of DNA binding sites for 4',6-diamidino-2-phenylindole and bisbenzimide (Hoechst 33258). A comparative footprinting study. *Biochimica et Biophysica Acta* 949: 158-168.

PRICE, H. J. 1976. Evolution of DNA content in higher plants. *The botanical review* 42: 27-52.
_____. 1988. Nuclear DNA content variation within angiosperm species. *Evolutionary trends in plants* 2: 53-60.

PRICE, H. J., AND K. BACHMANN. 1975. DNA content and evolution in the Microseridinae. *American Journal of Botany* 62: 262-267.

RODENBURG, C. M. 1960. Varieties of lettuce: An international monograph. Varietal Discription Nr. 3, I.V.T, Wageningen.

SIMS, L. E., AND H. J. PRICE. 1985. Nuclear DNA content variation in *Helianthus* (Asteraceae). *American Journal of Botany* 72: 1213-1219.

STEBBINS, G. L. 1937. Critical notes on *Lactuca* and related genera. *Journal of Botany*: 12-18.

SWOFFORD, D. L. 1999. PAUP*: phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland, Massachusetts, USA.

SWOFFORD, D. L., AND W. P. MADDISON. 1987. Reconstructing Ancestral Character States Under Wagner Parsimony. *Mathematical Biosciences* 87: 199-229.

TORRELL, M., AND J. VALLÈS. 2001. Genome size in 21 Artemisia L. species (Asteraceae, Anthemideae): Systematic, evolutionary, and ecological implications. *Genome* 44: 231-238.

VINOGRADOV, A. E. 1994. Measurement by Flow Cytometry of Genomic AT/GC Ratio and Genome Size. *Cytometry* 16: 34-40.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAN DE LEE, M. HORNES, A. FRIJTERS, J. POT, J. PELEMAN, M. KUIPER, AND M. ZABEAU. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407-4414.

WILSON, W. D., F. A. TANIOUS, H. J. BARTON, L. STREKOWSKI, AND D. W. BOYKIN. 1989. Binding of 4',6-diamidino-2-phenylindole (DAPI) to GC and mixed sequences in DNA: Intercalation of a classical groove-binding molecule. *Journal of the American Chemical Society* 111: 5008-5010.

7

Phylogenetic signal in AFLP data sets

W. J. M. Koopman[1]

[1] Biosystematics Group, Nationaal Herbarium Nederland - Wageningen University branch, Wageningen University, Generaal Foulkesweg 37, 6703 BL Wageningen, The Netherlands

Summary

The application of AFLP markers in cladistic studies requires the presence of phylogenetic signal in AFLP data sets. However, general drawbacks related to the use of presence/absence of restriction fragments as data may obscure this signal. These drawbacks are: non-independence of fragments, non-homology of fragments, asymmetry in the probability of loosing and gaining fragments, and failure to distinguish heterozygotes from homozygotes. Two AFLP data sets of *Lactuca* s.l. were examined on the presence of phylogenetic signal. The presence of phylogenetic signal could warrant cladistic analysis of the AFLP data, in spite of the general drawbacks. I used three approaches: 1) direct testing on the data set using tree length distribution skewness (g_1) and relative apparent synapomorphy analysis (RASA); 2) determination of branch supports for a most-parsimonious tree (MPT) using bootstrap, jackknife, and decay analysis; 3) comparison of AFLP and ITS-1 data sets and MPTs using the partition homogeneity test (PHT), Templeton test, and regression analyses. All three approaches indicated the presence phylogenetic signal in the AFLP data sets, although AFLP and ITS MPT topologies conflicted on some points. An extensive literature survey demonstrated a general congruence of AFLP and ITS tree topologies across a wide range of taxonomic groups, indicating that the present results and conclusions have a general bearing.

Key words: AFLP markers, Congruence, g_1 statistic, Internal Transcribed Spacer (ITS), Lactuca, Partition Homogeneity Test (PHT), Phylogenetic signal.

Introduction

AFLP® markers (Vos et al., 1995) were originally intended for use in marker assisted breeding, but their popularity in systematic studies rapidly increased over the past few years. AFLP markers are typically applied within populations, within species, or among closely related species, analyzed using clustering methods (usually UPGMA, e.g. Huys et al., 1996 (*Aeromonas*); Keim et al., 1997 (*Bacillus*); Kardolus et al., 1998 (*Solanum*); DeScenzo et al., 1999 (*Eutypa*); Kiers et al., 2000 (*Cichorium*)), Neighbor-Joining (e.g. Lu et al., 1996 (*Pisum*); Angiolillo et al., 1999 (*Olea*); Tredway et al., 1999 (Clavicipitaceae); Giannasi et al., 2001 (*Trimeresurus*)), or cladistics (e.g. Keim et al., 1997; Kardolus et al., 1998; Tredway et al., 1999; Hodkinson et al., 2000 (*Phyllostachys*)).

To be suitable for cladistic analysis, restriction fragment data (such as AFLP markers) have to meet two basic requirements (Backeljau et al., 1995; Swofford and Olsen, 1990): 1) the fragments must have evolved independently (Karp et al., 1996); and 2) fragments of equal length must be homologous (Black, 1993; Karp et al., 1996). As was pointed out by Karp et al. (1996), non-independence and non-homology are general theoretical drawbacks of methods using presence/absence of restriction fragments as data. For AFLP markers, the homology problem seems limited with respect to intraspecific variation, as demonstrated by Rouppe van der Voort et al. (1997) for relatively unrelated *S. tuberosum* genotypes. Considering interspecific variation, O' Hanlon and Peakall (2000) showed that the level of non-homology rapidly increases with taxonomic divergence, mounting up to 100% for species from different subtribes of Carduinae thistles (Asteraceae). Non-homologous fragments of equal length may be present in different genotypes (Rouppe van der Voort et al., 1997; O' Hanlon and Peakall, 2000), as well as within one genotype (Rouppe van der Voort et al., 1997; Hansen et al., 1999; Meksem et al., 2001). To my knowledge, the occurrence of non-independent fragments in AFLP studies has not yet been extensively studied. Besides non-homology and non-independence, asymmetry in the probability of loosing and gaining fragments (loss of a fragment is much more probable than gain), and the fact that AFLP markers are usually scored dominantly (i.e. without distinction between homozygotes and heterozygotes) may also limit their validity as cladistic characters (see Karp et al. (1996) and references therein).

Taken together, the presence of non-independent bands, non-homologous comigrating bands, the asymmetry in loss and gain of fragments, and the dominant scoring of codominant AFLP markers may obscure the phylogenetic signal in the data set, rendering a reliable cladistic analysis impossible. Cladistic analyses of AFLP data have been published for a number of species groups, but only Giannasi et al. (2001) briefly addressed the presence of phylogenetic signal in his data set. To my knowledge, no comprehensive study into the matter has been performed so far. Given the popularity of AFLP markers for relationship studies, the increasing number of cladistic analyses among these studies, and the limitations of AFLP markers as cladistic characters, a more thorough examination of phylogenetic signal in AFLP data sets is warranted.

In the present paper I examine this signal in data sets of *Lactuca* s.l., employing three different approaches: 1) testing for phylogenetic structure in the data sets, using various statistical procedures; 2) calculating most-parsimonious trees (MPTs), and determining their branch supports; 3) determining congruence between AFLP based and ITS-1 (Internal Transcribed Spacer-1) based MPT topologies. The validity of most procedures to test phylogenetic signal is disputed in cladistic literature, and therefore I explicitly justify my choices of methods in the materials and methods section. In an extensive literature study, the congruence of ITS and AFLP based tree topologies was verified and confirmed for a wide range of taxa, corroborating the results for *Lactuca* s.l.

Materials and methods

The AFLP data sets in the present study were selected from larger sets from Koopman et al. (2001), and contained 84 accessions (1 plant per accession) from 19 species of *Lactuca* and related genera. The accessions are identical to those from a previous ITS-1 sequence study (Koopman et al., 1998), but with the exclusion of *Lactuca sativa* CGN 5045, *Mycelis muralis* CGN 9367 and CGN 5005, *Prenanthes purpurea* W9534, and all accessions of *Sonchus*, *Taraxacum*, and *Chondrilla*. The first data set was generated with primer combination (pc) E35/M48 (EcoRI + ACA/MseI + CAC), and contained 530 polymorphic bands. The second data set was generated with pc E35/M49 (EcoRI + ACA/MseI + CAG), and contained 500 polymorphic bands. The ITS-1 data set used in the present study was a subset from the data set of Koopman et al. (1998), containing the same accessions as the AFLP data sets. The original multiple sequence alignment was adjusted by hand where necessary. The ITS data set contained several duplicate sequences (see Koopman et al., 1998). Calculation of g_1, RASA, and most parsimonious trees (MPTs) was performed without these duplicate sequences (the pruned data set contained 46 unique sequences). Tests of congruence were performed including duplicate sequences. To make the ITS MPT suitable for this purpose, accessions with duplicate sequences were added to the tree manually.

As a first approach, I examined whether the data sets contain phylogenetic signal at all. Three techniques for testing phylogenetic signal are prominently present in systematic literature: tree length distribution skewness (TLD, Hillis, 1991; Huelsenbeck, 1991; Hillis and Huelsenbeck, 1992), Permutation Tail Probability (PTP) testing (Archie, 1989; Faith and Cranston, 1991), and Relative Apparent Synapomorphy Analysis (RASA, Lyons-Weiler et al. (1996)). Data Decisiveness (Goloboff, 1991a,b) is a related technique, but without the explicit claim that phylogenetic signal is measured (Carpenter, 1992).

I used TLD to measure phylogenetic signal because, although sometimes criticized (Källersjö et al., 1992; Lyons-Weiler et al., 1996), it is still widely employed and generally accepted. To determine phylogenetic signal based on TLD, a length distribution of randomly generated phylogenetic trees is assembled based on the observed data set. The skewness of this

distribution is described by the g_1 statistic (Sokal and Rohlf, 1969). Negative g_1 values below a certain critical value (derived from tree length distributions based on random data sets) indicate significant phylogenetic signal (Hillis and Huelsenbeck, 1992). I determined TLD and g_1 of 100 000 random trees for the separate and combined AFLP data sets and for the ITS-1 data set, using PAUP* 4.0b8 (PPC/Altivec) (Swofford, 1999).

The PTP tests of Archie (1989) and Faith and Cranston (1991) determine phylogenetic signal by comparing the length of a MPT based on an observed data set with the lengths of MPTs based on randomizations (usually 99 or 999) of that data set. The data set is randomized by randomly permuting the states of each of the characters across all taxa. Test statistic is the fraction of MPTs from the permuted data sets that is equally short or shorter than the MPT from the observed data set. Significant phylogenetic signal is concluded when this fraction is below a critical value. The PTP test is extensively criticized (Carpenter, 1992; Källersjö et al., 1992; Steel et al., 1993; Lyons-Weiler et al., 1996; Carpenter et al., 1998), mainly because the null hypothesis of random distribution of character states is considered invalid. Given the conclusions of Carpenter (1992) and Carpenter et al. (1998) that the application of character randomization in cladistics is "ill conceived", and that PTP "seems to have no phylogenetic interpretation at all", I decided not to use PTP tests in the present study.

Presented as an alternative to TLD and PTP testing, RASA is a relatively new but already widely employed procedure to test for phylogenetic signal (see list of papers at http://bio.uml.edu/LW/rasaEnhanced.html). I note that RASA was recently heavily criticized (Simmons et al., 2002), but employ it here awaiting further discussion on its applicability. To perform RASA, two measures are calculated for each taxon pair: a Relative Apparent Synapomorphy score RAS (representing the number of times that a taxon pair shares a character state to the exclusion of another taxon, summed over all characters), and the number of characters involved in the computation of RAS, called E. RAS is plotted against E, and the observed slope of RAS on E is the measure of phylogenetic signal. A null slope is determined from a plot of RAS against E after reciprocal equiprobable redistribution of RAS and E. The test statistic for homogeneity of slopes (Myers, 1990) is used to compare the slopes, and an observed slope significantly steeper than the null slope indicates the presence of phylogenetic signal. I performed RASA analyses for the separate and combined AFLP data sets and for the ITS-1 data set, using the RASA Web Tool at http://bioinformatics.uml.edu/RASA.shtml. Gaps were treated as missing data.

As a second approach, I calculated the supports for MPTs based on the separate and combined AFLP data sets, using PAUP* 4.0b8 (PPC/Altivec). MPTs were calculated in heuristic searches comprising 10 000 random addition sequences with TBR branch swapping, and "multrees" switched off. Parsimony settings were: acctran and "collapse of zero-length branches". Based on previous results (Koopman et al., 1998), *Prenanthes purpurea* was used as outgroup. Support for the MPT topologies was determined using the three most widely employed methods: nonparametric bootstrapping (Felsenstein, 1985), jackknifing (Farris et al.,

1996), and branch support (Bremer, 1988, 1994) the latter also known as decay index (Donoghue et al., 1992) or Bremer support (Källersjö et al., 1992). The methods and their drawbacks are discussed in Lee (2000). Bootstrap values were calculated in 2500 replicates of a full heuristic search, with 10 random addition sequences in each replicate, and remaining settings as above. Jackknife values were calculated in a fast heuristic search with 25000 replicates, nominal deletion of 37% of the characters in each replicate (according to Farris et al. (1996)), and "Jac" resampling. Starting trees were obtained using random addition sequences without branch swapping. Branch support values were calculated with AutoDecay version 3.03 (Eriksson and Wikström, 1996), using the heuristic search option with 10 random addition sequences, TBR branch swapping, and "multrees" switched on.

As a third approach, I examined the congruence of the AFLP data and ITS-1 sequence data, and the congruence of MPTs based on these data. ITS sequences are generally accepted phylogenetic markers (see Baldwin (1992) and numerous papers thereafter), and I therefore consider congruence of ITS and AFLP data or MPTs to indicate that AFLP data track phylogeny. ITS sequences are the reference of choice in this case, since both AFLP markers and ITS sequences can be applied at similar taxonomic levels, although as a rule AFLP markers are somewhat more variable.

I determined the congruence of AFLP and ITS-1 data sets using the Partition Homogeneity Test (PHT) of Farris et al. (1995a,b), based on the incongruence length difference of Mickevich and Farris (1981). The test comprises the following steps: 1) determine the sum L_{x+y} of the lengths of the MPTs from both data sets; 2) randomly partition all characters into new data sets of the original sizes, and do this W times (e.g. 100); 3) determine the lengths of the MPTs from the partitioned data sets; 4) count the number S of MPTs that are longer than L_{x+y}; 5) the error rate on rejecting the null hypothesis of congruency is $P = 1 - (S/(W+1))$. I used the test as implemented in PAUP* 4.0b8 (PPC/Altivec), performing 500 replicates (= 499 repartitions). Trees for each replicate were generated in an heuristic search with 1000 random addition sequences, TBR branch swapping, and "multrees" switched off. Parsimony settings were: acctran and "collapse of zero-length branches".

To serve as a reference for the AFLP MPTs, I calculated ITS-1 MPTs in PAUP* 4.0b8 (PPC/Altivec) using a heuristic search with simple taxon addition, TBR branch swapping, and "multrees" switched on. Parsimony settings were: acctran and "collapse of zero-length branches". *Prenanthes purpurea* was used as outgroup. An additional analysis with 10000 random addition sequences and "multrees" switched off was performed to identify possible islands of shorter trees. However, no such islands were found. Support for the ITS-1 MPT topologies was determined as described for AFLP MPTs.

Congruence of AFLP and ITS-1 MPTs was determined in three different ways, using one arbitrarily selected MPT for each of the data sets.

Firstly, topologies were compared by eye, taking into account the branch supports.

Secondly, I compared the fit of both trees to either of the data sets, using the Templeton test (Templeton, 1983) as implemented in PAUP. The Templeton test is the procedure of choice for this purpose, because it is a robust nonparametric test, requiring no assumptions about the distribution of the test statistic under the null hypothesis. The original Templeton test is a one-tailed Wilcoxon matched-pairs signed-ranks test (Siegel, 1956) that compares the number of changes required for each character on each of the trees (excluding ties). The difference for each character gets a signed rank number, and the negative rank sum is used as test statistic. The test statistic and the total number of ranks are converted into a probability statement based on the statistical tables from the Wilcoxon test, or, for a total number of ranks > 25, using a normal approximation (Siegel, 1956). The Templeton test implemented in PAUP is a two tailed version. Note that the likelihood based Kishino-Hasegawa test (Hasegawa and Kishino, 1989; Kishino and Hasegawa, 1989) or Shimodaira-Hasegawa test (Shimodaira and Hasegawa, 1999) are unfit for comparing the tree topologies because they either require that both trees are derived independent of the data set used in the test (the Kishino-Hasegawa test) or that the Maximum Likelihood tree is included in the set of tested tree topologies (the Shimodaira-Hasegawa test)(Goldman et al., 2000).

Thirdly, I compared the tree topologies by examining the associations between the pairwise phylogenetic distances in an AFLP MPT and an ITS-1 MPT. Pairwise phylogenetic distances for both trees were determined for all pairs of taxa as the total length of the branches connecting the taxa, and assembled in two distance matrices. Significance of matrix associations is usually calculated by testing the significance of the Pearson product-moment correlation (Rademaker et al., 2000), or using the Mantel test (Mantel, 1967; Rohlf, 1993). The Pearson product-moment correlation assumes a linear association, whereas the Mantel test is a quadratic assignment method. However, preliminary analyses of the phylogenetic distance matrices showed a curvilinear, non-quadratic association of AFLP and ITS distances, which can not be reliably assessed using either Pearson or Mantel's coefficient. Therefore, I examined the associations using a series of regression analyses. This approach not only enables testing a variety of fits, but also enables a visual inspection of the results in scatterplots. Linear fits and polynomial fits up to the tenth degree were calculated and tested for significance ($\alpha = 0.01$) in Excel 97, using the "regression" option from the "data analysis" platform. Regression lines were calculated and plotted using the "scatterplot" and "add trendline" options from the "chart" platform. I determined associations for three sets of distances. Firstly, for distances between all accessions in the present study. Secondly, for intraspecific distances and interspecific distances among the species from the moderately supported clades in the AFLP MPT. Together these categories comprise all supported parts of the AFLP MPT (bootstrap or jackknife support of at least 75%). Thirdly, for intraspecific distances and distances between *L. sativa*, *L. serriola*, *L. dregeana*, and *L. altaica*. These species are probably conspecific (Koopman et al., 2001), and therefore I will refer to this third set of distances as "intraspecific

distances". I use the general term "association" rather than "correlation", since "correlation" only applies to linear associations.

Results

Phylogenetic Signal

The g_1 statistic was -0.52 for AFLP pc E35/M48, -0.46 for pc E35/M49, and -0.50 for the combined data sets. The lettuce data sets contain over 25 taxa and 500 or more variable characters, and therefore the critical value of -0.08 (P = 0.01) was used (Hillis and Huelsenbeck, 1992). All three g_1 values are considerably lower than this critical value, indicating the presence of significant phylogenetic signal in the AFLP data sets. The g_1 statistic for the ITS-1 data set was - 0.59. The g_1 is considerably lower than the critical value of -0.12 (> 25 taxa, 100 variable characters, P = 0.01), indicating significant phylogenetic signal in the ITS-1 data set.

The RASA test for pc E35/M48 showed an observed slope (β_{obs}) of 19.62, an expected slope (β_{null}) of 8.69, and a test statistic t_{RASA} of 31.62, with 3399 degrees of freedom (df). The test for E35/M49 showed a β_{obs} of 20.33, a β_{null} of 8.53, and a t_{RASA} of 34.04 (df = 3399). The combined AFLP data sets showed a β_{obs} of 20.10, a β_{null} of 8.61, and a t_{RASA} of 33.16 (df = 3399). The ITS data set showed a β_{obs} of 13.20, a β_{null} of 6.58, and a t_{RASA} of 22.41, with 986 degrees of freedom. In all cases, t_{RASA} indicates significant phylogenetic signal ($\alpha = 0.05$).

