

ORIGINS OF DOMESTICATION AND POLYPLOIDY IN OCA (*Oxalis tuberosa*: OXALIDACEAE). 2. CHLOROPLAST- EXPRESSED GLUTAMINE SYNTHETASE DATA¹

EVE EMSHWILLER^{2,3,4} AND JEFF J. DOYLE³

²Department of Botany, The Field Museum of Natural History, 1400 S. Lake Shore Drive, Chicago, Illinois 60605-2496 USA;

³L. H. Bailey Hortorium and Department of Plant Biology, Cornell University, 466 Mann Library Building, Ithaca, New York 14853-4301 USA

In continuing study of the origins of the octoploid tuber crop oca, *Oxalis tuberosa* Molina, we used phylogenetic analysis of DNA sequences of the chloroplast-active (nuclear encoded) isozyme of glutamine synthetase (ncpGS) from cultivated oca, its allies in the “*Oxalis tuberosa* alliance,” and other Andean *Oxalis*. Multiple ncpGS sequences found within individuals of both the cultigen and a yet unnamed wild tuber-bearing taxon of Bolivia were separated by molecular cloning, but some cloned sequences appeared to be artifacts of polymerase chain reaction (PCR) recombination and/or *Taq* error. Nonetheless, three classes of nonrecombinant sequences each joined a different part of the *O. tuberosa* alliance clade on the ncpGS gene tree. Octoploid oca shares two sequence classes with the Bolivian tuber-bearing taxon (of unknown ploidy level). Fixed heterozygosity of these two sequence classes in all oca samples suggests that they represent homeologous loci and that oca is allopolyploid. A third sequence class, found in eight of nine oca plants sampled, might represent a third homeologous locus, suggesting that oca may be autoallopolyploid, and is shared with another wild tuber-bearing species, tetraploid *O. picchensis* of southern Peru. Thus, ncpGS data identify these two taxa as the best candidates as progenitors of cultivated oca.

Key words: crop evolution; glutamine synthetase; ncpGS; oca; Oxalidaceae; *Oxalis tuberosa*; PCR recombination; polyploidy.

Oxalis tuberosa Molina, commonly known as oca, is an octoploid tuber crop cultivated in the central Andean highlands. We have initiated studies to determine the progenitors of domestication and origins of polyploidy of this cultigen, and we previously reported the taxonomic and cytological background and the use of DNA sequence data from the internal transcribed spacer of nuclear ribosomal DNA (nrDNA ITS) as a first step in this project (Emshwiller and Doyle, 1998). That study confirmed the naturalness of a group of *Oxalis* species first recognized on cytological grounds by de Azkue and Martínez (1990) as the $x = 8$ “*Oxalis tuberosa* alliance.” However, the ITS results indicated that the alliance

includes several additional species, not studied by de Azkue and Martínez, that still lack cytological data. The ITS data support the origins of cultivated *O. tuberosa* from within the $x = 8$ alliance. Specifically, the predominant ITS sequence of the cultigen was found within the clade that included the members of “*O. tuberosa* alliance,” with faint signs of a second ITS sequence (seen as a minor secondary band on manual sequencing films in one of the three plants sampled), which also groups within the same clade (Emshwiller and Doyle, 1998). Because it was unknown whether oca is autopolyploid or allopolyploid, one explanation for this second faint sequence is the presence of multiple, possibly homeologous loci, with one predominating due to differential amplification or real copy number differences. Such differences could be due to concerted evolution among homeologous loci (Wendel, Schnabel, and Seelman, 1995). In addition to these complicating factors, the ITS region was not variable enough among the species of this group to resolve their relationships, in spite of high levels of variation across divergent *Oxalis* species. The ITS data could not identify the progenitor genomes of octoploid *O. tuberosa* with precision, as there were three wild Bolivian taxa that shared the identical ITS sequence with cultivated oca, among them a yet unnamed wild tuber-bearing taxon.

We have continued the study of the origins of *O. tuberosa* using DNA sequence data from another independently evolving locus, the nuclear gene encoding chloroplast-expressed glutamine synthetase (ncpGS). This locus is single copy in *Oxalis*, as in most taxa studied to date, and it diverged long ago from the cytosolic-expressed isozymes (Pesole et al., 1991), so that primers have been designed that amplify only the chloroplast-expressed form (Emshwiller and Doyle, 1999). In a pilot study, the gene tree of ncpGS was generally congruent with that of ITS, with somewhat more variation among the ncpGS sequences from the species studied than among

¹ Manuscript received 30 August 2001; revision accepted 5 February 2002.

The authors thank INRENA and IBTA for plant collection permits for Peru and Bolivia, respectively; the Centro Internacional de la Papa (CIP), Centro de Investigación de Recursos Naturales y Medio Ambiente (CIRNMA), Servicios Múltiples de Tecnología Apropiada (SEMATA), Centro de Investigación en Cultivos Andinos (CICA), Instituto Nacional de Investigación Agraria (INIA) E. E. Andenes, and the McKnight Foundation Project on Andean Root and Tuber Crops for logistical support. EE also thanks the Programa de Investigación de la Papa (PROINPA) for bringing her to areas where the wild tuber-bearing *Oxalis* of Bolivia, which are supported in this study as one of the possible genome donors of domesticated oca, could be found; R. Estrada for facilitating an affiliation with UNMSM during her Fulbright year in Peru; and the following people (in chronological order) for logistical help in Bolivia and Peru: M. G. Asbun Claros; J. Miller; M. L. Ugarte; F. Terrazas; S. Guzmán; T. Villarroel; and J. Almanza; R. Vargas; B. Eriksen and U. Molau; A. Valladolid; G. Meza and P. Cruz; A. Castelo; A. Tupayachi; P. Núñez; and R. Urrunaga; H. Flores; R. Estrada; R. Ortega; C. Arbizu; M. Ramírez; and A. Andia; F. Vivanco; N. Arce and L. Torrez; M. Vallenás; and I. Sánchez. JJD acknowledges funding from the NSF Systematic Biology Program, most recently DEB-0089483. EE thanks the A. W. Mellon Foundation, a J. W. Fulbright Travel Grant, an NSF Dissertation Improvement Grant, U.S. Department of Education, L. H. Bailey Hortorium, and Cornell Graduate School. A grant from Abbott Laboratories to the Field Museum supported EE during the final editing. This work is part of a Ph.D. dissertation presented by EE to Cornell University.

⁴ Author for reprint requests (e-mail: eemshwiller@fmnh.org).

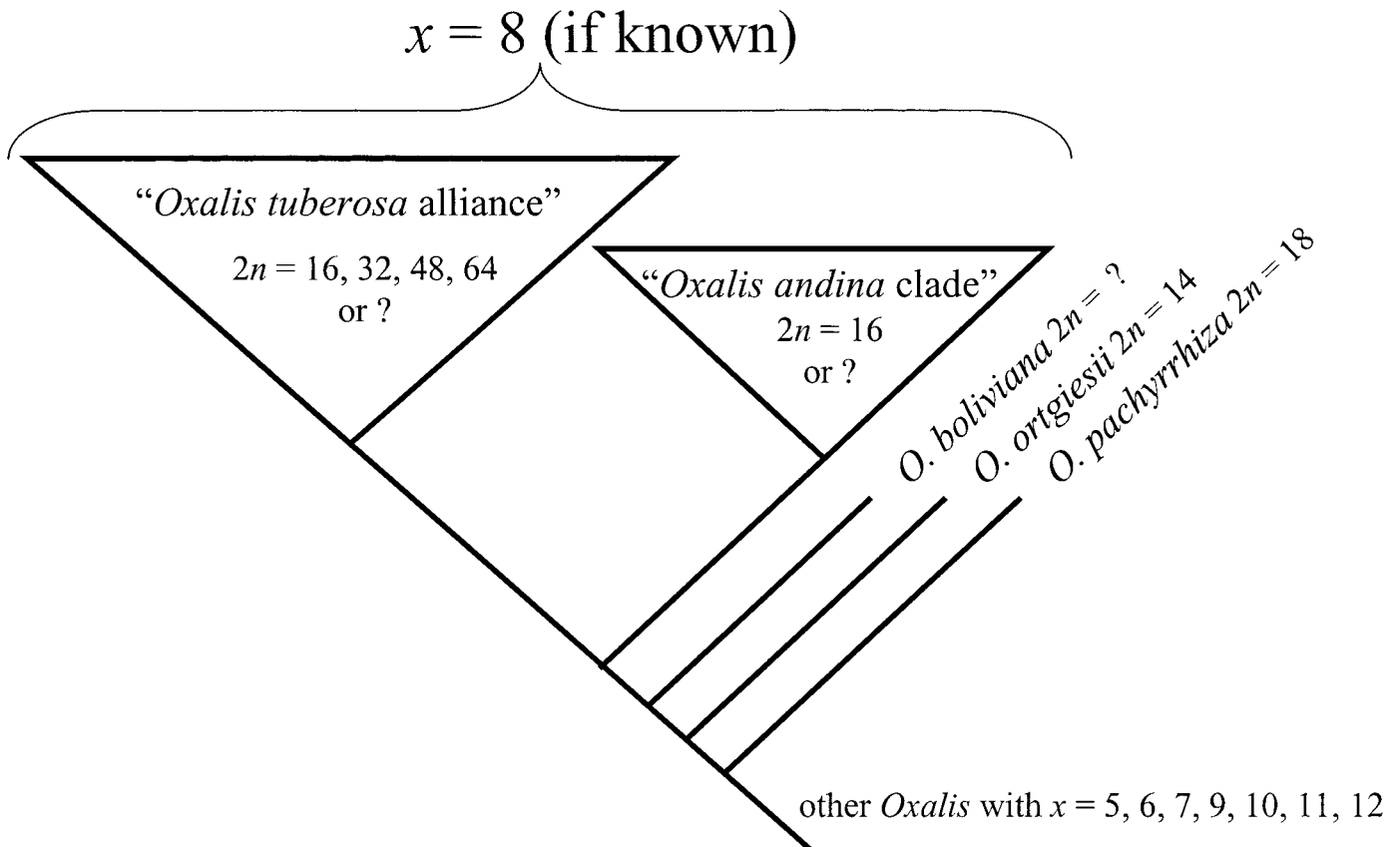


Fig. 1. Summary of relationships, inferred from ITS data, of the *Oxalis tuberosa* alliance with outgroup taxa whose ITS sequences were alignable with those of alliance species (modified from Emshwiller and Doyle, 1998), compared with currently available cytological information (for which references are provided in the text, in that prior study, or at <http://ajbsupp.botany.org/v89/emshwiller/table1>). Question marks indicate those for which chromosome counts are not yet available.

their ITS sequences (Emshwiller and Doyle, 1999). Initial attempts to sequence ncpGS from cultivated *O. tuberosa* directly from polymerase chain reaction (PCR) amplification products showed clear signs of multiple sequences within individual oca plants. This intraindividual sequence heterogeneity of ncpGS suggested that this locus could provide evidence of the origins of all of oca's genomes. Here we report the results of analysis of ncpGS data from cultivated oca and wild Andean *Oxalis* taxa as they contribute to the elucidation of the origins of the crop.

A note is in order to define what is meant here by the informal name “*Oxalis tuberosa* alliance,” first proposed by de Azkue and Martínez (1990) for a dozen morphologically similar $x = 8$ species, but here including additional species, as mentioned above. The alliance probably comprises at least 40–50 species. In the past, we have used the “*O. tuberosa* alliance” and “ $x = 8$ group” interchangeably. However, *O. andina* has recently been reported to have 16 chromosomes (de Azkue, 2000), indicating that the “ $x = 8$ group” may also include the clade that was sister to the alliance (“*O. andina* clade” in Fig. 1) in our prior ITS study (Emshwiller and Doyle, 1998). We use “*O. tuberosa* alliance” here in a narrower sense than “ $x = 8$ group,” to refer to the clade, on both the ITS and ncpGS gene trees (see RESULTS and Emshwiller and Doyle, 1998, 1999; Emshwiller, 1999), that includes all of the sequences of cultivated *O. tuberosa*, along with other

taxa that either are reported to be based on $x = 8$ or lack cytological data but excluding the “*O. andina* clade.”

