Chapter 10. Relationship between Chinese chive (*Allium tuberosum*) and its putative progenitor *A. ramosum* as assessed by random amplified polymorphic DNA (RAPD).

Frank R. Blattner and Nikolai Friesen

A. Introduction

Chinese chive, also called Chinese leek, is the second-most economically important crop species of the onion genus Allium in Eastern Asia, and is widely cultivated throughout China, Korea, Vietnam, and Japan. It has been introduced to most other Asian countries and more recently to the Caribbean Islands, the United States, and some parts of Europe (Hanelt 2001). Leaves and flower scapes are used as vegetable or in salad (often called ‘garlic sprouts’ in Chinese restaurants), and young inflorescences make a tasty soup.

Hanelt (2001) proposed that domestication of Chinese chive took place in northern China more than 3000 years ago because this crop is mentioned in the classic Chinese ‘Book of Poetry’, which was compiled during the Chou dynasty at the beginning of the first millennium B.C. Populations of an *Allium* that closely resembles the modern crop species and, thus, are thought to be ancestral to the cultigen, occur in steppes and dry meadows in southern Siberia, Mongolia, northern China, and North Korea. These wild plants are treated by some taxonomists as a distinct species, *A. ramosum* L., with the crop being referred to as *A. tuberosum* Rottl. ex Spreng. (for additional taxonomic details see Stearn 1944). However, the division of these taxa is controversial, and Hanelt (1988, 2001) in his latest accounts on these *Allium* species subsumed all forms within *A. ramosum*. Wild and cultivated forms are slightly distinct with respect to morphology and

differ in life history traits. The wild populations flower in summer (June to July) and possess narrow tepals and short filaments, whereas the cultivated forms flower later in the year (August to October) and have broader tepals as well as long filaments. Although Hanelt (1988) reports substantial morphological variability and, particularly in Mongolia, the occurrence of morphologically transitional types in the wild forms, we will informally treat wild and cultivated forms here as two species, just to simplify our reference to the two forms. The taxonomic consequences of our research, i.e., if it is desirable to formally divide the plants into two species or merge them into one, will be discussed at the end of this chapter.

The archeological record of early *Allium* crops is scarce and does not provide a reliable estimation of domestication areas and wild progenitors of the crop species. Therefore, modern collections of the cultivated species and their putative wild relatives are the major sources of information to reveal location, time, and mode of domestication. Here, we describe an analysis of cultivated *A. tuberosum* germplasm from Eastern Asia and *A. ramosum* genotypes from the wild in our attempt to define the geographical vicinity of Chinese chive domestication and to elucidate the mode of this process (i.e., if a single domestication event occurred or if separate domestications took place in parallel in multiple regions). To study relationships among the wild and cultivated accessions, we used an anonymous genetic marker approach called random amplified polymorphic DNA analysis (RAPD), which enables one to detect small differences among the genomes of the individuals surveyed.

A. Prior Research

Studies designed to reveal the origins of *Allium* crops are rare, possibly due to the relatively limited economic importance of these cultigens. In addition, the morphological similarities of closely related *Allium* taxa hampers identification at a glance. Thus, several incorrectly identified plants are maintained in living collections at botanical gardens throughout the world. Furthermore, there is the potentiality for inter-taxon hybridization due to the allowance of open pollination in gardens. Consequently, scientists should be suspicious of seeds obtained from these institutions. The inclusion of garden material often leads to peculiar and conflicting results in investigations of closely related species in studies of *Allium* phylogeny (Friesen et al. 1999; Fritsch et al. 2001; Klaas and Friesen 2002). The origin of the studied material, therefore, is crucial for the interpretation of data in this genus. Consequently, we have created a living collection of *Allium* taxa, where most plants are propagated vegetatively from material originally collected in the wild.

Molecular studies on major *Allium* crops have been performed for *A. ampeloprasum* L. (leek; Kik et al. 1997), *A. cepa* L. (common onion; Havey 1992; van Raamsdonk et al. 2000; Fritsch et al. 2001), *A.* % *cornutum* Clementi et Visiani (triploid onion) and *A. oschaninii* B. Fedtsch. (French gray shallot; Friesen and Klaas 1998), *A. fistulosum* L. (Japanese bunching onion; Friesen et al. 1999), *A. sativum* L. (garlic; Maass and Klaas 1995), and *A. schoenoprasum* L. (chives; Friesen and Blattner 2000). In these studies, different molecular marker techniques were used to identify progenitor species (or populations) and to reveal the geographic origins of the wild plants that were initially involved in the domestication process. Clear identification of crop progenitor was only

possible for two species. For *A. fistulosum*, the wild bunching onion (*A. altaicum* Pall.) from southern Siberia and Mongolia was found to be the wild progenitor to the crop. For the French gray shallot, wild populations of *A. oschaninii* were revealed as ancestral to domesticated germplasm in the species. In most other species, no clear results could be obtained, partly due to identification problems with the plant material involved (e.g., Dubouzet et al. 1997), and partly because only one or a small number of accessions of the putative wild progenitor were included in the studies (e.g., van Raamsdonk et al. 2000). When few accessions are used, it is not always possible to distinguish between a wild species being the closest relative (i.e., sister group) to the crop or being the direct progenitor. To resolve this problem, it is necessary to include several accessions of the wild species, with these accessions representing the majority of genetic variation within the taxon. Only then can a phylogenetic tree reveal the relationships among a crop species and its closest relatives, allowing one to draw conclusions concerning the process of domestication (Heun et al. 1997; Friesen et al. 1999; Badr et al. 2000). One caveat, however, is that it is possible that the wild species (or ancestral populations) from which domestication started may be extinct or might not yet be known to the scientific community (Fritsch et al. 2001).

A. Recognizing evidence of domestication in Allium

B. Morphological and physiological changes

In crops where seeds were the objects of human interest, strong selection took place on traits related to yield and harvesting. Thus, unintentional domestication and selection for favorable traits occurred automatically as soon as hunter-gatherer

communities started to sow seeds. The major cereals are a good example of this mode of domestication. The main trait related to domestication in cereals is the shift from fragile to tough rachis, which allows harvesting of all seeds of a spike at once (Zohary and Hopf 2000). Plants with a fragile rachis contributed notably fewer seeds to the next plant generation due to the loss of seeds through broken ears before harvesting. Thus, this unfavorable trait was eliminated automatically from populations as soon as the crop was under cultivation. The process of cereal domestication consisted of automatic coevolutionary changes without meticulous planning. These are the types of changes that occur in other mutualistic relationships, such as the relationships between flowers and their pollinators.