MPT Support

The search on AFLP data set E35/M48 yielded 240 MPTs of 1902 steps, a CI of 0.279, a RC of 0.191, and a RI of 0.685. The search with data set E35/M49 yielded 1238 MPTs of 1835 steps, a CI of 0.272, a RC of 0.181, and a RI of 0.666. The search using the combined data from E35/M48 and E35/M49 yielded 6 MPTs of 3783 steps, a CI of 0.272, a RC of 0.183, and a RI of 0.671. The six trees based on the combined primer combinations (pcs) differed only in a few terminal branches.

One of the MPTs based on the combined pcs is depicted in Fig. 1. The MPT shows two moderately supported clades of species that are in accordance with the generally applied morphology-based classification of Feráková (1977). The first clade comprises all subsection *Lactuca* species in the present study: *L. sativa*, *L. serriola*, *L. dregeana*, *L. altaica*, *L. aculeata*, *L. virosa*, and *L. saligna*. The second clade comprises the section *Mulgedium* species *L. tatarica* and *L. sibirica*, and section *Lactucopsis* species *L. quercina*. Within the first clade, a subclade with *L. sativa*, *L. serriola*, *L. dregeana*, *L. altaica* is well supported (94% bootstrap support, 95% jackknife support, and 9 steps Branch support, respectively), as is a larger clade including *L. sativa*, *L. serriola*, *L. dregeana*, *L. altaica*, and *L. aculeata* (100%, 100%, 16 steps). Two clades within *L. virosa*, possibly identifying intraspecific taxa, also have high supports (100%, 100%, 23 steps; and 100%, 100%, 38 steps). Within the second clade, the subclade with *L. tatarica* and *L. sibirica* is well supported (92%, 95%, 11 steps). For all but a

few species, the bootstrap and jackknife supports were 100%, and the Branch support values exceeded 18 steps (not shown on the MPT). The only exceptions were *L. virosa* (79%, 89%, 3 steps), and *L. sativa*, *L. serriola*, *L. dregeana*, *L. altaica*. The latter four species are probably conspecific (Koopman et al., 2001), and accessions of these species are intermixed in the MPT.

The 50% majority rule consensus trees from the analyses of the separate AFLP data sets also showed the clades and subclades described above, with one exception: in the E35/M48 consensus tree, *L. quercina* did not group with *L. tatarica* and *L. sibirica*.

Congruence

In the PHT test, 17 out of 500 trees were longer than the sum of tree lengths for the original data sets. The corresponding error rate on rejecting the hypothesis of congruence between the AFLP and ITS data sets is 1 - (17/500) = 0.966, meaning that the data sets show significant congruence at $P = 0.034$.

The search with ITS-1 sequences yielded 558 trees of 279 steps, a CI of 0.667, a RC of 0.584, and a RI of 0.876. One of the MPTs is shown in Fig. 2. Visual comparison of the ITS MPT with the AFLP MPT shows a general congruence for the supported parts of the trees.

Similar to the AFLP MPT, the ITS MPT shows a clade with all subsection *Lactuca* species. This clade has 69% Bootstrap, 76% Jackknife, and 3 steps Branch support, while the support for its subclades varies: the subclade with *L. sativa*, *L. serriola*, *L. dregeana*, *L. altaica* has a 63% bootstrap support, 67% jackknife support, and 1 step Branch support, the larger clade including *L. sativa*, *L. serriola*, *L. dregeana*, *L. altaica*, and *L. aculeata* is supported with 96%, 97%, and 4 steps, and the two clades within *L. virosa* are supported with 88%, 88%, 2 steps, and 62%, 61%, 1 step, respectively. The second of the moderately supported clades in the AFLP MPT is only partially reflected by the ITS MPT. In the AFLP MPT this clade comprises *L. tatarica*, *L. sibirica*, and *L. quercina*, whereas in the ITS MPT a clade is present comprising *L. tatarica*, *L. sibirica*, and *L. viminea*. The support values for the clade are low: 61%, 51%, and 0 steps, and there is no supported *L. tatarica/L. sibirica* subcluster (the ITS 50% majority rule consensus tree showed a polytomy with *L. sibirica* and *L. viminea*).

The Templeton test showed significant conflict in tree topologies. Using the AFLP data set to compare the topologies, the AFLP MPT measured 3783 steps, the ITS MPT 4467. The AFLP MPT is significantly shorter (and thus incongruent) at $P < 0.0001$ (test statistic $T = 5185.5$, number of signed-ranks $N = 440$). Using the ITS-1 data set, the ITS-1 MPT measured 279 steps, the AFLP MPT 318. The AFLP MPT is significantly longer than the ITS tree (and thus incongruent) at $P < 0.0001$ ($T = 65$, $N = 34$).
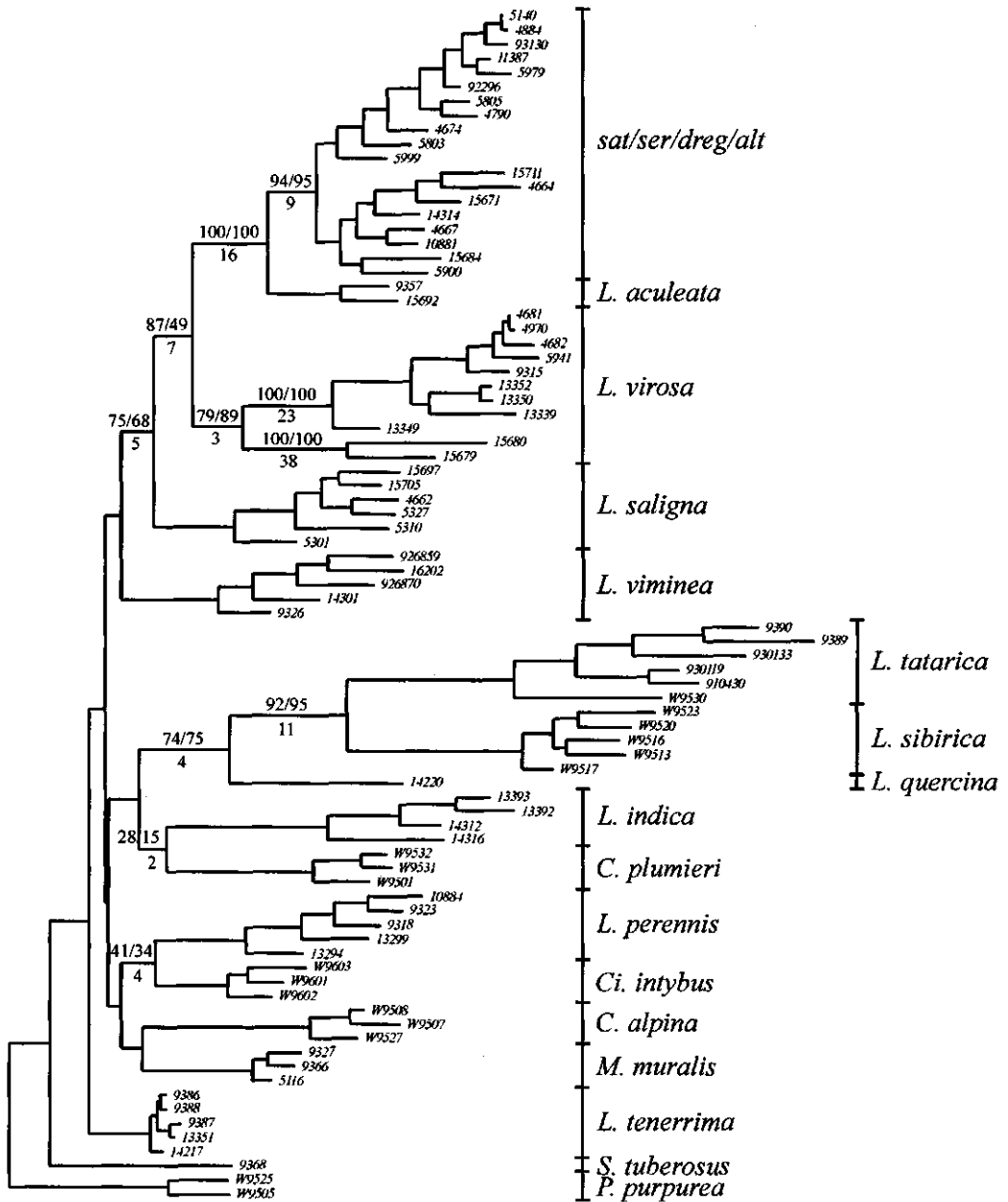
Fig. 1. One of the MPTs based on combined AFLP primercombinations E35/M48 and E35/M49. Above branches: bootstrap values / jackknife values. Below branches: branch supports. *sat/ser/dreg/alt*: group of intermixed and closely related species *Lactuca sativa*, *L. serriola*, *L. dregeana*, and *L. altaica*. Genus abbreviations: *L = Lactuca, C = Cicerbita, Ci = Cichorium, M = Mycelis, S = Steptorhamphus, P = Prenanthes*.
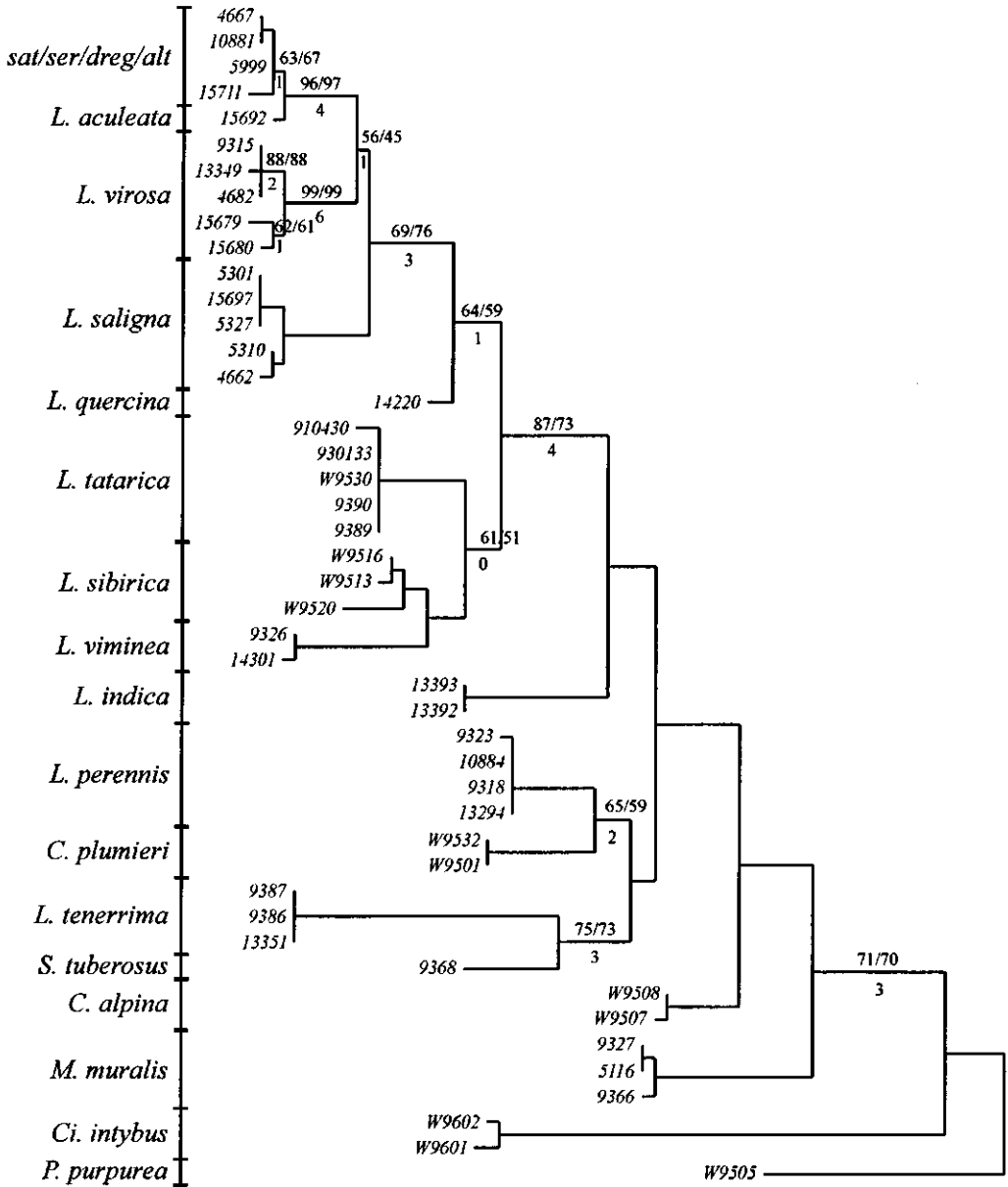
Fig. 2. One of the MPTs based on ITS-1sequences. Above branches: bootstrap values / jackknife values. Below branches: branch supports. *sat/ser/dreg/alt*: group of intermixed and closely related species *Lactuca sativa, L. serriola, L. dregeana,* and *L. altaica.* Genus abbreviations: *L = Lactuca, C = Cicerbita, Ci = Cichorium, M = Mycelis, S = Steptorhamphus, P = Prenanthes.*

Comparison of AFLP phylogenetic distances and ITS phylogenetic distances using regression analyses showed that the distances are significantly associated. For the set of all distances (Fig. 3), the linear association (= correlation) showed an R-squared (R^2) value of 0.021, the second degree polynomial showed $R^2 = 0.257$, and the third degree polynomial showed $R^2 = 0.488$. The fit increased significantly up to the ninth order polynomial ($P = 3.831 * 10^{-39}$, $R^2 = 0.595$), but the higher degree polynomials showed only marginal increase in R^2 values. Therefore, the third order polynomial is shown in Fig. 3 as the "best fit" on the data.
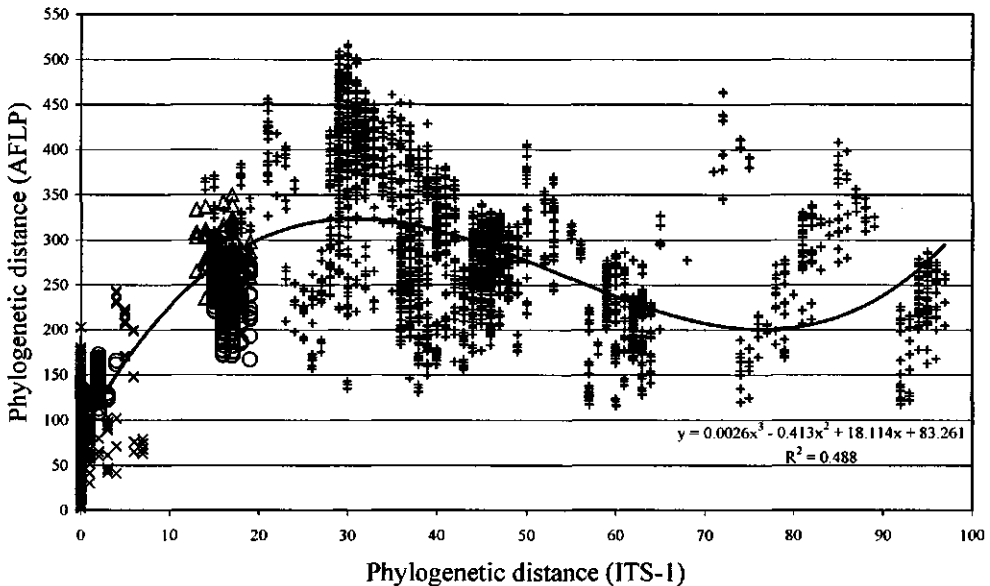
Fig. 3. Association between pairwise phylogenetic distances based on ITS-1 sequences, and pairwise phylogenetic distances based on AFLP markers, for distances among all accessions. × intraspecific distances; ○ interspecific distances within subsect. *Lactuca*; Δ interspecific distances among *L. tatarica*/*L. sibirica*/*L. quercina*; + interspecific distances for remaining (not supported) clades.

For the set of intraspecific distances and distances from moderately supported clades (Fig. 4), the linear association showed $R^2 = 0.782$ ($P = 3.252 * 10^{-272}$), and the second degree polynomial $R^2 = 0.809$ ($P = 4.15 * 10^{-25}$). Higher degree polynomials showed no significant increase in association ($\alpha = 0.01$), and therefore the second degree association was considered the "best fit" on the data. For the set of intraspecific distances (data points marked with × in the left half of Fig. 4) the linear association showed $R^2 = 0.164$, the second degree polynomial $R^2 = 0.190$, and the third degree polynomial $R^2 = 0.244$ ($P = 3.962 * 10^{-6}$). Higher degree polynomials show no significant increase in fit ($\alpha = 0.01$).
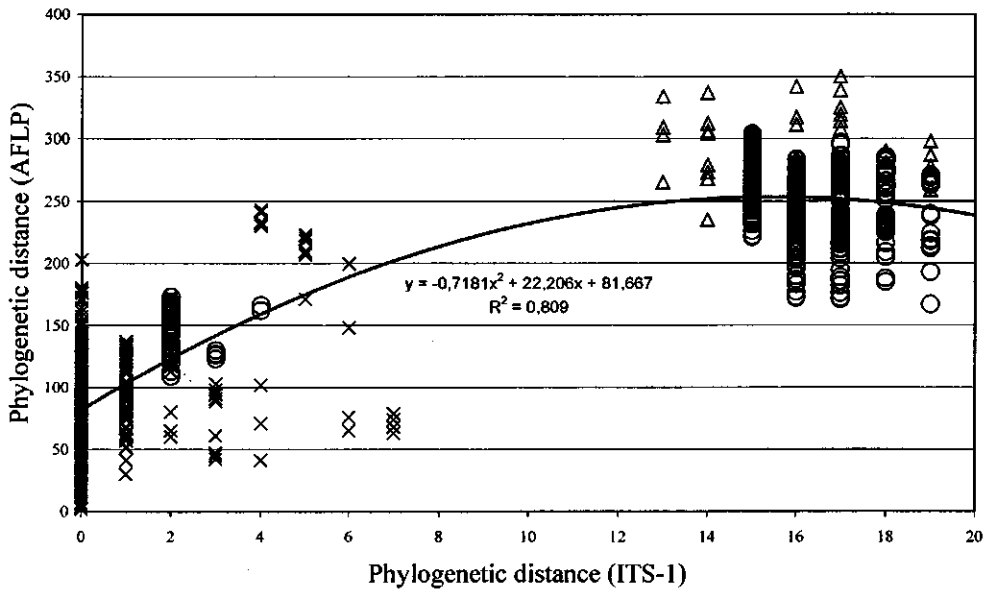
Fig. 4. Association between pairwise phylogenetic distances based on ITS-1 sequences, and pairwise phylogenetic distances based on AFLP markers, for intraspecific distances and distances within supported groups. × intraspecific distances; ○ interspecific distances within subsect. *Lactuca*; Δ interspecific distances among *L. tatarica/L. sibirica/L. quercina*.

Discussion

I examined the presence of phylogenetic signal in AFLP data sets using three approaches. Firstly, I tested for phylogenetic signal in the data sets. Using TLD and RASA, significant phylogenetic signal was detected in the data sets generated with both pcs. The signal for the combined data sets was also significant, indicating that the signal in the separate data sets was not conflicting. An extensive literature search into cladistic AFLP studies (for practical reasons I limited the search to studies involving three or more species) showed that testing AFLP data sets on phylogenetic signal prior to cladistic analysis is not common practice. However, Giannasi et al. (2001) tested TLD of a data set comprising 27 *Trimeresurus* accessions (four species), and reported a g_1 value of -0.66. This values corresponds to a $P < 0.001$, indicating abundant phylogenetic signal in the data set. Apparently, the presence of phylogenetic signal as detected by TLD is not limited to the test data sets used in the present study.

Secondly, I determined branch supports for MPTs. The MPT of the combined *Lactuca* s.l. data sets showed two large clades with moderate support. Within these clades, various smaller clades showed high supports, as did the clades for the individual species. The presence of these clades indicates that at least some parts of the data sets must contain strong phylogenetic

signal. The fact that the moderately supported clades from the MPT based on the combined data set were also present in the 50% majority rule consensus trees of the separate data sets (although with one exception), indicates that the signal in both data sets is not conflicting. My literature survey showed that in most of the cladistic AFLP analyses some kind of support is determined, usually bootstrap values. The general picture in these studies is that at least parts of the trees show high support, but usually poorly supported parts are also present. The studies demonstrate that the presence of well supported topologies is a general phenomenon in MPTs based on AFLP data.

