Plant Physiol. (1977) 60, 89-94

Evolutionary Conservation of Chloroplast Genes Coding for the Large Subunits of Fraction 1 Protein

Received for publication December 31, 1976 and in revised form February 25, 1977

SHAIN-DOW KUNG, AND CHING-ING LEE
Department of Biological Sciences, University of Maryland Baltimore County, Catonsville, Maryland 21228
DENISE D. WOOD AND MARIO A. MOSCARELLO
Department of Biochemistry, Research Institute, The Hospital for Sick Children, Toronto, Canada

ABSTRACT

Crystalline fraction 1 protein, obtained from four species of *Nicotiana*, have identical polypeptide compositions and isoelectric points. However, the tryptic peptide map of the large subunit of this protein from *N. knightiana* and *N. paniculata* differs from that of *N. tomentosa* and *N. tomentosiformis*. Since the large subunits of fraction 1 protein are coded by chloroplast DNA, the difference in their primary structure reflects the structural changes of the chloroplast genes containing the coding information. This indicates that the rate of mutation of chloroplast DNA seems to be higher than predicated from the analysis of isoelectric points of this protein.

Chloroplast DNA in higher plants has a constant base composition in the region of 37 to 38% GC (11) suggesting the possibility of conservation of this organellar DNA during evolution. Recently, restriction-enzyme fragmentation experiments have revealed differences in chloroplast DNA not previously apparent (1). These studies demonstrate that while chloroplast DNAs are similar in base composition, they have no fragmentation pattern in common unless they come from closely related species. In the genus of *Nicotiana*, the chloroplast genes coding for the large subunits of fraction 1 protein (F1 protein) (5, 10) exhibit a high degree of evolutionary stability (6). Only two mutations have survived through the entire life span. This stability is suggested by the isoelectric points of the polypeptides of the large subunits of F1 protein.

Because isoelectric points measure only the over-all net charge of the polypeptides, more detailed information concerning the amino acid composition and the primary structure of these polypeptides is needed to confirm such conclusions. We have, therefore, undertaken the investigation of the amino acid composition and the tryptic peptide patterns of the large and small subunits of F1 protein from several species of *Nicotiana: N. tomentosa, N. tomentosiformis, N. paniculata,* and *N. knightiana*. They have been selected for identical isoelectric points in the polypeptides of their F1 proteins and for morphological similarities.

MATERIALS AND METHODS

Preparations of F1 Protein and Separation of Subunits. F1 proteins from leaves of *N. tomentosa, N. tomentosiformis, N. paniculata,* and *N. knightiana* were prepared by a direct crystallization procedure (4). The three-times recrystallized proteins were dissociated in 0.5% SDS and the subunits were separated by Sephadex G-100 column according to the method of Rutner and Lane (17).

Isoelectric Focusing. The crystalline F1 proteins were *S*-carboxymethylated (12) before applying to a prefocused 4.5% polyacrylamide slab gel containing 1% ampholine (pH 5 to 7) and 8 M urea. After 18 hr of electrofocusing, the slab gel was stained in bromophenol blue. All operations were carried out as described by Kung *et al.* (12). The isoelectric points were determined as previously described (9).

Tryptic Peptide Mapping. After separation, the large and small subunits of F1 proteins were subjected to trypsin digestion. Three mg of lyophilized protein was dissolved in 1 ml of 0.1 N NH_4HCO_3 (pH 8.0) and digestion with chymotrypsin-free trypsin (Calbiochem) was carried at 30 C for 30 hr. The final enzyme to substrate ratio was 1:50 (13). After digestion, the sample was lyophilized and spotted on Whatman No. 3MM paper. Paper chromatography was run for 24 hr in 1-butanol-acetic acid-water (4:1:5). The thoroughly dried paper was subsequently electrophoresed in pyridine-acetic acid-water (1:10:289) (pH 3.6) at 3,500 v for 60 min. The chromatographs were dipped through buffered ninhydrin solutions, air-dried, and then placed in a cool oven which was then heated to 70 to 80 to develop the spot (13). Samples containing only trypsin were occasionally used as a control.

