Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 3880–3884, June 1986
Evolution

Evidence for selection as a mechanism in the concerted evolution of *Lycopersicon esculentum* (tomato) genes encoding the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase

(plant nuclear genes/multigene family/Solanaceae)

E. PICHERSKY*, R. BERNATZKY†‡, S. D. TANKSLEY†‡, AND A. R. CASHMORE*

*Laboratory of Cell Biology, The Rockefeller University, New York, NY 10021; and †Department of Horticulture and Plant Genetic Engineering Laboratory, New Mexico State University, Las Cruces, NM 88003

Communicated by Philip Siekevitz, January 27, 1986

ABSTRACT **The nuclear gene sequences encoding RBCS, the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (EC 4.1.1.39) from several plants show extensive interspecific divergence but little intraspecific divergence, suggesting that these genes are evolving in concert within a genome. In this study, the nucleotide sequences of two tomato (*Lycopersicon esculentum*) RBCS genes and a cDNA clone containing the entire coding region of a third tomato RBCS gene were determined. The three genes, designated *Rbcs-1*, *Rbcs-2A*, and *Rbcs-3A*, each belong to a different one of the three RBCS loci in the tomato genome. The nucleotide sequence of *Rbcs-1* differs from that of *Rbcs-2A* and *Rbcs-3A* by 13.9% and 13.1%, respectively. *Rbcs-2A* and *Rbcs-3A* differ from each other by 10.7%. A recently published RBCS gene sequence from tobacco (*Nicotiana tabacum*) [Mazur, B. J. & Chui, C.-F. (1985) *Nucleic Acids Res.* 13, 2373–2386] differs by 10.6% and 11.3% from *Rbcs-2A* and *Rbcs-3A*, respectively, and by 15.0% from *Rbcs-1*. Thus the tobacco gene seems to be phylogenetically as closely related to the tomato genes *Rbcs-2A* and *Rbcs-3A* as the latter two are to each other, and more closely related to them than *Rbcs-1* is. However, the mature part of the polypeptide encoded by the tobacco RBCS gene differs by five and six amino acids from the corresponding region in the polypeptides encoded by *Rbcs-2A* and *Rbcs-3A*, respectively, while these two tomato RBCS polypeptides differ from each other in the mature part by a single amino acid. *Rbcs-1*, whose nucleotide sequence shows higher divergence from both the tobacco RBCS gene and *Rbcs-2A* and *Rbcs-3A*, encodes a polypeptide whose mature part differs by eight amino acids from the corresponding region in the tobacco polypeptide but only by three and four amino acids from the corresponding regions of *Rbcs-2A*- and *Rbcs-3A*-encoded polypeptides, respectively. Thus, it appears that in the tomato selection has maintained near uniformity of the coding information in the portion of the RBCS genes encoding the mature polypeptides.**

Several reports have shown that the small subunit (RBCS) of the chloroplast enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO; EC 4.1.1.39) is encoded by a small nuclear gene family in the genomes of higher plants (1–4). The RBCS gene products are made as precursors containing an N-terminal "transit" peptide, which is removed after transport into the chloroplast (5–7). The time of origin of the duplications that gave rise to members of the RBCS gene family is not known.

Interspecific comparisons of RBCS sequences show extensive divergence of the genes and proteins (8, 9). In contrast, intraspecific comparisons show a remarkable conservation of the amino acid sequence of the mature part of the polypeptide (2, 3, 10–13). However, in the few comparisons that have been made, the RBCS genes from the same organism also show higher nucleotide sequence homology to each other than to any RBCS gene sequence from the most closely related species available. This result leaves open the possibility that the RBCS genes from the same organism are more closely related to each other because they originated by recent gene duplications or because they have gone through recent gene correction events.

We present here the nucleotide sequences of three tomato RBCS genes, one from each of the three RBCS loci in this plant (*Lycopersicon esculentum*) (14). The regions in the three tomato genes that encode the mature RBCS polypeptides show a high ratio of "silent" to "nonsilent" nucleotide substitutions when compared with each other, with the result that the predicted polypeptide sequences are almost identical. This observation suggests that selection is the mechanism responsible for the concerted evolution of the RBCS in tomato. An RBCS gene sequence recently reported for tobacco, *Nicotiana tabacum* (9), a species in the same family, Solanaceae, shows higher homology to one tomato RBCS gene than the other two tomato RBCS genes do. This observation indicates that the three tomato RBCS loci diverged from each other before the split of the tomato and tobacco lineages. In spite of the relatedness of one of the tomato RBCS DNA sequences with the tobacco sequence, the corresponding polypeptide comparison shows more divergence than any intraspecific divergence observed for the tomato RBCS polypeptides. We interpret this observation as confirmation that tomato RBCS polypeptide near-uniformity results from selection.