Different mechanisms of plant domestication may be operating when plant parts other than seeds are the target of human interest. In the case of *Allium* taxa, it is essential to consider separately the production of plant material for human sustenance and the need to sustain the crop through propagation units (e.g., seeds or bulbs). In *A. tuberosum*, where both leaves and young inflorescences are eaten, we find an uncoupling of leaf production from the production of inflorescences and, hence, seeds. Harvesting of plant parts during a vegetative period and then allowing the plants to set seeds for reproduction later on in the year is not an unusual scenario for various crops (e.g., asparagus, lettuce, most herbs). Therefore, the different flowering times of *A. tuberosum* and its putative wild progenitor, *A. ramosum*, might be directly related to human impact on the earliest cultivated plants. Or, maybe the trait that allowed *A. tuberosum* to survive the gathering of vegetative parts by humans more easily resulted in domestication of an already preadapted plant species.

Differences in flowering time are mostly undetectable in the archeological record (Zohary and Hopf 2000). Apart from these phenological differences, *A. tuberosum* and *A. ramosum* are morphologically very similar and difficult to distinguish. Consequently, we cannot expect to get relevant information about the domestication process from preserved plant remains. Phylogenetic and population genetic data are needed to better understand the domestication of *A. tuberosum*.

B. Genomic changes

Differences in the processes of domestication leave specific marks on the genome of a crop. These differences can be predicted (Figure 10.1) and compared with the outcome of molecular analyses. Three major assertions for domestication via a single domestication event from a population in a defined geographical area can be made: 1. The crop should be the closest relative of its putative wild progenitor; 2. Accessions of the crop species should be nested as a single clade (i.e., branch) within the populations of its wild progenitor in a phylogenetic tree and should show highest similarity to the wild populations from the area where domestication took place; and 3. Genetic diversity within the crop species should be lower than in the wild species because domestication generally results in a severe genetic bottleneck due to the inclusion of only a small part of the naturally-occurring genetic variation of the species. Though the above-mentioned scenario is the most common type for crop evolution (Diamond 1997; Zohary 1999), deviations from the predicted results can point to alternative mechanisms of domestication, incorrect assumptions, or post-domestication events. For example, deviant results may occur when: 1. Incorrect assumptions are made about the phylogeny of close

relatives of the crop; 2. The crop has hybrid origins (van Raamsdonk 1995); 3. The crop has a polytypic origin, i.e., parallel domestication of the same wild species in different areas (Salgado et al. 1995); or 4. Post-domestication gene flow occurs between wild and domesticated populations where these grow in close proximity (Blattner and Badani Méndez 2001).

To analyze the genetic structure of *A. tuberosum* and *A. ramosum*, we used an anonymous marker approach to screen large parts of the genome for taxon differences and similarities. Anonymous markers are a method within the class of molecular techniques that make use of the polymerase chain reaction (PCR; Mullis and Faloona 1987). The protocol of PCR is relatively simple. It basically comprises four steps: 1. Melting of double-stranded genomic DNA at high temperatures, resulting in single strands that act as templates; 2. Reduction of temperature and annealing of two short oligonucleotide primers (about 20 nucleotides in length) to complementary regions of the single strands on both ends of the targeted sequence; 3. Raising of temperature to about 70°C where the primers are elongated by a thermostable DNA polymerase; and 4. Denaturation of the newly synthesized double-stranded target DNA. Steps 2 to 4 are repeated 30 to 40 times, which results in a nearly exponential increase of the target DNA as every newly synthesized DNA can act as a template in the following cycle. The major drawback of the method is the requirement of prior sequence information for the targeted region in order to design the complementary primers. Depending on genome region and plant group, the primer sequences can vary enormously even among related species, which often prevents successful amplification reactions. Without specific sequence information, it is only possible to amplify relatively conserved parts of the genome with

‘universal primers’, i.e., primers that bind to evolutionary conserved genes and thus are useful in most plant families. PCR of variable regions without prior primer sequence information became possible with the invention of universal anonymous markers, which bind with multiple matching sequences arbitrarily all over the genome.

The most familiar anonymous marker techniques are random amplified DNA (RAPD; Williams et al. 1990; along with a similar technique described by Welsh and McClelland 1990 as arbitrary primed PCR (AP-PCR)); inter simple sequence repeats (ISSR; Gupta et al. 1994; Zietkiewicz et al. 1994); random amplified microsatellite polymorphisms (RAMP; Wu et al. 1994), which is a combination of the first two methods; and amplified fragment length polymorphisms (AFLP; Vos et al. 1995). The differences among these methods rely mostly on the nature of the primer binding sites, which might be completely arbitrary (RAPD), or mediated by short nucleotide repeat motifs (ISSR) or by restriction enzyme recognition sites (AFLP). All methods have in common that they produce a variable number of PCR fragments, which originate from loci randomly distributed in the genome. The presence or absence of a specific PCR product is usually scored as an independent binary character, based on the assumption that the difference represents a mutation in the primer-binding site or the restriction site (but see discussion below of the possibility that they vary by insertion or deletion mutations). Here we will describe briefly the RAPD procedure.

The RAPD method uses a single PCR primer comprising only ten nucleotides, which is relatively short. The sequence of the primer is arbitrary and due to its short length it binds often in the genome. At positions where two primer-binding sites with inverse orientation are at a convenient distance of up to 2000 base pairs on opposite DNA

strands, the PCR procedure will result in an amplification product. Depending on genome size, genome composition, and primer sequence, a single RAPD reaction will typically produce 5–25 DNA fragments of different lengths. After PCR, this fragment mixture is sorted according to fragment size via electrophoresis on an agarose gel, stained with a fluorescent dye that specifically attaches to DNA, and visualized under UV-light. Band scoring is mostly done by eye with an enlarged photograph of the gel. RAPDs have been used in a large number of studies (see reviews in Bachmann 1994, 1997; Karp et al. 1996; Rieseberg 1996; Wolfe and Liston 1998). The advantages of RAPD analyses are time and cost efficiency and that only minimal laboratory equipment is required. In addition, only small amounts of DNA are needed and detection of the DNA fragments does not involve radioactivity. Thus, a relatively high number of polymorphisms can be easily detected when several commercially-available RAPD primers are used on a specific sample of accessions.