Amino Acid Analysis. The large and small subunits were lyophilized to a salt-free white powder. Duplicate or triplicate samples of lyophilized large and small subunits were hydrolyzed in 1 ml of constant boiling HCl (5.7 M) at 110 C for 20 hr under vacuum. Amino acid analyses were performed on a Technicon TSM amino acid analyzer.

RESULTS

Isoelectric Points of Fraction 1 Protein. Isoelectric focusing of *S*-carboxymethylated F1 proteins shows an identical polypeptide pattern for all four species of *Nicotiana:* three polypeptides in the large subunit and one in the small subunit. Figure 1 represents a typical separation pattern of the four polypeptides of F1 proteins for *N. tomentosa, N. tomentosiformis, N. paniculata,* and *N. knightiana*. This result agrees with that previously reported by Stavely *et al.* (19) and Chen *et al.* (6). Chen *et al.* (6) demonstrated that F1 proteins from a group of six species of *Nicotiana,* including the above four species, have an identical pattern with respect to their isoelectric points. The small subunit consists of a single polypeptide with an isoelectric point of 5.30. The isoelectric points measured for the three polypeptides of the large subunits are 6.05, 6.00, and 5.95, respectively.

Amino Acid Composition of the Subunits. The large and small subunits of F1 protein were separated on Sephadex G-100 in the presence of SDS (Fig. 2). The separated subunits were homogeneous as judged by SDS gel electrophoresis and amino acid analysis. The amino acid compositions of the large and small

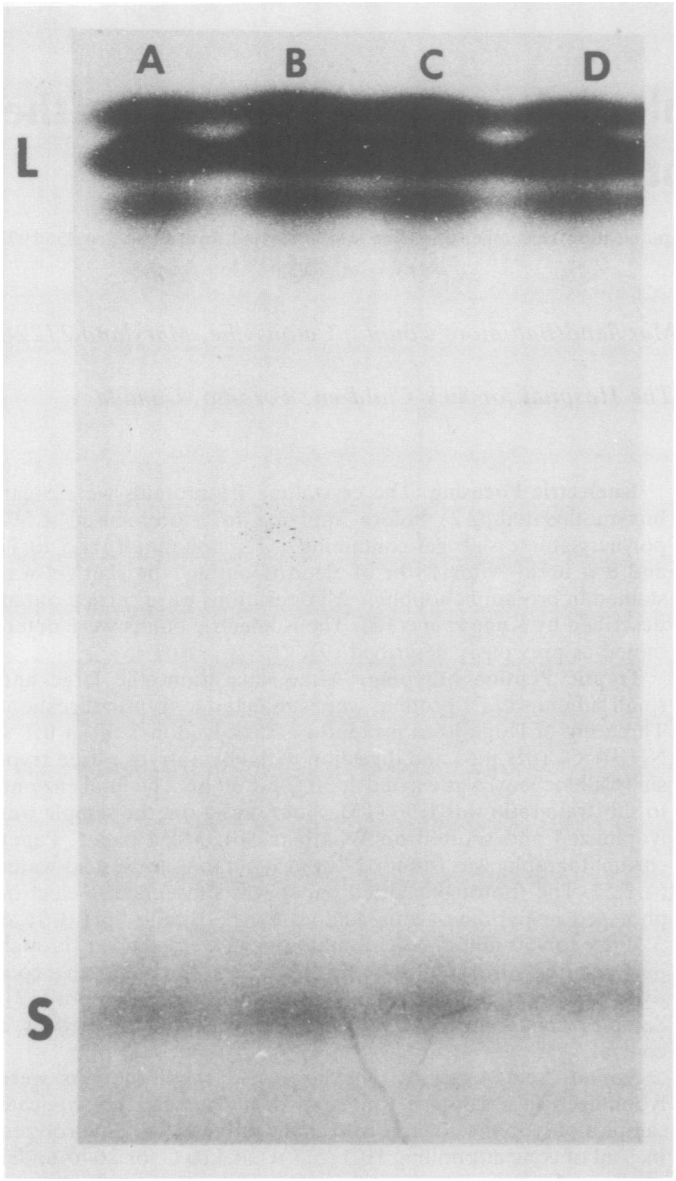

FIG. 1. Slab-gel isoelectric focusing of S-carboxymethylated fraction 1 protein from *N. tomentosa* (A), *N. tomentosiformis* (B), *N. paniculata* (C), and *N. knightiana* (D). They all have three polypeptides in the large (L) subunit and one in the small (S) subunit. The isoelectric point of each polypeptide is in the text.