MATERIALS AND METHODS

Cloning and Characterization of Tomato RBCS Genes. The construction of a tomato genomic library in λ Charon 4 vector has been previously described (15). The tomato RBCS genes were isolated from this library by using hybridization conditions identical to those described for the isolation of the tomato chlorophyll a/b-binding proteins (CAB) genes (15). The probe used was an RBCS cDNA clone from *Pisum sativum* (pea), pSS15 (1). Characterization of the RBCS genomic clone was by the same procedures and techniques (Southern blots, construction of restriction maps, subcloning in pUC9 and pUC18 plasmid vectors) described in ref. 15.

Tomato RBCS cDNA Clones. Construction of a tomato cDNA library and isolation of RBCS cDNA clones is described elsewhere (21).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: kb, kilobase(s); RuBisCO, ribulose-1,5-bisphosphate carboxylase/oxygenase; RBCS, RuBisCO small subunit.

‡Present address: Department of Plant Breeding and Biometry, 252 Emerson Hall, Cornell University, Ithaca, NY 14853.

DNA Sequencing. The Maxam and Gilbert (16) DNA sequencing method was employed.

RESULTS

Isolation, Molecular Characterization, and Locus Assignment of the Tomato RBCS Genes. Screening the tomato genomic library by hybridization to a pea RBCS cDNA probe yielded two recombinant phages, designated 20B and 24A, which contain sequences that hybridize to RBCS probes. Phage 20B contains a 9.6-kilobase (kb) tomato DNA fragment with one RBCS gene (Figs. 1*C* and 2). Phage 24A contains a 16.7-kb tomato DNA insert that also contains one RBCS gene (Figs. 1*A* and 2). The nucleotide sequences of these genes, including the coding region (contained in three exons), two introns, and 5′ and 3′ flanking regions are shown in Fig. 3.

The RBCS gene in phage 24A is located on two *Eco*RI fragments, 1.7 and 15.0 kb (Fig. 2). These fragments have been previously mapped to the *Rbcs-1* locus on chromosome 2 (14). Hybridization experiments involving subfragments of phage 24A DNA with various RBCS probes together with sequence data revealed only one RBCS gene on phage 24A; together with the data of Vallejos *et al.* (14) this indicates that there is only one RBCS gene in the *Rbcs-1* locus.

The RBCS gene in phage 20B is located on a 3.8-kb *Hin*dIII fragment (Figs. 1*C* and 2) that is part of the *Rbcs-3* locus, also on chromosome 2 but genetically unlinked to *Rbcs-1* (14). This fragment contains the entire gene, which we designate *Rbcs-3A*, and no part of another RBCS gene. Since the other two *Hin*dIII fragments (a 4.0-kb fragment and a 3.0-kb fragment) that also map to this locus (14) each hybridize to the various RBCS probes with an intensity approximately equal to that of the 3.8-kb fragment (Fig. 1), it is reasonable to conclude that the *Rbcs-3* locus contains three RBCS genes.

RBCS cDNA Clones and Their Locus Assignments. Two tomato RBCS cDNA clones have been sequenced (Fig. 3).