MATERIALS AND METHODS

Taxonomic sampling—The 60 *Oxalis* accessions sampled included nine plants of cultivated *O. tuberosa*, five outgroup taxa, and 46 wild plants in the *Oxalis tuberosa* alliance, which represent from 22 to 35 species, depending on delimitation (see below). Source, voucher, etc. information has been archived at the Botanical Society of America website (<http://ajbsupp.botany.org/v89/emshwiller/table1>). Because our focus was on the origins of *O. tuberosa*, we used our previous ITS results (Emshwiller and Doyle, 1998), as well as our ncpGS results as they accumulated, as a guide to sampling ncpGS more intensively in the ingroup (the *O. tuberosa* alliance) and somewhat less in the outgroup. At least one individual from each species found within the $x = 8$ clade on the ITS tree (Emshwiller and Doyle, 1998) was resampled for ncpGS. Additional ingroup sampling included Peruvian *Oxalis* taxa that were either reported to be members of the *O. tuberosa* alliance (de Azkue and Martínez, 1990) or were morphologically similar to alliance species.

Outgroup sampling for ncpGS included three representative species from among those included in the main ITS analyses, including one representative (*O. andina*) of the clade that was sister to the known *O. tuberosa* alliance clade on the ITS tree (see above). Two outgroup taxa from among Peruvian *Oxalis* accessions were added, one of which, *O. laxa* var. *hispidissima* (EE744), is a member of a morphological group within *Oxalis* that had not been included in the previous study. The second additional outgroup, *O. megalorrhiza* (EE773), is type species of section *Carnosae* Reiche (*O. megalorrhiza* is often referred to by the misapplied name *O. carnosae* Molina; see

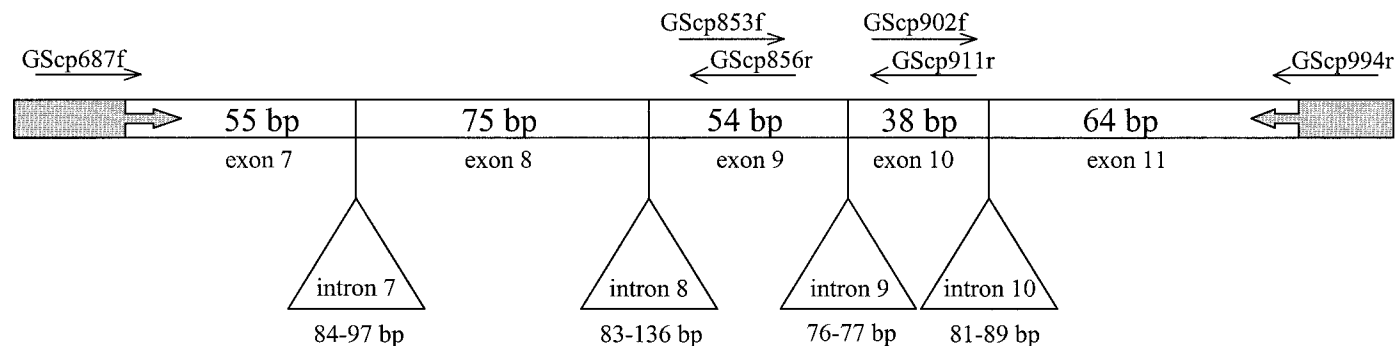


Fig. 2. Amplified region of chloroplast-expressed glutamine synthetase, indicating the positions of the primers used for amplification and sequencing and the sizes in base pairs (bp) of introns and exons (excluding the amplification primer annealing regions in the cases of exons 7 and 11). Figure adapted from Emshwiller and Doyle (1999) and revised to indicate the wider ranges of intron sizes found with the larger sampling of *Oxalis* taxa in the current study.

Dandy and Young [1959] for an explanation of the history of this name). Its inclusion was intended to test congruence of the molecular tree with cytological data (see below) and morphologically based sectional classifications, because it had published chromosome counts (albeit conflicting, $2n = 14$ or 18, see below) and it is morphologically very similar to *O. pachyrrhiza*, which nonetheless was classified by Knuth (1930) in a different section.

As we outlined earlier (Emshwiller and Doyle, 1998), the search for the origins of *oca* has been impeded by the confused state of *Oxalis* taxonomy. Hybridization and lack of breeding barriers may be obscuring species limits (Emshwiller and Doyle, 1998), and it is possible that Andean *Oxalis* populations experienced repeated cycles of divergence and contact during the history of mountain uplift and later glaciation, leading to a multitude of forms whose species membership is uncertain. In addition, the dependence of taxonomic workers on herbarium specimens has obscured species limits, because some important characteristics are lost upon drying (e.g., distribution of pigments, see also Salter, 1944; Emshwiller, 1999, 2002a) and others are subject to phenotypic plasticity (e.g., swollen petioles). The recent revision of the genus by Lourteig (1994, 2000) has greatly clarified its taxonomy. She recognizes 280 species in *Oxalis* (excluding the South African section *Cernuae* treated by Salter [1944]), classified in four subgenera and 28 sections. These sections are in closer agreement with the *O. tuberosa* alliance recognized here than the prior infrageneric taxonomy of Knuth (e.g., 1930). Specifically, the species of the alliance were placed in three sections (*Lotoideae* Lourt., *Herrerae* Knuth, and *Ortgieseeae* Knuth) by Lourteig (2000), whereas they were placed in seven sections by Knuth. However, there are some cases in which we at least tentatively retain the use of species names that Lourteig (2000) has reduced to synonymy, because the plants concerned have dissimilar ploidy levels or ncpGS sequences and/or they have morphological differences that persist in common garden conditions (see also Emshwiller, 1999, 2002a). In some cases fixed differences support the retention of species names (e.g., *O. picchensis*, *O. unduavensis*), but in other cases the morphologically different accessions may or may not ultimately be found to be distinct at the species level. Nonetheless, the provisional use of these subsumed names (e.g., *O. weberbaueri*, *O. oblongiformis*, *O. staffordiana*, and others in the group designated here as the "*O. peduncularis* clade" [see below]) reflects the observed differences among the populations.

Gene amplification and sequencing—DNA isolations either followed Doyle and Doyle (1990) or used DNeasy Plant Mini Kits (QIAGEN, Valencia, California, USA) according to the manufacturer's instructions. Amplification of a region of the ncpGS locus that includes four introns was performed using primers GScp687f and GScp994r (Fig. 2) and thermocycling conditions described in Emshwiller and Doyle (1999). Amplification products either were cloned (see below) or were sequenced directly, either by manual sequencing as described previously (Emshwiller and Doyle, 1998) or with an ABI 377 automated sequencer operated by the Cornell Biotechnology Center. Electropherograms were examined and edited using either Chromas 1.43 (McCarthy, 1997) or Sequencher 3.1 (Gene Codes Corporation, Ann Arbor, Michigan,

USA). Direct sequencing of PCR products used the amplification primers or internal primers GScp853f, GScp856r, or GScp911r (Emshwiller and Doyle, 1999), whereas sequencing of clones used standard primers that anneal to the plasmid. Sequences were determined in both directions, with the exception of individuals that were heterozygous for insertion/deletion differences (indels). Thus, sequences could only be determined in one direction for accessions EE190, EE871, EE960, and ORT1, and part of the sequence is missing for accessions EE511 and EE512, because the latter plants were heterozygous for two indels.

Multiple sequence types within individuals: molecular cloning—The plants selected for molecular cloning of ncpGS sequences included three individual plants of different morphotypes of cultivated *oca* (accessions MHG884, MHG913, and 35-04) and one plant from among wild tuber-bearing populations found in Bolivia (EE259). Molecular cloning of amplification products used either the Original TA cloning kit or the TOPO TA cloning kit with One Shot competent cells (all from Invitrogen, Carlsbad, California, USA) according to the manufacturer's instructions. Results were improved when the amplification products were first cut out of an agarose gel and cleaned in QIAquick gel extraction kit columns (QIAGEN) and the "overhanging" 3' adenines replaced (Anonymous, 1996).

Multiple sequence types within individuals: screening by direct sequencing—Once the various sequence classes of cultivated *oca* and the wild tuber-bearing plant had been identified by cloning and sequencing, a strategy was developed to test whether these same sequence classes were present in a larger sample of individual plants without sequencing large numbers of clones from many individuals. The sample for direct sequencing of amplification products included the same plants that had been used in the molecular cloning experiments, along with six additional morphologically distinct accessions of cultivated *oca* and three of the wild Bolivian taxon (<http://ajbsupp.botany.org/v89/emshwiller/table1>). There are relatively few changes that distinguish the various sequence classes found among the clones from these individuals (see below). Thus, electropherogram traces were examined to determine whether there were double peaks (two nucleotides at a single site) at the particular sites that distinguish the sequence classes, or in cases of sequence classes that differ by an indel character, whether the sequence became unreadable (mostly double peaks) past the location of the indel (see Cronn et al. [2002] for a similar strategy). Because the plants whose sequences were cloned were all heterozygous for several indels, much of the sequence was unreadable if the amplification primers were used for sequencing (although GScp994r was used in one case). Primer GScp911r was designed in exon 10 (Fig. 2) to screen for heterozygosity at a particular site in intron 9, without interference from the indel in intron 10. In order to reduce the effect of "PCR drift" (Wagner et al., 1993) the products of several (usually 3–5) 25- μ L reactions were pooled, rather than performing a single larger (100- μ L) reaction.

Sequence alignment and phylogenetic analysis—As noted previously (Emshwiller and Doyle, 1999) alignment of the ncpGS sequences of the *O. tuberosa* alliance is generally straightforward and unambiguous. Therefore, DNA sequences were aligned by visual inspection, using Microsoft WordPad (Windows 95 accessory, Microsoft, California, USA), and further data entry and editing used Dada version 12 (Nixon, 1998). Ambiguity exists in placement of some gaps in areas of sequence repeats, but because the different placements do not overlap other informative characters, the different placements of gaps should have equal effect on the results of analysis. Overlapping gaps in intron 8 were treated as multistate characters, while binary characters were added to the matrix to represent other gap characters. Heterozygous sites and indels in individuals that were not cloned were coded as subset polymorphisms. Discussion of indel and substitution characters below will follow the numbering of aligned sites as presented in an alignment of ncpGS sequences (Emshwiller, 1999, Appendix 5.1), which is available as: (1) an aligned “popset” in GenBank including all of the sequence between the amplification primers (729 aligned nucleotide sites), associated with the individual accessions GBAN-AF470234 to GBAN-AF470317 and GBAN-AF098977 to GBAN-AF098984 (the prefix “GBAN-” has been added to Genbank accession numbers to link the online version of the *American Journal of Botany* with GenBank but is not part of the actual accession number) and (2) an alignment that also includes 23 gap characters, archived at <http://ajbsupp.botany.org/v89/emshwiller/appendix1>.

Phylogenetic analyses were performed using Nona (version 1.6 for Windows NT; Goloboff, 1998), using the search strategy hold/50; hold*; mult*100; max* (i.e., initially 50 trees are held from each of 100 replicate analyses, followed by tree bisection-reconnection (TBR) branch swapping on all of the trees found). Clados (version 17; Nixon, 1996) or WinClada (version 0.9.99m24[beta]; Nixon, 2000) was used to examine the maximally parsimonious trees (MPTs) and character optimizations. The sources of homoplasy and causes of multiple topologies were explored by running additional analyses to determine the effect of the inclusion and exclusion of particular sequences on the results (e.g., cloned sequences of cultivated oca that appeared to be recombined and some direct sequences from heterozygous accessions, which caused an increase in the number of trees and loss of resolution in the consensus trees; see RESULTS). Trees were rooted with *O. laxa* var. *hispidissima* and/or *O. pachyrrhiza* and *O. megalorrhiza*, which were suggested to be “basal” (earliest diverging) among the taxa sampled here by the results of independent analysis of 5.8S and ITS2 sequences of a diverse sample of *Oxalis* species (data not shown) as discussed in Emshwiller and Doyle (1998).