Thirdly, I examined the congruence of AFLP and ITS data sets and MPT topologies. The PHT showed a significant congruence of the AFLP and ITS data sets. Comparison of the AFLP and ITS MPTs demonstrated that the moderately supported parts of the MPTs showed a general similarity, although differences also existed. These differences were reflected in the Templeton test, showing significant topological incongruence. A more detailed comparison of the MPTs using the association between the AFLP and ITS pairwise phylogenetic distances (determined as branch lengths from the MPTs) showed high and significant associations between AFLP and ITS distances, especially when only the supported groups in the AFLP MPT were considered. The association for intraspecific distances was much lower.

The results of the present study are corroborated by previous studies comparing AFLP based and ITS sequence based MPTs in Clavicipitaceae (Tredway et al., 1999), Ustilaginomycetes (Bakkeren et al., 2000), and Bambusoideae (Hodkinson et al., 2000). The studies indicated a general congruence of ITS and AFLP based MPTs, although sometimes accompanied by local conflicts in topology for some of the species. Tredway et al. (1999) examined seven species of *Epichloë*, *Neotyphodium* (the anamorph of *Epichloë*), and *Balansia*. AFLP and ITS MPTs were entirely congruent as to the relationships among *Epichloë* and *Balansia* accessions, but the AFLP MPT was slightly more resolved. The MPTs were entirely in conflict regarding the relationships among the *Neotyphodium* accessions. According to Tredway et al. (1999) this conflict may result from vegetative hybridization between *Neotyphodium* and *Epichloë*. Such hybridization events can result in an evolutionary history for ITS sequences that is not necessarily similar to that of the genome as a whole (see Tredway et al. (1999) for a more detailed discussion). Bakkeren et al. (2000) determined ITS sequences of 13 species of *Ustilago*, *Sporisorium*, and *Tilletia*, and examined a subset of eight species using AFLP markers. The eight species form a single clade in the ITS tree, consisting of two subclades. The one subclade consists of two species, and is basal to the other subclade that shows a polytomy of three smaller clades. The three smaller clades in the polytomy are a branch with one species, a resolved clade with two species, and a polytomy with three species, respectively. The AFLP tree shows the same clades as the ITS tree, but both polytomies present in the ITS tree are resolved in the AFLP tree. Hodkinson et al. (2000) determined ITS sequences of 12 species of *Phyllostachys* and 4 species from related genera, and AFLP patterns for 23 species of *Phyllostachys*. The ITS tree shows two major clades, representing two sections of

Phyllostachys: section *Heteroclada* (two species), and section *Phyllostachys* (10 species). The *Phyllostachys* clade consists of two subclades: one with two species, and one with eight. The AFLP tree shows a similar subdivision in two major clades, but the *Heteroclada* clade also contains one species that occupied the *Phyllostachys* clade in the ITS tree. The subclades within sect. *Phyllostachys* are not reflected in the AFLP tree.

It is interesting to note that for wide range of species, AFLP based phenograms also show a general congruence with ITS based MPTs. Although AFLP phenograms represent overall similarities rather than true phylogenetic relationships, their congruence with ITS based MPTs adds to the general picture that AFLP and ITS data sets contain a similar type of relationship information. Congruence of AFLP phenograms with ITS sequence based MPTs was found in *Eutypa* strains (DeScenzo et al., 1999), *Datura* and *Brugmansia* (Mace et al., 1999a), *Solanum* (Mace et al., 1999b), *Trichophyton* (Gräser et al., 1999), *Phytophthora* species (Brasier et al., 1999), *Cichorium* (Kiers et al., 1999, 2000), *Oxalis* (Tosto and Hopp, 1996, 2000; Emshwiller and Doyle, 1998), the *Microsporum canis* complex (Gräser et al., 2000), Peronosporales (Rehmany et al., 2000), and *Soldanella* (Zhang et al., 2001). Wang et al. (1998) reported congruence of AFLP and ITS results in *Cercospora* without further analysis of the ITS data. Congruence of an AFLP phenogram with ITS-RFLP results was reported by Dunkle and Levy (2000) for *Cercospora zeae-maydis*, and by Bonants et al. (2000) for *Phytophthora*.

Conclusions

All three approaches used in the present study indicated (statistically significant) phylogenetic signal in the *Lactuca* s.l. data sets, although significant conflict also existed in some parts of the AFLP and ITS MPTs. As stated in the introduction, restriction fragment markers have a number of drawbacks that theoretically could lead to a loss of phylogenetic signal in AFLP data sets. The presence of significant signal in the *Lactuca* s.l. test data sets indicates that in practice, the influence of these drawbacks is limited. It should be noted however that the present conclusions only apply to data sets with relatively closely related species, because AFLP markers are highly variable and the proportion of non-homologous fragments increases with taxonomic divergence (O' Hanlon and Peakall, 2000). In data sets including more distantly related taxa, proportions of non-homologous fragments among taxa may become so high that phylogenetic signal is lost. However, data sets can be tested on the presence of phylogentic signal, and (parts of) data sets without signal can be discarded. The exact level of divergence that can be studied varies among taxa, and should be determined for each group separately.

An extensive literature survey revealed topological congruence of (parts of) AFLP and ITS trees in a wide range of taxa, indicating the presence of phylogentic signal in all AFLP data sets. Gross topological incongruence of AFLP and ITS MPTs was reported only by Tredway et al. (1999), but this incongruence was limited to a very specific part of the data set (see

discussion). Thus, the results on *Lactuca* s.l. are corroborated by literature data on a wide range of taxa, indicating that the present study is representative for AFLP data sets in general.

Acknowledgements

References

ANGIOLILLO, A., M. MENCUCCINI, AND L. BALDONI. 1999. Olive genetic diversity assessed using amplified fragment length polymorphisms. Theor. Appl. Genet. 98:411-421.

ARCHIE, J. W. 1989. A randomization test for phylogenetic information in systematic data. Syst. Zool. 38:239-252.

BACKELJAU, T., L. DE BRUYN, H. DE WOLF, K. JORDAENS, S. VAN DONGEN, R. VERHAGEN, AND B. WINNEPENNINCKX. 1995. Random amplified polymorphic DNA (RAPD) and parsimony methods. Cladistics 11:119-130.

BAKKEREN, G., J. W. KRONSTAD, AND C. A. LÉVESQUE. 2000. Comparison of AFLP fingerprints and ITS sequences as phylogenetic markers in Ustilaginomycetes. Mycologia 92:510-521.

BALDWIN, B. G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the Compositae. Mol. Phylogenet. Evol. 1:3-16.

BLACK, W. C. 1993. PCR with arbitrary primers: approach with care. Insect Mol. Biol. 2:1-6.

BONANTS, P. J. M., M. HAGENAAR-DE WEERDT, W. A. MAN IN 'T VELD, AND R. P. BAAYEN. 2000. Molecular characterization of natural hybrids of *Phytophthora nicotianae* and *P. cactorum*. Phytopathology 90:867-874.

BRASIER, C. M., D. E. L. COOKE, AND J. M. DUNCAN. 1999. Origin of a new *Phytophthora* pathogen through interspecific hybridization. P. Natl. Acad. Sci. USA 96:5878-5883.

BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42:795-803.

BREMER, K. 1994. Branch support and tree stability. Cladistics:295-304.

CARPENTER, J. M. 1992. Random Cladistics. Cladistics 8:147-153.

CARPENTER, J. M., P. A. GOLOBOFF, AND J. S. FARRIS. 1998. PTP is Meaningless, T-PTP is Contradictory: A Reply to Trueman. Cladistics 14:105-116.

DESCENZO, R. A., S. R. ENGEL, G. GOMEZ, E. L. JACKSON, G. P. MUNKVOLD, J. WELLER, AND N. A. IRELAN. 1999. Genetic Analysis of *Eutypa* Strains from California Supports the Presence of Two Pathogenic Species. Phytopathology 89:884-893.

DONOGHUE, M. J., R. G. OLMSTEAD, J. F. SMITH, AND J. D. PALMER. 1992. Phylogenetic relationships of Dipsacales based on rbcL sequences. Ann. Mo. Bot. Gard. 79:333-345.

DUNKLE, L. D., AND M. LEVY. 2000. Genetic relatedness of African and United States populations of *Cercospora zeae-maydis*. Phytopathology 90:486-490.

EMSHWILLER, E., AND J. J. DOYLE. 1998. Origins of domestication and polyploidy in oca (*Oxalis tuberosa*: Oxalidaceae): nrDNA ITS data. Am. J. Bot. 85:975-985.

ERIKSSON, T., AND N. WIKSTRÖM. 1996. AutoDecay, version 3.03. Department of Botany, Stockholm University, Stockholm.

FAITH, D. P., AND P. S. CRANSTON. 1991. Could a cladogram this short have arisen by chance alone?: on permutation tests for cladistic structure. Cladistics 7:1-28.

FARRIS, J. S., V. A. ALBERT, M. KÄLLERSJÖ, D. LIPSCOMB, AND A. G. KLUGE. 1996. Parsimony jackknifing outperforms neighbor-joining. Cladistics 12:99-124.

FARRIS, J. S., M. KÄLLERSJÖ, A. G. KLUGE, AND C. BULT. 1995a. Constructing a Significance Test for Incongruence. Syst. Biol. 44:570-572.

FARRIS, J. S., M. KÄLLERSJÖ, A. G. KLUGE, AND C. BULT. 1995b. Testing significance of incongruence. Cladistics 10:315-319.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783-791.

FERÁKOVÁ, V. 1977. The genus *Lactuca* L. in Europe. 122 pp. Univerzita Komenského, Bratislava.

GIANNASI, N., R. S. THORPE, AND A. MALHOTRA. 2001. The use of amplified fragment length polymorphism in determining species trees at fine taxonomic levels: analysis of a medically important snake, *Trimeresurus albolabris*. Mol. Ecol.10:419-426.

GOLDMAN, N., J. P. ANDERSON, AND A. G. RODRIGO. 2000. Likelihood-Based Tests of Topologies in Phylogenetics. Syst. Biol. 49:652-670.

GOLOBOFF, P. A. 1991a. Homoplasy and the choice among cladograms. Cladistics 7:215-232.

GOLOBOFF, P. A. 1991b. Random data, homoplasy and information. Cladistics 7:395-406.

GRÄSER, Y., A. F. A. KUIJPERS, M. EL FARI, W. PRESBER, AND G. S. DE HOOG. 2000. Molecular and conventional taxonomy of the *Microsporum canis* complex. Med. Mycol. 38:143-153.

GRÄSER, Y., A. F. A. KUIJPERS, W. PRESBER, AND G. S. DE HOOG. 1999. Molecular taxonomy of *Trichophyton mentagrophytes* and *T. tonsurans*. Med. Mycol. 37:315-330.

HANSEN, M., T. KRAFT, M. CHRISTIANSSON, AND N.-O. NILSSON. 1999. Evaluation of AFLP in *Beta*. Theor. Appl. Genet. 98:845-852.

HASEGAWA, M., AND H. KISHINO. 1989. Confidence limits on the maximum-likelihood estimate of the hominoid tree from mitochondrial-DNA sequences. Evolution 43:672-677.

HILLIS, D. M. 1991. Discriminating Between Phylogenetic Signal and Random Noise in DNA Sequences. Pages 278-294 *in* Phylogenetic analysis of DNA sequences (M. M. Miyamoto, and J. Cracraft, eds.). Oxford University Press, New York.

HILLIS, D. M., AND J. P. HUELSENBECK. 1992. Signal, Noise, and Reliability in Molecular Phylogenetic Analyses. J. Hered. 83:189-195.

HODKINSON, T. R., S. A. RENVOIZE, G. N. CHONGHAILE, C. M. A. STAPLETON, AND M. W. CHASE. 2000. A Comparison of ITS Nuclear rDNA Sequence Data and AFLP Markers for Phylogenetic Studies in *Phyllostachys* (Bambusoideae, Poaceae). J. Plant Res. 113:259-269.

HUELSENBECK, J. P. 1991. Tree-length distribution skewness: an indicator of phylogenetic information. Syst. Zool. 40:257-270.

HUYS, G., R. COOPMAN, P. JANSSEN, AND K. KERSTERS. 1996. High-Resolution Genotypic Analysis of the Genus *Aeromonas* by AFLP Fingerprinting. Int. J. Syst. Bacteriol. 46:572-580.

KÄLLERSJÖ, M., J. S. FARRIS, A. G. KLUGE, AND C. BULT. 1992. Skewness and permutation. Cladistics 8:275-287.

KARDOLUS, J. P., H. J. VAN ECK, AND R. G. VAN DEN BERG. 1998. The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy (Solanaceae). Plant Syst. Evol. 210:87-103.

KARP, A., O. SEBERG, AND M. BUIATTI. 1996. Molecular Techniques in the Assessment of Botanical Diversity. Ann. Bot.-London 78:143-149.

KEIM, P., A. KALIF, J. SCHUPP, K. HILL, S. E. TRAVIS, K. RICHMOND, D. M. ADAIR, M. HUGH-JONES, C. R. KUSKE, AND P. JACKSON. 1997. Molecular evolution and diversity in *Bacillus anthracis* as detected by amplified fragment length polymorphism markers. J. Bacteriol. 179:818-824.

KIERS, A. M., T. H. M. MES, R. VAN DER MEIJDEN, AND K. BACHMANN. 1999. Morphologically defined *Cichorium* (Asteraceae) species reflect lineages based on chloroplast and nuclear (ITS) DNA data. Syst. Bot. 24:645-659.

KIERS, A. M., T. H. M. MES, R. VAN DER MEIJDEN, AND K. BACHMANN. 2000. A search for diagnostic AFLP markers in *Cichorium* species with emphasis on endive and chicory cultivar groups. Genome 43:470-476.

KISHINO, H., AND M. HASEGAWA. 1989. Evaluation of the Maximum Likelihood Estimate of the Evolutionary Tree Topologies from DNA Sequence Data, and the Branching Order in Hominoidea. J. Mol. Evol. 29:170-179.

KOOPMAN, W. J. M., E. GUETTA, C. C. M. VAN DE WIEL, B. VOSMAN, AND R. G. VAN DEN BERG. 1998. Phylogenetic relationships among *Lactuca* (Asteraceae) species and related genera based on ITS-1 DNA sequences. Am. J. Bot. 85:1517-1530.

KOOPMAN, W. J. M., M. J. ZEVENBERGEN, AND R. G. VAN DEN BERG. 2001. Species relationships in *Lactuca* s.l. (Lactuceae, Asteraceae) inferred from AFLP fingerprints. Am. J. Bot. 88:1881-1887.

LEE, M. S. Y. 2000. Tree Robustness and Clade Significance. Syst. Biol. 49:829-836.

LU, J., M. R. KNOX, M. J. AMBROSE, J. K. M. BROWN, AND T. H. N. ELLIS. 1996. Comparative analysis of genetic diversity in pea assessed by RFLP- and PCR-based methods. Theor. Appl. Genet. 93:1103-1111.

LYONS-WEILER, J., G. A. HOELZER, AND R. J. TAUSCH. 1996. Relative Apparent Synapomorphy (RASA) I: The Statistical Measurement of Phylogenetic Signal. Mol. Biol. Evol. 13:749-757.

MACE, E. S., C. G. GEBHARDT, AND R. N. LESTER. 1999a. AFLP analysis of genetic relationships in the tribe Datureae (Solanaceae). Theor. Appl. Genet. 99:634-641.

MACE, E. S., R. N. LESTER, AND C. G. GEBHARDT. 1999b. AFLP analysis of genetic relationships among the cultivated eggplant, *Solanum melongena* L., and wild relatives (Solanaceae). Theor. Appl. Genet. 99:626-633.

MANTEL, N. 1967. The Detection of Disease Clustering and a Generalized Regression Approach. Cancer Res. 27:209-220.

MEKSEM, K., E. RUBEN, D. HYTEN, K. TRIWITAYAKORN, AND D. A. LIGHTFOOT. 2001. Conversion of AFLP bands into high-throughput DNA markers. Mol. Gen. Genomics 265:207-214.

MICKEVICH, M. F., AND J. S. FARRIS. 1981. The Implications of Congruence in *Menidia*. Syst. Zool. 30:351-370.

MYERS, R. H. 1990. Classical and modern regression with applications. PWS-Kent, Boston.

O' HANLON, P. C., AND R. PEAKALL. 2000. A simple method for the detection of size homoplasy among amplified fragment length polomorphism fragments. Mol. Ecol. 9:815-816.

RADEMAKER, J. L. W., B. HOSTE, F. J. LOUWS, K. KERSTERS, J. SWINGS, L. VAUTERIN, P. VAUTERIN, AND F. J. DE BRUIJN. 2000. Comparison of AFLP and rep-PCR genomic fingerprinting with DNA-DNA homology studies: *Xanthomonas* as a model system. Int. J. Syst. Evol. Micr. 50:665-677.

REHMANY, A. P., J. R. LYNN, M. TOR, E. B. HOLUB, AND J. L. BEYNON. 2000. A comparison of *Peronospora parasitica* (downy mildew) isolates from *Arabidopsis thaliana* and *Brassica oleracea* using amplified fragment length polymorphism and internal transcribed spacer 1 sequence analyses. Fungal Genet. Biol. 30:95-103.

ROHLF, F. J. 1993. NTSYS-pc, Numerical Taxonomy and Multivariate Analysis System., version 1.80. Exeter Software, Setauket.

ROUPPE VAN DER VOORT, J. N. A. M., P. VAN ZANDVOORT, H. J. VAN ECK, R. T. FOLKERTSMA, R. C. B. HUTTEN, J. DRAAISTRA, F. J. GOMMERS, E. JACOBSEN, J. HELDER, AND J. BAKKER. 1997. Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Mol. Gen. Genet. 255:438-447.

SHIMODAIRA, H., AND M. HASEGAWA. 1999. Multiple Comparisons of Log-Likelihoods with Applications to Phylogenetic Inference. Mol. Biol. Evol. 16:1114-1116.

SIEGEL, S. 1956. Nonparametric statistics for the behavioral sciences. McGraw-Hill Book Company, New York.

SIMMONS, M. P., C. P. RANDLE, J. V. FREUDENSTEIN, AND J. W. WENZEL. 2002. Limitations of Relative Apparent Synapomorphy Analysis (RASA) for Measuring Phylogenetic Signal. Mol. Biol. Evol. 19:14-23.

SOKAL, R. R., AND F. J. ROHLF. 1969. Biometry. The principles and practice of statistics in biological research. W. H. Freeman and Company, San Francisco.

STEEL, M. A., P. J. LOCKHART, AND D. PENNY. 1993. Confidence in evolutionary trees from biological sequence data. Nature 364:440-442.

SWOFFORD, D. L. 1999. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates, Sunderland.

SWOFFORD, D. L., AND G. J. OLSEN. 1990. Phylogeny reconstruction. Pages 411-501 *in* Molecular Systematics (D. M. Hillis, and C. Moritz, eds.). Sinauer Associates, Sunderland.

TEMPLETON, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37:221-244.

TOSTO, D. S., AND H. E. HOPP. 1996. Sequence analysis of the 5.8S ribosomal DNA and internal transcribed spacers (ITS1 and ITS2) from five species of the *Oxalis tuberosa* alliance. DNA sequence 6:361-364.

TOSTO, D. S., AND H. E. HOPP. 2000. Suitability of AFLP markers for the study of genomic relationships within the *Oxalis tuberosa* alliance. Plant Syst. Evol. 223:201-209.

TREDWAY, L. P., J. F. J. WHITE, G. B.S., P. V. REDDY, M. D. RICHARDSON, AND B. B. CLARKE. 1999. Phylogenetic relationships within and between *Epichloe* and *Neotyphodium* endophytes as estimated by AFLP markers and rDNA sequences. Mycol. Res. 103:1593-1603.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAN DE LEE, M. HORNES, A. FRIJTERS, J. POT, J. PELEMAN, M. KUIPER, AND M. ZABEAU. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407-4414.

WANG, J., M. LEVY, AND L. D. DUNKLE. 1998. Sibling Species of *Cercospora* Associated with Gray Leaf Spot of Maize. Phytopathology 88:1269-1275.