However, several limitations of the RAPD technique exist and should be considered before starting an analysis. Two major concerns are unproven band homology and possible non-Mendelian inheritance of fragments (Bachmann 1994; Rieseberg 1996). In the analysis of RAPD data, it is assumed that co-migrating bands (i.e., bands of identical lengths) represent homologous loci in the genome. Limitations in length measurement are due to the resolution of agarose gels. As in every electrophoretic technique, physical limitations prevent precise scoring of long fragments. Slight length differences might go unnoticed or might not be visible during gel examination. Additionally, the possibility of the occurrence of non-homologous fragments with equal lengths rises with increasing genetic distances among the studied organisms (e.g., when

comparing species instead of infraspecific lines). Both drawbacks can be reduced by using another gel matrix (e.g., polyacrylamide gels provide a higher resolution) and by testing band homology via Southern hybridization (i.e., excising a band from the gel and using it as a hybridization probe). However, these processes lead to a more time-consuming RAPD approach and abolish the major advantages of the technique and, therefore, are rarely performed.

Concerns about the heritability of RAPD markers are caused mainly by two sources of errors. Variation in reaction conditions can produce faint or even invisible bands. This might be due to imperfect primer pairing that allows amplification of the fragment only under otherwise optimal reaction conditions (influenced by DNA purity, PCR protocol, brand of thermostable DNA polymerase and thermocycler). To avoid these problems, rigorous standardization of the RAPD procedure is necessary. Genetic reasons for non-Mendelian inheritance of RAPD markers is either their origin from extra nuclear DNA (chloroplast and mitochondrial genomes), which is mostly uniparentally inherited, or due to the dominant nature of RAPD markers. With dominant markers, heterozygous alleles can not easily be detected because the absence of one allele is masked by the presence of the second one. Also, polymorphisms due to length differences (e.g., via insertions or deletions in a specific fragment instead of primer binding site mutations) can not easily be scored. This will result in an incorrect estimation of relatedness because these characters are non-independent (Bachmann 1997; Isabel et al. 1999).

The limitations of the RAPD method has caused a heated debate on the usability of RAPDs. However, most of the disadvantages listed above hold for other anonymous markers as well. In all cases, problems can be reduced via improvement and

standardization of the reaction conditions (e.g., Benter et al. 1995 *[1994 in biblio pg 20]*), several repetitions of the analysis with the inclusion only of reproducible fragments in the data matrix, the use of a high number of primers to get enough characters (usually character numbers should be three times higher than the number of accessions studied), and band scoring only within a single gel. The last protocol particularly limits the number of accessions within a study and might thus diminish the value of RAPD analysis (Isabel et al. 1999). Internal size standards, included in every lane of a gel might help to overcome this handicap in the future. However, the basically non-cumulative quality of all anonymous markers prevents their use as universal tools opposite to DNA sequences. Sequence data have the great advantage of being stored in open data bases, which can be extended with every newly submitted sequence. For anonymous marker data, no such data base storage tools are available, nor are they workable.

In summary, we conclude that although not every RAPD fragment may be reproducible in different laboratories, the overall outcome of two studies will be comparable *if* RAPD analyses are carefully conducted. When ample resources are available, we would always opt for a cumulative, sequence-based marker technique for comparative taxon studies. But, when this is economically not reasonable, RAPDs provide a possible alternative.

A. Materials and Methods

A total of 29 accessions of *A. tuberosum* and *A. ramosum* from the living collection of the Department of Taxonomy of the IPK Gatersleben were investigated (Table 10.1). The analyzed accessions represent the entire geographical ranges of the

species and mostly belong to the predominantly-occurring tetraploid plant types (2n = 32), though we also included diploids (2n = 16) and triploids (2n = 24). Two accessions of *A. oreiprason* Schrenk were included as outgroups. An analysis of a non-coding part of the nuclear ribosomal DNA for the entire genus *Allium* (Friesen et al. in prep. *[not in biblio pg 22]*) revealed that *A. ramosum* is the closest relative of *A. tuberosum* and that *A. oreiprason* is the sister group of both taxa. Consequently, we are sure that we analyzed a natural group of species.

DNA was isolated from one plant per accession with the NucleoSpin Plant kit (Macherey-Nagel, Düren, Germany) according to the instructions of the manufacturer. The concentration of the extracted DNA was checked on an agarose gel. Ten microliters (µL) of the isolated DNA were dissolved in 150 µL of water and 4 µL (approximately 50 ng) of this DNA solution were used for PCR amplification.

Prior to the analysis of the entire set of accessions, about 50 RAPD primers were tested on a small set of plants. Final amplifications were carried out using 11 arbitrary decamer primers (A19, AB04, AB18, AC02, C07, C09, C13, D01, D03, G13, and G19 obtained from Operon Technologies, Alameda, CA) that provided clear and reproducible bands in the initial screening. The amplification conditions were optimized according to Friesen et al. (1997). Twelve microliters of each RAPD reaction mixture were separated on 1.5% agarose gels in 0.5x TBE buffer, followed by staining with ethidium bromide (Sambrook et al. 1989). Clearly visible RAPD bands (an example of a RAPD gel is given in Figure 10.2) were scored manually for presence (1) or absence (0) using enlarged photographs of the gels. Differing band intensities were not taken into account to avoid errors introduced by competition among priming sites during the initial PCR cycles

(Bachmann 1997). Only reproducible bands of two independent amplification reactions were included in the data analyses.

From the resulting binary data matrix, distances and character-based analyses were calculated. Pairwise genetic distances were calculated using the Nei's coefficient that only considers shared presence of a fragment as a character (Nei and Li 1979). The absence of a band could be due to various reasons and, therefore, was not taken into account. Phenograms were prepared applying UPGMA (unweighted pair group method using arithmetic averages) and neighbor-joining (NJ) cluster analyses (Saitou and Nei 1987) of the genetic distance matrix (with PAUP*; Swofford 2002). Furthermore, the genetic distance matrix was subjected to a principal coordinate analysis (PCoA) where, starting from the distances, new independent axial coordinates representing most of the variability of the original data were calculated (with NTSYSpc; Rohlf 1998). The accessions were then plotted as points in a three-dimensional continuous space defined by the first three coordinates. Maximum parsimony analyses (MP) of the binary data matrix were also calculated (in PAUP*; Swofford 2002), using either branch-and-bound or the heuristic (with 200 random addition sequences) search options, MULPARS, ACCTRAN, and TBR branch swapping. Bootstrap analysis (Felsenstein 1985) with 1000 resamples was used to examine the statistical support of branches in the MP and NJ trees found.

