Theor Appl Genet (1986) 72:665–670

The use of seed protein electrophoresis in the study of phylogenetic relationships in Chili pepper (*Capsicum* L.)

R. C. Panda, O. Aniel Kumar and K. G. Raja Rao

Department of Botany, Andhra University, Waltair-530 003, India

Received November 2, 1985; Accepted January 5, 1986
Communicated by G. S. Khush

Summary. The seed protein profile of eight taxa of Chili peppers obtained by disc electrophoresis was found to be a diagnostic character in the study of phylogenetic relationships. The distinctness of each species and the wild and cultivated nature of concerned taxa has been confirmed. While the clustering of wild *C. annuum* var. 'glabriusculum' with *C. baccatum* types indicated that the former is the progenitor of the latter group, the marked differences discernible in the seed protein profile of all other taxa suggest a polyphyletic origin for the genus *Capsicum*.

Key words: Chili pepper – *Capsicum* – Electrophoresis – Phylogeny

Introduction

The basic criterion of phylogenetic relationships is gene homology which, in general, can not be measured directly between species because of reproductive barriers (Johnson and Thein 1970). In recent years protein and isoenzymatic analysis by polyacrylamide gel electrophoresis (PAGE) has been considered as a unique and powerful technique for ascertaining gene homology at the molecular level because of its superior capability for component resolution. Further, PAGE provides an additional tool for species identification and delimitation and has been particularly helpful in deducing systematic relationships between groups where morphological and cytological data were not corollary. In view of this, innumerable chemotaxonomists have successfully established the phylogenetic relationships employing protein electrophoretic studies in major crops like rice, wheat, barley, soyabean, broadbean, chickpea and cotton etc. (see review by Ladizinsky and Hymowitz 1979).

Though both proteins and enzymes are equally important parameters in their applications to biochemical taxonomy, the data from the former discipline seems to give more accurate information on phylogenetic relationships than the latter. Proteins separated by electrophoretic methods are thought to undergo the process of evolution with relative slowness due to their "non-essential" nature (Margoliash and Fitch 1968) while enzymes are thought to be extremely sensitive to selection pressures in evolution and thus to the survival of the organism (McDaniel 1970).

Chili pepper (*Capsicum*), a member of Solanaceae is a major spice crop with considerable economic importance and is almost cosmopolitan in distribution. However for the last several years the genus has been enshrouded with controversy relating to its taxonomy, origin and phylogenetic affinities.

To date, the electrophoretic patterns of soluble seed protein fractions in chili pepper has not been studied. Therefore, the present study was undertaken with the object of gaining further insight in elucidating the phylogenetic relationships in this genus.

Materials and methods

Seeds of the following eight taxa were employed in the present study. They are: *C. chacoense, C. baccatum* var. 'baccatum', *C. annuum* var. 'glabriusculum', *C. annuum* var. 'antigua', *C. annuum* var. 'cerasiformis', *C. baccatum* var. 'pendulum', *C. frutescens* and *C. chinense*. Of these, the first four are wild representatives while the remaining four are cultivated forms.

Employing the polyacrylamide gel electrophoresis technique outlined by Davis (1964) and Ornstein (1964), the soluble protein pattern was studied. Extraction was done by grinding

the seeds in few drops of 0.01 M TRIS-HCl buffer (pH 7.5) using a chilled mortar and pestle. The extracts were absorbed onto a filter paper disc of 0.5 cm diameter punched out to fit exactly into the gel tubes. Using Tris-glycine buffer (pH 8.3) in both reservoirs, electrophoresis was carried out for 90 min at a constant current of 3 mA per column at 4 °C. At the end of electrophoresis, the gels were removed with a fine stream of water and immediately stained with 0.25% Coomassie brilliant blue in a solution of HOAc-MeOH-H_2O (2 : 3 : 15) for 1 h to locate the protein bands, and destained with the same solution without the dye. The densitographs of the spectrum of soluble protein were prepared by scanning the gels on a Shimadzu UV-VIS Double Beam 240 Spectrophotometer. Electrophorograms were prepared on the basis of protein mobility and density expressed in R_f values. The R_f value for each band was computed from the mean of observations obtained from five independent electrophoretic runs and two separate extractions.

The percentage of similarities between different pairs of species and varieties were calculated as follows:

$$\text{Percentage similarity} = \frac{\text{No. of pairs of similar bands}}{\text{No. of different bands} + \text{No. of similar bands}} \times 100$$

The group affinity (GA) and isolation values (1. Vn) were calculated following the methods adapted by Ellison et al. (1962).