RESULTS

Characteristics of ncpGS sequences—Amplification of ncpGS in *Oxalis* generally resulted in a single amplification product. Fainter bands appeared inconsistently in some reactions, but their sequences were either unreadable or not similar to any form of glutamine synthetase or any other GenBank accessions as determined by a BLAST (Altschul et al., 1990) similarity search. Thus there is no evidence either of paralogous copies of ncpGS or of any unintended amplification of cytosolic glutamine synthetase. The general congruence with previous ITS results (see below and Emshwiller and Doyle, 1998, 1999) also supports the conclusion that ncpGS is single copy in these *Oxalis* taxa and therefore free of problems with paralogy. Base composition, transition/transversion ratios, and lengths of the total amplified region and of the exons and introns are similar to those presented previously (Emshwiller and Doyle, 1999), although additional indels add more variation to the intron lengths (Fig. 2).

Indel variation among ncpGS sequences—The indels in intron eight are more numerous, and many are larger, than those in the other three introns. Greater length variation in intron 8 has also been observed in sampled taxa of legumes (J. L. Doyle and

J. J. Doyle, Cornell University, unpublished data). Although this length variation could potentially create problems for alignment among more divergent sequences, it was a good source of characters in this study. The discussion of indel variation below refers to positions in the alignment archived at <http://ajbsupp.botany.org/v89/emshwiller/appendix1>, and the indels as designated at <http://ajbsupp.botany.org/v89/emshwiller/appendix2>.

Many of the indels (both small and large) seem to have resulted from slippage-like processes (Levinson and Gutman, 1987; Hancock, 1995), which can make alignment ambiguous, but not necessarily problematic as long as the indels do not overlap other informative characters. Examples include single base indels in small homopolymer runs (e.g., sites 117, 137, and 608), addition or deletion of dinucleotide repeats (sites 614–615 and 620–621), and a 20-base duplication (sites 314–333). Notably, two deletions that each involve a unique sequence flanked by repeated segments (137–148 and 379–398) remove most of the T-rich region near the 3' intron splice junction, which is putatively important for intron splicing (Csank, Taylor, and Martindale, 1990; Ko et al., 1998). Perhaps these latter deletions are only tolerated because the plants have other functional copies of ncpGS, as they are both probably polyploid (i.e., one cloned accession of oca and one of the wild tuber-bearing *Oxalis* of Bolivia).

Alignment is uncertain for indels in homopolymer runs, because the indel might be at any position along the run. This is the case for the sixth thymine inserted in a run of five (site 117), for which accession EE249 of *O. spiralis* is homozygous and accession EE500 of *O. picchensis* is heterozygous. Coding both plants that have six thymines with the same character state treats these individuals as sharing a homologous indel. However, the separate placement of these two species on the ncpGS trees (see below) indicates that these probably represent separate insertion events.

Sequence heterozygosity—In addition to the ncpGS sequence heterozygosity in cultivated *O. tuberosa* and in the wild tuber-bearing populations of Bolivia, which is discussed separately below, some of the other wild *Oxalis* sampled were heterozygous for either substitutions or indels (whereas 24 other sampled plants had a single sequence type for the amplified region of ncpGS). These other heterozygous plants, whose ncpGS sequences were not cloned, were coded as polymorphic for their heterozygous sites or indel characters. We recognize that this expedient would not have been acceptable if the objective of the study were to reconstruct a complete gene tree of all alleles for all accessions, as it did not accurately represent the situation of having multiple sequences in an individual plant. Some cases of polymorphism coding of heterozygous individuals had little or no effect on the analyses, whereas other cases had greater effect. Nonetheless, molecular cloning of the individuals concerned was still not performed because doing so was not expected to contribute to understanding the origins of oca.

This form of coding did not create problems when the heterozygosity occurred in noninformative characters. Unique changes would be autapomorphic in the phylogenetic analyses, but coded as polymorphic the sequences are treated as identical. When the characters were informative, polymorphic coding had variable effects on the phylogenetic analyses. Two polymorphic characters (sites 559 and 591) did not appear as steps on the trees at all. If the alleles in heterozygous EE797 had been coded separately, the apomorphic state at site 559

would have united one of them with those of EE807. Site 591, however, would have been homoplasious, as even if EE746 had been homozygous for the apomorphic state at site 591, it would have appeared on the MPTs as a separate origin from the alleles in EE916 and EE504. Polymorphic coding of four informative substitutions (sites 1, 7, 144, 390) in EE871, a plant from an area in which hybridization may have been occurring, caused additional rearrangements in the “*O. lucumayensis* group” (see below), but the consensus tree was unchanged.

Considerably more effect on the phylogenetic analyses was caused by EE511 and EE512, which were heterozygous for two indels (“s” and “u”). Part of the sequences between those two indels could not be read without cloning, which was not undertaken due to the focus of this project on the origins of oca. Indel “s” is shared with several other taxa and is a non-homoplasious character on the MPTs, whereas indel “u” was found only in EE511, EE512, and EE960 (this deletion was not observed in homozygous form in any of the plants sampled). EE511 and EE512 are also heterozygous for three informative substitutions, which can be used to infer the positions of the (presumably) two sequences on the MPTs (see **Phylogenetic results 2**, below). As expected, inclusion of these sequences, coded as polymorphic, caused a dramatic increase in the number of MPTs and loss of resolution in the *O. peduncularis* clade. On the other hand, although EE960 is heterozygous for six substitutions, it had only a single indel, so its entire sequence was read in at least one direction. Inclusion of EE960, even if coded as polymorphic at the pertinent sites, did not cause an increase in the number of trees, but joined the base of the clade that would have included both of its separate sequences (see **Phylogenetic results 2**, below), which are not as divergent as those of EE511 and EE512. While this resembles the behavior of hybrids in morphological analyses (e.g., McDade, 1990, 1992), it does not truly represent the situation of having two different alleles in the plant, since if the two sequences had been cloned, they might have each joined separate branches.

Most of the accessions in the examples above, or different plants of the same population, have been inferred to be diploids by flow cytometry (Emshwiller, 2002b), with the exception of EE960, which was no longer alive at the time of the flow cytometry study. Thus, their sequence heterozygosity seems to represent either normal allelic polymorphism or interspecific hybridization. Accession EE500, in contrast, is inferred to be tetraploid, but it may be autopolyploid, as its sequences differ only very slightly (see description of its heterozygosity for a homoplasious single base insertion in a homopolymer region, above).

Cloned sequences of cultivated oca and wild tuber-bearing *Oxalis*—Chloroplast-expressed GS sequences were determined for 36 molecular clones from three morphotypes of cultivated oca (11 clones from MHG884, 10 from MHG913, and 15 from 35-04) and eight clones from one accession (EE259) of the unnamed Bolivian wild tuber-bearing taxon. However, some of these clones (originally designated class “A”) were later determined to be contaminants in the PCR reactions (six of those of MHG884 and one of MHG913, but none of 35-04). Figure 3 shows only the variable sites (and indel characters) among the cloned sequences, excluding the contaminant sequence class (these sequences were similar or identical to that of accession EE184 of *O. spiralis* and EE359 of *O. mollissi-*

ma). A “hypothetical ancestor” sequence was included (showing the states at the base of the $x = 8$ group) so that the apomorphic states of each sequence can be seen more easily.

PCR recombination and Taq error—In the course of grouping the cloned sequences into similar classes that might represent the homeologous loci, it appeared that some of the sequences were the results of PCR artifacts such as “Taq error” and “PCR recombination” (Jansen and Ledley, 1990; Bradley and Hillis, 1997; Cronn et al., 2002). Among the multiple clones from each accession of cultivated oca there were relatively few that were identical (Fig. 3). Twelve clearly different sequences were recovered from accession 35-04, more than the eight that are theoretically possible for a single-copy gene in an octoploid plant, making it clear that some sequences must be artifacts.

Among these kinds of possible PCR artifacts, shuffling of sequences by PCR recombination was especially challenging for the interpretation of possible homeologous loci. PCR recombination is the formation of artifactual sequences in vitro that combine the features of different template sequences in a heterogeneous reaction mixture, such as different alleles, paralogues, or homeologues, and is thought to occur when uncompleted PCR products act as primers in subsequent cycles, re-annealing to different templates (reviewed in Cronn et al. [2002]). Rather than being a rare phenomenon, recombinants can make up a large proportion of PCR products, particularly when cloning is used as an intermediate step (Jansen and Ledley, 1990 [25% recombinants]; Bradley and Hillis, 1997 [43%]; Cronn et al., 2002 [up to at least 89% in different gene systems]). In theory, the proportion of recombinant artifacts might be even higher in an octoploid, because, with a higher number of different template sequences, there might be a correspondingly higher chance of re-annealing to one of the “wrong” sequence types. Thus, although recombination among sequence classes could truly occur in an octoploid, we suspect that the recombination observed here is probably an artifact of cloning PCR products.

Certain of the cloned ncpGS sequences appear to be recombined because they have a mixture of the character states found in other sequences (Fig. 3). A simple example is EE259 clone 1, which resembles class “B” clones 2, 3, 8, 9, 10, and 11 (from the same plant) at the 5' end and middle, but resembles class “D” clone 7 at the 3' end. In this case a single recombination event would be inferred, but other examples may represent the products of several recombinations (e.g., MHG884 clone m2; Fig. 3). In this study, unlike the situation in cotton (Cronn et al., 2002), the diploid progenitors are not known a priori. Nevertheless, in most cases the nonrecombinant conditions could be inferred through a comparison of diagnostic nucleotides in ncpGS sequences of wild *Oxalis* species, as well as information from the phylogenetic context (see below). In the example above, the character states at the 5' and middle of EE259 clone 1 would place it in one part of the tree, while those at the 3' end would place it in a different part of the tree. In some cases the inclusion of the putative recombined sequences in cladistic analyses increased the numbers of MPTs and caused collapse of some nodes in the consensus tree; in other cases their inclusion only increased homoplasy and tree length, as the recombined characters appeared as extra steps on the trees. Clones that do not appear to be recombined cause neither loss of resolution nor increase in homoplasy and join