Two different data sets were used to reveal the relationships among *A. tuberosum* and *A. ramosum* accessions. In the first data set, we included nine representative accessions of the two species together with the outgroups. We used this analysis to root the tree of the ingroups. A second data set included all 29 ingroup accessions but not *A. oreiprason*. This allowed us to score the differences among the ingroup populations in

more detail, while avoiding the inclusion of false-homologous fragments in the more distantly related outgroup.

A. Results

In the data set including the outgroup *A. oreiprason*, we found 136 polymorphic RAPD bands. Analyses of these bands resulted in a phylogenetic hypothesis that puts into question the putative progenitor–descendant relationship between the two ingroup species. All methods (NJ, UPGMA, and MP) resulted in very similar trees (only the NJ tree is shown Figure 10.3), which placed *A. tuberosum* as sister group of *A. ramosum*, corresponding to hypothesis C in Figure 10.1. Cladistic analysis (MP) resulted in one most parsimonious tree (not shown) of 224 steps length and a consistency index of 0.6071. The bootstrap support of this topology was relatively high, with values of 82% for the clade comprising *A. tuberosum* and 88% for *A. ramosum* in the cladistic analysis (slightly lower in the NJ analysis, see Figure 10.3). The nearly identical magnitudes of genetic variation within the crop and the wild species, as indicated by the branch lengths in Figure 10.3, is surprising. Within *A. tuberosum*, the diploid accession (Tax 2033) of unclear geographical origin occurs at a basal position.

In the second data set, which included all 29 ingroup accessions but not *A. oreiprason*, we found 127 polymorphic RAPD bands. Analyses of these data with phenetic (NJ and UPGMA), cladistic (MP), and statistical (PCoA) methods produced nearly identical results, similar to the analyses of the first data set—accessions of *A. ramosum* and *A. tuberosum* formed two genetically distinct and clearly separated groups (the NJ tree is shown in Figure 10.4a). Within the *A. ramosum* branch, a triploid

accession from China (Tax 2014) occurs at a basal position (Figure 10.4a). Within *A. tuberosum* the deepest split separates the putative Chinese diploid (Tax 2033) from all other accessions. The next branches are represented by material from Korea and China. Samples from other countries occur at derived positions in the tree (Figure 10.4a). Bootstrap support of the internal branches is higher for *A. tuberosum* than for *A. ramosum* even though genetic distances within both species were nearly equal.

The PCoA plot (Figure 10.4b) based on the second data set shows the distribution of genetic variation along the first three coordinal axes. Genetic diversity in *A. ramosum* is mostly seen along the Z-axis, whereas the accessions of *A. tuberosum* vary mainly in their positions along the X- and Y-axes. In *A. tuberosum*, the PCoA plot reflects geographical differentiation among the accessions—mostly Eastern Asian accessions are in the foreground (i.e., on the left side of the Y-axes) and samples from the Indian subcontinent are further back, grouping on the right side of the Y-axes (Figure 10.4b). This result is similar to that of the NJ analysis (Figure 10.4a). Neither analysis points to a severe genetic bottleneck in the crop, which is often associated with single domestication events.

A. Interpretation of results relative to domestication

The implications about the domestication history of Chinese chive are mostly attributed to two assumptions concerning our RAPD results: 1. Tree topologies reflect phylogenetic relationships, and 2. Branch lengths are reliable measures of genetic divergence. A comparison of our results to the hypothetical expectations of a phylogenetic analysis of crops and their ancestors (Figure 10.1) leads to the conclusion

that the wild progenitor of *A. tuberosum* can not be *A. ramosum*. The crop is not nested within the wild species, as would be expected in Figure 10.1a. Instead, we found a sister group relationship, meaning that both species shared a common ancestor from which they developed as two independent lineages (Figure 10.1c). This progenitor clearly was not included in our study and might be unknown or already extinct.

Another possible *[?]* interpretation of the results is that we missed the population of *A. ramosum* from which domestication began. However, our representative sampling covered the geographical distribution of the species. Also, this interpretation conflicts with the second line of evidence, the branch lengths. Some of the longest branches in the phylogenetic trees (Figures 10.3 and 10.4a) served to separate *A. tuberosum* from *A. ramosum*. These species are also clearly separated in the PCoA plot (Figure 10.4b). In spite of their morphological similarities, the clear genetic differentiation makes a conspecific status for these species highly unlikely. Furthermore, branch lengths differences are important in revealing genetic bottlenecks. If a species experienced a severe genetic reduction in its history (e.g., due to far-reaching extinction or founder events), we would expect to find relatively limited variation within the present-day gene pool of this species when compared to a species that did not experience such a genetic bottleneck (Eyre-Walker et al. 1998; Mimura et al. 2000; Iqbal et al. 2001). This generally results in shorter branches in population-based trees of species with reduced genetic diversity. In the trees produced from our analyses, *A. tuberosum* accessions did not differ markedly in their branch lengths from *A. ramosum*, which points to a relatively uniform distribution of genetic diversity in both species.

A. Discussion

Our results of phenetic, cladistic, and multivariate analyses of RAPD data revealed an unexpected relationship between wild *A. ramosum* and domesticated *A. tuberosum*. In spite of their great morphological similarity, which led Hanelt (2001) to merge both taxa into *A. ramosum*, molecular data clearly separated wild and crop populations. In our opinion, this is an indication that *A. ramosum* and *A. tuberosum* consist of two long-separated gene pools with no (or only very restricted) gene flow between interspecific populations. Therefore, we propose to maintain both taxa as separate species, i.e., to use *A. ramosum* only for the wild, early-flowering plants and *A. tuberosum* for the domesticate.

Additionally, the clear division of both species makes a progenitor–derivative relationship between them highly unlikely. Instead, we hypothesize that there was a common progenitor somewhere in the past from which both species evolved. The basal branches in the *A. tuberosum* group of the phylogenetic tree (Figure 10.4a) were formed by material from China and Korea, with the only diploid of this study occurring at the deepest branch. This pattern points to northern China as the most likely place where the (diploid?) progenitor populations of the crop were subjected to domestication. If we take into account that the flora of this area (as well as Mongolia) is far from being thoroughly studied, it even seems possible that some progenitor populations might still persist there. Yang et al. (1998) reported a (wild?) diploid *A. tuberosum* from Shangxi in China, which also indicates this area as a possible shelter of relatively basal types of the species. However, the large amount of genetic diversity within *A. tuberosum* calls into question a

single domestication from only one progenitor population. Most species with a unique domestication event are characterized by short internal branches in phylograms (Heun et al. 1997; Badr et al. 2000), i.e., relatively small genetic differences compared to their wild relatives. In Chinese chive, no such genetic bottleneck could be detected.