Results

The electrophoretic seed protein profiles of the representative species and varieties have been outlined in the form, of electrophorograms in Fig. 1. Altogether 30 proteins bands have been identified, though the maximum number of bands for any individual species has not exceeded 9. The R_f values ranged from 0.02 to 0.94 indicating that bands were present in all the regions of mobility, i.e. from slow to fast migrating bands. Two wild representatives, *C. chacoense* and *C. annuum* var. 'glabriusculum', have the maximum number of bands (9) while two cultivated taxa, *C. chinense* and *C. annuum* var. 'cerasiformis', have the minimum (7). Not even a single band is common to all the taxa employed in the present investigation. However, the band at R_f 0.66 is present in six of the eight taxa followed by the band at R_f 0.83 which is shared by four taxa (Fig. 1). The band at R_f 0.94, i.e. with very fast mobility, is observed in the wild species *C. chacoense* followed by the two wild varieties of *C. annuum* (*glabriusculum* and *antigua*), at R_f 0.87 which is unique to them. Bands at R_f 0.19 and 0.24 are also met with in the wild representatives (*C. chacoense, C. annuum* var. 'glabriusculum' and *C. annuum* var. 'antigua') while bands at R_f 0.15 and 0.51 are exclusive to cultivated forms *C. baccatum* and *C. frutescens,* respectively.

When the protein profile pattern of cultivated and wild representatives of any individual species was analysed it was found that bands at R_f 0.04, 0.49 and 0.66 were common to both cultivated and wild forms of *C. baccatum* while only a single band at R_f 0.66 was noticed in both wild and cultivated varieties of *C. annuum.*

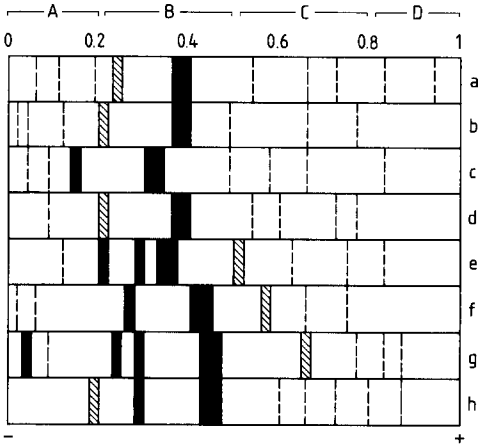

Fig. 1. Electrophorograms of soluble seed proteins in species and varieties of *Capsicum.* **a)** *C. chacoense;* **b)** *C. baccatum* var. 'baccatum'; **c)** *C. baccatum* var. 'pendulum'; **d)** *C. chinense;* **e)** *C. frutescens;* **f)** *C. annuum* var. 'cerasiformis'; **g)** *C. annuum* var. 'glabriusculum'; **h)** *C. annuum* var. 'antigua'

The densitographic analysis of the protein fractions also showed distinct quantitative differences among the species and varieties (Fig. 2). Whenever bands are present in the same position for some taxa under investigation, it does not necessarily mean that their activity is also the same. The activity of each band varies from species to species. For convenience, the protein fractions are broadly grouped into four zones depending upon their mobility, namely, (A) slow (0.0–0.2), (B) intermediate (0.21–0.5), (C) fast (0.51–0.8) and (D) very fast (0.81–1.0) (Table 1; Fig. 1). The data, in general, suggested that the wild species *C. chacoense* and the two wild varieties of *C. annuum* (*glabriusculum* and *antigua*) along with cultivated *C. baccatum* and *C. frutescens* possess very fast mobile protein fractions. Nevertheless, they also possess some slow mobility proteins which are present in other taxa. The proportion of intermediate mobility fraction varies from band to band and from species to species (Table 1). In general, the protein activity is more in this intermediate zone than in the rest.

Percentage similarities or paired affinity index between all possible pairs of species and varieties are presented in Table 2. The data indicates that none of the taxa has more than a 50% similarity with any one of the rest. The two wild varieties of *C. annuum* are related to 47.1% while both of them are almost equally isolated from the cultivated variety (12.5 and 13.33% similarity). The percentage similarity between the wild and cultivated varieties of *C. baccatum* is 37.5%. Sur-