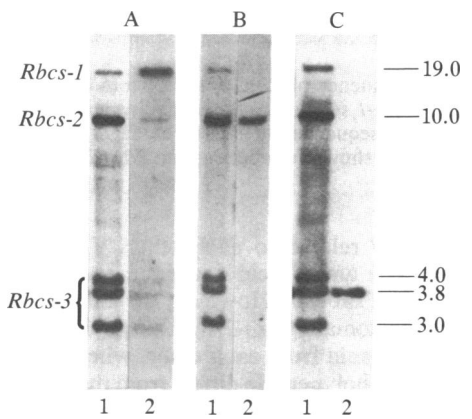

FIG. 1. Southern blot analyses. All hybridizations were carried out at 65–68°C in 6× SSC (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 6.8) according to ref. 15. (*A*) Lane 1, tomato DNA digested with *Hin*dIII. The probe used was p3-91, and washing was done at a final concentration of 1× SSC at 68°C. Lane 2, tomato DNA digested with *Hin*dIII. The probe used was the 1.7-kb *Eco*RI fragment from phage 24A. Washing conditions were the same as for lane 1. Because this fragment, in addition to approximately 1400 nucleotides of 3′ noncoding region of the *Rbcs-1* gene, also contains part of the coding region (codons 101 to 181), some hybridization is observed with the other tomato RBCS genes. (*B*) Lane 1, the same as lane 1 in *A*. Lane 2, the same as lane 1 in *A* but washing was done at a final concentration of 0.1× SSC at 68°C. (*C*) Lane 1, the same as lane 1 in *A*. Lane 2, phage 20B DNA digested with *Hin*dIII. The probe and conditions are the same as in lane 1. The tomato RBCS locus to which each of the *Hin*dIII fragments maps (14) is indicated on the left. The size, in kb, of each of the fragments is indicated on the right.

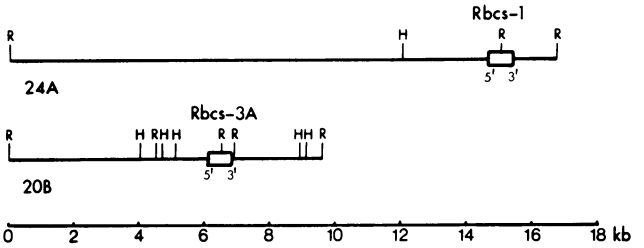

FIG. 2. Restriction map of the tomato DNA inserts in phages 24A and 20B and the location and orientation of the RBCS genes on the inserts. Sites are H, *Hin*dIII; R, *Eco*RI.

The sequence of one incomplete cDNA clone, p3-167, is identical to the corresponding sequence of *Rbcs-3A*. The other RBCS cDNA clone, p3-91, contains the entire coding region and short 5′ and 3′ noncoding sequences and is different from both *Rbcs-1* and *Rbcs-3A* (Fig. 3). The mRNA from which cDNA p3-91 was made was transcribed from a gene in the *Rbcs-2* locus. This was determined by hybridizing ^{32}P-labeled p3-91 to a Southern blot of tomato DNA digested with *Hin*dIII and then washing the blot under high-stringency conditions (68°C in 0.1× SSC) (Fig. 1*B*). Under these conditions, the labeled p3-91 probe remained hybridized almost exclusively to the 10.0-kb *Hin*dIII fragment that maps to the *Rbcs-2* locus on chromosome 3 (14). The *Rbcs-2* locus contains two *Eco*RI fragments of 2.0 and 6.0 kb, the former hybridizing greater than 3 times more intensely than the latter (14). This is consistent with our results, since p3-91 does not contain an *Eco*RI site, whereas *Rbcs-1* and *Rbcs-3A*, and apparently all the other tomato RBCS genes, do (this report, ref. 14). Thus the 2.0-kb *Eco*RI fragment from the *Rbcs-2* locus appears to contain an entire RBCS gene, which we designate *Rbcs-2A* (and from which p3-91 was transcribed), and part of another RBCS gene. This analysis suggests that there are two genes in the *Rbcs-2* locus, and thus the total number of RBCS genes in tomato appears to be six.

DISCUSSION

Sequences of Isolated Tomato RBCS Genes and the Encoded Polypeptides. A previous study identified three loci containing RBCS sequences in tomato (14). The RBCS sequences determined in the current study (Fig. 3) represent all three of these loci. We have isolated a cDNA clone, p3-167, whose sequence corresponds to *Rbcs-3A* (Fig. 3); this indicates that this gene is expressed. The RBCS cDNA clone p3-91 contains a complete coding sequence and was transcribed from *Rbcs-2A*. We do not yet know if *Rbcs-1* is expressed, but the fact that its open reading frame has been maintained strongly suggests that at least until recently it has been.