The question remains how domestication took place in Chinese chive. Without a known crop progenitor, most of the propositions about the mode of domestication remain rather speculative. However, some statements seem to be possible considering the current knowledge of *A. tuberosum*. High genetic variability within a crop can either be the result of multiple parallel domestications, which means that different populations contributed to the crop's gene pool, or from post-domestication hybridization events between the crop and adjacent wild populations. The fact that possibly no wild *A. tuberosum* populations exist, at least not on a large scale, could be an indication that many wild populations served as gene pool for the domesticate. This pattern occurs when wild plants are gathered for consumption throughout their entire range, which can lead to the extinction of the wild populations (Wawrosch et al. 2001). In this scenario, locally-kept garden populations, mostly founded when demand on wild plants became severe, are the diffuse starting point for the crop's evolution. Gene flow among domesticated and wild plants can also result in absorption of the wild gene pool into the crop (Freyre et al. 1996). In this case, extinction of the wild populations is possible, though not inevitable, via the transmission of traits not favorable in wild plants. Against the background of current gathering behavior of many native human tribes in Eastern Asia, the first hypothesis, i.e., extinction of the wild progenitor by human collecting, seems more likely to us.

Future research in *A. tuberosum* will concentrate on the inclusion of the newly found diploid material from Shangxi province. These plants might resemble the putative wild progenitors of the crop in morphology and karyotype. It would be interesting to test the position of these plants in a phylogeographic analysis (Schaal and Olsen 2000) relative to the domesticated germplasm. Furthermore, due to its ancestral ploidy level it might be a starting point to study the evolution of the complex ploidy patterns (di-, tri-, and tetraploids) in *A. tuberosum*.

A. References

Bachmann, K. 1994. Molecular markers in plant Ecology. *New Phytologist* 126: 403–418.

_____. 1997. Nuclear DNA markers in biosystematic research. *Opera Botan*ica 132: 137–148.

Badr, A.; K. Müller; R. Schäfer-Pregl; H. El Rabey; S. Effgen; H.H. Ibrahim; C. Pozzi; W. Rohde; and F. Salamini. 2000. On the origin and domestication history of barley (*Hordeum vulgare*). *Molecular and Biological Evolution* 17:499–510.

Benter, T.; S. Papadopoulos; M. Pape; M. Manns; and H. Poliwoda. 1994. Optimization and reproducibility of random amplified polymorphic DNA in humans. *Analytical Biochemistry* 230: 92–100. *[1994 in text pg 11]*

Blattner, F.R. and A.G. Badani Méndez. 2001. RAPD data do not support a second centre of barley domestication in Morocco. *Genetic Resources and Crop Evo*lution 48: 13–19.

Diamond, J. 1997. *Guns, germs and steel: the fates of human societies*. London: Jonathan Cape.

Dubouzet, J.G.; K. Shinoda; and N. Murata. 1997. Phylogeny of *Allium* L. subgenus *Rhizirideum* (G. Don ex Koch) Wendelbo according to dot-blot hybridization with randomly amplified DNA probes. *Theoretical and Applied Genetics* 95: 1223–1228.

Eyre-Walker, A.; R.L. Gaut; H. Hilton; D.L. Feldman; and B.S. Gaut. 1998. investigation of the bottleneck leading to the domestication of maize. *Proceedings of the National Academy of Sciences* 95:4441–4446.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39: 783–791.

Freyre, R.; R. Rios; L. Guzman; D.G. Debouck; and P. Gepts. 1996. Ecogeographic distribution of *Phaseolus* spp. (Fabaceae) in Bolivia. *Economic Botany* 50:195–215.

Friesen, N. and F.R. Blattner. 2000. RAPD analysis reveals geographic differentiations within *Allium schoenoprasum* L. (Alliaceae). *Plant Biology* 2: 297–305.

Friesen, N. and M. Klaas. 1998. Origin of some minor vegetatively propagated *allium* crops studied with rapd and gish. genetic resources and crop. *Evolution* 45: 511–523.

Friesen, N.; R.M. Fritsch; and K. Bachmann. 1997. Hybrid origin of some ornamental *Alliums* of subgenus *Melanocrommyum* verified with GISH and RAPD. *Theoretical and Applied Genetics* 95: 1229–1238.

Fritsch, R.M.; F. Matin; and M. Klaas. 2001. *Allium vavilovii* M. Popov et Vved. and a new Iranian Species are the closest among the known relatives of the common onion *A. cepa* L. (Alliaceae). *Genetic Resources and Crop Evolution* 48: 401–408.

Friesen, N.; S. Pollner; K. Bachmann; and F.R. Blattner. 1999. RAPDs and noncoding chloroplast DNA reveal a single origin of the cultivated *Allium fistulosum* from *A. altaicum* (Alliaceae). *American Journal of Botany* 86: 554–562.

[Friesen et al. in prep in text page 12]

Gupta, M.; Y.-S. Chyi; J. Romero-Severson; and J.L. Owen. 1994. Amplification of DNA markers from evolutionary diverse genomes using single primers of simple-sequence repeats. *Theoretical and Applied Genetics* 89: 998–1006.

Hanelt, P. 1988. Taxonomy as a tool for studying plant genetic resources. *Die Kulturpfla*nze 36: 169–187.

_____. 2001. Alliaceae. In *Mansfeld's Encyclopedia of Agricultural and Horticultural Crop*s, P. Hanelt and Institute of Plant Genetics and Crop Plant Research (eds.), pp. 2250–2269. Berlin: Springer-Verlag.

Havey, M.J. 1992. Restriction enzyme analysis of the chloroplast and nuclear 45s ribosomal DNA of *Allium* sections *Cepa* and *Phyllodon* (Alliaceae). *Plant Systematics and Evolution* 183: 17–31.

Heun, M.; R. Schäfer-Pregl; D. Klawan; R. Castagna; M. Accerbi; B. Borghi; and F. Salamini. 1997. Site of einkorn wheat domestication identified by DNA fingerprinting. *Science* 278: 1312–1314.

Iqbal, M.J.; O.U.K. Reddy; K.M. El-Zik; and A.E. Pepper. 2001. A genetic bottleneck in the 'evolution under domestication' of upland cotton *Gossypium hirsutum* L. examined using DNA fingerprinting. *Theoretical and Applied Genetics* 103: 547–554.