subunits of F1 protein from the four species of *Nicotiana* are shown in Tables I and II. They are averages of at least two, and in many cases, four or five runs. The results are expressed as probably numbers of residues/subunit, based on mol wt of 5.5×10^4 and 1.2×10^4 for the large and small subunits, respectively. The results of Kawashima *et al.* (8) for the amino acid composition of the large and small subunits of F1 protein from *N. tabacum* are included for comparison.

It is quite apparent that the over-all amino acid compositions of the large subunits from all five species of *Nicotiana* are very similar. *N. tabacum*, however, contains less alanine and histidine residues than the other species. This reflects the difference in polypeptide composition of F1 proteins between *N. tabacum* and the other species (6). It should be pointed out that the F1 protein from *N. peniculata* and *N. knightiana* contains approximately 20% more proline residues than that from *N. tomentosa*, *N. tomentosiformis*, and *N. tabacum*. Based on the comparison of their amino acid compositions, they fall into two categories: *N. paniculata* and *N. knightiana* in one, and *N. tomentosa*, *N. tomentosiformis*, and *N. tabacum* in another. These biochemical characteristics are reflected in their morphology and cytogenetics. According to Goodspeed (7), *N. paniculata* and *N. knightiana* are grouped in the section of paniculatae, and *N. tomentosa*, *N. tomentosiformis*, and *N. tabacum* are classified in the section of tomentosae.

The amino acid compositions of the small subunits from *N. tomentosa*, *N. tomentosiformis*, *N. paniculata*, and *N. knightiana* are also very similar (Table II). This similarity can be expected from their identical isoelectric points. On the other hand, *N. tabacum* seems to contain less glycine than the other four species studied here. This difference may be a reflection of the fact that there are two polypeptides in the small subunit of *N. tabacum*. One of them was originated from *N. tomentosiformis*, the other from *N. sylvestris* (6, 10). In the case of *N. excelsior*, an Australian species, the amino acid composition is notably different in the proline and histidine residues. In contrast to all species listed in Table II, *N. excelsior* has four polypeptides in the small subunit (Table III).

Tryptic Peptide Composition of the Subunits. Figure 3 shows the peptide maps representing a typical pattern of the tryptic peptides of the large subunit of F1 proteins from *N. knightiana*, *N. paniculata*, *N. tomentosa*, and *N. tomentosiformis*. It is quite evident that they exhibit similar over-all patterns. Some peptides are not well resolved (upper left corner) because of tailing. Therefore, they are not included for comparisons. However, all of the major peptides are clearly resolved and highly reproducible from preparation to preparation (Fig. 3, A and B). When the peptide maps of all four species are compared, they fall into two different patterns. The peptide maps of *N. knightiana* and *N. paniculata* are very similar, if not identical (Fig. 3, A and C). Similarly, no difference can be found between that of *N. tomentosa* (Fig. 3 D) and *N. tomentosiformis* (not shown). This is in close agreement with the results of amino acid composition. However, difference between these two groups is apparent. A group of six different peptides (marked a in Fig. 3C) is present in *N. paniculata* and *N. knightiana* only, whereas another group of six peptides (marked b in Fig. 3D) occurs only in *N. tomentosa* and *N. tomentosiformis*. Minor differences involving lightly staining components do not appear to be attributable to differences between species because comparable variations are also found in duplicate maps from the same protein preparation. In the area (upper left corner) where peptides are not well resolved, no difference can be positively identified. In all cases, the observed number of major tryptic peptides (~65) only slightly exceed the expected number as predicted by the arginine-lysine content (~60) of those subunits. The agreement is quite good. All peptides are released from the large subunits of F1 protein as the result of tryptic digestion. None of them has its origin from the trypsin used (results not shown). In the case of small subunits

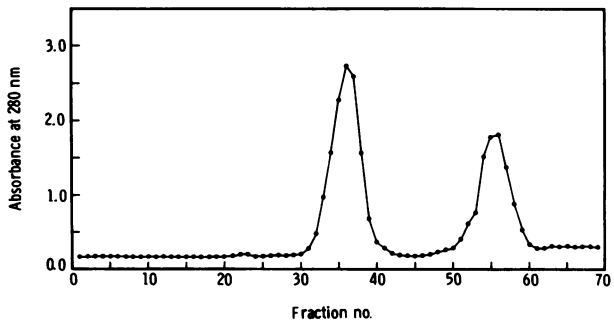

FIG. 2. Separation of large and small subunits of fraction 1 protein by Sephadex G-100 chromatography after dissociation by SDS.