The transit peptides encoded by the three tomato genes isolated differ by 8 to 12 amino acids and a deletion/insertion of one amino acid (Fig. 4). The predicted amino acid sequences of the mature polypeptides encoded by *Rbcs-3A* and *Rbcs-2A* differ only at position 114. *Rbcs-1*-encoded mature polypeptide differs from both polypeptides encoded by the other two genes at positions 106, 115, and 181. At position 114, *Rbcs-1* encodes the same amino acids as does *Rbcs-2A*. It is likely that *Rbcs-1* is the most divergent tomato RBCS gene relative to *Rbcs-2A* and *Rbcs-3A*; it hybridizes least intensely with gene probes from the latter two tomato RBCS genes (e.g., see Fig. 1). In addition, a probe derived from the 3′ noncoding region of *Rbcs-2A* detects some shared sequence homology with genes from the *Rbcs-3* locus but no homology with *Rbcs-1* (data not shown).

Evidence for Selection as a Mechanism for the Concerted Evolution of the Regions in the Tomato RBCS Genes that Encode the Mature Polypeptide. Multiple RBCS sequences

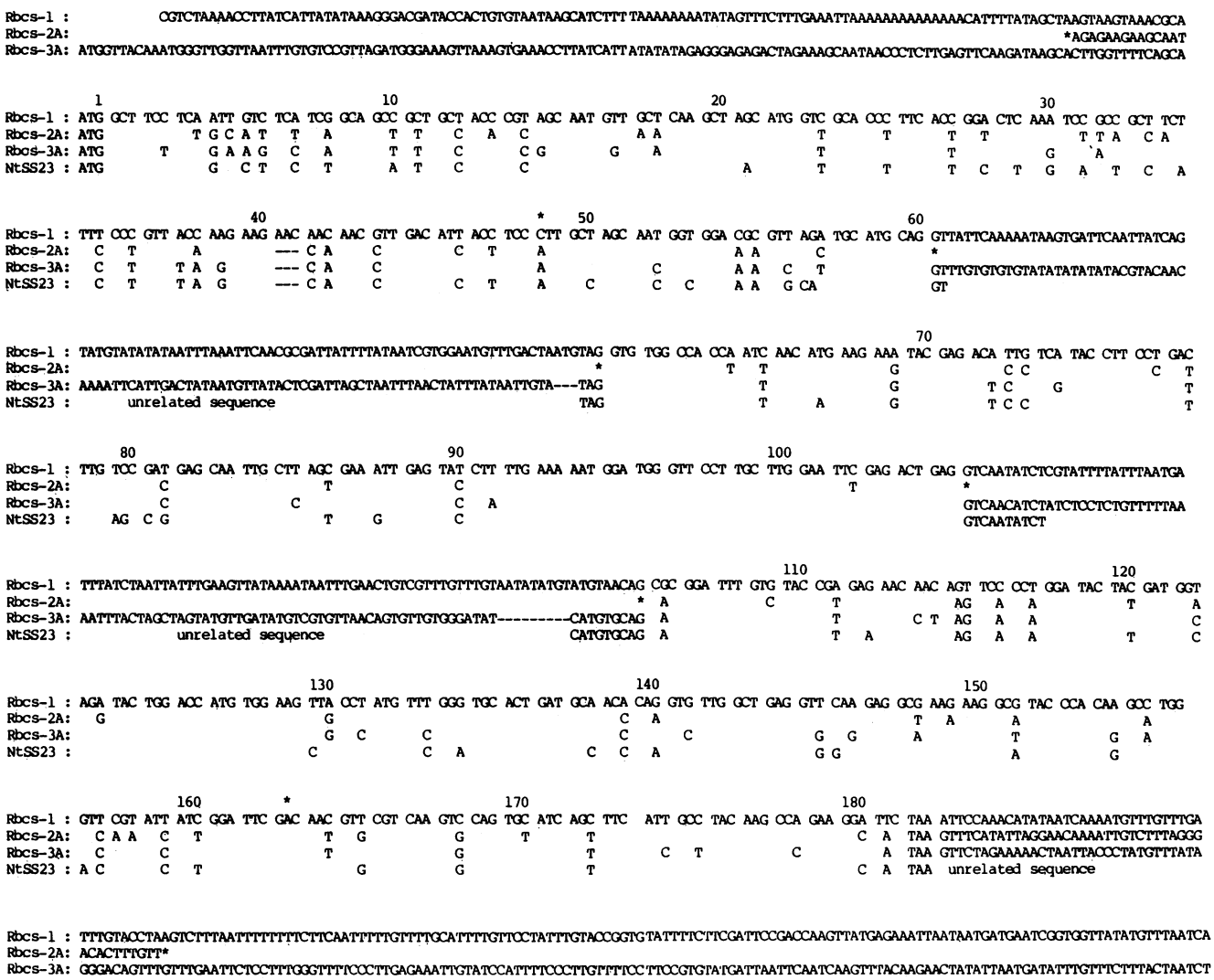

FIG. 3. Nucleotide sequences of *Rbcs-1*, *Rbcs-3A*, and p3-91 (*Rbcs-2A*). The coding sequence of *Rbcs-1* is shown in full, and the coding sequences of *Rbcs-2A* and *Rbcs-3A* are shown below it only when they differ from the *Rbcs-1* sequence. All noncoding sequences are shown in full. The sequence of p3-167 is identical to that of *Rbcs-3A*; asterisks above the *Rbcs-1* sequence indicate its extent. A dash represents a deletion introduced to maximize homology. The sequence of the tobacco gene *NtSS23* (9) is shown for comparison. *NtSS23* has a third intron that splits codon 123; this intron is not shown here.

from the same organism have been previously reported for three species: pea (2, 10, 17, 18), petunia (*Petunia hybrida*) (3, 11), and wheat (*Triticum aestivum*) (12, 13). In each of these species, the RBCS nucleotide sequences diverge by at most 10%, and the predicted amino acid sequences of the mature polypeptides differ at only a few positions. In each case, the most closely related organism from which RBCS sequences are available for comparisons [soybean (*Glycine max*) (8), tobacco (9), and duckweed (*Lemna