Isabel, N.; J. Beaulieu; P. Theriault; and J. Bousquet. 1999. Direct evidence of biased gene diversity estimates from dominant random amplified polymorphic DNA (RAPD) fingerprints. *Molecular Ecology* 8: 477–483.

Karp, A.; O. Seberg; and M. Buiatti. 1996. Molecular techniques in the assessment of botanical diversity. *Annals of Bota*ny 78: 143–148.

Kik, C.; A.M. Samoylov; W.H.J. Verbeek; and L.W.D. van Raamsdonk. 1997. Mitochondrial DNA variation and crossability of leek (*Allium porrum*) and its wild relatives from the *Allium ampeloprasum* complex. *Theoretical and Applied Genetics* 94:465–471.

Klaas, M. and N. Friesen. 2002. Molecular markers in *Allium*. In *Allium Crop Science — Recent Advances*, H. Rabinovich and L. Curah (eds.), pp. 159–185. Wallingford, UK: CABI Publishing.

Maass, H.I. and M. Klaas. 1995. Infraspecific Differentiation of Garlic (*Allium sativum* L.) by Isozyme and RAPD Markers. *Theoretical and Applied Genetics* 91: 89–97.

Mimura, M.; K. Yasuda; and H. Yamaguchi. 2000. RAPD Variation in wild, weedy and cultivated azuki beans in Asia. *Genetic Resources and Crop Evolution* 47: 603–610.

Mullis, K.B. and F.A. Faloona. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology* 155:335–350.

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* 76: 5269–5273.

van Raamsdonk, L.W.D. 1995. The cytological and genetical mechanisms of plant domestication exemplified by four crop models. *Botanical Rev*iews 61: 367-399.

van Raamsdonk, L.W.D.; M. Vrielink–van Ginkel; and C. Kik. 2000. Phylogeny reconstruction and hybrid analysis in *Allium* subgenus *Rhizirideum*. *Theoretical and Applied Genetics* 100: 1000–1009.

Rieseberg, L.H. 1996. Homology among RAPD fragments in interspecific comparisons. *Molecular Ecology* 5:99–105.

Rohlf, F.J. 1998. *NTSYSpc: Numerical axonomy system. Version 1.8.* Setauket, New York: Exeter Publishing Ltd.

Salgado, A.G.; P. Gepts; and D.G. Debouck. 1995. Evidence for two gene pools of the lima bean, *Phaseolus lunatus* L., in the Americas. *Genetic Resources and Crop Evolution* 42: 15–28.

Sambrook, J.; E.F. Fritsch; and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*. 2nd edition. Cold Spring Harbor, New York:Cold Spring Harbor Laboratory Press.

Saitou, N. and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Journal of Molecular Evolution* 4: 406–425.

Schaal, B.A. and K.M. Olsen. 2000. Gene genealogies and population variation in plants. *Proceedings of the National Academy of Sciences* 97: 7024–7029.

Stearn, W.T. 1944. Nomenclature and synonymy of *Allium odorum* and *A. tuberosum*. *Herbertia* 11: 227–245.

Swofford, D.L. 2002. *PAUP*. Phylogenetic analysis using parsimony (*and other methods),* Version 4. Sunderland, Massachusetts: Sinauer Associates.

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.

Wawrosch, C.; P.R. Malla; and B. Kopp. 2001. Micropropagation of *Allium wallichii* Kunth, a threatened medicinal plant of Nepal. *In Vitro Cellular and Developmental Biology–Plant* 37: 555–557.

Welsh, J. and J. McClelland. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* 18: 7213–7218.

Williams, G.K.; A.R. Kubelik; K.L. Livak; J.A. Rafalaski; and S.V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18: 6531–6535.

Wolfe, A.D. and A. Liston. 1998. Contribution of PCR-based methods to plant systematics and evolutionary biology. *In Molecular systematics of plants* II, D.E. Soltis, P.S. Soltis, and J.J. Doyle (eds.), pp. 43–86. Dordrecht, Netherlands: Kluwer Academic Publishers.

Wu, K.-S.; R. Jones; L. Danneberger; and P.A. Scolnik. 1994. Detection of microsatellite polymorphisms without cloning. *Nucleic Acids Research* 22: 3257–3258.

Yang, L.; J.-M. Xu; X.-L. Zhang; and H.-Q. Wan. 1998. Karyoptypical studies of six species of the genus *Allium*. *Acta Phytotaxonomica Sinica* 36: 36–46.

Zietkiewicz, E.; A. Rafalski; and D. Labuda. 1994. Genome fingerprinting by simple sequence (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20: 176–183.

Zohary, D. 1999. Monophyletic vs. polyphyletic origin of the crops on which agriculture was founded in the Near East. *Genetic Resources and Crop Evolut*ion 46: 133–142.

Zohary, D. and M. Hopf. 2000. *Domestication of plants in the Old World*. 3rd Edition. Oxford, UK: Oxford University Press.

Table 10.1. Accessions used in the comparison of *Allium ramosum* and *A. tuberosum*. 'Tax' refers to the accession numbers of the *Allium* collection of the Department of Taxonomy, Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Germany. All material was collected in the wild or, for the crop species, in the country given under 'Origin'. The only exception is accession Tax 2033, which is material from a germplasm collection of the Vavilov Institute (VIR, St. Petersburg, Russia) with unclear origin. Chromosome numbers are given under '2n', where 16 refers to diploid, 24 to triploid, and 32 to tetraploid samples.

#	Acc. No.	Species	Origin	2n
1	Tax 1419	*A. ramosum*	Russia, Tuva	32
2	Tax 1695	*A. ramosum*	Mongolia, Bayan-Chongor	32
3	Tax 1699	*A. ramosum*	Mongolia, Erdenesant	32
4	Tax 1836	*A. ramosum*	Russia, Yacutia	32
5	Tax 2014	*A. ramosum*	China	24
6	Tax 2115	*A. ramosum*	Northern Kazakhstan	32
7	Tax 2339	*A. ramosum*	Mongolia, Ulan-Bator	32
8	Tax 2347	*A. ramosum*	Mongolia, CherenBayan-Uul	32
9	Tax 2356	*A. ramosum*	Mongolia, Tumencogt-Uul	32
10	Tax 2363	*A. ramosum*	Mongolia, Erdenecagaan	32
11	Tax 2371	*A. ramosum*	Mongolia, Bayan-Changai	32
12	Tax 2378	*A. ramosum*	China	24
13	Tax 2735	*A. ramosum*	Kazakhstan	32
14	Tax 2755	*A. ramosum*	Russia, Buryatia	32
15	Tax 2759	*A. ramosum*	Russia, Tuva, Ersin	32
16	Tax 0582	*A. tuberosum*	Japan, Tsukuba	32
17	Tax 1482	*A. tuberosum*	Nepal	32
18	Tax 1969	*A. tuberosum*	Korea	32
19	Tax 1970	*A. tuberosum*	Korea	32
20	Tax 1971	*A. tuberosum*	Korea	32
21	Tax 2033	*A. tuberosum*	VIR, possibly from China	16
22	Tax 2426	*A. tuberosum*	Russia, Rusansky Chrebet (Pamir)	32
23	Tax 2453	*A. tuberosum*	India, Ladakh	32