Table I

Amino acid composition of the large subunits of Fraction 1 protein. Results are expressed as probable numbers of residues per subunit of 55,000 molecular weight.

Amino Acid	Species				
	N. tabacum[a]	N. tomentosa	N. tomentosiformis	N. paniculata	N. knightiana
Aspartic acid	44.2	46.4	45.5	44.6	48.0
Threonine	27.7	28.9	26.7	26.7	25.0
Serine	15.7	16.1	12.9	13.9	13.4
Glutamic acid	49.5	50.0	51.5	53.5	54.9
Proline	21.4	23.3	22.8	28.2	27.5
Glycine	49.3	49.5	52.5	50.0	51.5
Alanine	44.2	49.3	50.5	47.5	49.0
Valine	36.9	36.1	41.1	35.1	34.9
Methionine[b]	8.2	3.9	2.5	1.4	1.5
Isoleucine	21.4	23.0	23.8	20.8	19.9
Leucine	44.4	42.1	45.0	46.0	44.3
Tyrosine	19.3	18.3	16.3	15.8	15.6
Phenylalanine	21.9	23.3	23.3	21.8	23.5
Lysine	24.9	25.7	24.3	26.3	26.2
Histidine	13.8	15.8	15.9	19.3	15.6
Arginine	31.7	33.1	33.2	34.7	32.2
Cysteine		5.4	6.4	9.8	9.8

[a]Data from Kawashima et al. (8)
[b]Methionine was not corrected for methionine sulfoxone

Table II

Amino acid composition of the small subunits of Fraction 1 protein. Results are expressed as probable numbers of residues per subunit of 12,000 molecular weight.

Amino Acid	Species					
	N. tabacum[a]	N. tomentosa	N. tomentosiformis	N. paniculata	N. knightiana	N. excelsior
Aspartic acid	7.12	7.70	7.70	7.20	7.85	6.90
Threonine	4.06	4.35	4.20	4.10	3.85	4.10
Serine	3.58	3.40	3.90	3.30	2.90	4.80
Glutamic acid	14.63	15.50	15.50	15.85	15.15	15.70
Proline	6.52	7.00	7.00	6.30	7.50	8.20
Glycine	7.17	8.20	8.80	8.55	8.75	7.50
Alanine	5.72	5.95	6.00	5.55	6.15	5.70
Valine	6.40	5.10	6.40	6.70	6.50	7.30
Methionine[b]	1.61	0.40	0.46	0.45	0.30	0.50
Isoleucine	4.24	5.30	4.80	4.95	5.10	4.90
Leucine	8.26	8.50	8.70	9.50	8.60	8.60
Tyrosine	7.76	7.50	8.10	7.85	7.20	7.60
Phenylalanine	4.24	4.50	4.60	3.85	4.90	4.40
Lysine	7.90	7.60	7.00	7.95	7.95	7.10
Histidine	0.63	1.30	0.60	0.85	1.25	2.10
Arginine	4.52	4.10	4.50	4.40	4.65	3.50
Cysteine		1.60	1.90	1.60	1.60	1.80

[a]Data from Kawashima et al. (8)
[b]Methionine was not corrected for methionine sulfoxone

of *N. paniculata*, it releases 12 to 14 peptides upon tryptic digestion (Fig. 4). This is in good agreement with the arginine-lysine content of this small subunit (Table III).