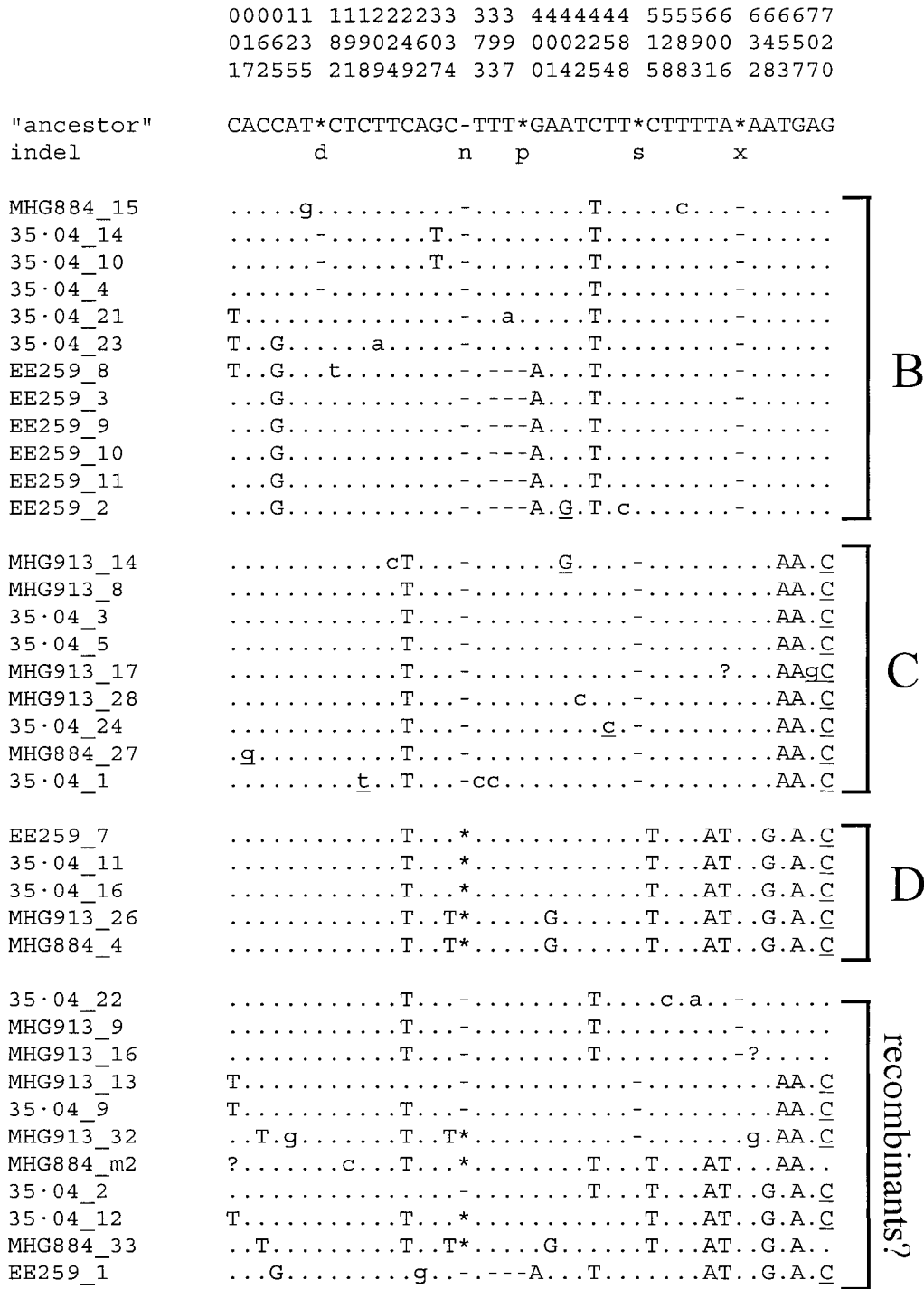


Fig. 3. Alignment of only the variable characters among cloned sequences of cultivated oca (accessions MHG884, MHG913, and 35·04) and the wild tuber-bearing *Oxalis* of Bolivia (accession EE259). Clone numbers follow accession numbers at the left side, and sequence classes B, C, and D, and the suspected PCR recombinants are indicated along the right. The hypothetical "ancestor" sequence (of the *O. tuberosa* alliance) is included so that the apomorphic states of each sequence are more visible, with the site number from the sequence alignment indicated above the ancestral state. Indel characters are positioned in the appropriate order among the substitution characters, with the indel character designations (from <http://ajbsupp.botany.org/v89/emshwiller/appendix2>) below the "ancestor" sequence. In the case of deletions, an asterisk appears in the "ancestor" sequence and dashes in the individual sequences, whereas the reverse is shown in the case of insertions. Singleton changes are shown in lowercase letters, while underlining indicates non-synonymous substitutions (see discussion of *Taq* error in the text).

the same three positions on the ncpGS tree that are discussed below (*Phylogenetic results 2*).

DNA polymerases that lack a proofreading ability, such as were used here, are prone to substitute incorrect nucleotides occasionally, a phenomenon commonly referred to as *Taq* error. Although these data alone cannot distinguish artifacts from real substitutions with certainty, singleton changes (shown in lowercase letters in Fig. 3) that occur in only one cloned sequence and are not shared by any sequences from wild species are more likely to be the results of *Taq* error than substitutions that are shared by different clones, plants, or species. Some singleton substitutions might be mutations that do exist in the plant, but the underlined bases would represent non-synonymous substitutions and so may be particularly suspect. Mistaken nucleotides that result from *Taq* error will usually be autapomorphic in the results of phylogenetic analysis, so *Taq* error is considered less of a problem for this study than PCR recombination.

Variation within cloned sequence classes—The cloned sequences that do not appear to be recombinants seem to fall into three different classes, designated at the right of Fig. 3 as classes “B,” “C,” and “D.” These sequence classes group in different places on the MPTs in phylogenetic analyses (see below). All three classes were present in each of the three plants of cultivated oca in the cloned sample, whereas class C was absent in one of nine oca plants screened by direct sequencing (see below). Two sequence classes, B and D, are present in the cloned accession (EE259) of the wild tuber-bearing taxon of Bolivia, and three additional accessions sequenced directly.

Variation was also observed among the sequences within each of the sequence classes. Although some of the differences among sequences may be artifactual, others probably represent true allelic variation. For example, accessions 35-04 and EE259 each have deletions that are not found in the other oca or Bolivian wild tuber-bearing accessions that were sequenced directly. Although length variation can also be an artifact of PCR (Fenton, Malloch, and Germa, 1998), the deletion in EE259 was confirmed in direct sequences performed under the same conditions as the other accessions. Some substitutions were shared by clones from more than one of the plants, so they were probably real variants, although they sometimes appear in more than one sequence class, making it uncertain to which class they really belong (e.g., it is unclear which sequence class was truly on the same strand with the transition substitution at site 1 in both the alignment and Fig. 3).

It is noteworthy to have encountered ncpGS sequence variation among plants in this very small sample of three cloned accessions. Oca cultivars are variable in traits such as tuber morphology and pigmentation, nutritional factors, insect resistance, phenology, and yield (Castillo, 1974; Poma Machaca, 1976; Bustinza López, 1979; Cortés Bravo, 1984; King and Gershoff, 1987; Arbizu et al., 1997). However, among molecular markers only low levels of variation are reported for isozymes (del Río, 1990), tuber proteins (Stegemann, Majino, and Schmiediche, 1988; Shah, Stegemann, and Galvez, 1993), and random amplified polymorphic DNAs ([RAPDs]; A. Donayre, Universidad Nacional Mayor San Marcos, Lima, Peru, personal communication; G. Piedra, Instituto Nacional de Investigación Agropecuaria, Quito, Ecuador, personal communication), but variability appears to be greater in AFLP markers in initial assessments (Tosto and Hopp, 2000; E. Emshwiller, un-

published data). At this level of sampling it is not possible to determine whether the ncpGS sequence variation among plants represents multiple origins of polyploidy or domestication, mutations that arose after the origin of the crop, or loss of alleles through sexual recombination.

Phylogenetic results 1: analyses excluding cloned sequences—Separate analyses were run with and without the cloned sequences of cultivated oca and the Bolivian wild tuber-bearing taxon and the accessions with sequence heterogeneity and missing data that caused loss of resolution when included, as discussed above (i.e., EE511, EE512, and EE960). One of the 20 MPTs that resulted from an analysis that excluded these sequences is shown in Fig. 4, which also indicates the branches that collapse in the consensus tree. For purposes of the following discussion three of the clades in Fig. 4 are designated as the “*O. lotoides* group,” the “*O. lucumayensis* group,” and the “*O. peduncularis* clade.” Although the latter clade is resolved as monophyletic in all analyses of ncpGS data, the *O. lotoides* group and *O. lucumayensis* group are resolved in various ways in the analyses that include the cloned sequences, so they are not necessarily monophyletic groups (in an analysis of combined ITS and ncpGS data these two groups join a single clade referred to as the “*O. lotoides* clade” in Emshwiller [2002a]).

The 20 topologies found in this analysis differ from each other in three easily identifiable ways. Firstly, when *O. laxa* var. *hispidissima* is included, the sequences of *O. ortgiesii* and *O. andina* either form a grade or a clade in alternative trees (see dashed lines in Fig. 4). Their resolution as a grade as shown in Fig. 4 was found in analyses of ITS data alone (Emshwiller and Doyle, 1998), combined ITS and ncpGS data (Emshwiller, 2002a), and analyses of ncpGS data that excluded *O. laxa* var. *hispidissima*. Secondly, *O. flagellata* either joins with *O. petrophila* or at the base of the *O. peduncularis* clade (it shares one substitution character with each of these). The resolution shown in Fig. 4 was chosen because of the morphological similarities between *O. flagellata* and *O. petrophila*. Thirdly, sequences within the *O. lucumayensis* group are arranged in five different ways, due to both character conflict within the group and polymorphic coding of heterozygosity in accession EE871 of *O. lucumayensis* ssp. *lucumayensis* (see above).

Phylogenetic results 2: analyses including cloned sequences—Analyses that included the three sequence classes of cultivated oca and two classes of the wild tuber-bearing plant EE259 found 208 MPTs, one of which is shown in Fig. 5. Cloned sequences that were judged to be PCR recombinants or contaminants (see above) were excluded to avoid problems associated with including putative recombined sequences in analyses. Heterozygous accession EE871 of *O. lucumayensis* ssp. *lucumayensis* was included in the analysis shown (Fig. 5). Analyses that excluded this sequence had only 72 MPTs, but the strict consensus was the same in either case. As above, this analysis excluded the sequences of EE511 and EE512, because when included, their missing data and polymorphic coding caused an increase in the number of MPTs (to 810 trees) and significant loss of resolution (results not shown). However, by assuming nonrecombination of characters found in the parts of these sequences that were determined, the positions on the tree that these sequences would join could be inferred, indicated by the asterisks in Fig. 5.

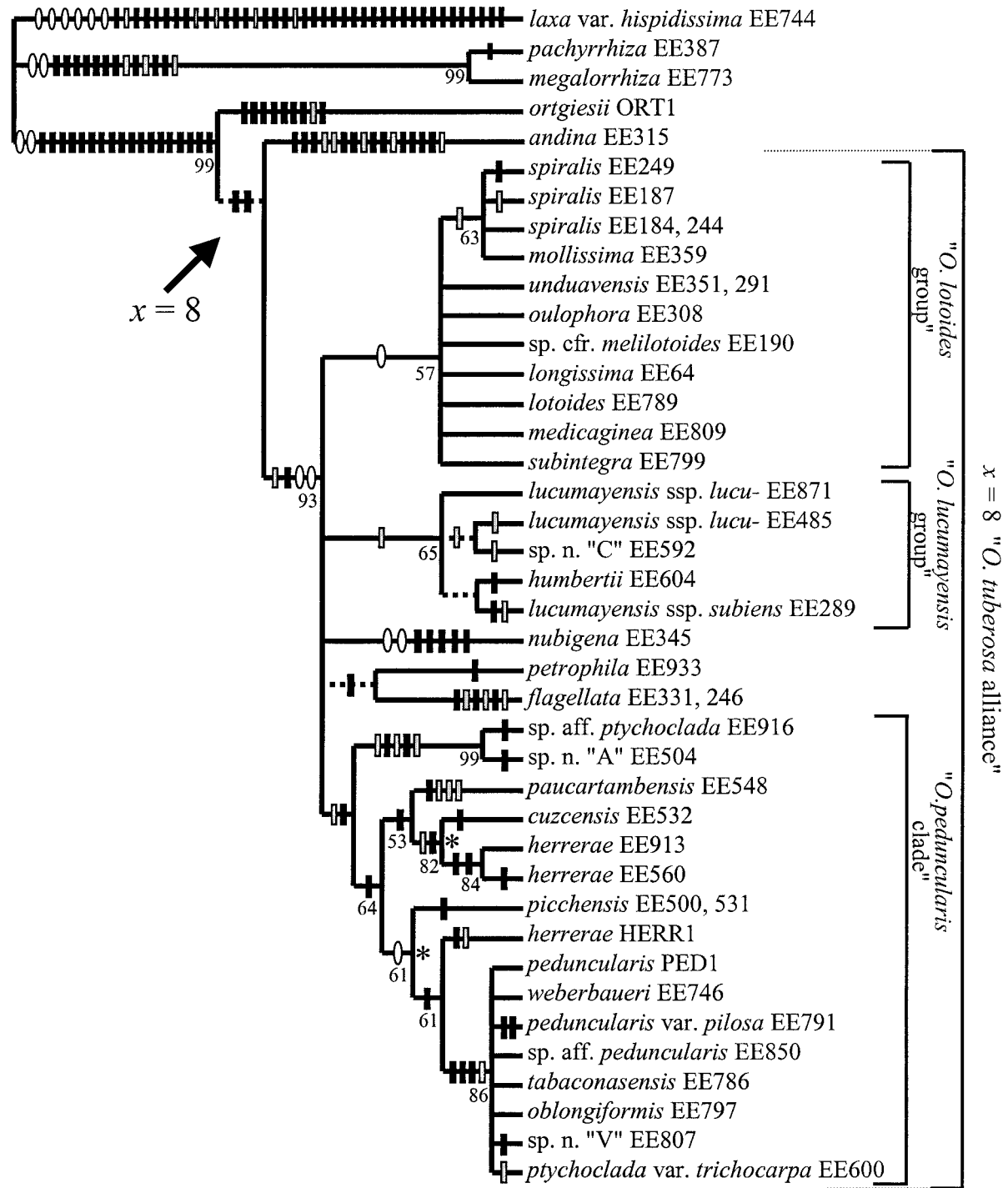


Fig. 4. One topology among the 20 maximally parsimonious trees found in analysis of ncpGS sequences of *Oxalis* that excluded cloned sequences of cultivated oca and wild tuber-bearing populations of Bolivia. This analysis resulted in trees of 124 steps, with a consensus index (CI) of 0.79 and an retention index (RI) of 0.85. Accession numbers follow the names of each species. Bracketed groups indicated at the right are discussed in the text. Black hashmarks indicate changes in nonhomoplasious characters, grey hashmarks indicate homoplasious changes, and white oval hashmarks indicate indel characters (none of these are homoplasious in this analysis). Arrow indicates branch where the base chromosome number $x = 8$ is inferred to have arisen. Branches that are not resolved in the strict consensus are indicated by dashed lines. Numbers below branches indicate bootstrap values from 1000 replicates.