24	Tax 2454	*A. tuberosum*	India, Agra	32
25	Tax 2499	*A. tuberosum*	China, NW Yunnan	32
26	Tax 3301	*A. tuberosum*	Pakistan, Gilgit	32
27	Tax 3866	*A. tuberosum*	India	32
28	Tax 4246	*A. tuberosum*	China	32
29	Tax 5557	*A. tuberosum*	Vietnam	24
30	Tax 3653	*A. oreiprason*	Kazakhstan, Transili-Alatau	16
31	Tax 5000	*A. oreiprason*	Kyrgyzstan, Tallas-Alatau	16

A. Figure Captions

Figure 10.1: Possible evolutionary relationships of crop species and their wild relatives (c = crop, p = progenitor, w = other species). a. The crop originated via a single domestication event from within its wild progenitor; b. Wrong prior assumptions about the direct progenitor result in an unexpected sister group of the crop species; c. The progenitor of the crop species is either unknown or extinct; thus, no direct progenitor can be found. Instead, the closest living relative of the crop species occurs as sister group to it; d. The crop originated via hybridization of two different progenitor species; e. Parallel domestication in different areas result in two (or more) independent crop lines; f. Gene flow among the crop and populations of its wild progenitor introduces additional genetic material into the crop's gene pool.

Figure 10.2: RAPD reaction of 29 *A. ramosum* and *A. tuberosum* accessions with Operon primer AB04, electrophoretically separated on a 1.5% agarose gel. In the first lane on the left side, a size standard (100 bp ladder) is included, which allows sizing of the RAPD fragments. The order of the samples on the gel (from left to right) is the same as given in Table 1.

Figure 10.3: Phenogram of a neighbor-joining analysis of 137 RAPD characters of accessions of wild *A. ramosum* and the crop plant *A. tuberosum* together with the closely related *A. oreiprason* as outgroup taxon. Numbers along the branches depict bootstrap

values (%). *Allium ramosum* and *A. tuberosum* occur as sister groups instead of the crop nested within its proposed wild progenitor.

Figure 10.4: a. Phenogram of an unrooted neighbor-joining analysis of 127 RAPD characters of 29 accessions of wild *A. ramosum* and the crop *A. tuberosum*. Branch lengths represent genetic distances. Bootstrap values (%) are given along the branches; b. Principal co-ordinate analysis of the same data set. Both analyses show that the crop and its putative wild progenitor are clearly differentiated and that genetic diversity within the domesticate is nearly of the same magnitude as within the wild species.