DISCUSSION

The evidence presented here demonstrates that identical isoelectric points of the polypeptides do not necessarily reflect the identity of their primary structure. This is because changes in amino acid composition or sequence without a change in over-all net charge of the polypeptides would not alter the isoelectric point. On the other hand, difference in amino acid composition and peptide map will certainly indicate difference in the primary structure of these polypeptides. *N. tomentosa* shares similar primary structure of the large subunits with *N. tomentosiformis*, and *N. paniculata* shares that with *N. knightiana*. Based on the difference in primary structure of the large subunits of F1 protein, the four species of *Nicotiana* fall into two distinct groups. This is clearly reflected in their morphology and cytogenetics. According to the classification of the genus *Nicotiana* (7), *N. tomentosa* and *N. tomentosiformis* are in the subgenus of tabacum and section of tomentosae whereas *N. paniculata* and *N. knightiana* are in the subgenus of rustica and section of paniculatae.

Since the large subunits of F1 protein are coded by chloroplast DNA (5, 10), the difference in their primary structure reflects the structural changes of the chloroplast genes containing the coding information. No information regarding the frequency of chloroplast DNA mutation between these two groups can be estimated from this study. However, the rate of mutation of chloroplast DNA seems to be higher than predicted from the analysis of isoelectric points of the large subunits of F1 proteins.

It has been frequently suggested that the chloroplast genes coding for the large subunit of F1 protein and 70S rRNA are highly conserved during evolution (6, 20). This conservation may be necessary to maintain their structure within very confined limits and to conserve biological activity. For example, F1 protein (21) is RuBP carboxylase-oxygenase (EC 4.1.1.39)