gibba*) (4), respectively] show much larger divergence in both nucleotide sequence (a minimum of 23%) and amino acid sequence (a minimum of 16%). Thus, it was not possible to distinguish between the possibility that the higher homology of RBCS genes within a species reflects recent gene duplication or gene correction events and the alternative explanation that selection at the protein level has resulted in conservation of nucleotide sequences within the genome.

A tobacco RBCS gene, designated *NtSS23*, whose sequence has recently been reported (9) is as homologous to the tomato genes *Rbcs-2A* and *Rbcs-3A* as these two tomato genes are to each other. In fact, of the three two-way comparisons, the highest homology (10.6%) is exhibited by the *NtSS23–Rbcs-2A* pair (Table 1). This means that, using parsimony on the raw data, the three genes are about equally phylogenetically related to each other. This conclusion is corroborated by amino acid sequence comparisons of the transit peptides (Table 2). However, the mature polypeptides encoded by the tomato *Rbcs-2A* and *Rbcs-3A* genes differ by only one amino acid from each other, whereas the tobacco-encoded mature polypeptide differs from them by five and six amino acids, respectively (Fig. 4). *Rbcs-1* is 15.0% divergent from the tobacco gene and 13.9% and 13.1% divergent from *Rbcs-2A* and *Rbcs-3A*, respectively. It encodes a mature polypeptide that differs by 8 residues from the tobacco polypeptide, but only by three and four residues from the mature polypeptides encoded by *Rbcs-2A* and *Rbcs-3A*, respectively. Moreover, of the three residues that differ in the *Rbcs-1*-encoded mature polypeptide relative to both *Rbcs-2A* and *Rbcs-3A* polypeptides, two (positions 106 and 115) are clustered around the already variable position in the *Rbcs-2A–Rbcs-3A* comparison (position 114) and also seem to compensate for each other by maintaining the positive charge. The third amino acid substitution is in the C-terminal amino acid (Fig. 4). Thus, for two contiguous stretches totaling over 90% of the length of the mature polypeptides, the three tomato sequences are identical. In contrast, the amino acid substitutions in the tobacco sequence relative to