ZHANG, L.-B., H. P. COMES, AND J. W. KADEREIT. 2001. Phylogeny and quarternary history of the European montane/alpine endemic *Soldanella* (Primulaceae) based on ITS and AFLP variation. Am. J. Bot. 88:2331-2345.

8

Identifying lettuce species (*Lactuca* subsect. *Lactuca*, Asteraceae): A practical application of flow cytometry

W. J. M. Koopman[1]

[1] Biosystematics Group, Nationaal Herbarium Nederland - Wageningen University branch, Wageningen University, Generaal Foulkesweg 37, 6703 BL Wageningen, The Netherlands

Summary

The wild lettuce species *L. serriola*, *L. saligna*, and *L. virosa* are important genitors in lettuce (*L. sativa*) breeding. Identifying these wild species can be problematic because in some cases they look very similar. Flow cytometry was tested for its reliability and general applicability as a tool to distinguish them. Three series of tests were conducted: (1) Tests with three accessions of *L. sativa* and one accession of each of the wild species, repeated three times throughout the year. In each repeat, the mean relative DNA amount of *L. serriola* was significantly higher than that of *L. saligna*, but significantly lower than that of *L. virosa*. The mean relative DNA amount of *L. sativa* did not differ from that of *L. serriola*. (2) Tests with each wild species represented by 10 accessions. Significant differences between the accessions within each species demonstrated the presence of intraspecific variation. Notwithstanding this intraspecific variation, the relative DNA amounts of all accessions of *L. serriola* were significantly higher than that of all *L. saligna* accessions, and significantly lower than that of all *L. virosa* accessions. Therefore, all accessions could be assigned to the appropriate species on the basis of their DNA amounts. (3) Tests with single plants from 10 accessions of each of the wild species. These tests revealed that individual plants of *L. serriola*, *L. saligna*, and *L. virosa* can be reliably identified with flow cytometry, when a *L. serriola* sample of established identity is used as internal reference.

Key words: Asteraceae, DNA content, flow cytometry, genitors, *Lactuca*, lettuce breeding.

Introduction

Cultivated lettuce (*Lactuca sativa* L.) is the world's most important leafy salad vegetable. To cope with the pests and diseases threatening this crop, plant breeders constantly introduce improved lettuce varieties. The wild lettuce species *L. serriola* L., *L. saligna* L., and *L. virosa* L. serve as important gene sources for these new varieties. So far, *L. serriola* provided resistance to downy mildew, corky root and big vein, and *L. saligna* to looper, downy mildew, infectious yellows, and cucumber mosaic. *L. virosa* has been used in lettuce breeding to obtain cultivars with improved color, root system, and texture (McGuire et al., 1993).

Due to their importance as genitors for cultivated lettuce, *L. serriola*, *L. saligna*, and *L. virosa* are extensively studied. The close resemblance of *L. serriola* to *L. saligna* or *L. virosa* gave rise to uncertain or incorrect identifications in some of these studies. Kesseli and Michelmore (1986) indicated uncertain identifications for *L. saligna* and *L. virosa* accessions. Frietema de Vries et al. (1994), and Frietema de Vries (1996) incorrectly identified *L. saligna* CGN 910414 as *L. serriola*. Figure 1 illustrates this close resemblance of *L. serriola* (Fig. 1a, c) to *L. saligna* (Fig. 1b) and *L. virosa* (Fig. 1d) on rosettes of plants from the present study. In recent practice, the close morphological resemblance of *L. serriola* to *L. saligna* hampered identification of these species in some cases (I.W. Boukema, Centre for Genetic Resources, The Netherlands; H.J. Van Eck, Wageningen University; personal communication). Given these identification problems, a quick, easy, and reliable taxonomic tool is needed to distinguish the wild lettuce genitors.

Flow cytometry (De Laat and Blaas, 1984) can be this tool. Koopman and De Jong (1996) demonstrated that flow cytometric determination of relative DNA amounts enabled the distinction of *L. serriola*, *L. saligna* and *L. virosa* in rosette stage. Because the species were represented by only one measurement, the general applicability of this method remained unclear. This paper reports on the follow-up experiments, demonstrating that flow cytometry is a generally applicable tool to reliably distinguish *L. serriola*, *L. saligna*, and *L. virosa*.

Materials and methods

All experiments
In the experiments, cultivated lettuce (*L. sativa*) was represented by three accessions. Each of the wild species *L. serriola*, *L. saligna*, and *L. virosa* was represented by 14 accessions (Tables 1-3). All material was obtained from the Centre for Genetic Resources, the Netherlands (CGN). Details on the accessions are available in the CGN database, currently at http://www.cpro.dlo.nl/CGN/database/. All plants were grown under standard greenhouse conditions. Voucher specimens of all accessions in rosette, bolting, and flowering stage were deposited at the Herbarium Vadense (WAG). The voucher specimens were supplemented with pappus preparations and fruit samples (all experiments), and photographs of the plants at all three stages (experiments 1 and 2).
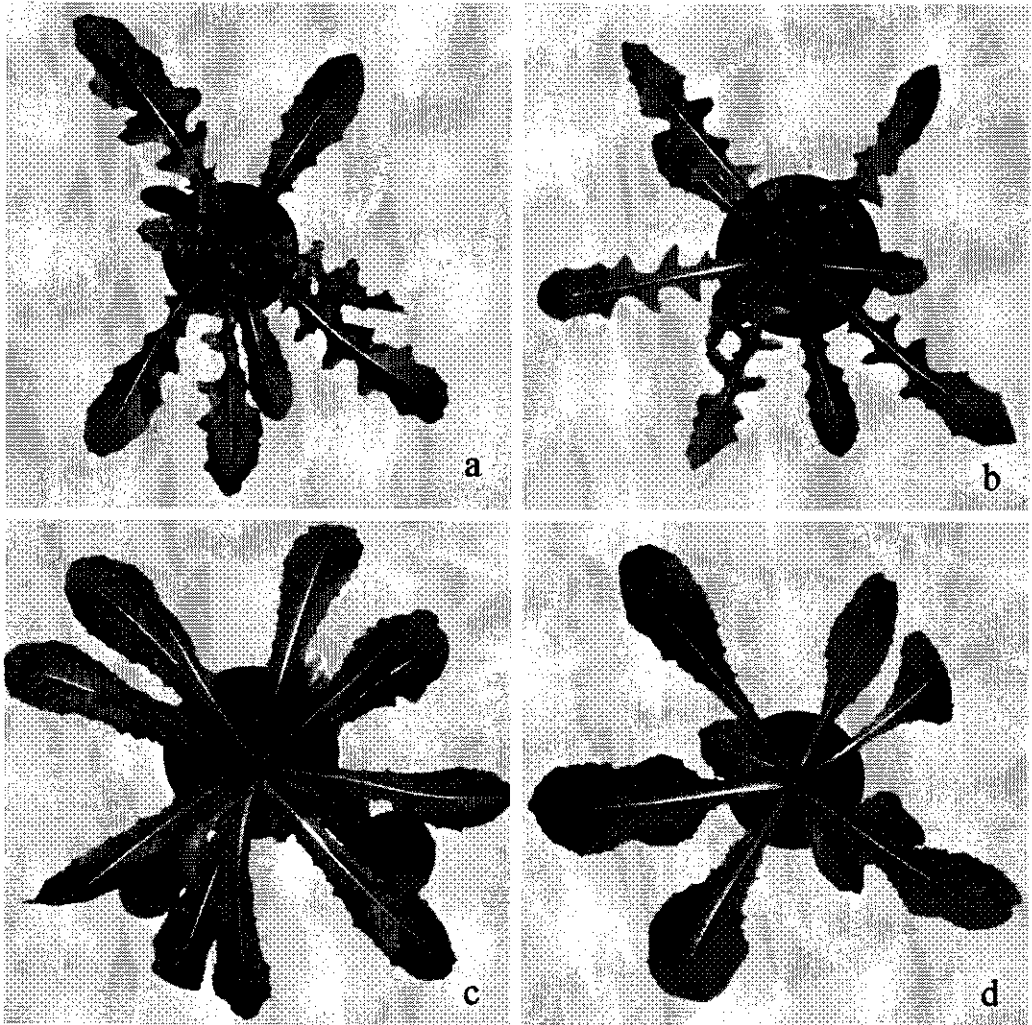
Fig. 1a-d. Rosettes of (**a**) *L. serriola* CGN 16211 resemble those of (**b**) *L. saligna* CGN 15726 quite closely, while rosettes of (**c**) *L. serriola* CGN 10881 look like (**d**) *L. virosa* CGN 9315.

DNA amounts were established by Plant Cytometry Services (Schijndel, The Netherlands) on a PAS II flow cytometer (Partec GmbH, Münster, Germany) equipped with an OSRAM HBO 103 W/2 high pressure mercury lamp, using heat protection filter KG-1, exciter filters UG-1 and BG-38, dichroic mirrors TK-420 and TK-560, and emission filter GG-435. A mixture of fresh leaf material from a 4 weeks old *Lactuca* plant (the "sample") and from a 4 weeks old *Lycopersicon esculentum* Mill. or *Lactuca* spp. reference plant (the "internal reference") was used for each measurement. The material was chopped in an ice-cold neutral buffer modified from De Laat and Blaas (1984), containing 15 mM Hepes, 1 mM EDTA, 80 mM KCl, 20 mM

NaCl, 0.5 mM spermine, 300 mM sucrose, 0.2% Triton X-100, 15 mM DTE, and 2 mg/l DAPI stain. The suspension was passed through a 40 μm nylon filter and analyzed after 15 - 60 min. staining. Note that the DAPI stain is AT preferential (Kapuscinski and Szer, 1979; Manzini et al., 1983). This AT preference makes that the relative DNA amounts reported in the present study not only depend on the absolute DNA amounts of the species, but also on the AT/CG ratio in their nuclear DNA. For the Tables 1-3a, the relative DNA amounts of the *Lactuca* samples were calculated from the obtained DNA histograms by dividing the median values of the G_1 peak of the samples by that of the internal references. To make the results in Tables 3a and 3b comparable, the relative DNA amounts for Table 3b were calculated by dividing the reference values by the sample value. The program package SAS 6.12 (SAS Institute Inc., NC, USA) was used to calculate the prediction interval in experiment 2. The program package JMP version 3.1.4 (SAS Institute Inc., NC, USA) was used to perform all other statistical procedures.

Experiment 1: Repeatability
The stability of the species differences was tested by repeatedly determining the relative DNA amount of three accessions of *L. sativa* and one accession of *L. serriola, L. saligna,* and *L. virosa* (Table 1). In order to include possible seasonal variation in the experiment, the repeats were conducted at different times throughout the year: in winter (23/1/1997), summer (12/6/1997), and autumn (17/10/1997). The three *L. sativa* accessions represent a high, middle, and low value of relative DNA amount within the range found for *L. sativa* by Koopman and De Jong (1996). Each accession was represented by four plants, and the DNA amounts were determined relative to the internal reference *Lycopersicon esculentum* 'Tiny Tim'. An ANOVA was conducted to test for repeat effects and differences among the accessions. A Tukey-Kramer HSD multiple comparison procedure was employed to test for pairwise differences between the accessions ($\alpha = 0.01$). The average coefficients of variation (cv's) of the histogram peaks produced by the flow cytometer were calculated separately for each species and for each repeat. The cv is defined as the standard deviation of a histogram peak expressed as percentage of the mean channel.

Experiment 2: Distinguishing the species
The objective of this experiment was a more extensive examination of the differences in relative DNA amount between the three wild *Lactuca* species demonstrated in Koopman and De Jong (1996). In order to include possible intraspecific differences, each species was represented by 10 randomly chosen accessions (Table 2). This sample of 10 accessions was assumed to adequately represent the variation within a species. Relative DNA amounts of four plants per accession were established using *L. esculentum* 'Tiny Tim' as internal reference. An ANOVA was conducted to test for interspecific and intraspecific differences. A prediction interval ($\alpha = 0.01$) was calculated for each species with SAS 6.12, procedure PROC MIXED,

allowing unequal variance among the species. The cv of the histogram peaks was calculated for each species as the average over all measurements. Estimates of the intraspecific variation were calculated from the lowest and highest mean DNA amounts measured within each species, using the formula: Intraspecific DNA variation = (highest value - lowest value)/lowest value*100. Estimates of the minimal interspecific variation were calculated accordingly, using the DNA values corresponding to the smallest differences between the species.

Experiment 3: Identifying individual plants

In this experiment, the possibility to distinguish individual plants of *L. serriola*, *L. saligna*, and *L. virosa* was examined. Three *Lactuca* accessions that showed an intermediate DNA amount in experiment 2 were used as internal reference, viz. *L. serriola* CGN 5900, *L. saligna* CGN 5147 and *L. virosa* CGN 9315. For each of the species, the test set comprised the three accessions from experiment 2 with the highest DNA amounts, the three accessions with the lowest DNA amounts, and four new accessions with unknown DNA amounts (Tables 3a,b). Two series of tests were conducted, using two plants per accession: (1) The DNA amounts of the *L. saligna* and the *L. virosa* accessions were determined with *L. serriola* CGN 5900 as internal reference (Table 3a). (2) The DNA amounts of the *L. serriola* accessions were determined with *L. saligna* CGN 5147 and with *L. virosa* CGN 9315 as internal reference (Table 3b). Additionally, two control experiments were conducted: (1) To determine the shape of the sample histogram, each accession was measured once without an internal reference. (2) To examine the influence of the internal reference and to detect possible intraspecific variation, each accession was measured once with an internal reference of the same species as the sample. The cv's for the histogram peaks from experiment 3 were calculated as the average of all peaks representing the respective species in all of the measurements.

In experiment 3, the species differences were derived directly from the DNA histograms resulting from the output of the flow cytometer. Because the position of the sample histogram relative to that of the internal reference reveals the identity of the sample directly (see Fig. 2 and discussion), statistical analysis of the results becomes superfluous.

Results

Experiment 1

Table 1 shows the mean relative DNA amounts of four plants per species/repeat combination, as determined for three accessions of *L. sativa*, and one of *L. serriola*, *L. saligna*, and *L. virosa*, respectively, in three replications. The average cv's of the histogram peaks in the first, second, and third repeat were 4.05%; 5.82%; 6.42% for *L. sativa*, 4.18%; 6.05%; 6.10% for *L. serriola*, 5.83%; 7.35%; 7.30% for *L. saligna*, and 4.28%; 5.60%; 6.20% for *L. virosa*, respectively.

Table 1. Means and standard errors of the means for relative DNA amounts in replicated measurements throughout the year. Four plants per accession were measured in each replication. The accessions are indicated by their CGN accession numbers.

Species	Accession	Repeat 23/1	Repeat 12/6	Repeat 17/10
Lactuca sativa	5979	2.657 ± 0.0282	2.821 ± 0.0244	2.716 ± 0.0112
	4707	2.724 ± 0.0052	2.780 ± 0.0175	2.711 ± 0.0017
	4600	2.696 ± 0.0148	2.723 ± 0.0205	2.695 ± 0.0033
Lactuca serriola	10881	2.681 ± 0.0211	2.805 ± 0.0170	2.646 ± 0.0200
Lactuca saligna	5310	2.267 ± 0.0122	2.319 ± 0.0212	2.264 ± 0.0117
Lactuca virosa	9315	3.500 ± 0.0328	3.682 ± 0.0092	3.483 ± 0.0176

The ANOVA showed significant effects of species (P < 0.00005), repeats (P < 0.00005), interaction of repeats and species (P = 0.0002), and interaction of accessions and repeats within species (P = 0.0051). The accessions within species effect was not significant at the 1% level (P = 0.0617). The results of the Tukey-Kramer HSD test ($\alpha = 0.01$) demonstrate the effects in more detail: Within the repeats, *L. saligna* and *L. virosa* differ significantly from each other, and from *L. sativa* and *L. serriola*, while there are no significant differences among the accessions of *L. sativa* and *L. serriola*. The significant difference between the repeats is caused by *L. serriola*, *L. virosa*, and one *L. sativa* accession: For *L. serriola* and *L. virosa*, the second repeat was significantly higher than the other two. For *L. sativa* CGN 5979, the second repeat was significantly higher than the first. No significant repeat differences were found for *L. saligna*. The interaction effects resulted from the presence of three *L. sativa* accessions in the experiment: For some of the *L. sativa* accessions, and between some of the repeats, significant differences were found with *L. serriola* (the interaction of repeats and species), and with other *L. sativa* accessions (the interaction of accessions and repeats within species).

Experiment 2

Table 2 shows the mean relative DNA amounts of four plants per accession for 10 accessions of *L. serriola*, *L. saligna*, and *L. virosa*, respectively. The average cv's of the histogram peaks were 4.02% for *L. serriola*, 3.98% for *L. saligna*, and 4.18% for *L. virosa*.

The ANOVA showed that both the species effect and the accessions within species effect were significant (P < 0.00005). The mean specific relative DNA amounts (mu) and the corresponding 99% prediction intervals are: *L. serriola*: mu = 2.813, (2.626, 3.000); *L. saligna*: mu = 2.385, (2.303, 2.468); *L. virosa*: mu = 3.580, (3.034, 4.125). The intraspecific variation is: 6.9% for *L. serriola*, 2.3% for *L. saligna*, and 14.4% for *L. virosa*. Note the exceptionally high relative DNA amounts of *L. virosa* accessions CGN 15679 and CGN 15680. The intraspecific variation for *L. virosa* drops to 1.5% when these anomalous values are excluded. The interspecific variation is 13.5% between *L. serriola* and *L. saligna*, 18.3% between *L. serriola* and *L. virosa* and 43,6% between *L. saligna* and *L. virosa*.

Table 2. Means and standard errors of the means for relative DNA amounts of 10 accessions of *L. serriola, L. saligna* and *L. virosa*. Four plants were measured for each accession. The accessions are indicated by their CGN accession numbers.

L. serriola	Relative DNA	L. saligna	Relative DNA	L. virosa	Relative DNA
10881	2.744 ± 0.0129	5310	2.398 ± 0.0170	9315	3.490 ± 0.0121
5900	2.805 ± 0.0096	5327	2.388 ± 0.0078	4682	3.490 ± 0.0232
4674	2.814 ± 0.0090	5301	2.373 ± 0.0068	4970	3.483 ± 0.0222
5803	2.886 ± 0.0053	4662	2.402 ± 0.0163	4681	3.467 ± 0.0111
14314	2.930 ± 0.0128	15705	2.376 ± 0.0083	15679	3.896 ± 0.0093
4667	2.741 ± 0.0071	15697	2.363 ± 0.0163	13349	3.520 ± 0.0071
15671	2.786 ± 0.0115	13371	2.414 ± 0.0152	13352	3.478 ± 0.0186
15684	2.841 ± 0.0138	5147	2.388 ± 0.0070	5941	3.520 ± 0.0195
4774	2.817 ± 0.0075	17444	2.389 ± 0.0235	13339	3.485 ± 0.0235
4769	2.765 ± 0.0164	15716	2.360 ± 0.0138	15680	3.965 ± 0.0096

Experiment 3

Table 3a shows the DNA amounts of individual plants of *L. saligna* and *L. virosa*. The DNA amounts of *L. saligna* were determined relative to the internal references *L. saligna* CGN 5147 (the control) and *L. serriola* CGN 5900. The DNA amounts of *L. virosa* were determined relative to the internal references *L. virosa* CGN 9315 (the control) and *L. serriola* CGN 5900.

Table 3a. DNA amounts of *L. saligna* and *L. virosa* accessions relative to controls and to *L. serriola*. 1st column: *L. saligna* CGN accession numbers; 2nd column: DNA amount of *L. saligna* accessions relative to *L. saligna* control; 3rd column: DNA amount of *L. saligna* accessions relative to *L. serriola* internal reference (repeated); 4th column: *L. virosa* CGN accession numbers; 5th column: DNA amount of *L. virosa* accessions relative to *L. virosa* control; 6th column: DNA amount of *L. virosa* accessions relative to *L. serriola* internal reference (repeated). The lowest *virosa/serriola* ratio is in boldface. This ratio corresponds to the smallest difference between *L. virosa* and *L. serriola* found in experiment 3.