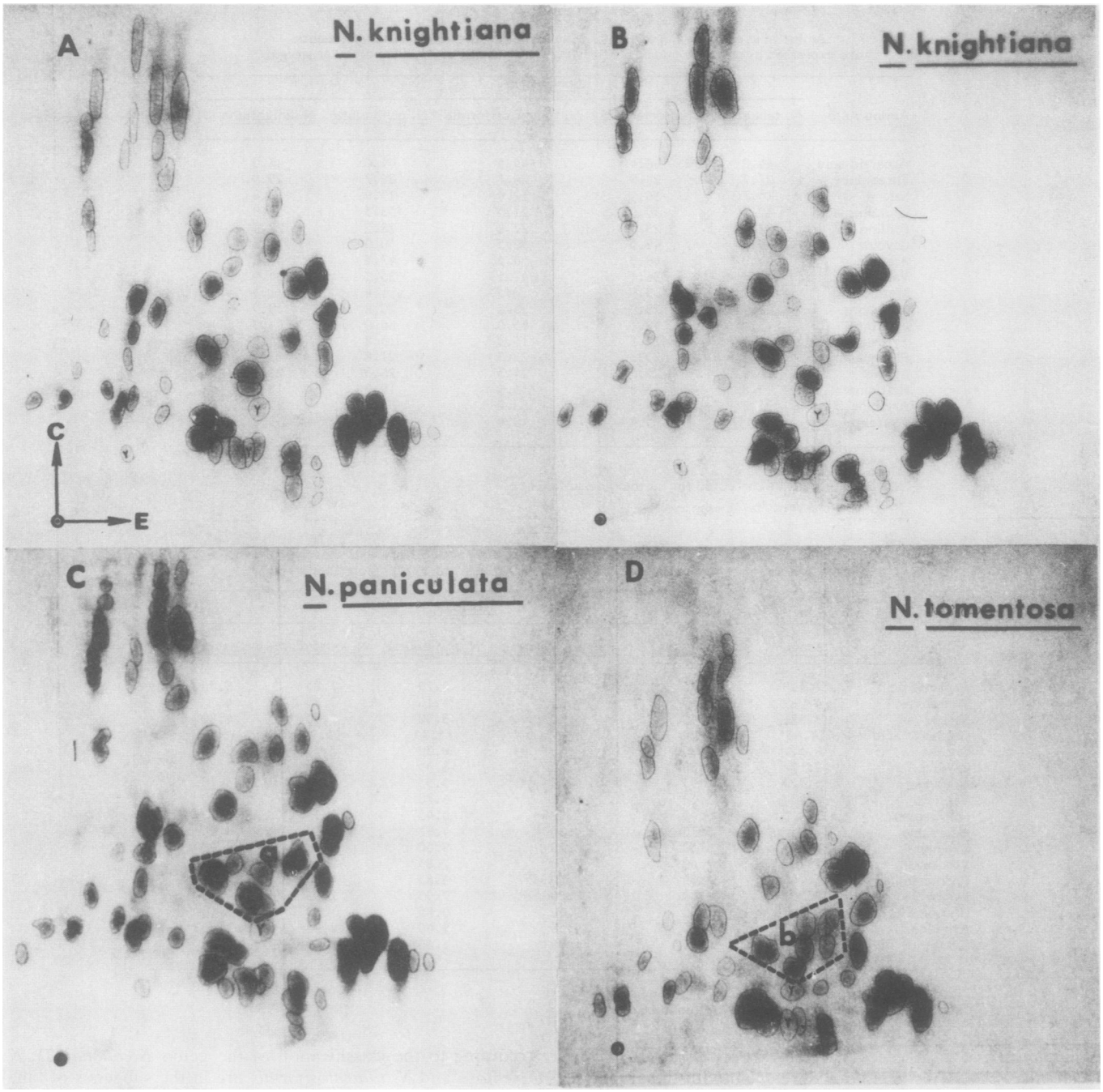

FIG. 3. Peptide maps from tryptic digests of the large subunits of fraction 1 protein from two independent preparations of *N. knightiana* (A and B), *N. paniculata* (C), and *N. tomentosa* (D). *N. tomentosiformis* (not shown here) has a map identical to that of *N. tomentosa*. Origins are at the lower left corners. Directions of chromatography (C) and electrophoresis (E) are indicated. Y = yellowish spot.

catalyzing the important reactions of both photosynthesis and photorespiration. It is composed of eight large and eight small subunits (2). The large subunits contain the catalytic sites (16) whereas the small subunits carry the regulatory sites (10). It has been suggested that the two subunits may have appeared at different times in evolution, the large subunit being the first to appear as a member of the primordial photosynthetic system (14). The genes coding for the large subunit were first established and expressed in anaerobic bacteria (15). As evolution advanced to a level of high complexity, the genes for the large subunits were compartmentalized into the organelle-chloroplast, probably by way of symbiosis (11). Most, but certainly not all, mutations in the chloroplast DNA coding for the large subunit of F1 protein are probably lethal. Recently, evidence has been presented to suggest the conservation of chloroplast rRNAs, which are known to be coded by chloroplast DNA (11, 20). Such conservation is probably a necessity because chloroplast mRNA can be recognized only by chloroplast ribosomes (3). The possible mechanism of such specific recognition could be achieved through specific base pairing between chloroplast mRNA and 16S rRNA of the 30S subunit of chloroplast ribosomes as in *Escherichia coli* (11). The question remaining to be answered is the degree of conservation required by these macromolecules for performing their activity.