The analyses that included the cloned ncpGS sequences found other rearrangements of taxa in addition to the alternative positions (discussed above) of *O. andina* and *O. ortgiesii*, *O. flagellata*, and the members of the *O. lucumayensis* group. These additional rearrangements merit further discus-

sion because of their bearing on the interpretation of the origin of the class "B" sequences of oca. When the cloned sequences were excluded, the *O. lucumayensis* group and *O. lotoides* group were each resolved as separate, albeit poorly supported, clades (Fig. 4). However, the cloned class B sequences of oca

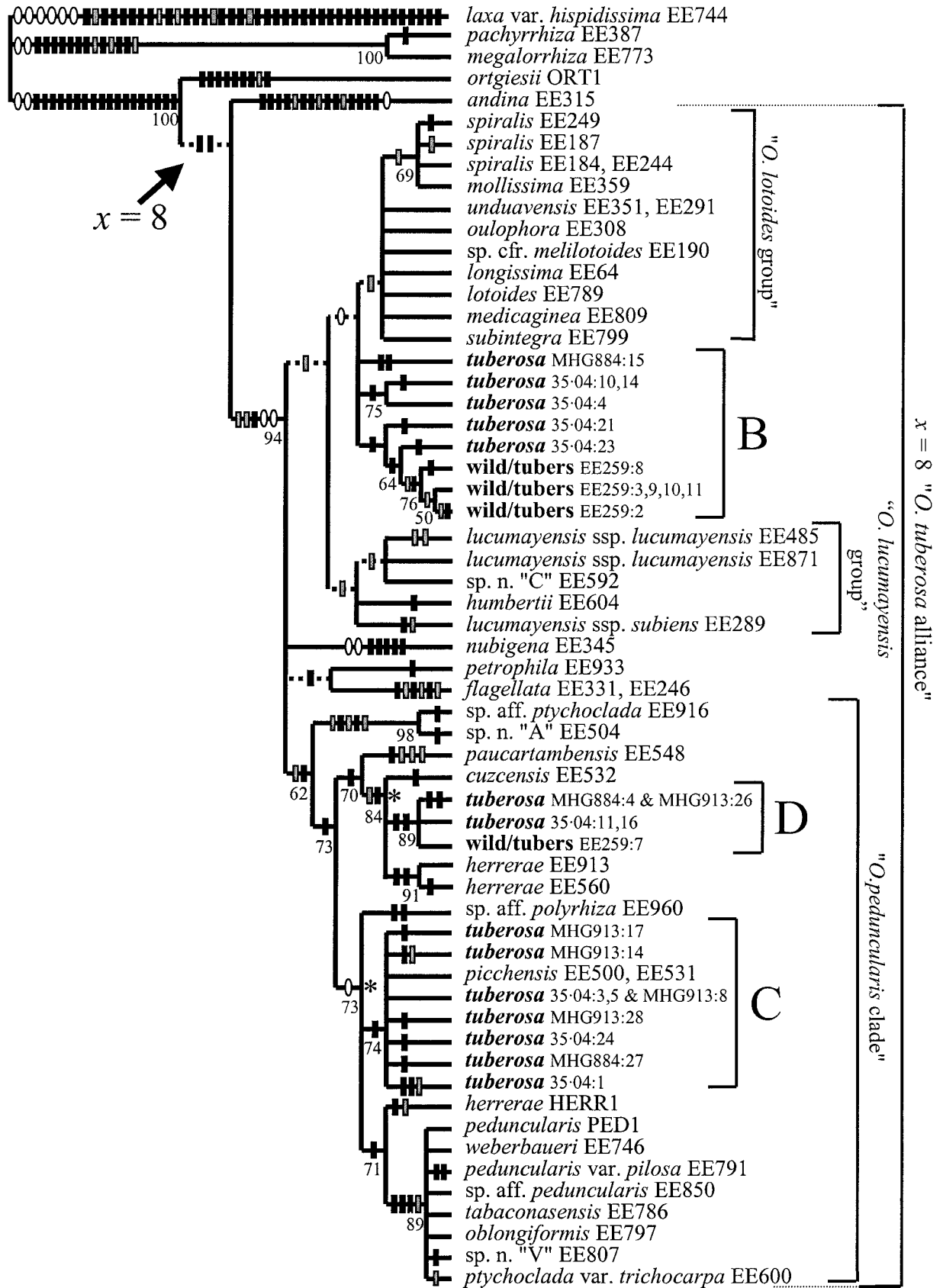


Fig. 5. One topology among the 208 maximally parsimonious trees found in analysis of ncpGS sequences of *Oxalis* that included cloned sequences of cultivated oca and wild tuber-bearing populations of Bolivia, with classes indicated by brackets around sequence classes B, C, and D. The CI with autapomorphic

and EE259 had the apomorphic states at each of the characters that supported these two clades. Hence, homoplasy arises between these characters when the cloned class B sequences are included, so that the *O. lucumayensis* group and *O. lotoides* group lose resolution in the strict consensus (only the clade that includes sequences of *O. spiralis* and *O. mollissima* remains intact; see Fig. 5). However, examination of the trees shows that the different topologies have four kinds of arrangements among these two groups (Fig. 6). In two of the four arrangements, the class B sequences of oca and EE259 are grouped together as a clade, whereas in the other two topologies, they are unresolved (i.e., they would form an internal node in network methods designed to account for surviving ancestral haplotypes, e.g., Templeton, Crandall, and Sing [1992]). It is not necessary that the sequences of cultivated oca and a progenitor candidate form a clade and share a unique synapomorphy. It is sufficient to have a paraphyletic group of sequences that share a set of character states not shared by any of the other sequences in the matrix. In this sense the class B sequences of oca and EE259 match each other, regardless of autapomorphic changes that are present in some of the clones. Thus, the variation among these topologies does not weaken the support of ncpGS data for the possible role of the wild tuber-bearing *Oxalis* of Bolivia in the origin of oca.

Although the four different arrangements among these groups are equally parsimonious, there may be some basis for favoring some of the MPTs over others. In the analyses excluding the cloned sequences, the *O. lotoides* group and the *O. lucumayensis* group are each supported by a single character: a two-base deletion and a transition substitution, respectively. Transition characters are more homoplasious than indels in this matrix in general and between these two characters in particular (data not shown). The tree chosen for Fig. 5 is like topology c in Fig. 6, in which there is a single change in the deletion character and a reversal in the transition character. On the other hand, the sharing of these characters in the class B sequences might derive from an ancient recombination event, rather than parallelism or reversal.

Comparison with previous ITS results and cytology—The results of phylogenetic analysis of ncpGS sequences are congruent overall with our previous ITS results (Emshwiller and Doyle, 1998). One character on the ITS gene tree conflicts with each of the ncpGS characters that unite the *O. lotoides* group, the *O. lucumayensis* group, and the sequences of *O. spiralis* and *O. mollissima* (Figs. 4 and 5; see also fig. 3 in Emshwiller and Doyle, 1999). In the latter case the ncpGS data seem to be more congruent with morphology and species boundaries than those of ITS.

There is somewhat more divergence overall among sequences of ncpGS than ITS, allowing more resolution of relationships, particularly in the *O. peduncularis* clade (Figs. 4 and 5). Only one uninformative substitution in the entire ITS region distinguished the sequences of the purchased plants of *O. peduncularis* (PED1) and *O. herrerae* (HERR1) (Emshwiller and Doyle, 1998), and the sequence of the latter was identical

to that found in *O. peduncularis*, *O. villosula*, *O. tabaconasensis*, and *O. herrerae* by Tosto and Hopp (1996). Additional sampling has not revealed any informative ITS variation among the taxa in this clade (Emshwiller, 2002a), whereas their ncpGS sequences form a clade supported by a total of nine characters. The greater divergence in ncpGS sequences than those of ITS is not seen among all taxa, however. The members of the *O. lotoides* group have little divergence in either ncpGS or ITS, in spite of considerable morphological diversity. Their ncpGS sequences are not identical, however, as the apparent branch lengths of zero for most of the species in this group (Figs. 4 and 5) masks heterozygosity in at least one autapomorphic site in nearly all of these plants. Nevertheless, although the accessions in the *O. lotoides* group were collected from a broad geographical area (from northern Peru to Bolivia) and are members of several distinct species, they have little molecular variation in the two loci studied so far.

The only member of the *O. lucumayensis* group included in our prior ITS study was accession EE289 of *O. lucumayensis* ssp. *subiens* (determined as *O. sp. aff. distincta* at that time; Emshwiller and Doyle, 1998). Subsequent sampling for ITS in this group (Emshwiller, 2002a) found two ITS sequence types (designated “B” and “C” in Emshwiller and Doyle, 1998), differing by a single substitution, that are also found in the *O. lotoides* group. The occurrence of hybridization among species in the *O. lucumayensis* group, suggested at first by the observation of morphologically intermediate individuals and supported by multiple ITS sequences in accession EE294 of *O. lucumayensis* ssp. *subiens* (Emshwiller and Doyle, 1998), is further supported by heterozygosity of ncpGS in accession EE871 of *O. lucumayensis* ssp. *lucumayensis*. Further study might clarify whether these observations are due to hybridization or simply to highly polymorphic species in this group.

Our previous ITS results (Emshwiller and Doyle, 1998) supported the monophyly of the cytologically based *Oxalis tuberosa* alliance (de Azkue and Martínez, 1990) with the inclusion of additional species for which cytological data are as yet unavailable. The ncpGS data add more support to the alliance, not only because of the congruence of the gene trees and the addition of more molecular synapomorphies of the alliance (including a 31 bp deletion), but also because of the addition of more species reported to share $x = 8$, such as *O. lotoides*, *O. medicaginea*, *O. tabaconasensis*, *O. oblongiformis*, and *O. ptychoclada* (Favarger and Huynh, 1965; Huynh, 1965; de Azkue and Martínez, 1990). The $x = 8$ clade is retained and enlarged when sequences of these taxa are added.

As mentioned above, *O. andina* has recently been reported to have 16 chromosomes (de Azkue, 2000), and thus it also shares $x = 8$. However, neither ITS (Emshwiller and Doyle, 1998) nor ncpGS (see below) indicate that *O. andina* or its allies were involved in the origins of oca. Thus, we consider the “*O. andina* clade” (Fig. 1) to be the sister group of the *O. tuberosa* alliance, rather than part of the alliance itself. Nevertheless, the discovery that *O. andina* also has $x = 8$ is consistent with a single origin of this base chromosome number within *Oxalis*, with the modification that the $x = 8$ clade

←

characters removed is 0.78, the RI is 0.87. Accession numbers follow the names of each species. Bracketed groups indicated at the right are discussed in the text. Black hashmarks indicate changes in nonhomoplasious characters, grey hashmarks indicate homoplasious changes, and white oval hashmarks indicate indel characters. Branches that are not resolved in the strict consensus are indicated by dashed lines. Numbers below branches indicate bootstrap values from 1000 replicates.