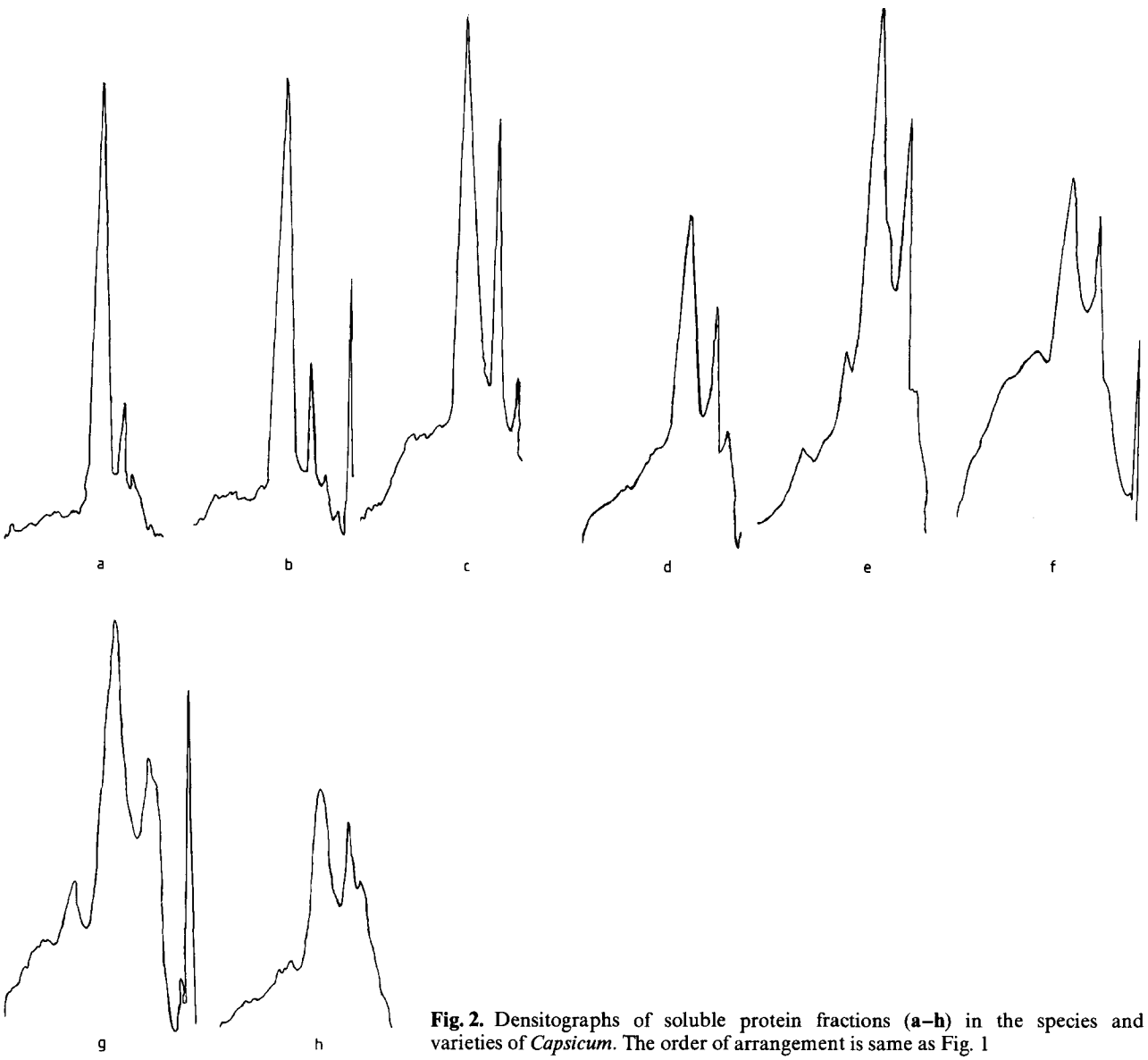

Fig. 2. Densitographs of soluble protein fractions (**a–h**) in the species and varieties of *Capsicum*. The order of arrangement is same as Fig. 1

Table 1. Distribution of soluble proteins in chili peppers on the basis of electrophoretic mobility (in %)

Selection no.	Taxa	Slow mobility (0.0–0.2) (A)	Intermediate mobility (0.21–0.5) (B)	Fast mobility (0.51–0.8) (C)	Very fast mobility (0.81–1.0) (D)
1.	*C. chacoense*	2.96	92.92	3.28	0.84
2.	*C. baccatum* var. 'baccatum'	7.24	82.75	10.0	–
3.	*C. baccatum* var. 'pendulum'	28.14	64.97	6.33	0.56
4.	*C. chinense*	7.90	81.26	10.83	–
5.	*C. frutescens*	6.10	73.27	19.52	1.10
6.	*C. annuum* var. cerasiformis'	5.01	67.06	27.94	–
7.	*C. annuum* var. 'glabriusculum'	11.71	72.51	12.64	3.15
8.	*C. annuum* var 'antigua'	12.0	74.62	11.64	1.74