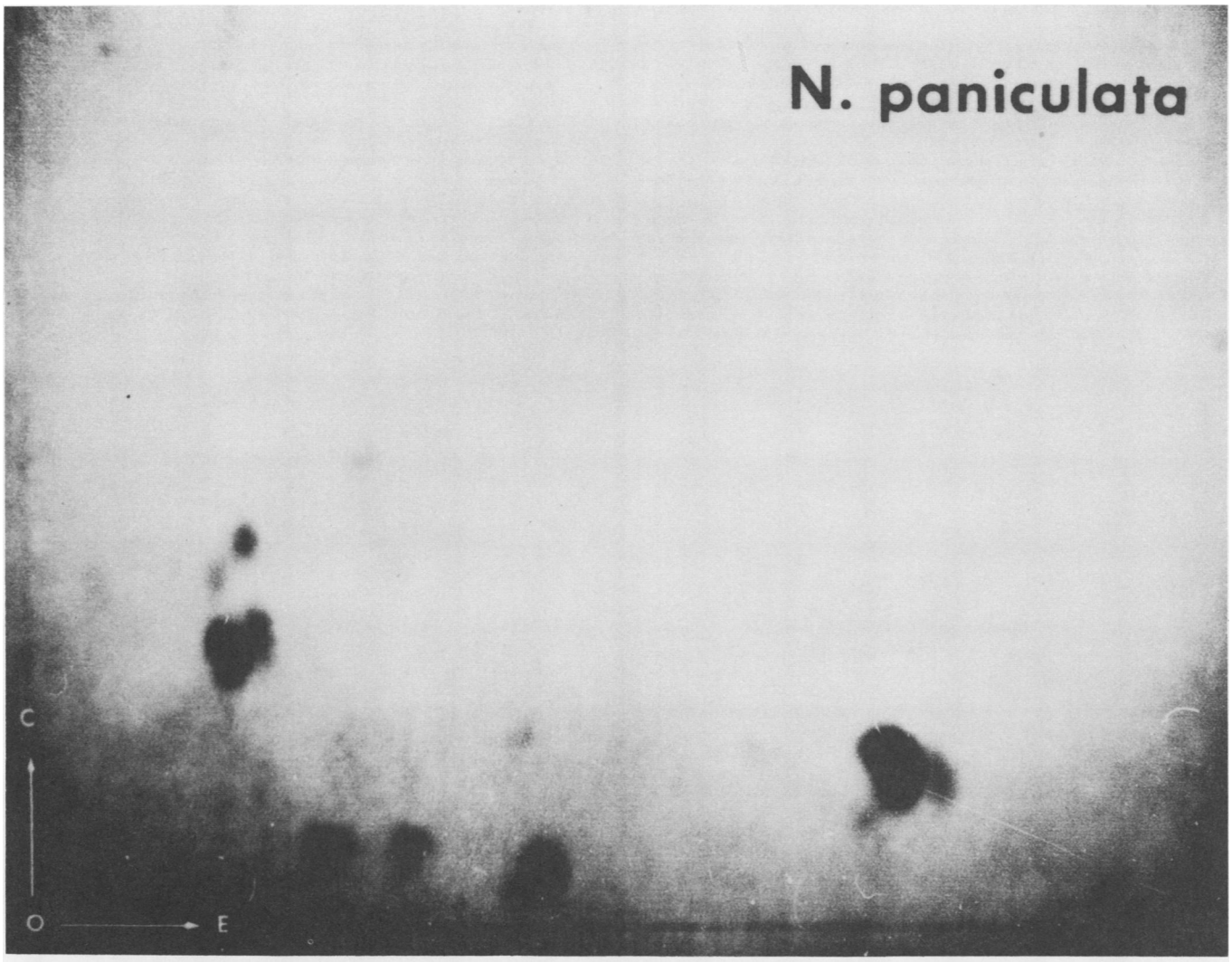

FIG. 4. Peptide map from tryptic digest of the small subunit of fraction 1 protein from *N. paniculata*.

Table III

Relationship between the numbers of polypeptides and the ratio of arginine-lysine content and the number of tryptic peptides in the small subunit.

Species	Number of polypeptides	Arginine-lysine content	Tryptic peptides	Ratio of arginine-lysine & tryptic peptides
N. paniculata	1	12	12-14	1
N. tabacum[a]	2	13-14	30-32	<1
N. excelsior	4	11-12	28[b]	<1

[a] Data from Kung et al. (12)

[b] Data from Singh and Wildman (18)

The interspecific variation in the isoelectric points of the polypeptides of F1 proteins has been successfully applied as a simple genetic marker (10). However, this application is limited only to the situation where such variation is apparent. In cases where variation does not exist, the present finding offers an alternative. The difference in amino acid composition and peptide map can serve also as a genetic marker (5).

There are 11 to 12 arginine and lysine residues per small subunit (mol. wt. 1.2×10^4) in the F1 protein of *N. paniculata*. The observed number of tryptic peptides is consistent with the expected number. This may represent a general situation in which there is only a single polypeptide in the small subunit. When there is more than one polypeptide, the observed number of tryptic peptides exceeds that expected from the arginine and lysine and content (Table III). It seems reasonable to assume that the ratio of the number of arginine and lysine to the number of tryptic peptides in each subunit is an indication of how many polypeptides exist in this subunit. If the ratio is 1, there is probably only a single polypeptide as in the case of the small subunit of F1 protein from *N. paniculata*. If the ratio is less than 1, multiple polypeptides may exist as in the case of the small subunit of *N. tabacum* and *N. excelsior*.