Accession	saligna/saligna	saligna/serriola	Accession	virosa/virosa	virosa/serriola
L. saligna			L. virosa		
5310	1.000	0.856 / 0.854	4681	1.000	1.292 / 1.298
4662	1.000	0.849 / 0.837	4970	1.000	1.305 / 1.313
5301	1.000	0.853 / 0.855	5941	1.000	1.300 / 1.305
13371	1.000	0.853 / 0.854	13352	1.000	1.293 / 1.290
15697	1.000	0.852 / 0.852	15679	1.158	1.481 / 1.469
15716	1.000	0.854 / 0.858	15680	1.160	1.473 / 1.485
5329	1.000	0.850 / 0.851	13356	1.000	1.275 / 1.282
15726	1.000	0.862 / 0.851	13361	1.000	1.302 / 1.319
16245	1.000	0.861 / 0.847	14310	1.000	1.292 / 1.295
19049	1.000	0.854 / 0.854	19045	1.000	1.302 / 1.295

Table 3b shows the DNA amounts of individual plants of *L. serriola* relative to the *L. serriola* CGN 5900 internal reference (the control), and relative to the internal references *L. saligna* CGN 5147 and *L. virosa* CGN 9315, respectively. The average cv's of all histogram

peaks in experiment 3 were 4.47% for *L. serriola*, 4.52% for *L. saligna*, and 4.30% for *L. virosa*.

Table 3b. DNA amounts of *L. serriola* accessions relative to a *L. serriola* control, and to *L. saligna* and *L. virosa*. 1st column: *L. serriola* CGN accession numbers; 2nd column: DNA amount of *L. serriola* accessions relative to *L. serriola* control; 3rd column: DNA amount of *L. serriola* accessions relative to *L. saligna* internal reference (repeated); 4th column: DNA amount of *L. serriola* accessions relative to *L. virosa* internal reference (repeated). The highest *saligna/serriola* ratio is in boldface. This ratio corresponds to the smallest difference between *L. saligna* and *L. serriola* found in experiment 3.

Accession	*serriola/serriola*	*saligna/serriola*	*virosa/serriola*
L. serriola			
10881	1.000	0.847 / 0.852	1.310 / 1.307
4667	1.000	0.867 / 0.855	1.316 / 1.313
4769	1.000	0.862 / 0.862	1.305 / 1.292
5803	1.000	0.852 / 0.855	1.292 / 1.297
14314	1.000	0.849 / 0.846	1.290 / 1.308
15684	1.000	0.843 / 0.846	1.302 / 1.295
5119	1.000	0.858 / 0.853	1.313 / 1.320
13374	1.000	0.850 / 0.852	1.305 / 1.305
16211	1.000	0.854 / 0.855	1.302 / 1.299
18664	1.000	0.862 / 0.861	1.321 / 1.323
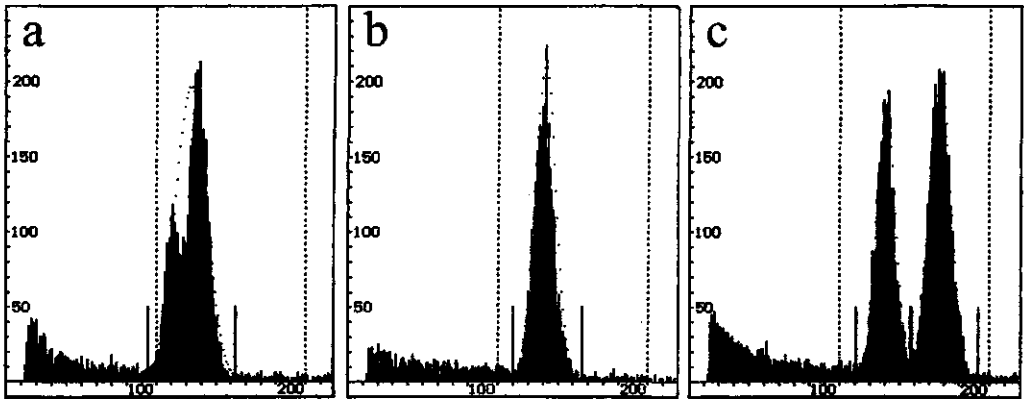
Fig. 2 a-c. Histograms resulting from the output of the flow cytometer, showing the smallest species differences found in experiment 3. On the x-axis: channel number in the measurement. On the y-axis: number of nuclei measured per channel. **a** Peak difference between *L. saligna* CGN 5147 (left peak at ca. 120) and *L. serriola* CGN 4667 (right peak at ca. 140). **b** The peaks of *L. serriola* CGN 5900 and *L. serriola* CGN 4667 coincide at ca. 140. **c** Peak difference between *L. serriola* CGN 5900 (left peak at ca. 140) and *L. virosa* CGN 13356 (right peak at ca. 180).

The values in Tables 3a,b were derived from histograms of DNA amount resulting from the output of the flow cytometer. In all cases, peaks of *L. saligna* appeared to the left of the *L.*

serriola peaks, resulting in ratio's smaller than 1.000, and peaks of *L. virosa* appeared to the right of the *L. serriola* peaks, resulting in ratio's larger than 1.000. Because of these peak positions, the identity of the *Lactuca* samples could be derived directly from the histograms. In cases where a sample and a reference belonged to the same species, the histograms showed coinciding peaks, and the resulting ratio was 1.000 (except for *L. virosa* CGN 15679 and CGN 15680; see discussion). Examples of the peak positions are given in Fig. 2a-c. The histograms in these examples correspond to the smallest species differences found in experiment 3, i.e. the highest *L. saligna/L. serriola* ratio, and the lowest *L. virosa/L. serriola* ratio (boldface values in Tables 3a,b).

Discussion

In this study, the possibility was examined to use flow cytometry as a tool to distinguish wild lettuce species. In experiment 1, various effects were found to be significant at the 1% level: (1) The species effect. The *L. serriola* sample had a relative DNA amount that was significantly higher than that of *L. saligna*, and significantly lower than that of *L. virosa*. (2) The repeat effect. For some accessions, the result of the second repeat differed significantly from that of the other two. (3) Interaction effects. Some *L. sativa* accessions showed significant differences with other *L. sativa* accessions and with *L. serriola*, between some of the repeats. Notwithstanding the repeat and interaction effects, the species differences between *L. serriola*, *L. saligna*, and *L. virosa* were consistently present in all three repeats. In experiment 2, two effects were found to be significant at the 1% level. (1) The species effect. All *L. serriola* accessions had relative DNA amounts that were significantly higher than that of all *L. saligna* accessions, and significantly lower than that of all *L. virosa* accessions. (2) The accession within species effect. Various accessions within each of the species were significantly different from each other.

Two possible sources of variation may have caused the repeat and interaction effects demonstrated in experiment 1. Firstly, the flow cytometer itself. The settings of the flow cytometer may have been slightly different in the second repeat than in the other two. Therefore, comparisons of results from different flow cytometer runs should be avoided in future experiments. Secondly, the plant material. The *Lycopersicon esculentum* reference and the *Lactuca* samples may have responded differently to environmental conditions in the greenhouse during the different seasons. The possibility of such environmental influences on DNA content has been shown for e.g. sunflower (Price and Johnston, 1996; Johnston et al., 1996) and cottonwood (Dhillon, 1988). The risk of such a differential response to environmental conditions can be minimized by using a *Lactuca* species as an internal reference instead of *Lycopersicon esculentum*. Two additional sources of variation are present in both experiment 1 and 2. These are the variation between plants within the accessions, and the variation between accessions within species. The variation between plants within the

accessions can be avoided by measuring only one plant per accession. This is warranted because the variation within the accessions is only small compared to the variation between the accessions (see means and standard errors in Tables 1 and 2). The variation between the accessions cannot be avoided, since it represents the variation range for each of the species. Obviously, these ranges should be included in the experiments.

In experiment 3, all four sources of variation were accounted for. To avoid effects of the flow cytometer settings, each species comparison was done in one run. To minimize the risk of differential responses to environmental factors, *L. serriola* was chosen as internal reference instead of *Lycopersicon esculentum*. *L. serriola* is most suitable as internal reference, because its relative DNA content is intermediate between that of *L. saligna* and *L. virosa* (see Fig. 2a-c). The influence of variation between plants within accessions was tested by measuring two plants for each accession. The results were similar for both plants, showing that one plant per accession suffices for a reliable identification. The influence of variation between accessions within species was tested with control measurements. In these measurements, the accessions of each species were measured relative to a control of the same species. In all cases (except for *L. virosa* CGN 15679 and 15680; see below) the histogram peaks of the accessions and the species reference coincided (Table 3a,b). These coinciding peaks show that the variation between the accessions is so small that it is obscured by the noise in the histogram peaks. In contrast to the accession differences, the species differences are detectable in all cases. This is illustrated by the histogram peaks in Fig. 2a-c. These histogram peaks correspond to the smallest species differences found in experiment 3, and thus to the worst results in that experiment. The fact that these worst results still enable an unequivocal identification of the species proves the reliability of the method. The accession differences, although significant at the 1% level, did not hamper the species identification.

In the three experiments together, each of the wild species was represented by a sample of 82 plants originating from 14 different accessions. This is a sample size larger than in most DNA content studies, and it was assumed to adequately represent the intraspecific variation of each species. Estimates for the intraspecific variation were calculated from experiment 2. The intraspecific variation was 6.9% for *L. serriola*, and 2.3% for *L. saligna*. The intraspecific variation for *L. virosa* is 14.4% when the anomalous accessions CGN 15679 and CGN 15680 are included, and 1.5% when they are excluded. Intraspecific variation in DNA amounts has been reported for numerous species, as was reviewed by Bennett (1985) and Cavallini and Natali (1991). The intraspecific variation in lettuce species as demonstrated in our study is among the lowest reported by them. The interspecific variation found in experiment 2 was 13.5% between *L. serriola* and *L. saligna*, and 18.3% between *L. serriola* and *L. virosa*. Because these interspecific differences are sufficiently larger than the intraspecific differences, all accessions in experiment 3 could be reliably identified. The reliability of the identifications in experiment 2 is illustrated by the 99% prediction intervals. The fact that these intervals do

not overlap indicates that additional plants measured in the future can be assigned to each of the species in a reliable way.

The intraspecific variation of *L. virosa* is higher than that of *L. saligna* or *L. serriola*. This is caused by the relatively high DNA amounts of *L. virosa* accessions CGN 15679 and CGN 15680, both wild collections from Daghestan. These accessions showed a separate peak in the control measurements with approximately 1.16 times the DNA value of the *L. virosa* control (Table 3a). The DNA amount of all other *L. virosa* accessions was identical to that of the *L. virosa* control (value 1.000). Chromosome counts on Giemsa stained root tip squashes of CGN 15679 and CGN 15680 showed that the diploid chromosome number of these accessions was the usual 2n = 18. This excludes the possibility of extra chromosomes contributing to the large genome size. Therefore, the deviating DNA amounts must be caused by unusually large chromosome sizes. These larger chromosome sizes of CGN 15679 and CGN 15680 relative to the other *L. virosa* accessions are probably caused by a larger amount of repetitive DNA sequences in CGN 15679 and CGN 15680 (Flavell et al., 1974; see Flavell, 1986 for a discussion). The anomalous peaks of CGN 15679 and 15680 caused no identification problems, because in the histograms they are clearly visible to the right of the *L. serriola* reference. However, the presence of these anomalous peak values indicates that exceptional karyotypes can be a possible source of identification problems.

Relative and absolute DNA amounts have been determined in numerous plant species over the years. DNA amounts have become a reliable character to establish ploidy levels, and as such have been extensively used for taxonomic purposes (e.g. Huff and Palazzo, 1998). Interspecific differences in DNA amounts within one ploidy level have been reported for numerous genera (see Bennett and Smith (1976, 1991), Bennett et al., (1982) and Bennett and Leitch (1995, 1997) for an overview). In many cases clear differences were found between species with identical chromosome numbers (e.g. Price and Bachmann, 1975; Labani and Elkington, 1987; Hammatt, 1991; Nandini and Murray, 1997). Within *Lactuca*, 2C DNA values have been reported only for cultivated lettuce *L. sativa* (Michaelson et al., 1991; Arumuganathan and Earle, 1991) and for the wild lettuce species *L. serriola* (Bennett and Smith, 1976). Koopman and de Jong (1996) were the first to report on differences in relative DNA amounts between wild lettuce species. Up till now no elaborate study had been carried out to examine the possibility of distinguishing lettuce species by their DNA amounts. The present paper reports on such a study. The lettuce species used are all diploids with a chromosome number of 2n = 18. It was shown that the intraspecific variation in relative DNA amount within each of these species is relatively low. The differences in relative DNA amount between the species are larger. The difference between intra- and interspecific variation enables distinction of the species by their relative DNA amount. It is demonstrated in experiment 3 that one single measurement suffices to make this distinction, when *L. serriola* is used as internal reference. Since *L. serriola*, *L. saligna*, and *L. virosa* are commonly used

genitors in lettuce breeding and distinguishing them is sometimes problematic, the method will prove useful in research and plant breeding practice (Koopman, 1999).

Acknowledgements

References

Arumuganathan, K. & E.D. Earle, 1991. Nuclear DNA content of some important plant species. Plant Mol Biol Reporter 9: 208-218.

Bennett, M.D., 1985. Intraspecific variation in DNA amount and the nucleotypic dimension in plant genetics. In: Freeling, M. (Ed.), Plant genetics: proceedings of the third annual ARCO Plant Cell Research Institute-UCLA symposium on plant biology held in Keystone, Colorado April 13-19 1985, pp. 283--302. Alan R. Liss, Inc., New York.

Bennett, M.D. & J.B. Smith, 1976. Nuclear DNA amounts in angiosperms. Philos Trans, Ser B 274: 227-274.

Bennett, M.D., J.B. Smith & J.S. Heslop-Harrison, 1982. Nuclear DNA amounts in angiosperms. Proc Roy Soc London, Ser B, Biol Sci 216: 179-199.

Bennett, M.D. & J.B. Smith, 1991. Nuclear DNA amounts in angiosperms. Philos Trans, Ser B 334: 309-345.

Bennett, M.D. & I.J. Leitch, 1995. Nuclear DNA amounts in angiosperms. Ann Bot 76: 113-176.

Bennett, M.D. & I.J. Leitch , 1997. Nuclear DNA amounts in angiosperms-583 new estimates. Ann Bot 80: 169-196.

Cavallini, A. & L. Natali, 1991. Intraspecific variation of nuclear DNA content in plant species. Caryologia 44: 93-107.

De Laat, A.M.M. & J. Blaas, 1984. Flow-cytometric characterization and sorting of plant chromosomes. Theor Appl Genet 67: 463-467.

Dhillon, S.S., 1988. DNA analysis during growth and development. In: Hanover, J.W. & D.E. Keithley (Eds.), Genetic manipulation of woody plants, pp. 265--274. Plenum, New York.

Flavell, R.B., 1986. Repetitive DNA and chromosome evolution in plants. Phil Trans R Soc Lond. B: 227-242.

Flavell, R.B., M.D. Bennett, J.B. Smith & D.B. Smith, 1974. Genome size and the proportion of repeated nucleotide sequence DNA in plants. Biochem Genet 12: 257-269.

Frietema de Vries, F.T., R. Van der Meijden & W.A. Brandenburg, 1994. Botanical Files on Lettuce (*Lactuca sativa*). Gorteria suppl. 2.

Frietema de Vries, F.T., 1996. Cultivated plants and the wild flora. PhD Thesis, Rijksherbarium/Hortus Botanicus, Leiden, The Netherlands.

Hammatt, N., N.W. Blackhall & M.R. Davey, 1991. Variation in the DNA content of *Glycine* species. J Exp Bot 42: 659-665.

Huff, D.R. & A.J. Palazzo, 1998. Fine fescue species determination by laser flow cytometry. Crop Sci 38: 445-450.

Johnston, J.S., A. Jensen, D.G. Czeschin, Jr. & H.J. Price, 1996. Environmentally induced nuclear 2C DNA content instability in *Helianthus annuus* (Asteraceae). Amer J Bot 83: 1113-1120.

Kapuscinski, J. & W. Szer, 1979. Interactions of 4', 6-diamidine-2-phenylindole with synthetic polynucleotides. Nucl Acids Res 6: 3519-3534.

Kesseli, R.V. & R.W. Michelmore, 1986. Genetic variation and phylogenies detected from isozyme markers in species of *Lactuca*. J Heredity 77: 324-331.

Koopman, W.J.M., 1999. Plant systematics as a useful tool for plant breeders: examples from lettuce. In: Lebeda, A. & E. Kristkova. Eucarpia Leafy Vegetables '99. Proceedings of the Eucarpia meeting on leafy vegetables genetics and breeding, Olomouc, Czech Republic, 8-11 June, 1999, pp. 95--105. Palacky University, Olomouc.

Koopman, W.J.M. & J.H. De Jong, 1996. A numerical analysis of karyotypes and DNA amounts in lettuce cultivars and species (*Lactuca* subsect. *Lactuca*, Compositae). Acta Bot Neerl 45:211-222.

Labani, R.M. & T.T. Elkington, 1987. Nuclear DNA variation in the genus *Allium* L. (Liliaceae). Heredity 59: 119-128.

Manzini, G., M.L. Barcellona, M. Avitabile & F. Quadrifoglio, 1983. Interaction of diamidino-2-phenylindole (DAPI) with natural and synthetic nucleic acids. Nucl Acids Res 11: 8861-8876.

McGuire, P.E., E.J. Ryder, R.W. Michelmore, R.L. Clark, R. Antle, G. Emery, R.N. Hannan, R.V. Kesseli, E.A. Kurtz, O. Ochoa, V.E. Rubatzky & W. Waycott, 1993. Genetic Resources of Lettuce and *Lactuca* Species in California. An Assessment of the USDA and UC Collections and Recommendations for Long-term Security. Report No. 12, Genetic Resources Conservation Program, Division of Agriculture and Natural Resources, University of California, Davis, California.

Michaelson, M.J., H.J. Price, J.R. Ellison & J.H. Johnston, 1991. Comparison of plant DNA contents determined by feulgen microspectrophotometry and laser flow cytometry. Amer J Bot 78: 183-188.

Nandini, A.V. & B.G. Murray F.L.S., 1997. Intra- and interspecific variation in genome size in *Lathyrus* (Leguminosae). Bot J Lin Soc 125: 359-366.

Price, H.J. & K. Bachmann, 1975. DNA content and evolution in the Microseridinae. Amer J Bot 62: 262-267.

Price, H.J. & J.S. Johnston, 1996. Influence of light on DNA content of *Helianthus annuus* Linnaeus. Proc Natl Acad Sci USA. 93: 11264-11267.

9

Plant systematics as a useful tool for plant breeders: examples from lettuce

W. J. M. Koopman[1]

[1] Biosystematics Group, Nationaal Herbarium Nederland - Wageningen University branch, Wageningen University, Generaal Foulkesweg 37, 6703 BL Wageningen, The Netherlands

Summary

Systematics covers the areas of classification (and its associated activities of nomenclature and identification), phylogeny, and evolutionary processes. Recent examples from lettuce research in these three areas demonstrate the use of plant systematics for plant breeders. Regarding identification: flow cytometry can be used to distinguish the lettuce genitors *L. serriola*, *L. saligna*, and *L. virosa*. Regarding phylogeny: from a phylogenetic study including cultivated lettuce (*L. sativa*) and 18 wild relatives it was concluded that *L. quercina*, *L. tatarica*, and *L. sibirica* are more promising future gene sources than is *L. perennis*. Regarding evolutionary processes: from lettuce karyograms and supplemental literature data it was derived that *L. virosa* is possibly a hybrid species, and that there may be an unknown additional species in the primary or secondary gene pool of cultivated lettuce. Elaborating upon these examples it can be concluded that molecular systematic research offers opportunities for both systematists and plant breeders.

Key words: *Lactuca* spp., classification, phylogeny, evolutionary processes, flow cytometry, karyograms, molecular systematics.