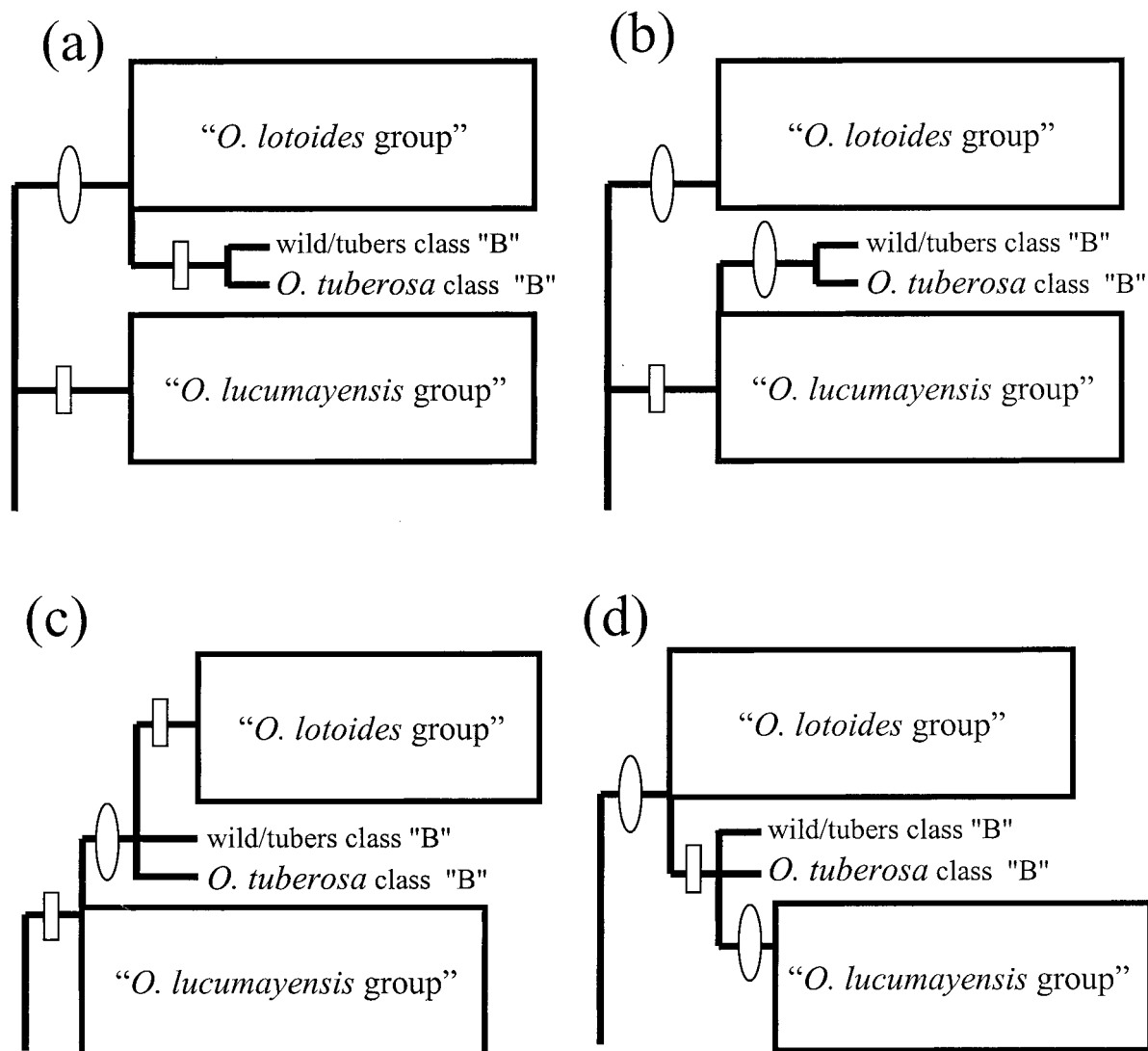


Fig. 6. Four kinds of arrangements among the class B sequences of cultivated *O. tuberosa* and the wild tuber-bearing populations of Bolivia and the members of the *O. lucumayensis* group and the *O. lotoides* group (see text). The single characters that supported each of the latter groups as clades in analyses that excluded cloned sequences are shown. The indel character that supported the *O. lotoides* group (designated "x" in <http://ajbsupp.botany.org/v89/emshwiller/appendix2> and Fig. 3, sites 620–621 in the alignment at <http://ajbsupp.botany.org/v89/emshwiller/appendix1>) is shown as an oval, the transition character that supported the *O. lucumayensis* group (site 425 in alignment and Fig. 3) as a rectangle. Trees (a) and (c) have a single origin of the two-base deletion, with homoplasy in the transitional substitution whereas trees (b) and (d) have homoplasy (either two origins or a reversal) in the indel character and a single change in the substitution character. The relationship among these groups is unresolved in the strict consensus.

circumscribes a larger group than the *O. tuberosa* alliance sensu stricto. Although cytological data are still lacking for many of the taxa whose sequences group in this clade, there are as yet no members of this larger clade known to have a base chromosome number other than eight. The obverse is also true, in that no *Oxalis* taxa outside of this clade are reported to have $x = 8$. Although fewer highly divergent taxa were included in the ncpGS sample than that of ITS, the monophyly of the $x = 8$ alliance is further upheld with the inclusion of the additional outgroups *O. laxa* var. *hispidissima* and *O. megalorrhiza*, neither of which is based on $x = 8$. Although there are no chromosome number reports for *Oxalis laxa* var. *hispidissima*, there is a report of $2n = 18$ for the morphologically similar *O. micrantha* (Naranjo et al., 1982). Species boundaries between these taxa have been delimited in various ways by different workers (as inferred from specimen annotations), and

O. micrantha var. *setifera* is considered a synonym of *O. laxa* var. *hispidissima* by Lourteig (1988, 2000). A similar confusion surrounds chromosome number reports for *O. megalorrhiza*. A count of $2n = 18$ is reported by de Azkue (2000) for *O. pachyrrhiza* and for *O. carnosa* (= *O. megalorrhiza*, see above). The same number was reported by Diers (1961) for *O. solarensis* Knuth, also considered to be a synonym of *O. megalorrhiza* (Pool in Brako and Zarucchi, 1993; Lourteig, 2000). However, Heitz (1926; cited in Federov, 1974) reported a count of $2n = 14$ for *O. carnosa*. Nevertheless, neither of these conflicting counts would include *O. megalorrhiza* in the $x = 8$ alliance.

Screening of oca and wild tuber-bearing accession EE259 by direct sequencing—Direct sequencing of ncpGS from accessions of cultivated oca and the wild tuber-bearing taxon

using internal primers was employed at first to test whether all of the sequence classes (possible homeologues) found among the cloned sequences were present in a larger sample and later to confirm that the cloned class A sequences had derived from contamination. In the course of this screening of direct sequences, one other anomalous result was encountered. In eight out of nine oca accessions it was possible to confirm the presence of sequence classes B, C, and D. However, direct sequences of accession 02-08 showed no sign of the two characters, an indel and a substitution (positions 38 and 49, respectively, in Fig. 3; and positions 503 and 654 in alignment), that distinguish the class C sequences. The absence of this sequence class from accession 02-08 was confirmed by direct sequencing of two separate preparations of PCR products (from the same template genomic DNA), each of which comprised pooled products of 3–5 amplification reactions. Thus this plant has sequence classes B and D, but does not have class C.

Variation was also observed among direct sequences of ncpGS of three additional plants from the Bolivian wild tuber-bearing populations, only one of which (EE260) was collected from a locality relatively close to that of EE259. All three plants lacked two autapomorphies found in the cloned individual EE259 (a 20-base deletion and a nearby substitution) that distinguish the class B sequences of EE259 from those of oca. Thus, the class B sequences of these other three plants are better matches with those of oca than are the cloned class B sequences of EE259. Additional variation appears in the form of an apomorphic character state that occurs (as heterozygous) in two of the four plants. Little can be concluded from this variability at this level of sampling. However, it does indicate that diversity among the wild populations may allow future studies to identify which were involved in the origins of cultivated oca and to study the possibility of multiple origins of polyploidy or domestication.

DISCUSSION

One of the advantages of molecular methods in plant systematics is their potential for the study of the origins of polyploid genomes: to determine the number of independently derived genomes that make up the polyploid and to identify the “progenitors” of the polyploid, or more properly, the extant diploid taxa that are derived from the same ancestral diploid genomes. By using nuclear loci, one hopes to see evidence of all the genomes of the polyploid through the presence of homeologous loci. That is, the expectation that the polyploid will be “additive with respect to its diploid progenitors” and that sequences of loci in the polyploid will be “phylogenetically sister to their counterparts (orthologues) from the respective diploids” form a convenient “null hypothesis” (Wendel, 2000, quotes all p. 238) for study of polyploid origins. Recent studies have shown, however, that genome evolution in many polyploids is considerably more complex than this simple scenario and can be complicated by various processes such as multiple origins of polyploids (Soltis and Soltis, 1993, 1999), heterozygosity in diploids that may be carried over to polyploids to greater or lesser extent depending on the mode of polyploidization (Watanabe and Peloquin, 1991), and nonadditivity due to non-Mendelian changes in recently formed polyploids, which can include loss of homeologous loci, gene conversion, and concerted evolution (reviewed in Soltis and Soltis, 1999;

Wendel, 2000). Thus, the interpretation of DNA sequence data in polyploids is not always straightforward.

The paucity of information about oca makes the interpretation of its sequence data even more complicated. Unlike the situation in more thoroughly studied polyploid crops, there are no prior hypotheses of its origin derived from morphological or cytological data for the molecular data to test and some of the cytological reports are conflicting. Most recent workers have found oca to have $2n = 8x = 64$ in well over 100 oca accessions from diverse areas of the Andes (de Azkue and Martínez, 1990; Medina Hinojosa, 1994; Valladolid, Arbizu, and Talledo, 1994; Valladolid, 1996; Vinueza Vela, 1997). Vinueza Vela (1997) grouped oca chromosomes into eight homologous sets according to their form and banding pattern, but homeologous genomes were not distinguished. However, there have been conflicting chromosome counts reported in both older and more recent reports (e.g., Heitz, 1927; Talledo and Escobar, 1995; Hayano Kanashiro, 1998). Lower euploid chromosome numbers have been reported for some cultivated oca accessions, notably from communities in areas of Bolivia where wild tuber-bearing populations occur (Guamán, 1997). Screening of ploidy levels in ten cultivated oca accessions by flow cytometry found that these plants had DNA contents roughly four times that of diploid species in the alliance, thus confirming that they are in the octoploid range (Emshwiller, 2002b). This does not preclude the possibility of aneuploidy, however. Vegetative propagation and dispersal by humans mean that oca may not be under the same selection pressures to regain fertility by reducing meiotic abnormalities that operate in seed-propagated species. It is also possible that populations of different ploidy levels were domesticated or that the octoploid level was reached after original domestication at lower ploidy levels. Further screening of other oca cultivars may resolve whether some have different ploidy levels.

Genomes of *Oxalis tuberosa*—In the absence of prior information from chromosome pairing studies, the results of analysis of ncpGS sequences of oca and wild *Oxalis* provide the first evidence of the different genomes present in *Oxalis tuberosa*.

Among the three ncpGS sequence classes found within each of the three genotypes of oca, two classes, B and D, were found in all nine oca accessions sampled by either molecular cloning or direct sequencing. Thus these two sequence classes exhibited fixed heterozygosity, which is consistent with the hypothesis that oca is allopolyploid. There can be problems in using fixed heterozygosity in a single gene, by itself, to infer allopolyploidy. Clonal propagation, which is the rule in cultivated oca, has been shown to maintain fixed heterozygosity in diploid parasitic protozoa (Tibayrenc, Kjellberg, and Ayala, 1990). Autopolyploids can also appear to exhibit fixed heterozygosity because polysomic segregation leads to rare homozygotes, which can escape detection if sampling is insufficient (Vogel et al., 1999). However, in the case of ncpGS data from oca, we have information about the orthologous sequences present in diploids, as well as phylogenetic information from the ncpGS gene tree. Sequence classes B and D grouped in morphologically divergent subclades within the *O. tuberosa* alliance (Figs. 4 and 5), and no diploids were found with sequences from both of these clades. With this phylogenetic information the occurrence of fixed heterozygosity is stronger evidence of allopolyploidy, an interpretation that is also reasonable because higher level polyploids (such as octoploids)

in nature are rarely complete autopolyploids (Stebbins, 1947, 1950). Future studies are planned to test whether fixed heterozygosity is also present at other nuclear loci in oca.