LITERATURE CITED

1. ATCHISON BA, PR WHITFELD, W BOTTOMLEY 1976 Comparison of chloroplast DNAs by specific fragmentation with EcoR1 endonuclease. Mol Gen Genet 148: 263-269
2. BAKER TS, D EISENBERG, FA EISERLING, L WEISSMAN 1975 The structure of form I crystals of D-ribulose-1,5-diphosphate carboxylase. J Mol Biol 91: 391-399
3. BOTTOMLEY W, TJV HIGGINS, PR WHITFELD 1976 Differential recognition of chloroplast and cytoplasmic messenger RNA by 70S and 80S ribosomal systems. FEBS Lett 63: 120-124
4. CHAN PH, K SAKANO, S SINGH, SG WILDMAN 1972 Crystalline fraction 1 protein: preparation in large yield. Science 176: 1145-1146
5. CHAN PH, SG WILDMAN 1972 Chloroplast DNA codes for the primary structure of large subunit of fraction 1 protein. Biochim Biophys Acta 277: 677-680
6. CHEN K, S JOHAL, SG WILDMAN 1976 Role of chloroplast and nuclear DNA genes during evolution of fraction 1 protein. *In* T. Bucker, W. Newport, W. Sebald, S. Werner, eds.

Genetics and Biogenesis of Chloroplasts and Mitochondria. North Holland Publishing Co. In press

7. GOODSPEED TH 1954 The Genus Nicotiana. Chronica Botanica, Waltham Mass
8. KAWASHIMA N, SY KWOK, SG WILDMAN 1971 Studies on fraction 1 protein III. Comparison of the primary structure of the large and small subunits obtained from five species of *Nicotiana*. Biochim Biophys Acta 236: 578–586
9. KUNG SD 1976 Isoelectric points of the polypeptide components of tobacco fraction 1 protein. Bot Bull Acad Sinica 17: 185–191
10. KUNG SD 1976 Tobacco fraction 1 protein: a unique genetic marker. Science 191: 429–434
11. KUNG SD 1977 Expression of chloroplast genomes in higher plants. Annu Rev Plant Physiol 28: 401–437
12. KUNG SD, K SAKANO, SG WILDMAN 1974 Multiple peptide composition of the large and small subunits of *Nicotiana tabacum* fraction 1 protein ascertained by fingerprinting and electrofocusing. Biochim Biophys Acta 365: 138–147
13. KUNG SD, JP THORNBER, SG WILDMAN 1972 Nuclear DNA codes for the photosystem II chlorophyll-protein of chloroplast membranes. FEBS Lett 24: 185–188
14. KWOK SY, SG WILDMAN 1974 Evolutionary divergence in the two kinds of subunits of ribulose diphosphate carboxylase isolated from different species of *Nicotiana*. J Mol Evol 3: 103–108
15. MCFADDEN BA, FR TABITA 1974 D-Ribulose-1,5-diphosphate carboxylase and the evolution of autotrophy. Biosystems 6: 93–112
16. NISHIMURA M, T AKAZAWA 1973 Further proof for the catalytic role of the large subunit in the spinach leaf ribulose-1,5-diphosphate carboxylase. Biochem Biophys Res Commun 54: 842–845
17. RUTNER AC, DM LANE 1967 Nonidentical subunits of RuDP carboxylase. Biochem Biophys Res Commun 28: 531–537
18. SINGH S, SG WILDMAN 1973 Chloroplast DNA codes for the RuDP carboxylase catalytic site on fraction 1 protein of *Nicotiana* species. Mol Gen Genet 124: 187–196
19. STAVLEY JR, SD KUNG, LJ SLANA 1977 Comparison of polypeptide compositions of fraction 1 protein of certain *Nicotiana* species and hybrids related to root knot resistent tobacco. Tobacco Sci 21: 31–32
20. THOMAS JR, KK TEWARI 1974 Conservation of 70S ribosomal RNA genes in the chloroplast DNAs of higher plants. Proc Nat Acad Sci USA 71: 3147–3151
21. TOLBERT NE, FJ RYAN 1975 Glycolate biosynthesis by RuDP carboxylase/oxygenase. *In* M Avron, ed. Proc III Intern Congr on Photosynthesis Vol III. Elsevier, Amsterdam pp 1303–1319