Introduction

In the view of Stuessy (21), the field of systematics covers three related areas, viz. classification with its associated activities of nomenclature and identification, the study of phylogeny, and the study of evolutionary processes. Classification aims at the grouping of individuals into so called taxa, and the assignment of these taxa to the appropriate levels (ranks) in the taxonomic hierarchy. In practice, this means the grouping of individuals into species, species into genera, genera into families, and so on. Intermediate ranks such as sections (within genera) and tribes (within families) may also be applied. In the classification process, similarities and/or differences among individuals and groups are used as grouping criteria, a practice generally known as phenetics. Directly associated with classification are nomenclature, i.e. the naming of the taxa, and identification, i.e. the assignment of individuals to already established taxa. The combined activities of classification, nomenclature, and identification can be regarded as synonymous with taxonomy in a narrow sense.

The study of phylogeny is dealing with the way different groups evolved, and as such aims to establish the patterns resulting from evolutionary processes. In a phylogenetic study, evolutionary relationships among taxa and groups of taxa are determined from data matrices, using special algorithms. These relationships are then depicted in hierarchical structures, called phylogenies or phylogenetic trees. In such trees, evolutionary related taxa are grouped together, the groups are arranged into larger groups, and so on. Each group is connected to an (often putative) common ancestor. Phylogenetic studies are often used to evaluate existing classifications, and in case of discrepancies the classifications may be changed to match the phylogeny.

The third area of systematics, the study of evolutionary processes, can be seen as an extension of the study of phylogeny. The objective of research in this area of systematics is to understand the processes behind the evolutionary patterns. For example: different types of characters can indicate different evolutionary patterns for the same group of taxa. Such a conflict in characters may originate in evolutionary processes like parallel evolution or hybridization. A study of these evolutionary processes in the group of taxa concerned can help to reconstruct the actual course of their evolution.

Considering the three different areas of systematics, it can be postulated that the central goal of systematics is to generate insight in the structure of biodiversity. By generating this insight, the systematist contributes to making the biodiversity more accessible. At the Plant Taxonomy Group of the Wageningen Agricultural University, in a project called "Cytogenetic and molecular genetic characterization of *Lactuca* subsect. *Lactuca* and related species", we aim to do just that for the genetic resources of lettuce. In the present paper, examples from this project will be presented. When these examples are discussed in relation to the relevant literature, the value of plant systematic research for future plant breeding programs becomes clear.

Materials and Methods

Three cases will be discussed, related to the three different areas of systematics. The classification of European *Lactuca* species according to Feráková (5) was used as a starting point. The species *L. sativa*, *L. serriola*, *L. saligna*, and *L. virosa* were examined in cases I (the latter three species) and III (all four species). They are classified in *Lactuca* sect. *Lactuca* subsect. *Lactuca*, and are by far the most important genitors for cultivated lettuce (*L. sativa*). For case II, these four species were supplemented with 10 other *Lactuca* species and with five species from closely related genera (Fig. 1). Most of the material was provided by the Centre for Genetic Resources, The Netherlands (CGN). Details on the accessions used can be found in Koopman et al. (17).

Sect. *Phaenixopus* (Cass.) Benth. — *L. viminea* (L.) J. et C. Presl

Sect. *Mulgedium* (Cass.) C.B. Clarke — *L. tatarica* (L.) C.A. Mey.
— *L. sibirica* (L.) Benth. ex Maxim.

Sect. *Lactucopsis* (Schultz-Bip. ex Vis. et Panc.) Rouy — *L. quercina* L.

Subsect. *Lactuca* — *L. sativa* L.
— *L. serriola* L.
— *L. dregeana* DC.
— *L. aculeata* Boiss. et Kotschy ex Boiss.
— *L. altaica* Fisch. et C.A. Mey.
— *L. saligna* L.
— *L. virosa* L.

Sect. *Lactuca*

Subsect. *Cyanicae* DC. — *L. tenerrima* Pourr.
— *L. perennis* L.

Lactuca indica L.
Mycelis muralis (L.) Dumort .
Cicerbita plumieri (L.) Kirschl.
Cicerbita alpina (L.) Wallr.
Steptorhamphus tuberosus (Jacq.) Grossh.
Prenanthes purpurea L.

Fig. 1. Species used in the various lettuce examples. The European species are classified according to Feráková (5). The non-European species *L. dregeana* and *L. aculeata* are added to subsect. *Lactuca* because of their close relationship with *L. serriola* (25). The additional species are the Asiatic *L. indica* (According to Iwatsuki et al. (11) classified in the non-European section *Tuberosae* Boiss.), and species from genera closely related to *Lactuca*. *L. sativa* is the cultivated lettuce (underlined).

Case I: Identification

Identification of lettuce species using flow cytometry

The close morphological resemblance of *L. serriola* to *L. virosa*, and especially to *L. saligna*, gives rise to identification problems in literature (Kesseli and Michelmore (12) for *L. saligna* and *L. virosa*; Frietema de Vries et al. (7), and Frietema de Vries (6) for *L. serriola* CGN 910414), and in practice (I.W. Boukema, CGN, The Netherlands, and H.J. van Eck, WAU, The Netherlands; personal communication). Previous research using one accession per species, showed that young rosette plants of these accessions could be distinguished by their total DNA amount relative to the tomato internal standard *Lycopersicon esculentum* Mill. 'Tiny Tim' (16). In the follow-up experiment discussed in the current case, the general applicability of flow-cytometric identification of the three lettuce species was examined. In this experiment, each species was represented by 10 randomly chosen accessions, and each accession was represented by four plants. The DNA amount of each plant was determined relative to the 'Tiny Tim' internal standard by Plant Cytometry Services (Schijndel, The Netherlands) using flow cytometry. The differences in relative DNA amounts among the accessions were tested on significance with the statistical program package JMP 3.1.4 (SAS Institute Inc., NC, USA), using a Tukey-Kramer HSD multiple comparison procedure ($\alpha = 0.01$). The differences in relative DNA amounts among the species were quantified by calculating the prediction interval for each species with the statistical program package SAS 6.12 (SAS Institute Inc., NC, USA), procedure PROC MIXED ($\alpha = 0.01$). Additional details on the plant material and the methods are described in Koopman (14).

Case II: Phylogeny

Evaluation of the classification of cultivated lettuce and its (potential) genitors using ITS-1 sequences

In this case, the connection between the position of species in the phylogeny, the gene pool (9) of cultivated lettuce, and the classification of Feráková (5) was examined. Sequences of the Internal Transcribed Spacer 1 (ITS-1) were determined and used as molecular characters for this examination. For most of the species in the experiment, sequences were determined from more than one accession. In the analysis, each species was represented by only one consensus sequence, constructed from the individual sequences in three steps. 1) Align all sequences for all species. 2) For each species, merge the sequences from the accessions representing it, using the ambiguity codes of the NC-IUB for both ambiguous and variable positions. 3) For each species, introduce into the consensus sequence all gaps present in any of the sequences representing that species. The consensus sequences were analyzed with the phylogenetic analysis program PAUP version 3.1.1. (22), using a Branch and Bound search that finds all optimal phylogenetic trees. Additional details are described in Koopman et al. (17).

Case III: Evolutionary processes

Genome evolution within subsect. Lactuca, and postulated hybrid origin of L. virosa

Two chromosome studies were carried out on the subsect. *Lactuca* species *L. sativa*, *L. serriola*, *L. saligna*, and *L. virosa* (15, 16). The first study was carried out on metaphase chromosome complements in spread preparations from root tip cells. The chromosomes were differentially stained with the C-banding, N-banding, and Ag-NOR staining techniques. Using the C- and N-banding techniques, the chromosomes show a pattern of bands and dots that indicate specific classes of DNA. The Ag-NOR staining reveals a metabolically active chromosome region called the nucleolar organizer region (NOR). Studying the banding patterns and the Ag-NOR staining yields information on chromosome constitution. The second study was carried out on metaphase chromosome complements in squash preparations from root tip cells. The chromosomes were stained undifferentially, and their overall morphology was studied. Information about the chromosome constitution and morphology in combination with literature data yielded a general picture of chromosome evolution in the four species studied. Based on this general picture, a scenario for the evolution of the four species was proposed.

Results

Case I: Identification
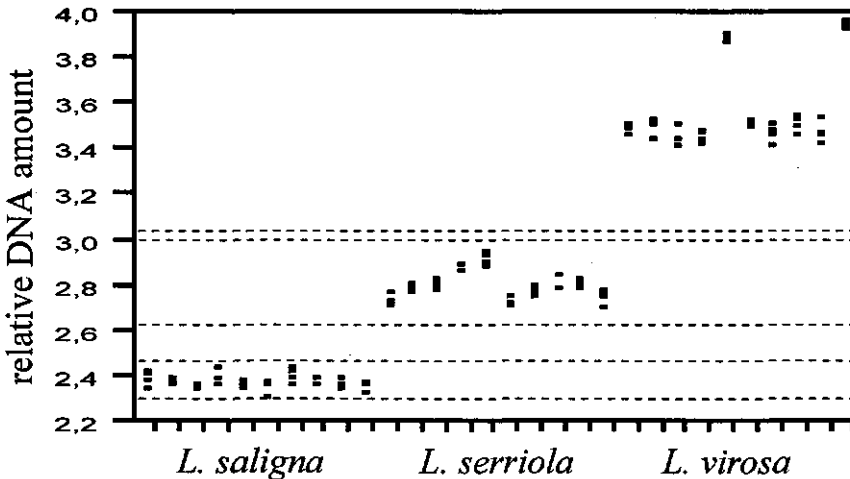
Fig. 2. DNA amounts of the wild lettuce genitors *L. saligna*, *L. serriola*, and *L. virosa*, relative to tomato 'Tiny Tim'. Black squares represent the results for individual plants (four plants per accession were used). The dotted lines above and below each group of accessions indicate the boundaries of the corresponding 99% prediction intervals. The upper boundary for the *L. virosa* accessions is located outside the figure, and is therefore not shown.

Figure 2 shows the relative DNA amounts of *L. serriola, L. saligna,* and *L. virosa.* The mean relative DNA amounts per species (μ) and the corresponding 99% prediction intervals are: *L. serriola:* μ = 2.813, (2.626, 3.000); *L. saligna:* μ = 2.385, (2.303, 2.468); *L. virosa:* μ = 3.580, (3.034, 4.125). As can be seen in Fig. 2, the relative DNA amounts of various accessions within the species overlap, but the three species have clearly different relative DNA amounts. The results of the Tukey-Kramer HSD test showed that the relative DNA amount of all *L. serriola* accessions was significantly higher than that of all *L. saligna* accessions and significantly lower than that of all *L. virosa* accessions. Moreover, the prediction intervals of the species do not overlap, notwithstanding the large variation within *L. virosa* (caused by two accessions with extremely high relative DNA amounts).

Case II: Phylogeny
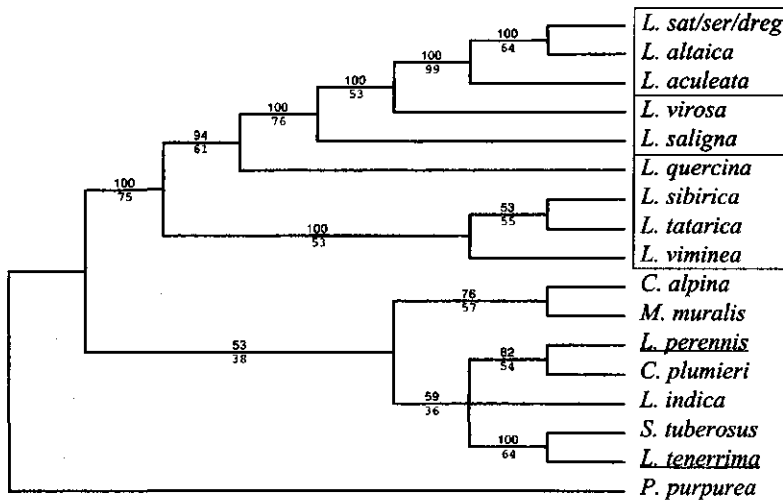
Fig. 3. 50% majority rule consensus tree, depicting the phylogenetic relationships of 19 species from *Lactuca* and related genera. The numbers above the branches indicate the percentage of original trees that showed the corresponding species group. The numbers below the branches are the bootstrap values, indicating the relative support of each group by the data. The box indicates the lettuce gene pool. Upper part of the box: primary gene pool; middle part: secondary gene pool; lower part: tertiary gene pool. Species outside the box are considered to be outside the lettuce gene pool. *L. sat/ser/dreg* = *L. sativa, L. serriola, L. dregeana.*

The PAUP Branch and Bound search resulted in 17 shortest trees, representing 17 slightly different "best guesses" for the actual evolutionary pattern among the 19 species examined. Figure 3 shows a summation of these trees in the form of a 50% majority rule consensus tree. Such a tree only contains those dichotomous splits that lead to groups that were present in more than 50% (= 9 or more) of the original trees. Dichotomous splits leading to groups that were

present in less than 50% of the original trees are condensed into polytomies. In Fig. 3 the only polytomy is the trichotomy leading to the species cluster *L. perennis/C. plumieri/L. indica/S. tuberosus/L. tenerrima*, indicating that the original trees were contradictory as to the evolutionary relationships within this cluster.

Case III: Evolutionary processes
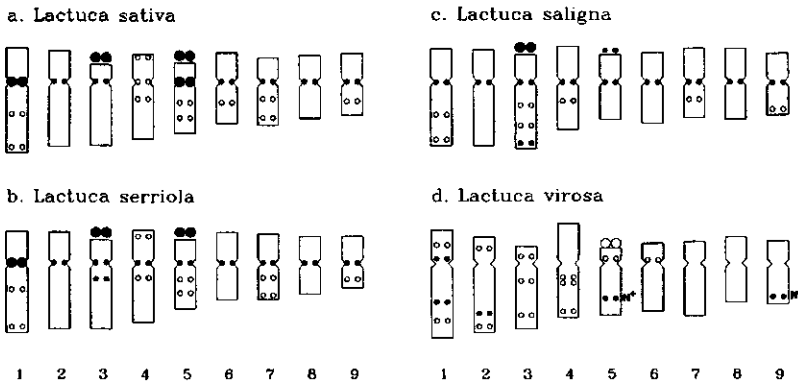
Fig. 4. Chromosome banding idiograms of *L. sativa*, *L. serriola*, *L. saligna*, and *L. virosa*. The numbers below the chromosomes refer to the chromosome pair number, in order of decreasing length. Filled circles indicate obvious C-bands, open circles variable C-bands, N^+ indicates *L. virosa* bands visible with both C- and N-banding. The satellites (circles on top of some of the short chromosome arms) were also visible with Ag-NOR staining.

Figure 4 shows the results of the C-banding, N-banding, and Ag-NOR staining from Koopman et al., (15). As can be seen from the figure, the patterns of *L. sativa* and *L. serriola* are very similar, while the patterns of *L. saligna* and *L. virosa* differ from the *L. sativa/L. serriola* pattern and from each other. The banding patterns of *L. sativa/L. serriola* resemble that of *L. saligna* more closely than that of *L. virosa*. Measurements on the undifferentially stained chromosomes (for details, see 16) confirmed this picture. In these measurements, the chromosomes of *L. sativa* and *L. serriola* were found to be very similar. The chromosomes of *L. virosa* were relatively large (in both length and area) and asymmetrical (large difference between long and short chromosome arm), while the length differences among the chromosomes in one metaphase cell were relatively low. The chromosomes of *L. saligna* were relatively small compared to those of *L. virosa*, more symmetrical, and with higher differences in length among individual chromosomes from one cell. Among the four species studied, *L. sativa/L. serriola* showed the most symmetrical chromosomes. For the remaining characters, the chromosomes of *L. sativa/L. serriola* were morphologically intermediate between those of *L. virosa* and *L. saligna*. The species differences in symmetry within and among the chromosomes can be seen in Fig. 4.

Discussion

Case I: Identification

The results showed that *L. saligna*, *L. serriola*, and *L. virosa* have distinct relative DNA amounts. The prediction intervals showed no overlap, indicating that relative DNA amounts can be used to assign accessions to the species in a reliable way. Supplementary research, using *L. serriola* as internal reference instead of tomato, showed that individual plants of the three species can be identified with a single measurement (Koopman, 2000).

Because *L. saligna*, *L. serriola*, and *L. virosa* are the main wild genitors for cultivated lettuce, they are very important species for lettuce breeding. However, the types of desirable traits as well as the crossability with cultivated lettuce vary considerably among the species, and each will need its own specific approach when used in breeding programs. Therefore, misidentification of plant material can seriously hamper breeding programs and their supporting research. Although up till now no dramatic effects have been reported in literature, uncertain identification of *L. saligna*, *L. serriola*, and *L. virosa* is an actual problem. For instance, Kesseli and Michelmore (12) presented results with uncertain determinations for *L. saligna* and *L. virosa*, while Frietema de Vries et al. (7), and Frietema de Vries (6) incorrectly identified *L. saligna* CGN 910414 as *L. serriola*. The determination of relative DNA amounts could have provided the necessary knowledge in these cases. Even to date, the distinction of especially *L. saligna* and *L. serriola* still poses problems in some cases (I.W. Boukema, CGN, The Netherlands; H.J. van Eck, WAU, The Netherlands; personal communication), while occasionally young plants of *L. serriola* and *L. virosa* can also be hard to distinguish. In the future, these identification problems can be solved with flow cytometric determination of relative DNA amounts: a striking example of the use of plant systematics for research and plant breeding practice.

Case II: Phylogeny

The phylogeny as shown in Fig. 3 depicts the evolutionary relationships among *Lactuca* and related genera, as established by the analysis of ITS-1 sequences. The phylogeny was used to trace evolutionary structure in the gene pool of cultivated lettuce. A subsequent examination of the relation phylogeny/gene pool/classification enabled a prediction of future lettuce genitors, and an evaluation of the classification given in Fig. 1. According to Harlan and de Wet (9) three gene pools can be distinguished: 1) The primary gene pool, containing species that are easy to cross with the crop and yield fully fertile hybrids. 2) The secondary gene pool, containing species that are difficult to cross with the crop and yield hybrids that tend to be sterile. 3) The tertiary gene pool, from which gene transfer is only possible with elaborate technical measures such as embryo culture or bridging species.

According to the available hybridization data (e.g. 2, 18, 23, Koopman, unpublished results), the species *L. serriola*, *L. dregeana*, *L. altaica* and *L. aculeata* belong to the primary gene pool

of cultivated lettuce. From the phylogeny, it can be seen that these species are also the most closely related to cultivated lettuce (upper part of the box in Fig. 3). The phylogeny also shows that *L. saligna* and *L. virosa* are somewhat less related (middle part of the box in Fig. 3) to cultivated lettuce than the primary gene-pool species are. Hybridization data indicate that *L. saligna* and *L. virosa* belong to the secondary gene pool of cultivated lettuce (see 17 for a discussion). Next to the species from the secondary gene pool, a group can be distinguished (lower part of the box in Fig. 3) that is even less related to cultivated lettuce. Literature data indicate that the species in this group, viz. *L. quercina* (sect. *Lactucopsis*), *L. sibirica*, *L. tatarica* (sect. *Mulgedium*), and *L. viminea* (sect. *Phaenixopus*) are (candidates for) the tertiary gene pool. *L. tatarica* can be somatically hybridized with *L. sativa* to produce a fertile hybrid (1, 19), and it is therefore clearly a tertiary gene-pool species. *L. viminea* and *L. sibirica* can be hybridized with *L. virosa* (8) and with *L. tatarica* (Koopman, unpublished results), respectively. Using the latter two as bridging species for crosses with cultivated lettuce would make *L. viminea* and *L. sibirica* accessible as tertiary gene-pool species, too. For the group of species outside the box in Fig. 3, literature data indicate that they are outside the lettuce gene pool (see 17 for a discussion). Considering these results, there seems to be a clear connection between the place of a species in the phylogeny and in the lettuce gene pool.