The third sequence class, C, is present in eight out of nine oca plants sampled by either cloning or direct sequencing. The class C sequences appear to be another homeologous locus, representing a third genome type in octoploid oca, an interpretation that would be straightforward if this class were present in all plants sampled. However, its absence from one oca accession (02-08) leaves open several possible explanations (see also Fig. 5.7 in Emshwiller, 1999): (1) It is possible that 02-08 is not octoploid, and so it does not have all the genomes present in other accessions, which would imply that there is ploidy level variation in cultivated oca. This plant was no longer available alive at the time that the flow cytometry study was conducted, so its ploidy level is unknown. (2) All oca accessions may be octoploid, but may have had multiple origins of polyploidy, in which some of the octoploids were formed without the contribution of the genome donor with the class C ncpGS sequence. (3) The class C sequence may have been lost in at least one oca lineage. Recent studies have demonstrated rapid sequence elimination in some polyploids even within a few generations after their formation (Parokony et al., 1994; Song et al., 1995; Escalante et al., 1998; Liu, Vega, and Feldman, 1998; Liu et al., 1998; Ozkan, Levy, and Feldman, 2001; Shaked et al., 2001; see also reviews by Soltis and Soltis, 1999; Wendel, 2000; and Pikaard, 2001). (4) The class C sequence may represent gene flow (e.g., introgression between wild and cultivated populations) after the origins of the octoploid. The class C sequence of oca is not geographically restricted to the range of the wild species, *O. picchensis*, that also has this sequence class (see below). Thus, it would be necessary to suppose that the oca genotypes that contain this sequence were extensively selected and dispersed, by either natural or human means, to explain their predominance (eight out of nine) in the sampled accessions. More problematic is the requirement of this hypothesis for gene flow across differing ploidy levels. (5) The class C and class D sequences join different branches within the same subclade (the *O. peduncularis* clade) on the ncpGS gene tree. Some diploid taxa have been found to be polymorphic for sequence types that fall in similarly separated parts of that subclade (see asterisks in Figs. 4 and 5), suggesting the possibility that classes C and D represent alleles at homologous loci. However, other wild *Oxalis* taxa have been found that have better matches for each of these two sequence classes, and none have been found to be polymorphic for sequence classes C and D or sequences that join them on the gene tree, so it seems more likely that these two sequence classes in oca derive from separate species.

The first three possibilities above are consistent with the idea that the class C sequences are indeed homeologous loci and thus that at least some oca cultivars have three different genomes, represented by the B, C, and D sequence classes. Given current data, these possibilities seem less problematic than the latter two. As an octoploid, the crop may theoretically be derived ultimately from four diploid progenitor species, or at least have four homeologous paired sets of chromosomes. However, if the ncpGS sequence classes that have been distinguished among the cultivated oca clones do indeed represent the homeologous loci, there appear to be three classes, rather than the four homeologous loci that might theoretically be possible. Thus oca seems to be an autoallopolyploid, but the mode

of origin and which of the genomes might be present in greater copy number than the others is yet unknown.

Putative progenitors of *O. tuberosa*—Among the wild *Oxalis* populations that were sampled for ncpGS, two taxa have sequences that match those of the different sequence classes of cultivated oca. One of these is the unnamed wild tuber-bearing taxon from Bolivia, in which the different populations sampled (one accession whose sequences were cloned and three that were sequenced directly) all have sequence classes B and D. Two of the cloned class D sequences of oca (accession 35-04, clones 11 and 16) are identical to one sequence from the wild taxon (accession EE259, clone 7). There is intraspecific variation among the class B sequences of both oca and the wild tuber-bearing populations, so they are not necessarily identical, but they do share a set of characters (see above and Fig. 6) that do not occur together in any other *Oxalis* sampled. As in the case of cultivated oca, the fixed heterozygosity of these two sequence classes, which join morphologically different subclades within the *O. tuberosa* alliance, provides evidence to support the conclusion that the wild Bolivian tuber-bearing populations are probably also allopolyploids. However, cytological information for these populations is as yet unknown, because living material was not available for analysis. Although triploid numbers have been reported for some Bolivian wild tuber-bearing *Oxalis* (Guamán, 1997), these counts have not been independently confirmed. Some of the wild tuber-bearing *Oxalis* sampled in this study (i.e., EE259 and EE260) were collected from populations with all three style morphs present (most *Oxalis* species are tristylous), suggesting that the plants in these populations are reproducing by seed (Emshwiller and Doyle, 1998). Thus, it is unlikely that they could be odd polyploids, which are usually sterile (Allard, 1960).

The class C sequences of oca, on the other hand, were shared with *O. picchensis*, another wild tuber-bearing species found in the department of Cusco, Peru (the sequences of MHG913 clone 8 and 35-04 clones 3 and 5 are identical to that of *O. picchensis*). Estimation of DNA content by flow cytometry indicates that this taxon is tetraploid (Emshwiller, 2002b). It is probably autotetraploid because the two plants sequenced had a single sequence class (one was heterozygous for a single one-base indel).

One interpretation of these data is that these wild tuber-bearing taxa (the populations of Bolivia on the one hand and *O. picchensis* on the other) may both be progenitors of domesticated oca. The Bolivian wild tuber-bearing *Oxalis* taxon may itself be a hybrid of two as yet unknown progenitors and possibly may be either tetraploid or hexaploid. Further hybridization with *O. picchensis* may have resulted in octoploid *O. tuberosa*.

In an autoallopolyploid, one of the homeologous genomes is present in greater copy number than the other(s). In the absence of information on chromosome pairing behavior, the ploidy level of the wild tuber-bearing taxon of Bolivia, or the mode of polyploidization (e.g., “asexual polyploidization,” “unilateral sexual polyploidization,” or “bilateral sexual polyploidization” sensu Mendiburu and Peloquin, 1976), we can only speculate about the dosage of each genome. The relative intensity of the class C peaks in direct sequences of oca is much lower than the others, which might argue for the possibility that the octoploid could have been formed by unilateral sexual polyploidization (i.e., if the wild tuber-bearing taxon of

Bolivia were hexaploid, it might have contributed a $2n [=6x]$ gamete that joined with a normal $1n [=2x]$ gamete from *O. picchensis*). This cannot be considered definitive evidence of gene dosage, however, because PCR amplification conditions can favor one sequence type over another (Wagner et al., 1993).

Even with the caveats discussed above, current data and sampling support both of the wild tuber-bearing taxa as the best candidates as progenitors of domesticated *O. tuberosa*. These two taxa were the only ones sampled that had sequences that matched those of the various sequence classes of oca and grouped in the same places with the oca sequences on the ncpGS tree. These are also the only members of the *O. tuberosa* alliance that bear tubers. Tubers have also been observed in accessions of *O. boliviana* Britton (or perhaps *O. rigidicaulis* Knuth, which usually considered a synonym of *O. boliviana*, e.g., Lourteig, 2000, but which may be distinct from that taxon) from Oxapampa, in the department of Pasco, Peru (AAV5413, housed in living collections of the International Potato Center, Lima, Peru). However, ITS sequences of other accessions of *O. boliviana* grouped outside of the $x = 8$ group in our previous study (Emshwiller and Doyle, 1998), indicating that these are probably a separate origin of tuber formation. Most other reports of tuber-bearing *Oxalis* actually refer to nonhomologous structures, such as bulbs, corms, tuberous roots, or thickened aboveground stems. It is interesting that both of the candidates as progenitors are tuber-bearing, since only one such ancestral species could be sufficient to pass this trait on to the polyploid domesticate.

The identification of *O. picchensis* and the yet unnamed wild tuber-bearing taxon of Bolivia as putative progenitors of the octoploid cultigen has provided a working hypothesis for further studies of their role in the origins of the crop. Chloroplast DNA variation may identify the maternal genome donor, although initial limited tests did not reveal any such variation among alliance species (data not shown). Continued sampling of other independent sources of data and additional cultivated and wild accessions will test the working hypothesis and explore unresolved questions regarding the absence of the class C sequences in one oca accession, the mode of polyploidization, whether there were multiple origins of polyploidy or domestication, and other aspects of polyploid evolution in this important Andean food crop.

LITERATURE CITED

- ALLARD, R. W. 1960. Principles of plant breeding. John Wiley and Sons, New York, New York, USA.
- ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS, AND D. J. LIPMAN. 1990. Basic local alignment tool. *Journal of Molecular Biology* 215: 403–410.
- ANONYMOUS. 1996. PCR Cloning Tip No. 6. *Expressions (Newsletter of In-vitrogen, Inc. Carlsbad, CA)* 4.7: 14.
- ARBIZU, C., R. BLAS, M. HOLLE, F. VIVANCO, AND M. GHISLAIN. 1997. Advances in the morphological characterization of oca, ulluco, mashua, and arracacha collections. International Potato Center. Program Report 1995–1996. International Potato Center, Lima, Peru.
- BRADLEY, R. D., AND D. M. HILLIS. 1997. Recombinant DNA sequences generated by PCR amplification. *Molecular Biology and Evolution* 14: 592–593.
- BRAKO, L., AND J. L. ZARUCCHI. 1993. Catalogue of the flowering plants and gymnosperms of Peru. Missouri Botanical Garden, St. Louis, Missouri, USA.
- BUSTINZA LÓPEZ, T. 1979. Rendimiento en almidón, harina y materia seca en 50 clones de la colección ocas Cusco. Ingeniero Agrónomo thesis, Universidad Nacional de San Antonio Abad del Cusco, Cusco, Peru.
- CASTILLO, G. 1974. Estudio de la variabilidad morfológica de 142 clones de la colección de ocas del Cusco. Ingeniero Agrónomo thesis, Universidad Nacional de San Antonio Abad del Cusco, Cusco, Peru.
- CORTÉS BRAVO, H. 1984. Alcances de la investigación en tres tubérculos andinos: oca (*Oxalis tuberosa*), ulluco (*Ullucus tuberosus*), maswa, isaño o aña (*Tropaeolum tuberosum*). In M. E. Tapia [ed.], Avances en las Investigaciones sobre Tubérculos Alimenticios de los Andes, 62–83. Proyecto Investigación de los Sistemas Agrícolas Andinos (PISCA) IICA-CIID, Lima, Peru.
- CRONN, R., M. CEDRONI, T. HASELKORN, C. OSBORNE, AND J. F. WENDEL. 2002. PCR-mediated recombination in amplification products derived from polyploid cotton. *Theoretical and Applied Genetics* 104: 482–489.
- CSANK, C., F. M. TAYLOR, AND D. W. MARTINDALE. 1990. Nuclear pre-mRNA introns: analysis and comparison of intron sequences from *Tetrahymena thermophila* and other eukaryotes. *Nucleic Acids Research* 18: 5133–5141.
- DANDY, J. E., AND D. P. YOUNG. 1959. *Oxalis megalorrhiza*. *Proceedings of the Botanical Society of the British Isles* 3: 174–175.
- DE AZKUE, D. 2000. Chromosome diversity of South American *Oxalis* (Oxalidaceae). *Botanical Journal of the Linnean Society* 132: 143–152.
- DE AZKUE, D., AND A. MARTÍNEZ. 1990. Chromosome number of the *Oxalis tuberosa* alliance (Oxalidaceae). *Plant Systematics and Evolution* 169: 25–29.
- DEL RÍO, A. H. 1990. Analisis de la variación isoenzimática de *Oxalis tuberosa* Molina “oca” y su distribución geográfica. Licenciado en Biología thesis, Universidad Ricardo Palma, Lima, Peru.
- DIERS, L. 1961. Der Anteil an Polypliden in den Vegetationsgürteln der Westkordillere Perus. *Zeitschrift für Botanik* 49: 437–488.
- DOYLE, J. J., AND J. L. DOYLE. 1990. Isolation of plant DNA from fresh tissue. *Focus (Gibco BRL)* 12: 13–15.
- EMSHWILLER, E. 1999. Origins of domestication and polyploidy in the Andean tuber crop *Oxalis tuberosa* Molina (Oxalidaceae). Ph.D. dissertation, L. H. Bailey Hortorium, Cornell University, Ithaca, New York. (University Microforms, no. AAT 9914659.)
- EMSHWILLER, E. 2002a. Biogeography of the “*Oxalis tuberosa* alliance.” *Botanical Review* 68: 128–152.
- EMSHWILLER, E. 2002b. Ploidy levels among species in the “*Oxalis tuberosa* alliance” as inferred by flow cytometry. *Annals of Botany* 89: 741–753.
- EMSHWILLER, E., AND J. J. DOYLE. 1998. Origins of domestication and polyploidy in oca (*Oxalis tuberosa*: Oxalidaceae): nrDNA ITS data. *American Journal of Botany* 85: 975–985.
- EMSHWILLER, E., AND J. J. DOYLE. 1999. Chloroplast-expressed glutamine synthetase (ncpGS): potential utility for phylogenetic studies with an example from *Oxalis* (Oxalidaceae). *Molecular Phylogenetics and Evolution* 12: 310–319.
- ESCALANTE, A., S. IMANISHI, M. HOSSAIN, N. OHMIDO, AND K. FUKUI. 1998. RFLP analysis and genomic in situ hybridization (GISH) in somatic hybrids and their progeny between *Lycopersicon esculentum* and *Solanum lycopersicoides*. *Theoretical and Applied Genetics* 96: 719–726.
- FAVARGER, C., AND K. HUYNH. 1965. Chromosome number reports IV. *Taxon* 14: 86–92.
- FEDEROV, A. A. [ED.]. 1974. Chromosome numbers of flowering plants. Academy of Sciences, Leningrad, Russia.
- FENTON, B., G. MALLOCH, AND F. GERMA. 1998. A study of variation in rDNA ITS regions shows that two haplotypes coexist within a single aphid genome. *Genome* 41: 337–345.
- GOLOBOFF, P. 1998. Nona. Distributed by the author, Buenos Aires, Argentina.
- GUAMÁN, S. 1997. Conservación in situ caracterización y evaluación de la biodiversidad de oca (*Oxalis tuberosa*) y papalisa (*Ullucus tuberosus*) en Candelaria (Chapare) y Pocanche (Ayopaya). Ingeniero Agrónomo thesis, Universidad Mayor de San Simón, Cochabamba, Bolivia.
- HANCOCK, J. M. 1995. The contribution of slippage-like processes to genome evolution. *Journal of Molecular Evolution* 41: 1038–1047.
- HAYANO KANASHIRO, A. C. 1998. Citogenética de *Oxalis tuberosa* Mol. “Oca”: número cromosómico y análisis cariotípico. Licenciado en Biología thesis, Universidad Ricardo Palma, Lima, Peru.
- HEITZ, E. 1926. Der Nachweis der Chromosomen: vergleichende Studien über ihre Zahl, Grösse und Form im Pflanzenreich. I. *Zeitschrift für Botanik* 18: 625–681.
- HEITZ, E. 1927. Ueber multiple und aberrante Chromosomenzahlen. *Abhandlungen aus dem Gebiete der Naturwissenschaften* (Hamburg) 21: 47–57.
- HUYNH, K.-L. 1965. Contribution à l'étude caryologique et embryologique des Phanérogames du Pérou. *Denkschriften der Schweizerischen naturforschenden Gesellschaft* 85: 1–178.