Table 2. Percentage similarities between species and varieties of chili peppers based on soluble protein component homologies

Selection no.	Taxa	1	2	3	4	5	6	7	8
1.	*C. chacoense*	100.0	23.53	23.53	37.5	11.76	25.00	33.33	35.30
2.	*C. baccatum* var. 'baccatum'		100.0	37.50	40.0	25.00	26.67	35.30	12.50
3.	*C. baccatum* var. 'pendulum'			100.00	13.33	12.50	26.67	47.10	12.50
4.	*C. chinense*				100.00	13.33	0.0	25.00	26.67
5.	*C. frutescens*					100.00	13.33	23.53	12.50
6.	*C. annuum* var. 'cerasiformis'						100.00	12.50	13.33
7.	*C. annuum* var. 'glabriusculum'							100.00	47.10
8.	*C. annuum* var. 'antigua'								100.00

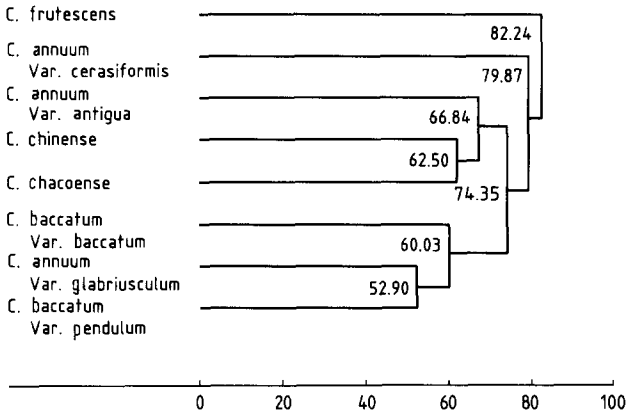

Fig. 3. Phenogram constructed on the basis of percentage of dissimilarities in the protein profile pattern

prisingly, the percentage similarity between cultivated *C. baccatum* var. 'pendulum' and wild *C. annuum* var. 'glabriusculum' is high (47.1%). The cultivated *C. annuum* and *C. chinense* have no band in common, i.e. the percentage similarity is nil.

The phenogram (Fig. 3) drawn from the percentage dissimilarities showed that the *C. annuum* var. 'glabriusculum' (wild) clusters with *C. baccatum* group, being in between the two wild and cultivated varieties. *C. chacoense* clusters with *C. chinense* and *C. annuum* var. 'antigua' (wild). The cultivated forms of *C. annuum* and *C. frutescens* do not cluster with any other or even between themselves and have more dissimilarities with the rest.

The group affinity (GA) values ranges from 211.95 (*C. frutescens*) to 323.86 (*C. annuum* var. 'glabriusculum') expressing a less than 40% relationship between them. The isolation values (I. Vn) shows more unique bands for *C. frutescens* followed by cultivated forms of *C. baccatum* and *C. annuum* while such unique bands were not present in wild *C. baccatum* var. 'baccatum', *C. annuum* var. 'glabriusculum' and *C. chinense.*

Discussion

The degree of homology in protein fractions appears to be a potential tool for measuring intra and interspecific relationships in chili peppers, as in many other crop plants. The protein profile pattern for each of the species and varieties studied here is specific to the taxon concerned. Such a constant and unique pattern of protein fraction is probably a consequence of specific gene arrangement, structure and activity in different species, as in *Vicia faba* (Ladizinsky 1975).

Duke and Glassman (1968) working on 29 species of *Drosophila* stated that the electrophoretic mobility of isozymes tended to be reduced as the species became advanced. This relationship has been further confirmed in rice (Siddiq et al. 1972). Although the present study is not based on any structurally or functionally definite protein bands, the gross fraction suggests that the species differ considerably in their mobility pattern. *C. chacoense* and both the wild varieties of *C. annuum* (*glabriusculum* and *antigua*) possess more fast mobility protein fractions, thus suggesting the view that they are of ancestral forms. The presence of some very fast mobility bands in *C. frutescens* suggests that it also possesses some traits of ancestral forms to some extent. The cultivated varieties of *C. annuum* and *C. chinense* do not possess very fast mobility protein fractions indicating that they are of recent origin or derived ones.

The percentage similarities suggest that there is considerable genetic variability among the taxa although there is not necessarily congruence with patterns of morphological variability. *C. annuum, C. frutescens* and *C. chinense* can be recognised fairly reliably on morphological grounds (Pickersgill et al. 1979) and the present data strengthens their view while the allozymic data depicts that they are inseparable (McLeod et al. 1979; Jensen et al. 1979). While the group affinity value exhibiting a less than 40% relationship justifies the distinctness of each taxon concerned, the isolation values which are potentially valuable in taxonomic consideration (Ellison et al.