The classification in Fig. 1 shows that the species that are most closely related to cultivated lettuce (the species from the primary and secondary gene pool) are classified in subsect. *Lactuca*. This indicates a connection between the position of a species in the phylogeny, in the gene pool, and in the classification. Given this connection, it does not seem unreasonable to expect that any additional species classified in this subsection would also contribute to the primary or secondary gene pool. For the group of species next to the primary and secondary gene-pool species, the available hybridization data indicate that they can be considered as (candidate) members of the tertiary gene pool. The connection of phylogeny, gene pool, and classification also seems to hold for this group, since all of these species are related to cultivated lettuce on the section level. Given this connection, additional species classified in the sections *Phaenixopus*, *Lactucopsis*, and *Mulgedium* (see 5) can also be considered candidates for the tertiary gene pool. In deviation from the classification, the phylogeny showed that subsect. *Cyanicae* (*L. perennis* and *L. tenerrima* in Fig. 3) is relatively unrelated to cultivated lettuce. Since it is also outside the lettuce gene pool, the connection between phylogeny and gene pool seems to hold for this subsection. However, the position of subsect. *Cyanicae* in the classification, close to subsect. *Lactuca* (and thus to cultivated lettuce), is not reflected in the phylogeny. Therefore, Koopman et al. (17) proposed an exclusion of subsect. *Cyanicae* from the *Lactuca* classification.

As will be apparent from the results mentioned above, a systematic study in the field of phylogeny reveals the evolutionary structure in the gene pools. The results demonstrated a connection between phylogeny, gene pools, and classification. This connection enabled the identification of new candidates for the tertiary gene pool of cultivated lettuce. Although this

tertiary gene pool is not yet utilized for lettuce breeding, it may well be the gene source of the future. The insights generated with the phylogenetic study will enable a directed choice of genitors from this future gene source: another example of the value of plant systematic research for plant breeders.

Case III: Evolutionary processes

In this case, karyotype information was used in combination with literature data to reconstruct the evolutionary history of species. The results showed that the karyotypes of *L. sativa* and *L. serriola* are very similar. They are intermediate between those of *L. saligna* and *L. virosa*, but resemble the *L. saligna* karyotype more closely than the *L. virosa* karyotype. Based on these insights, the following evolutionary scenario can be postulated: All species were derived from a common ancestor, but *L. virosa* split off much earlier than *L. saligna*. The species *L. serriola* and *L. sativa* are extremely closely related, or even conspecific. This evolutionary scenario is supported by data on DNA amounts (16) and crossability (2, 18, 23). However, data on plant morphology (4), SDS-electrophoresis patterns of seed proteins (3), isozyme analysis of foliar esterases (20), and nuclear AFLPs (10) indicate an alternative scenario, while yet another scenario is suggested by nuclear RFLP data (13) (see 17 for a discussion). This kind of discrepancies between studies based on different types of characters may very well indicate hybridization processes. In the case of lettuce species, a mtDNA RFLP study (24) demonstrated that *L. virosa* and *L. serriola* are maternally closely related. This close maternal relationship of *L. virosa* and *L. serriola*, combined with the fact that the *L. virosa* karyotype and fruit have a deviant morphology compared to *L. sativa*, *L. serriola*, and *L. saligna*, could be interpreted as to indicate that *L. virosa* is a hybrid species, and *L. serriola* (or a *L. serriola* like species) is its female parent.

This evolutionary scenario has two implications for *L. virosa* as a genitor. 1) As a species of hybrid origin, *L. virosa* could be expected to be more variable than a non-hybrid species would be. As a consequence, more interesting traits can be expected from it, and wild resources should be more extensively collected. 2) If *L. serriola* or a closely related species is the female parent of *L. virosa*, it must be compatible with the male parent. Since there is no indication that this parent is present among the well-known *Lactuca* species, an examination of lesser known relatives could yield a totally new lettuce genitor. As becomes clear from these two implications, a plant systematic study of evolutionary processes can provide a better understanding of familiar gene sources. Moreover, it can reveal unexpected opportunities for new gene sources, even in a species group that is considered to be thoroughly known. The value for plant breeders should be clear.

Conclusions

As is demonstrated by the three cases presented in this paper, the various fields of plant systematics provide valuable knowledge for plant breeding practice. Neither the crop nor the

cases have been specially selected to promote plant systematics, they are just examples that were available from the author's recent work. Systematic research provides useful information for other crops as well, as is proven by a pile of literature. With the development of new molecular techniques, plant systematics has become a rapidly evolving field of research. This brings about new chances for both breeders and systematist. They are there, just to be taken.

Acknowledgements

References

1. Chupeau, M-C., Maisonneuve, B., Bellec, Y. & Chupeau, Y. 1994. A *Lactuca* universal hybridizer, and its use in creation of fertile interspecific somatic hybrids. Mol. Gen. Genet. 245: 139-145.
2. De Vries, I.M. 1990. Crossing experiments of lettuce cultivars and species (*Lactuca* sect. *Lactuca*, Compositae). Pl. Syst. Evol. 171: 233-248.
3. De Vries, I.M. 1996. Characterization and identification of *Lactuca sativa* cultivars and wild relatives with SDS-electrophoresis (*Lactuca* sect. *Lactuca*, Compositae). Genet. Resour. Crop Evol. 43: 193-202.
4. De Vries, I.M. & Van Raamsdonk, L.W.D. 1994. Numerical morphological analysis of Lettuce cultivars and species (*Lactuca* sect. *Lactuca*, Asteraceae). Pl. Syst. Evol. 193:125-141.
5. Feráková, V. 1977. The genus *Lactuca* L. in Europe. Univerzita Komenského Press, Bratislava, Czechoslovakia.
6. Frietema de Vries, F.T. 1996. Cultivated plants and the wild flora. Ph.D. Thesis, Rijksherbarium/Hortus Botanicus, Leiden, The Netherlands.
7. Frietema de Vries, F.T., Van der Meijden, R. & Brandenburg, W.A. 1994. Botanical files on lettuce (*Lactuca sativa*). Gorteria suppl. 2.
8. Groenwold, R. 1983. Onderzoek van de relatie tussen *Lactuca* en Bremia *Lactucae*. Verslag van een voorlichtingsbijeenkomst voor slaveredelaars (vervolg). Zaadbelangen 37: 132.
9. Harlan, J.R. & De Wet, J.M.J. 1971. Towards a rational classification of cultivated plants. Taxon 20(4): 509-517.

10. Hill, M., Witsenboer, H., Zabeau, M., Vos, P., Kesseli, R. & Michelmore, R. 1996. PCR-Based fingerprinting using AFLPs as a tool for studying genetic relationships in *Lactuca* spp. Theor. Appl. Genet. 93: 1202-1210.

11. Iwatsuki, K., Yamazaki, T., Boufford, D.E. & Ohba, H. 1995. Flora of Japan, Volume IIIb. Kodansha Ltd., 12-21 Otowa 2-chome, Bunkyo-ku, Tokyo 112-01, Japan.

12. Kesseli, R.V.& Michelmore, R.W. 1986. Genetic variation and phylogenies detected from isozyme markers in species of *Lactuca*. J. Heredity 77: 324-331.

13. Kesseli, R.V., Ochoa, O. & Michelmore, R.W. 1991. Variation at RFLP loci in *Lactuca* spp. and origin of cultivated lettuce (*L. sativa*). Genome 34: 430-436.

14. Koopman, W.J.M. 2000. Identifying lettuce species (*Lactuca* subsect. *Lactuca*, Asteraceae): A practical application of flow cytometry. Euphytica 116: 151-159.

15. Koopman, W.J.M., De Jong, J.H. & De Vries, I.M. 1993. Chromosome banding patterns in lettuce species (*Lactuca* sect. *Lactuca*, Compositae). Pl. Syst. Evol. 185: 249-257.

16. Koopman, W.J.M. & De Jong, J.H. 1996. A numerical analysis of karyotypes and DNA amounts in lettuce cultivars and species (*Lactuca* subsect. *Lactuca*, Compositae). Acta Bot. Neerl. 45(2): 211-222.

17. Koopman, W.J.M., Guetta, E., Van de Wiel, C.C.M., Vosman, B. & Van den Berg, R.G. 1998. Phylogenetic relationships among *Lactuca* (Asteraceae) species and related genera based on ITS-1 DNA sequences. Am. J. Bot. 85(11): 1517-1530.

18. Lindqvist, K. 1960. Cytogenetic studies in the *serriola* group of *Lactuca*. Hereditas 46: 75-151.

19. Maisonneuve, B., Chupeau, M-C., Bellec, Y. & Chupeau, Y. 1995. Sexual and somatic hybridization in the genus *Lactuca*. Euphytica 85: 281-285.

20. Roux, L., Chengjiu, Z. & Roux, Y. 1985. Caractérisation de *Lactuca sativa* L. et des espèces apparentées par électrofocalisation des estérases. Agronomie 5(10): 915-921.

21. Stuessy, T.F. 1990. Plant Taxonomy: The systematic evaluation of comparative data. Columbia University Press, New York.

22. Swofford, D.L. 1993. PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1.1. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.

23. Thompson, R.C., Whitaker, T.W. & Kosar, W.F. 1941. Interspecific genetic relationships in *Lactuca*. J. Agric. Res. 63: 91-107.

24. Vermeulen, A., Desprez, B., Lancelin, D. & Bannerot, H. 1994. Relationships among Cichorium species and related genera as determined by analysis of mitochondrial RFLPs. Theor. Appl. Genet. 88: 159-166.

25. Zohary, D. 1991. The wild genetic resources of cultivated lettuce (*Lactuca sativa* L.) Euphytica 53: 31-35.

Discussion

Species relationships in *Lactuca* s.l.

The first major objective of this thesis was to determine the boundaries and species relationships in *Lactuca* s.l., focusing on: 1) the boundaries and phylogenetic relationships among the closely related species *L. sativa*, *L. serriola*, *L. dregeana*, *L. altaica*, and *L. aculeata* (the "*serriola*-like species"); 2) the evolutionary relationships among *L. sativa*, *L. serriola*, *L. saligna*, and *L. virosa*; 3) the evolutionary relationships in *Lactuca* s.l. in relation to the classification of Feráková (1977).

Boundaries and relationships among the serriola-like species

Considering the *serriola*-like species, the results indicate a close relationship between *L. sativa* and *L. serriola*, apparent from their similar chromosome morphology, and the fact that accessions of these species are intermixed in the ITS-1, AFLP, and DNA content analyses. These results corroborate the results from previous studies on, for example, crossability, morphology, and seed proteins, as discussed in chapters 4 and 5. These previous studies indicated a small overlap between *L. sativa* and *L. serriola*, but the interpretation of this overlap varies among authors. Opposite views are represented by De Vries and Van Raamsdonk (1994), who regard *L. sativa* and *L. serriola* as separate species, and Frietema de Vries (1996), who regards them as conspecific. Frietema de Vries (1996) however, retains *L. sativa* and *L. serriola* as separate subspecies within *L. sativa*, based on four morphological differences (number and distribution of prickles, achene color, shape of inflorescence, and position of involucrum after fruit set). However, these characters are related to selection by man (although this is not sure for inflorescence shape), and in my opinion such characters are not suitable to distinguish taxa. The main reason is that taxa are the result of evolutionary processes, and as such should be distinguished by characters resulting from evolutionary processes. Cultivated forms are man-made, and therefore do not fit the concept of a taxon. A more practical reason is that the use of cultivation-related characters results in highly unstable classifications: the continuous recombination and introduction of characters in crossing and selection programs would require (or at least allow) an ongoing recognition and merger of taxa. Therefore, I favor the opinion of Frietema de Vries (1996) that *L. sativa* and *L. serriola* are conspecific, but I do not corroborate her distinction of *L. sativa* and *L. serriola* as subspecies within *L. sativa*. A separate paper on this subject is in preparation.

Regarding *L. dregeana*, the ITS results showed a close relationship with *L. sativa*, *L. serriola*, and *L. altaica*, the AFLP results showed a close relationship with *L. sativa* and *L. serriola*, and the DNA content and base composition results showed that it is most closely related to *L. sativa*. These results are in line with the conclusion from chapter 5, that *L. dregeana* probably is a 17th century escape from cultivation. However, a definitive conclusion regarding the taxonomic status of *L. dregeana* requires study of the type specimen, which was not available during the present study.

L. altaica was closely related to *L. sativa*, *L. serriola*, and *L. dregeana* in the ITS experiment and in the DNA content and base composition experiment. *L. altaica* was closely related to *L. serriola* in the AFLP experiment. Thus, the results in the present thesis refute suggestions by previous authors (discussed in chapter 4) that *L. altaica* is an intermediate between *L. serriola* and *L. saligna*. The taxonomic position of *L. altaica* was examined more elaborately in an additional study on 23 accessions of wild *Lactuca* species from Uzbekistan, collected by Van Soest (1997). Among these accessions, we distinguished 57 morphotypes of *L. altaica*, *L. serriola*, and *L. saligna*. The study demonstrated that *L. altaica* is probably conspecific with *L. serriola*. A paper on this research is in preparation.

The DNA content and base composition of *L. aculeata* was similar to that of *L. sativa*, *L. serriola*, *L. dregeana*, and *L. altaica*, reflecting the close relationships among the *serriola*-like species. The ITS and AFLP experiments depicted the *L. aculeata* accessions grouping together on a well supported clade, basal to a *L. sativa/L. serriola/L. dregeana/L. altaica* clade. Given the DNA content and base composition data, and the position of *L. aculeata* in the phylogenies, *L. aculeata* must be considered a separate species, but closely related to *L. sativa/L. serriola/L. dregeana/L. altaica*. Given this position, *L. aculeata* is the outgroup of choice for studies on other *serriola*-like species.

Evolutionary relationships among L. sativa, L. serriola, L. saligna, and L. virosa
Considering the evolutionary relationships among *L. sativa*, *L. serriola*, *L. saligna*, and *L. virosa*, all results in the present thesis indicate that *L. sativa* and *L. serriola* are extremely closely related, while *L. saligna* and *L. virosa* are more distantly related to *L. sativa/L. serriola*. Chromosome morphology and DNA amounts (chapters 2, 3, 6, 8, and 9) showed that *L. saligna* is more closely related to *L. sativa/L. serriola* than *L. virosa*, while *L. virosa* and *L. saligna* are most distantly related. The most likely evolutionary scenario is that the species arose in two steps: *L. virosa* split off from a common ancestor first, while the split between *L. saligna* and *L. sativa/L. serriola* occurred later. *L. sativa/L. serriola* evolved into a crop-weed complex, while *L. saligna* evolved as a separate wild species (see chapter 2). The sequence of splits, however, could not be confirmed by the ITS and AFLP studies (chapters 4 and 5). Both studies showed poorly supported branches dividing *L. sativa/L. serriola*, *L. saligna*, and *L. virosa*, uninformative as to the position of the species relative to each other. The uncertainty of the relative positions of *L. sativa/L. serriola*, *L. saligna*, and *L. virosa* is also apparent from other studies, discussed in chapters 2 to 5. The most obvious explanation for the equivocal position of the species is a hybrid origin of *L. virosa* (postulated in chapter 4). The presence of two distinct evolutionary lines within *L. virosa* (chapters 4, 5, 6, 8, 9) may reflect the heterogeneity of the ancestral hybrid population.

Evolutionary relationships in Lactuca s.l. in relation to the classification of Feráková
The ITS and AFLP results revealed two well supported monophyletic groups that are in accordance with the classification of Feráková (1977): subsection *Lactuca* (represented by the *serriola*-like species, *L. saligna*, and *L. virosa*), and section *Mulgedium* (*L. tatarica* and *L. sibirica*). The positions in the cladograms of *L. viminea* (section *Phaenixopus*) and *L. quercina* (section *Lactucopsis*) vary, but the species are always associated with a subsect. *Lactuca* clade or with a sect. *Mulgedium* clade. In contrast, the subsect. *Cyanicae* species *L. perennis* and *L. tenerrima* are always associated with species outside *Lactuca*, but never with *Lactuca* species or with each other. The ITS results showed that *Lactuca* sensu Feráková, excluding subsect. *Cyanicae*, is a well supported monophyletic group. The AFLP results are consistent with the ITS results, but show insufficient resolution to recognize *Lactuca* sensu Feráková (excluding subsect. *Cyanicae*) as a single clade. The DNA amounts are too variable to reliably indicate relationships among the species. In summary, the results in the present thesis corroborate that *Lactuca* sensu Feráková is a monophyletic group, when subsect. *Cyanicae* (we examined *L. tenerrima* and *L. perennis*, but not *L. graeca*) is excluded. However, the sample of species in the present study is too limited to determine the proper position of the subsect. *Cyanicae* species within *Lactuca* s.l..

According to Bremer (1994), *Cichorium intybus* is classified in tribe Lactuceae, but the subtribal affinities of the genus *Cichorium* are uncertain. The view of Bremer (1994) is supported by the results of Kiers et al. (1999) and by the results of studies reviewed in Kiers et al. (1999). In the ITS analysis in chapter 4 of the present thesis, *C. intybus* was more closely related to *L. sativa* than *P. purpurea* was. In the AFLP analyses in chapter 5, *C. intybus* clustered with *L. perennis*. Both analyses indicate a close relationship of *C. intybus* to *Lactuca* s.l., and thus to subtribe Lactucinae. However, the position of *C. intybus* in the ITS and AFLP trees is poorly supported, and the sampling of species from subtribes outside Lactucinae was very limited. Therefore, the ITS and AFLP results must be regarded inconclusive as to the subtribal position of *C. intybus*.

Phylogenetic relationships and the lettuce gene pool

The second main objective of the present thesis was to relate the phylogenetic position of *Lactuca* s.l. species to the position of these species in the gene pool of cultivated lettuce.

The gene-pool concept of Harlan and De Wet (1971) describes three gene pools for a cultivated species. Species from the primary gene pool cross easily with the cultivated species, and the hybrids are generally fertile with good chromosome pairing. For species from the secondary gene pool, gene transfer to the cultivated species is possible, but hybrids tend to be sterile and chromosome pairing is poor or absent. Gene transfer from the tertiary gene pool requires drastic technical measures, and usually yields anomalous or sterile hybrids. Crossability data were available for only a limited number of species, but indicated that the

serriola-like species, closely related to cultivated lettuce, occupy the primary gene pool. The less closely related *L. saligna* and *L. virosa* occupy the secondary gene pool, and form a monophyletic group with the *serriola*-like species in the ITS and AFLP phylogenies (subsect. *Lactuca*; see chapters 4, 5, and 6). These results demonstrate the close connection between the phylogenetic position of a species and its position in the lettuce gene pool, validating the use of *Lactuca* phylogenies to predict the position of *Lactuca* species in the lettuce gene pool.

The added value of the ITS and AFLP phylogenies becomes apparent when the tertiary gene pool is considered. Information on hybridization of *L. sativa* with tertiary gene-pool species at present is limited to *L. viminea* and *L. tatarica* (discussed in chapter 4). However, the ITS and AFLP results showed that all species from sections *Phaenixopus*, *Mulgedium*, and *Lactucopsis* included in the experiments are related to *L. sativa* at a level comparable to that of *L. viminea* and *L. tatarica*. Therefore, these sect. *Phaenixopus*, *Mulgedium*, and *Lactucopsis* species can also be expected to be in the tertiary gene pool. Given the association of these sections to the tertiary gene pool, the *Phaenixopus*, *Mulgedium*, and *Lactucopsis* species not considered in the present research can be expected to be part of the tertiary gene pool as well. Therefore, these species are the most promising future genitors for lettuce. *L. tenerrima* and *L. perennis* clearly are more distantly related to *L. sativa*. They are not part of the lettuce gene pool.

The gene-pool concept of Harlan and De Wet (1971) was developed well in advance of the genomics era. The tertiary gene pool is defined as containing species that require "rather extreme or radical measures" to enable gene transfer to the cultivated species. Examples of these measures given by Harlan and De Wet (1971) include embryo culture, grafting, chromosome doubling, and the use of bridging species. Modern genetic modification procedures such as DNA transformation comply to the description "rather extreme or radical measures", but could not have been anticipated by Harlan and De Wet (1971). Inclusion of species that allow gene transfer to cultivated lettuce using these techniques would include practically all living organisms. Clearly, this can not have been the intention of Harlan and De Wet (1971). Moreover, the resulting definition of a tertiary gene-pool species is not a very useful one. Therefore, the definition of "rather extreme or radical measures" as employed in the present thesis excludes DNA transformation techniques. The advantage of this limitation is that the resulting set of tertiary gene-pool genitors can be employed using techniques generally accepted by environmentalists, without risking GMO related problems.

Secondary objectives

Secondary objectives of the present thesis were to contribute to the development of a theoretical framework for the use of AFLP markers in systematics, and to look for additional practical applications of the research.