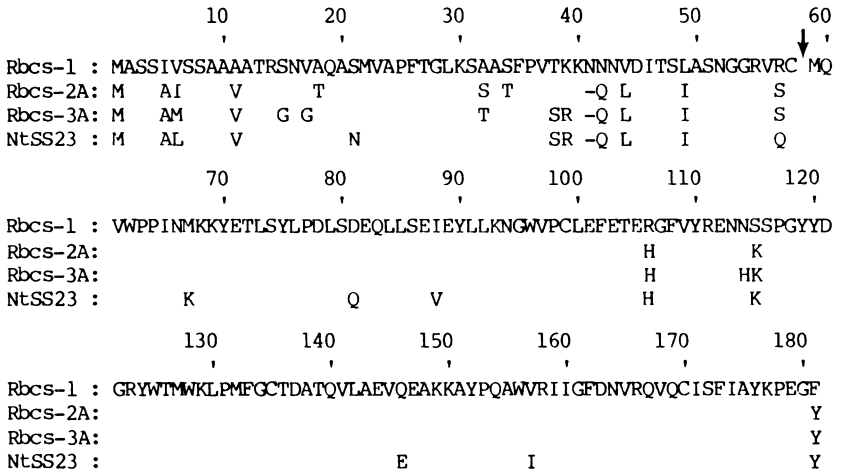

FIG. 4. Protein sequence comparisons. The predicted amino acid sequences of the RBCS proteins encoded by *Rbcs-1*, *Rbcs-2A*, and *Rbcs-3A*. The standard one-letter symbols for amino acids are used. The sequence of the *Rbcs-1* protein is shown in full; the other sequences are shown only when they differ from the *Rbcs-1* protein sequence. The sequence of the tobacco protein encoded by *NtSS23* (9) is shown for comparison. A dash represents a deletion introduced to maximize homology. The arrow indicates site of cleavage of precursor in the chloroplast.

the tomato sequences are more evenly distributed along the mature polypeptide chain (Fig. 4).

The higher amino acid homology of the mature polypeptides encoded by the three tomato genes, despite the nucleotide sequences being just as divergent or more so than the tobacco gene sequence is, results from a different distribution of the nucleotide substitutions. When compared with each other, the ratio of the silent to nonsilent substitutions between the tomato genes ranges from 7:1 (this number is somewhat misleadingly low because two substitutions are involved in one amino acid change) to 33:1, whereas in the tobacco–tomato comparisons this ratio drops to a uniform 5:1 (Table 3). In the sequences encoding the transit peptides, this ratio is always between 2:1 and 3:1 regardless of the type of comparison (inter- or intraspecific). It therefore appears that selection is the mechanism responsible for the maintenance of the (near) uniformity of amino acid sequence of the mature polypeptides encoded by members of the RBCS gene family in tomato. A selection of the same magnitude is apparently not in effect on the gene region encoding the transit peptide.

Gene Phylogeny Based on Silent Nucleotide Substitutions. If selection is the mechanism responsible (at least in part) for the concerted evolution of the tomato RBCS genes, gene phylogenies derived from comparisons of total nucleotide divergence will be misleading by underestimating the number of nucleotide substitutions that have occurred in intraspecific comparisons. We therefore analyzed the average number of nucleotide substitutions per site in pairwise comparisons of substitutions that do not lead to a changed amino acid (silent substitutions) in the mature polypeptide (Table 4). The data in Table 4 clearly show that, of the four genes involved in our comparisons, *Rbcs-2A* and *NtSS23* are the most closely related pair, and that *Rbcs-1* is least related to the other three RBCS genes. Thus, it seems that the tomato genes *Rbcs-1*, *Rbcs-2A*, and *Rbcs-3A* belong to ancient gene lineages that had diverged from each other considerably earlier than the date of the *L. esculentum–N. tabacum* split. This conclusion is consistent with the lack of substantial sequence homology in the noncoding regions of these genes. We do not yet have any data regarding the time of origin of the gene duplications within the complex loci *Rbcs-2* and *Rbcs-3*.

Taxonomic Relationships Between Tobacco, Petunia, and Tomato. The level of divergence at the mature polypeptide between tobacco and tomato RBCS sequences, 4.1%–6.5%, is the smallest so far for an interspecific comparison. The corresponding sequences in the petunia RBCS proteins show about 20% divergence from both the tobacco and tomato sequences (9, 11). Taxonomically, *N. tabacum* and petunia

Table 1. Pairwise comparisons of nucleotide sequence differences among the three tomato RBCS genes and the *N. tabacum* RBCS gene *NtSS23* (9)