- JANSEN, R., AND F. D. LEDLEY. 1990. Disruption of phase during PCR amplification and cloning of heterozygous target sequences. *Nucleic Acids Research* 18: 5153–5156.
- KING, S. R., AND S. N. GERSHOFF. 1987. Nutritional evaluation of three underexploited Andean tubers: *Oxalis tuberosa* Oxalidaceae, *Ullucus tuberosus* Basellaceae, and *Tropaeolum tuberosum* Tropaeolaceae. *Economic Botany* 41: 503–511.
- KNUTH, R. 1930. Oxalidaceae. In A. Engler [ed.], *Das Pflanzenreich; Regni vegetabilis conspectus*, vol. 13, 1–481. W. Engelmann, Leipzig, Germany.
- KO, C. H., V. BRENDEL, R. D. TAYLOR, AND V. WALBOT. 1998. U-richness is a defining feature of plant introns and may function as an intron recognition signal in maize. *Plant Molecular Biology* 36: 573–583.
- LEVINSON, G., AND G. A. GUTMAN. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution* 4: 203–221.
- LIU, B., J. M. VEGA, AND M. FELDMAN. 1998. Rapid genomic changes in newly synthesized amphiploids of *Triticum* and *Aegilops*: II. Changes in low-copy coding DNA sequences. *Genome* 41: 535–542.
- LIU, B., J. M. VEGA, G. SEGAL, S. ABBO, M. RODOVA, AND M. FELDMAN. 1998. Rapid genomic changes in newly synthesized amphiploids of *Triticum* and *Aegilops*. I. Changes in low-copy noncoding DNA sequences. *Genome* 41: 272–277.
- LOURTEIG, A. 1988. Oxalidaceae. *Flora Patagonica* 8: 1–29.
- LOURTEIG, A. 1994. *Oxalis* L. subgénero *Thamnoxyis* (Endl.) Reiche emend. Lourteig. *Bradea* 7: 1–199.
- LOURTEIG, A. 2000. *Oxalis* L. subgénero *Monoxalis* (Small) Lourt., *Oxalis* y *Trifidus* Lourt. *Bradea* 7: 201–629.
- MCCARTHY, C. 1997. Chromas computer program, Version 1.43. Brisbane, Queensland, Australia.
- MCDADE, L. A. 1990. Hybrids and phylogenetic systematics I. Patterns of character expression in hybrids and their implications for cladistic analysis. *Evolution* 44: 1685–1700.
- MCDADE, L. A. 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution* 46: 1329–1346.
- MEDINA HINOSTROZA, T. C. 1994. Contaje cromosómico de la oca (*Oxalis tuberosa* Molina) conservada in vitro. Ingeniero Agrónomo thesis, Universidad Nacional del Centro del Perú, Huancayo, Peru.
- MENDIBURU, A. O., AND S. J. PELOQUIN. 1976. Sexual polyploidization and diployploidization: some terminology and definitions. *Theoretical and Applied Genetics* 48: 137–143.
- NARANJO, C. A., L. M. MOLA, L. POGGIO, AND M. MULGURA DE ROMERO. 1982. Estudios citotaxonomicos y evolutivos en especies herbaceas sudamericanas de *Oxalis* (Oxalidaceae). I. *Boletín de la Sociedad Argentina de Botánica* 20: 183–200.
- NIXON, K. C. 1996. Clados computer program, Version 17. Distributed by the author, Ithaca, New York, USA.
- NIXON, K. C. 1998. Dada computer program, Version 12. Distributed by the author, Ithaca, New York, USA.
- NIXON, K. C. 2000. WinClada computer program, Version 0.9.99m24[beta]. Distributed by the author, Ithaca, New York, USA.
- OZKAN, H., A. A. LEVY, AND M. FELDMAN. 2001. Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops-Triticum*) group. *Plant Cell* 13: 1735–1747.
- PAROKONNY, A. S., A. KENTON, Y. Y. GLEBA, AND M. D. BENNETT. 1994. The fate of recombinant chromosomes and genome interaction in *Nicotiana* asymmetric somatic hybrids and their sexual progeny. *Theoretical and Applied Genetics* 89: 488–497.
- PESOLE, G., M. P. BOZZETTI, C. LANAVE, G. PREPARATA, AND C. SACCONI. 1991. Glutamine synthetase gene evolution: a good molecular clock. *Proceedings of the National Academy of Sciences, USA* 88: 522–526.
- PIKAARD, C. S. 2001. Genomic change and gene silencing in polyploids. *Trends in Genetics* 17: 675–677.
- POMA MACHACA, B. 1976. Evaluación de resistencia al “gusano de la oca” (Chrysomelidae) en 260 clones de la colección de ocas Cusco. Ingeniero Agrónomo thesis, Universidad Nacional de San Antonio Abad del Cusco, Cusco, Peru.
- SALTER, T. M. 1944. The genus *Oxalis* in South Africa. A taxonomic revision. *Journal of South African Botany* (suppl. 1): 1–355.
- SHAH, A. A., H. STEGEMANN, AND M. GALVEZ. 1993. The Andean tuber crops mashua, oca and ulluco: optimizing the discrimination between varieties by electrophoresis and some characters of the tuber proteins. *Plant Varieties and Seeds* 6: 97–108.
- SHAKED, H., K. KASHKUSH, H. OZKAN, M. FELDMAN, AND A. A. LEVY. 2001. Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* 13: 1749–1759.
- SOLTIS, D. E., AND P. S. SOLTIS. 1993. Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences* 12: 243–273.
- SOLTIS, D. E., AND P. S. SOLTIS. 1999. Polyploidy: recurrent formation and genome evolution. *Trends in Ecology and Evolution* 14: 348–352.
- SONG, K., P. LU, K. TANG, AND T. C. OSBORN. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proceedings of the National Academy of Sciences, USA* 92: 7719–7723.
- STEBBINS, G. L. 1947. Types of polyploids: their classification and significance. *Advances in Genetics* 1: 403–429.
- STEBBINS, G. L. 1950. Variation and evolution in plants. Columbia University Press, New York, New York, USA.
- STEGEMANN, H., S. MAJINO, AND P. SCHMEDICHE. 1988. Biochemical differentiation of clones of oca (*Oxalis tuberosa*, Oxalidaceae) by their tuber proteins and the properties of these proteins. *Economic Botany* 42: 37–44.
- TALLEDO, D., AND C. ESCOBAR. 1995. Citogenética de *Oxalis tuberosa*: ciclo celular y número cromosómico. *Biotempo (Universidad Ricardo Palma, Lima, Peru)* 2: 33–46.
- TEMPLETON, A. R., K. A. CRANDALL, AND C. F. SING. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132: 619–633.
- TIBAYRENC, M., F. KJELLBERG, AND F. J. AYALA. 1990. A clonal theory of parasitic protozoa: the population structures of *Entamoeba*, *Giardia*, *Leishmania*, *Naegleria*, *Plasmodium*, *Trichomonas*, and *Trypanosoma* and their medical and taxonomical consequences. *Proceedings of the National Academy of Sciences, USA* 87: 2414–2418.
- TOSTO, D. S., AND H. E. HOPP. 1996. Sequence analysis of the 5.8S ribosomal DNA and internal transcribed spacers (ITS1a and ITS2) from five species of the *Oxalis tuberosa* alliance. *DNA Sequence—The Journal of Sequencing and Mapping* 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 Systematics and Evolution* 223: 201–209.
- VALLADOLID, A. 1996. Niveles de ploidía de la oca (*Oxalis tuberosa* Mol.) y sus parientes silvestres. Master's thesis, Universidad Nacional Agraria La Molina, Lima, Peru.
- VALLADOLID, A., C. ARBIZU, AND D. TALLEDO. 1994. Niveles de ploidía de la oca (*Oxalis tuberosa* Mol.) y sus parientes silvestres. *Agro Sur* 22 (número especial): 11–12.
- VINUEZA VELA, J. F. 1997. Evaluación y caracterización citogenética de 20 entradas de oca (*Oxalis tuberosa* Molina) recolectadas en Ecuador. Ingeniero Agrónomo thesis, Universidad Central del Ecuador, Cutuglahua-Pichincha, Ecuador.
- VOGEL, J. C., J. A. BARRETT, F. J. RUMSEY, AND M. GIBBEY. 1999. Identifying multiple origins in polyploid homosporous peridophytes. In P. M. Hollingsworth, R. M. Bateman, and R. J. Gornall [eds.], *Molecular systematics and plant evolution*, 101–107. Taylor and Francis, London, UK.
- WAGNER, A., N. BLACKSTONE, P. CARTWRIGHT, M. DICK, B. MISOF, P. SNOW, G. P. WAGNER, B. J. BARTELS, M. MURTHA, AND J. PENDELTON. 1993. Surveys of gene families using polymerase chain reaction: PCR selection and PCR drift. *Systematic Biology* 43: 250–261.
- WATANABE, K., AND S. J. PELOQUIN. 1991. Genetic significance of mode of polyploidization: somatic doubling or 2n gametes? *Genome* 34: 28–34.
- WENDEL, J. F. 2000. Genome evolution in polyploids. *Plant Molecular Biology* 42: 225–249.
- WENDEL, J. F., A. SCHNABEL, AND T. SEELMAN. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proceedings of the National Academy of Sciences, USA* 92: 280–284.