In chapter 7, I address the theoretical issue of phylogenetic signal in AFLP data sets. Four features that may make AFLP data unsuited for cladistic analysis were detected, viz. non-

independence of AFLP fragments, non-homology of comigrating AFLP fragments, asymmetry in the probability of a fragment to be gained or lost, and the fact that AFLP markers are usually scored dominantly. Using various statistical procedures, I demonstrated that the *Lactuca* AFLP data sets from chapter 5 contain significant phylogenetic signal. Using the ITS-1 data from chapter 4, I also showed that trees based on the AFLP results do indicate phylogenetic relationships. Comparison with other studies on plants, animals, and fungi indicated that AFLP data contain phylogenetic information as a rule. Therefore, it can be concluded that AFLP data sets generally contain sufficient phylogenetic signal to warrant cladistic analysis, notwithstanding the possible presence of non-independent fragments, non-homologous fragments, the asymmetry in the probability of gaining or loosing fragments, and the fact that AFLP markers are usually scored dominantly.

As a practical application, I examined whether flow cytometry can be used to distinguish the wild lettuce genitors *L. serriola*, *L. saligna*, and *L. virosa*. In practice the distinction of these species is sometimes problematic (discussed in chapter 8), which hinders their use. In chapter 3, differences in relative DNA amounts were detected among *L. serriola*, *L. saligna*, and *L. virosa*. However, the experiments were limited to one accession per species. In chapter 8, it was demonstrated that the specific differences in DNA amounts were consistently present in a sample of 14 accessions per species. This sample was assumed to adequately represent the intraspecific variation of the species. The final experiments showed that leaf samples of arbitrarily selected single plants of each of the species can be identified by their DNA amount relative to a *L. serriola* internal reference. This demonstrates that flow cytometry can be used as a tool to distinguish these wild lettuce genitors.

The use of flow cytometry for species identification, the use of phylogenies for tracing new wild genitors, and the reconstruction of evolutionary processes for a better understanding of the lettuce gene pool are discussed more extensively in chapter 9, as examples of practical applications of the research in the present thesis.

Genome evolution in *Lactuca* s.l.

Although it was not a specific objective at the start of the project, combination of DNA amount (chapters 3, 6, 8) and base composition data (chapter 6) with data on ITS sequences (chapter 4) and AFLP markers (chapter 5) enabled a study of genome evolution in *Lactuca* s.l. (chapter 6). Tracing of DNA and base composition on a combined ITS/AFLP phylogeny indicated a general trend towards increasing genome size in *Lactuca* s.l. This trend was accompanied by a general decrease of the AT/GC nucleotide ratio in the genome, indicating that GC nucleotides are preferentially amplified during *Lactuca* s.l. evolution. However, opposite trends were also apparent, notably in *L. tatarica* and *L. sibirica*. A correlation of genome size and AT content in angiosperms was already suggested by Vinogradov (1994), but recently refuted by Meister and Barow (2001). The results in the present thesis show that the correlation does exist, at least in

Lactuca s.l. Further insight in the genomes of *Lactuca* s.l. was obtained from the association of genome size and number of AFLP bands. Vos et al. (1995) assumed an "almost linear" positive correlation between genome size and number of AFLP bands for relatively small genomes. For larger genomes, the correlation would be lost due to an increasing genome complexity, associated with increasing amounts of repetitive sequences. The results in chapter 6 showed that for *Lactuca* s.l., the lower boundary of a "complex genome" is at a 2C DNA content of 8.5 pg, equaling a haploid genome size of ca. 4165 Mbp. Among the species with DNA contents above 8.5 pg, *L. tatarica* shows an excess of AFLP bands relative to what is expected based on a linear correlation, while *L. indica* and *C. alpina* show a deficit. This may indicate different types of repeat sequences or different processes giving rise to the increase in genome size of *L. tatarica* and *L. indica/C. alpina*, respectively.

Apart from data on DNA content and base composition, data on chromosome morphology were obtained for the subsection *Lactuca* species *L. sativa*, *L. serriola*, *L. saligna*, and *L. virosa* (chapters 2 and 3). The combined data indicate a general lack of genome differentiation between *L. sativa* and *L. serriola*. The karyotypes of *L. sativa* and *L. serriola* are not significantly different in chromosome banding pattern (chapter 2), symmetry, chromosome length, chromosome area (chapter 3), DNA content (chapters 3, 6, 8, 9), or base composition (chapter 6). These features are mainly determined by repetitive sequences in the genome, which therefore must be highly similar (at least quantitatively). RFLP data (Kesseli, Ochoa, and Michelmore (1991); discussed in chapter 2) indicated similarity in unique sequences as well.

Considering genome evolution in subsect. *Lactuca*, the combined data indicate that relative to their common ancestor, the genome size of *L. saligna* decreased, while that of *L. virosa* increased. The genome size of *L. serriola/L. sativa* is intermediate (but more similar to *L. saligna* than to *L. virosa*), showing only a slight increase. In *L. virosa* the changes in genome size were accompanied by large scale chromosomal rearrangements: the chromosomes became longer, more asymmetrical, and two satellites were lost. The disappearance of centromeric C bands and the appearance of N bands indicates that the process was accompanied by large scale qualitative and quantitative changes in heterochromatin. The disappearance of the C bands may indicate that the increase in genome size mainly resulted from interspersed repetitive DNA sequences. Alternatively, it may indicate extensive intrachromosomal rearrangements, breaking up large heterochromatin blocks that were present in the common ancestor. As discussed previously, the general pattern in *Lactuca* s.l. is an increase in genome size through preferential amplification of GC nucleotides. Surprisingly, the subsect. *Lactuca* species do not follow this general pattern. With exception of *L. virosa* CGN 15679/CGN 15680, the changes in genome size among *L. sativa/L. serriola*, *L. saligna*, and *L. virosa* were not accompanied by a preferential amplification of either AT or GC nucleotides (chapter 6, Fig. 1). In *L. saligna*, the direction of karyotype changes was opposite to that in *L. virosa* (i.e. towards more symmetrical chromosomes and more length variation among chromosomes).

Concluding remarks

The present study shows that combining various sources of information on the composition of the nuclear genome provides the opportunity to explore the evolution of the genome itself. In groups with a complex evolutionary history and disputed boundaries among species and genera (such as Lactuceae), the study of genome evolution reveals evolutionary relationships where the study of sequences alone may fail to do so. Our results on the subsection *Lactuca* species demonstrate the success of this "genome" approach on a small scale.

The information obtained on genome evolution of subtribe Lactucinae as a whole was limited by the poor availability of Lactucinae species. However, the research presented in this thesis is a proper basis for a systematic genomics study on a larger scale. This study should include a representative sample of species from Lactucinae or Lactuceae, and use of additional techniques to explore the evolution of genes and genomes. It would yield a wealth of information for both practical application and for a fundamental understanding of the complicated evolution of this group of species.

References

Bremer, K. 1994. Asteraceae: cladistics and classification. Timber Press, Portland, Oregon, USA.

De Vries, I. M., and L. W. D. Van Raamsdonk. 1994. Numerical morphological analysis of Lettuce cultivars and species (*Lactuca* sect. *Lactuca*, Asteraceae). *Plant Systematics and Evolution* 193: 125-141.

Feráková, V. 1977. The genus *Lactuca* L. in Europe. Univerzita Komenského, Bratislava.

Frietema de Vries, F. T. 1996. Cultivated plants and the wild flora. Effect analysis by dispersal codes. PhD thesis, Leiden.

Harlan, J. R., and J. M. J. De Wet. 1971. Towards a rational classification of cultivated plants. *Taxon* 20: 509-517.

Kesseli, R., O. Ochoa, and R. Michelmore. 1991. Variation at RFLP loci in *Lactuca* spp. and origin of cultivated lettuce (*L. sativa*). *Genome* 34: 430-436.

Kiers, A. M., T. H. M. Mes, R. Van der Meijden, and K. Bachmann. 1999. Morphologically defined *Cichorium* (Asteraceae) species reflect lineages based on chloroplast and nuclear (ITS) DNA data. *Systematic Botany* 24: 645-659.

Meister, A., and M. Barow. 2001. Lack of correlation between AT frequency and genome size in higher plants and the effect of non-randomness of base sequences on dye binding. *Cytometry* 47: 1-7.

Van Soest, L. J. M. 1997. Report of expedition to Uzbekistan in 1997, Itinerary, collected materials and data. Centre for Plant Breeding and Reproduction Research (CPRO-DLO), Centre for Genetic Resources, The Netherlands (CGN), Wageningen.

Vinogradov, A. E. 1994. Measurement by Flow Cytometry of Genomic AT/GC Ratio and Genome Size. *Cytometry* 16: 34-40.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407-4414.

Summary

Lactuca sativa (cultivated lettuce) is the world's most important leafy salad vegetable. Apart from *L. sativa*, the genus *Lactuca* contains ca. 75 wild species, potentially useful to improve, for example, taste, texture, and disease resistance of cultivated lettuce. The wild species *L. serriola* (Prickly Lettuce), *L. saligna* (Least Lettuce), and *L. virosa* (Great Lettuce) are commonly used for lettuce improvement.

In preliminary experiments, we established that there is a close connection between evolutionary distances of wild species relative to cultivated lettuce, and their position in the lettuce gene pool (i.e., the possibility to hybridize them with cultivated lettuce). In the present thesis, we established evolutionary relationships among *L. sativa* and 22 wild species in order to predict this position.

We determined that *L. sativa*, *L. serriola*, *L. dregeana*, and *L. altaica* are closely related, and can be regarded as conspecific. *L. aculeata* is closely related to them, but is a distinct species. *L. serriola*, *L. dregeana*, *L. altaica*, and *L. aculeata* occupy the primary gene pool of cultivated lettuce. They can be easily hybridized with cultivated lettuce, and thus are readily accessible gene sources for lettuce improvement. *L. saligna* and *L. virosa* are less closely related to *L. sativa*, and occupy the secondary gene pool (i.e. hybridization with *L. sativa* is possible, but difficult). All primary and secondary gene-pool species can be classified in *Lactuca* sect. *Lactuca* subsect. *Lactuca*. We found that all tertiary gene-pool species (hybridization with *L. sativa* only possible with radical techniques) can be classified in the remaining sections of the genus *Lactuca* (sections *Phaenixopus*, *Mulgedium*, and *Lactucopsis*). These sections are the most promising sources of wild species for future improvement of cultivated lettuce. In the experiments, the tertiary gene-pool species were represented by *L. viminea*, *L. tatarica*, *L. sibirica*, and *L. quercina*. Surprisingly, the species classified in *Lactuca* sect. *Lactuca* subsect. *Cyanicae* are not evolutionary close to cultivated lettuce. They are not part of the lettuce gene pool, and should be excluded from *Lactuca*.

To determine the evolutionary relationships among *L. sativa* and its wild relatives, we examined the genomes of the species at various levels, which provided additional information on genome evolution. We established, that in general the genome sizes in the group increased during evolution, while the ratio of AT/GC nucleotides decreased. Genome complexity for species with 2C DNA amounts below 8.5 pg was similar, but species with 2C DNA amounts exceeding 8.5 pg had more complex and less similar genomes. The species from the primary gene pool share a common ancestor, but the genomes of *L. sativa*/*L. serriola*, *L. saligna*, and *L. virosa*, evolved in different directions.

The present thesis demonstrates that with the proper combination of techniques, a plant systematic study can provide both practically applicable results and fundamental evolutionary insights, thus bridging the gap between fundamental and applied research.

Samenvatting

Lactuca sativa, de cultuursla, is 's werelds belangrijkste verse bladgroente. Naast *L. sativa* bevat het geslacht *Lactuca* circa 75 wilde soorten, die nuttig kunnen zijn om bijvoorbeeld de smaak, eetbaarheid en ziekteresistentie van cultuursla te verbeteren. De wilde soorten *L. serriola*, *L. saligna* en *L. virosa* worden algemeen gebruikt voor de verbetering van cultuursla.

In verkennende experimenten hebben we vastgesteld dat er een nauw verband bestaat tussen de evolutionaire afstand van een soort tot cultuursla, en de positie van die soort in de sla gene-pool (dat wil zeggen, de toegankelijkheid van die soort voor de verbetering van cultuursla). In dit proefschrift hebben we de evolutionaire afstand tussen cultuursla en 22 wilde soorten bepaald, met als doel om de positie van die wilde soorten in de sla gene-pool te voorspellen.

We stelden vast dat *L. sativa*, *L. serriola*, *L. dregeana* en *L. altaica* zeer nauw verwant zijn, en als één soort gezien kunnen worden. *L. aculeata* is een nauw verwante -, maar duidelijk aparte soort. De wilde soorten *L. serriola*, *L. dregeana*, *L. altaica* en *L. aculeata* zitten in de primaire gene-pool van cultuursla (dat wil zeggen, ze zijn er gemakkelijk mee te hybridiseren). *L. saligna* en *L. virosa* zijn minder nauw verwant aan cultuursla, en behoren tot de secundaire gene-pool (hybridisatie met cultuursla is mogelijk, maar moeilijk). Alle soorten uit de primaire en secundaire gene-pool kunnen worden ingedeeld in *Lactuca* sectie *Lactuca* subsectie *Lactuca*. We vonden dat alle soorten uit de tertiaire gene-pool (deze soorten zijn alleen met *L. sativa* te hybridiseren met behulp van radicale technieken) kunnen worden ingedeeld in de overige secties van het geslacht *Lactuca* (secties *Phaenixopus*, *Mulgedium* en *Lactucopsis*). Deze secties zijn daarmee de meest veelbelovende bronnen van wilde verwanten voor de verbetering van cultuursla in de toekomst. In de experimenten werden de soorten uit de tertiaire gene-pool vertegenwoordigd door *L. viminea*, *L. tatarica*, *L. sibirica*, en *L. quercina*. Tot onze verrassing staan de soorten die kunnen worden ingedeeld in *Lactuca* sectie *Lactuca* subsectie *Cyanicae* in evolutionair opzicht niet dicht bij de cultuursla. De subsectie *Cyanicae* soorten behoren niet tot de sla gene-pool, en zouden niet in het geslacht *Lactuca* moeten worden ingedeeld.

Om de evolutionaire relaties tussen cultuursla en zijn wilde verwanten te bepalen, hebben we de genomen (vrij vertaald het DNA) van de soorten op verschillende niveaus onderzocht. Die aanpak leverde, naast informatie over de evolutionaire verwantschappen, aanvullende informatie op over de evolutie van de genomen zelf. De genoom grootte in *Lactuca* en verwante geslachten nam toe in de loop van de evolutie, terwijl de verhouding van AT versus GC nucleotiden afnam. Blijkbaar was er in de loop van de evolutie een selectieve toename van GC nucleotiden. Soorten met een 2C DNA gehalte in het genoom van minder dan 8.5 pg hebben een vergelijkbare genoom complexiteit. Soorten met een 2C DNA gehalte boven 8.5 pg DNA hebben meer complexe genomen, met meer verschillen tussen de soorten. De soorten uit de primaire – en secundaire gene-pool hebben een gemeenschappelijke voorouder, maar de genomen van *L. sativa/L. serriola*, *L. saligna* en *L. virosa* ontwikkelden zich in verschillende richtingen.

De resultaten in dit proefschrift laten zien dat, met een goed gekozen combinatie van technieken, een plantensystematische studie tegelijkertijd praktisch bruikbare resultaten en fundamentele evolutionaire inzichten op kan leveren.

Acknowledgements

Curriculum vitae

Wilhelmus Johannes Maria Koopman was born on 26 oktober 1963 in Grootebroek, The Netherlands. He studied Plant Breeding at Wageningen Agricultural University from 1982 till 1988. M.Sc. projects included plant systematics ("Cytotaxonomy of the *Lactuca serriola* complex"), in vitro culture ("In vitro vegetative propagation of *Hippeastrum* hybrids from flower-stalk explants"), fytopathology ("From *Sporodesmium* to *Alternaria*, everything you always wanted to know about *Alternaria dauci*"), and informatics ("Adjustment of the program "IMAGE" to binary format"). A practical training period was spent at Royal Sluis (now Seminis Vegetable Seeds) in Enkhuizen, conducting research on carrot and gherkin. In the first years after graduation, he worked in tissue culture: in 1988-1989 he was head of the tissue culture laboratory of "Enthoven Geranium" in Wateringen, and in 1989-1990 he worked as a researcher for the EU project "The genetic improvement of *Linum usitatissimum* L. (flax and linseed)" at the Foundation for Agricultural Plant Breeding (SVP, now part of Plant Research International). In 1991 he started work at the Department of Plant Taxonomy of Wageningen Agricultural University (now the Biosystematics Group of Wageningen University). From 1991 till 1994 he examined the crossability of wild tuber-bearing *Solanum* species, testing the so called "Endosperm Balance Number" hypothesis. From 1995 till 2002 he worked on the Ph.D. project "Cytogenetic and molecular genetic characterization of lettuce (*Lactuca* subsect. *Lactuca*) and related species", that was co-funded by seven Dutch plant breeding companies. At present, he holds a research position at the Business Unit Biodiversity and Identity of Plant Research International, studying the genetic diversity of indigenous plant material.

Publications

Koopman, W. J. M., J. Hadam, and J. Doležel. In Prep. Evolution of genome size and base composition in *Lactuca* (Asteraceae) and related genera.

Koopman, W. J. M. Submitted. Phylogenetic signal in AFLP data sets.

Koopman, W. J. M. Submitted. A format for Asteraceae model voucher collections, developed in *Lactuca* L.

Abdel Khalik, K., L. J. G. van der Maesen, W. J. M. Koopman, and R. G. van den Berg. Submitted. Numerical taxonomic study of some tribes of Brassicaceae from Egypt.

Sosef, M. S. M., Y. Issembe, H. P. Bourobou Bourobou, and W. J. M. Koopman. Submitted. Botanical diversity of the Pleistocene forest refuge Monts Doudou (Gabon).

Koopman, W. J. M., M. J. Zevenbergen, and R. G. van den Berg. 2001. Species relationships in *Lactuca* s.l. (Lactuceae, Asteraceae) inferred from AFLP fingerprints. *American Journal of Botany* 88: 1881-1887.

Koopman, W. J. M. 2000. Identifying lettuce species (*Lactuca* subsect. *Lactuca*, Asteraceae): A practical application of flow cytometry. *Euphytica* 116: 151-159.

Koopman, W. J. M. 1999. Plant systematics as a useful tool for plant breeders: examples from lettuce. *In*: Lebeda, A. & E. Křístková. Eucarpia Leafy Vegetables '99. Proceedings of the Eucarpia meeting on leafy vegetables genetics and breeding, Olomouc, Czech Republic, 8-11 June, 1999, pp. 95-105. Palacký University, Olomouc.

Koopman, W. J. M., E. Guetta, C. C. M. van de Wiel, B. Vosman, and R. G. van den Berg. 1998. Phylogenetic relationships among *Lactuca* (Asteraceae) species and related genera based on ITS-1 DNA sequences. *American Journal of Botany* 85: 1517-1530.

Koopman, W. J. M., and J. H. de Jong. 1996. A numerical analysis of karyotypes and DNA amounts in lettuce cultivars and species (*Lactuca* subsect. *Lactuca*, Compositae). *Acta Botanica Neerlandica* 45: 211-222.

Koopman, W. J. M., J. H. de Jong, and I. M. de Vries. 1993. Chromosome banding patterns in lettuce species (*Lactuca* subsect. *Lactuca*, Compositae). *Plant Systematics and Evolution* 185: 249-257.

Koopman, W. J. M., H. H. M. Steegmans, en R. L. M. Pierik. 1987. Stengel *Hippeastrum* produceert vele bollen. *Bloembollencultuur* 31: 18-20.

Koopman, W. J. M., M. B. M. Bruins, en E. Flipse. 1987. Het belang van stage tijdens de opleiding. *Prophyta* 9: 188-189.

On the cover: *Lactuca serriola* L. CGN 10881: rosette plant, chromosomes at metaphase, histogram of relative DNA content, AFLP fingerprint (right-most lanes), ITS-1 sequence. Cover design: W. J. M. Koopman. Photo rosette plant: S. Massalt.