	Rbcs-1	*Rbcs-2A*	*Rbcs-3A*
Rbcs-2A	35/171 (20.5%) 40/369 (10.8%) 75/540 (13.9%)		
Rbcs-3A	32/171 (18.7%) 39/369 (10.6%) 71/540 (13.1%)	24/171 (14.0%) 34/369 (9.2%) 58/540 (10.7%)	
NtSS23	39/171 (22.8%) 42/369 (11.4%) 81/540 (15.0%)	27/171 (15.8%) 30/369 (8.1%) 57/540 (10.6%)	26/171 (15.2%) 35/369 (9.5%) 61/540 (11.3%)

The first pair of numbers for each comparison are the fraction and percent of nucleotide sequence differences over the gene region encoding the transit peptide (171 nucleotides). The second pair of numbers are the fraction and percent of nucleotide sequence differences over the gene region encoding the mature polypeptide (369 nucleotides). The third pair of numbers are the fraction and percent of nucleotide sequence differences over the entire coding region (540 nucleotides). The insertion of one codon in *Rbcs-1* is ignored in these comparisons.

Table 2. Amino acid differences in pairwise comparisons between polypeptides encoded by the tomato RBCS genes and the *N. tabacum* RBCS gene *NtSS23* (9)

	Rbcs-1	*Rbcs-2A*	*Rbcs-3A*
Rbcs-2A	10/57 (17.5%) 3/123 (2.4%)		
Rbcs-3A	12/57 (21.1%) 4/123 (3.3%)	8/57 (14.0%) 1/123 (0.8%)	
NtSS23	10/57 (17.5%) 8/123 (6.5%)	8/57 (14.0%) 5/123 (4.1%)	6/57 (10.5%) 6/123 (4.9%)

The first pair of numbers in each comparison is the fraction and percent of amino acid differences in the transit peptide (57 amino acids long). The second pair of numbers is the fraction and percent of amino acid differences in the mature polypeptide (123 amino acids long). The insertion of one amino acid in the transit peptide encoded by *Rbcs-1* is ignored in these comparisons.

Table 3. Silent vs. nonsilent nucleotide substitutions in pairwise comparisons among the three tomato RBCS sequences and the tobacco RBCS sequence

	Rbcs-1	*Rbcs-2A*	*Rbcs-3A*
Rbcs-2A	25:10 36:4		
Rbcs-3A	20:12 34:5	16:8 33:1	
NtSS23	28:10 35:9	19:8 25:5	20:6 29:6

The first number in each comparison is the silent to nonsilent ratio in the region encoding the transit peptide. The second number is the ratio in the gene region encoding the mature polypeptide. The insertion of one codon in the region encoding the transit peptide in *Rbcs*-1 is ignored in these comparisons.

are placed in the same tribe, Nicotianeae, in the family Solanaceae, and *L. esculentum* is in another tribe, Solaneae (19). However, both RBCS sequences (refs. 9 and 11; this report) and CAB gene sequences and organization in tomato (15), tobacco (C. Castresana, R. Stanalone, and A.R.C., unpublished data) and petunia (20) demonstrate higher homology between corresponding tobacco and tomato genes and proteins than between either tomato and petunia or tobacco and petunia. This suggests a closer affinity between tomato and tobacco than between tobacco and petunia. Consistent with this suggestion is the observation that *L. esculentum* and most tobacco species share the same basic chromosome number (12 pairs), whereas petunia is different (7 pairs) (19).

Significance of RBCS Sequence Uniformity Within Species. The RuBisCO holoenzyme is a multimeric complex composed of eight large subunits (encoded by chloroplast DNA) and eight small subunits. At present, the function of the RBCS subunits is not known. As noted before, although RBCS proteins show extensive divergence among species (more than the RuBisCO large subunits do), within a species the RBCS mature polypeptides are identical or almost identical. It could be that the nature of the molecular interactions involving the RBCS proteins in the plant cell requires symmetry or interchangeability of RBCS proteins somewhat independent of a specific amino acid sequence. Alternatively, a specific amino acid sequence might be selected in different genomes and in different environmental conditions.

To test experimentally for adverse effects on RuBisCO caused by RBCS sequence heterogeneity, one would like to produce "hybrid" RuBisCO molecules in which two or more significantly different types of RBCS subunits are present.