Figure 10.1

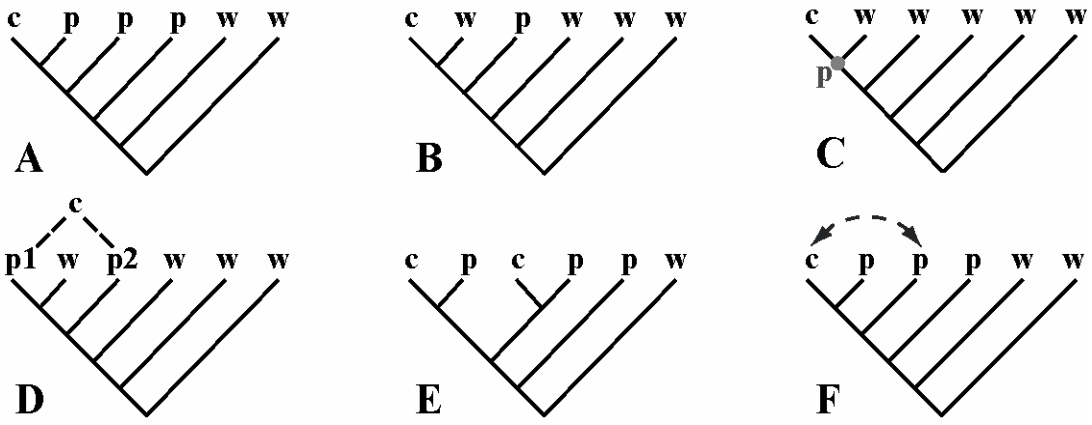


Figure 10.2

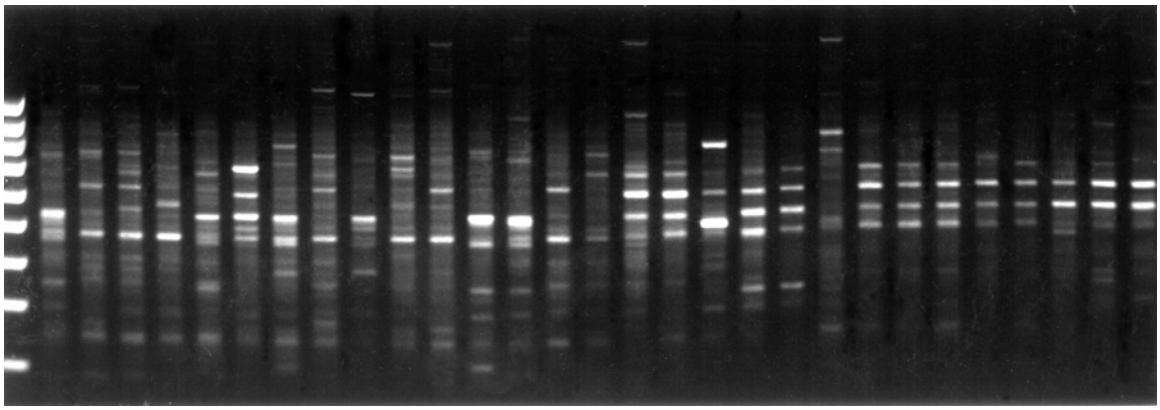


Figure 10.3

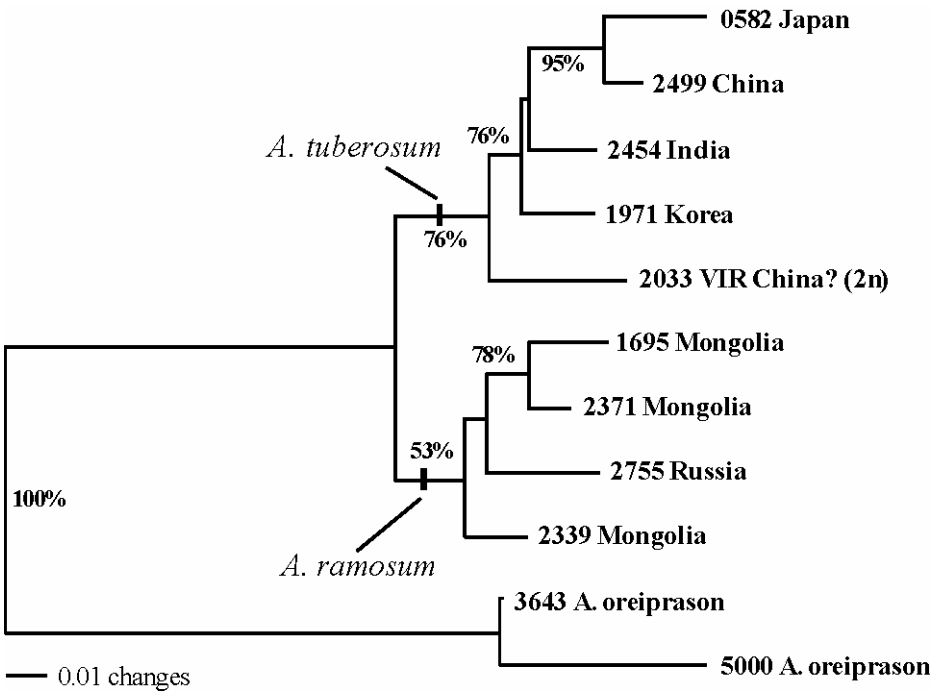
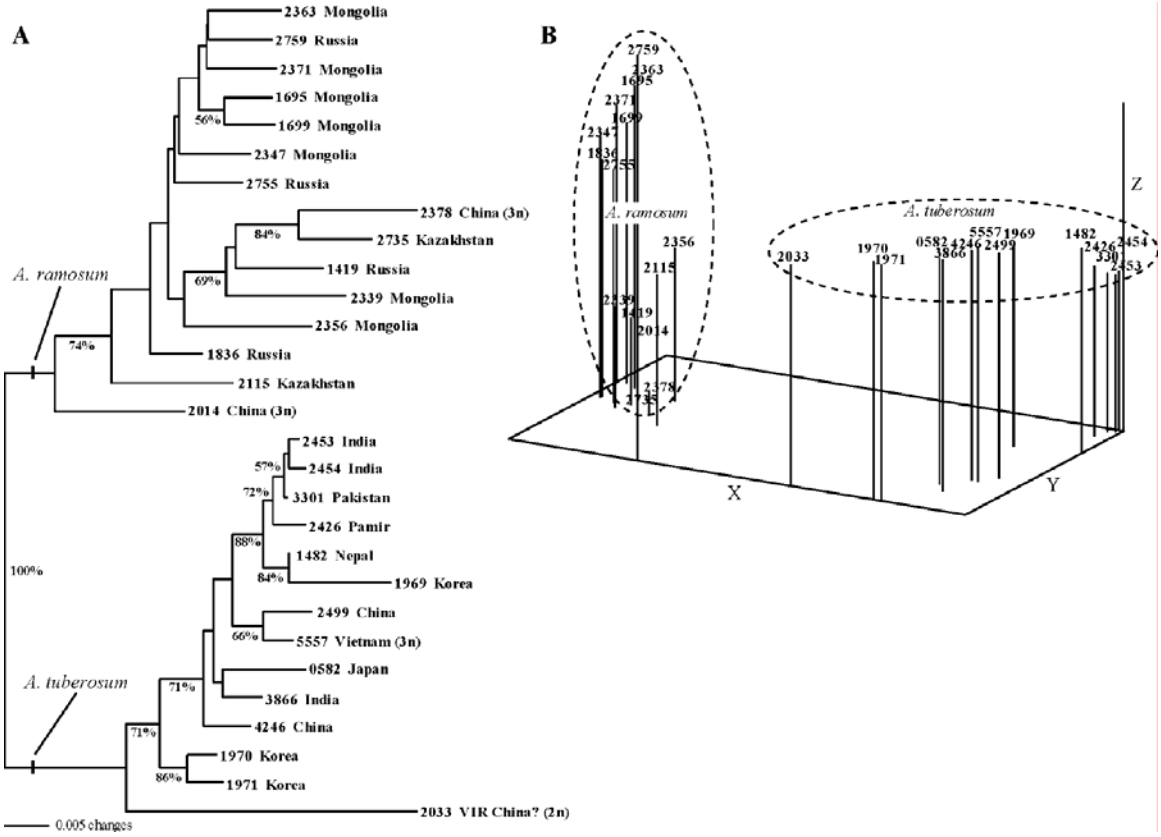

0582 Japan
95%
2499 China
A. tuberosum
76%
2454 India
1971 Korea
76%
2033 VIR China? (2n)
1695 Mongolia
78%
2371 Mongolia
53%
2755 Russia
100%
2339 Mongolia
A. ramosum
3643 A. oreiprason
5000 A. oreiprason
0.01 changes

Figure 10.4


A
2363 Mongolia
2759 Russia
2371 Mongolia
1695 Mongolia
1699 Mongolia
2347 Mongolia
2755 Russia
2378 China (3n)
2735 Kazakhstan
1419 Russia
2339 Mongolia
2356 Mongolia
1836 Russia
2115 Kazakhstan
2014 China (3n)
2453 India
2454 India
3301 Pakistan
2426 Pamir
1482 Nepal
1969 Korea
2499 China
5557 Vietnam (3n)
0582 Japan
3866 India
4246 China
1970 Korea
1971 Korea
2033 VIR China? (2n)
A. ramosum
A. tuberosum
56%
84%
69%
74%
100%
88%
84%
66%
71%
71%
86%
57%
72%
0.005 changes
B
A. ramosum
A. tuberosum
X
Y
Z