Table 4. Average number of silent nucleotide substitutions per site in pairwise comparisons of the tomato and tobacco RBCS gene regions encoding the mature polypeptides

	Rbcs-1	*Rbcs-2A*	*Rbcs-3A*
Rbcs-2A	0.30		
Rbcs-3A	0.28	0.27	
NtSS23	0.29	0.21	0.25

The dissociation and reassociation of RuBisCO from higher plants has proved difficult. Such hybrids, however, might possibly be produced in large quantities and as a large fraction of the total RuBisCO *in vivo* through the introduction of heterologous RBCS genes into the plant genome by DNA transformation procedures. The results of such experiments are of interest, especially in light of the much-discussed possibility of "improving" RuBisCO and thus increasing the efficiency of photosynthesis by genetic engineering techniques.

This work was supported by a research grant to A.R.C. from the National Institutes of Health and a grant to S.D.T. from the Plant Genetic Engineering Laboratory, New Mexico State University. E.P. is a recipient of a National Institutes of Health postdoctoral training grant.

1. Coruzzi, G., Broglie, R., Cashmore, A. R. & Chua, N.-H. (1983) *J. Biol. Chem.* **258**, 1399–1402.
2. Cashmore, A. R. (1983) in *Genetic Engineering of Plants*, eds. Kosuge, T., Meredith, C. P. & Hollaender, A. (Plenum, New York), pp. 29–38.
3. Dunsmuir, P., Smith, S. M. & Bedbrook, J. (1983) *Nucleic Acids Res.* **11**, 4177–4183.
4. Stiekema, W. J., Wimpee, C. F. & Tobin, E. M. (1983) *Nucleic Acids Res.* **11**, 8051–8061.
5. Chua, N.-H. & Schmidt, G. W. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 6110–6114.
6. Cashmore, A. R., Broadhurst, M. K. & Gray, R. E. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 655–659.
7. Highfield, P. E. & Ellis, R. J. (1978) *Nature (London)* **271**, 420–424.
8. Berry-Lowe, S. L., McKnight, T. D., Shah, D. M. & Meagher, R. B. (1982) *J. Mol. Appl. Gen.* **1**, 483–498.
9. Mazur, B. J. & Chui, C.-F. (1985) *Nucleic Acids Res.* **13**, 2373–2386.
10. Timko, M. P., Kausch, A. P., Hand, J. M., Herrera-Estrella, L., Van der Broeck, G., Van Montagu, M. & Cashmore, A. R. (1985) in *Molecular Biology of the Photosynthetic Apparatus*, eds. Arntzen, C., Bogorad, L., Bonitz, S. & Steinback, K. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 381–396.
11. Dean, C., van der Elzen, P., Dunsmuir, P. & Bedbrook, J. (1985) in *Molecular Form and Function of the Plant Genome*, eds. Vloten-Doting, L., Groot, G. S. P. & Hall, T. C. (Plenum, New York), pp. 167–174.
12. Broglie, R., Coruzzi, G., Lamppa, G., Keith, B. & Chua, N.-H. (1983) *Bio/Technology* **1**, 55–61.
13. Smith, S. M., Bedbrook, J. & Spiers, J. (1983) *Nucleic Acids Res.* **11**, 8719–8734.
14. Vallejos, C. E., Tanksley, S. D. & Bernatzky, R. (1986) *Genetics* **112**, 93–105.
15. Pichersky, E., Bernatzky, R., Tanksley, S. D., Breidenbach, W., Kausch, A. P. & Cashmore, A. R. (1985) *Gene* **40**, 247–258.
16. Maxam, A. D. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
17. Bedbrook, J. R., Smith, S. M. & Ellis, R. J. (1980) *Nature (London)* **287**, 692–697.
18. Coruzzi, G., Broglie, R., Edwards, C. & Chua, N.-H. (1984) *EMBO J.* **3**, 1671–1679.
19. D'Arcy, W. G. (1969) in *The Biology and Taxonomy of the Solanaceae*, eds. Hawkes, J. G., Lester, R. N. & Skelding, A. D. (Academic, London), Linnean Society Symposium Series, No. 7, pp. 3–48.
20. Dunsmuir, P. (1985) *Nucleic Acids Res.* **13**, 2503–2518.
21. Bernatzky, R. & Tanksley, S. D. (1986) *Mol. Gen. Genet.*, in press.