1962) emphasize the distinctness of cultivated *C. annuum, C. frutescens* and *C. baccatum.* The close relationship found between the wild and cultivated varieties of *C. baccatum* indicate that both are genetically related. Therefore, the inclusion of *C. microcarpum* (*C. baccatum* var. 'baccatum') and *C. pendulum* (*C. baccatum* var. 'pendulum') under one species *C. baccatum* by Eshbaugh (1970) from breeding data and Jensen et al. (1979) from allozymic data is supported by the present electrophoretic protein profile.

The high stability of the seed protein profile and its additive nature make seed protein electrophoresis a powerful tool in elucidating the origin and evolution of cultivated plants (Ladizinsky and Hymowitz 1979). A cultivated plant and its immediate wild progenitor still form a common gene pool (Harlan and de Wet 1971) and can be considered from the genetic point of view as members of the same species. Therefore, despite conspicuous morphological differences between them, they will share, more or less, the same protein profile. Indeed, similarity in the protein profile of cultivated and their wild counterparts has been reported in wheat (Johnson et al. 1967), barley (McDaniel 1970), Cotton (Johnson and Thein 1970), soyabean (Mies and Hymowitz 1973), peanut (Cherry 1975) and corn (Paulis and Wall 1977). Employing this principle, Ladizinsky and Alder (1975a) identified *Cicer reticulum* as the wild progenitor of chickpea which gained support through a cytogenetic evidence (Ladizinsky and Alder 1975b). In contrast, the hypothesis that the broadbean *Vicia faba* has been evolved from *V. narbonensis* (Zohary and Hopf 1973) was not supported from the study of seed protein profiles (Ladizinsky 1975).

The application of the above principle to the taxa of the present study sheds considerable light as to which of the taxa are putative and which are of recent origin. For example, the clustering of *C. annuum* var. 'glabriusculum' in between wild and cultivated *C. baccatum* types indicates that the former probably represents an archtype of the latter. Since both wild *C. annuum* and cultivated *C. baccatum* are allopatric in their distribution (Anonymous 1983; Pickersgill 1971) it is likely that both might have separated in pre-historic times since the archelogical remains of *C. annuum* dates back to 7000 B.C. while that of *C. baccatum* only to 2500 B.C. (Anonymous 1983). The presence of flowers with yellow corolla spots among wild *C. annuum* cultivars (Pickersgill et al. 1979) and our breeding experiments (Panda, unpublished) further strengthen this view point. In view of the above evidence it is reasonable to conceive that the evolution may have proceeded parallelly in both cultivated and wild forms of *C. baccatum* from a common ancestor, possibly a wild representative of *C. annuum.* While *C. baccatum* var. 'pendulum' is extensively cultivated the wild *C. baccatum* var. 'baccatum' has escaped cultivation to a great extent. However, Eshbaugh (1970) and McLeod et al. (1983) do not share this idea. The complete absence of very fast mobility protein fractions in *C. baccatum* var. 'baccatum' supports such a suggestion. The marked differences observed in the protein profile pattern of wild and cultivated forms of *C. annuum* suggest that the cultivated form might have evolved from a gene pool not related with the wild forms under reference.

The clustering of *C. chacoense* with *C. chinense* indicates some homology while the interspecific hybridization between these two (Aniel Kumar 1984) shows distant relationship. However, *C. chacoense* and *C. baccatum* group cluster differently suggests lack of homology between them and the evidence from crossability relationships supports such a contention (Panda, unpublished). Cultivated forms of *C. annuum* and *C. frutescens* did not cluster with each other or with any other species which indicates that they are genetically distinct from each other as well as from the rest of the taxa studied here. If Gottleib's (1977) hypothesis that electrophoretic difference rather than electrophoretic similarity is a valid criterion and if the quantum of variation is a function of genetic divergence over evolutionary time (Vaidyanath 1981) then the findings of the present study prompt us to suggest that the origin and evolution of chili peppers is polyphyletic.

Acknowledgements. We are thankful to Prof. Dr. G. M. Reddy for providing necessary facilities to carry out the whole work in his laboratory. The first two authors are grateful to U.G.C., New Delhi for award of fellowships.

References

Aniel Kumar O (1984) Cytogenetical and palynological studies in the genus *Capsicum* L. PhD Thesis, Andhra University, Waltair, India

Anonymous (1983) Genetic resources of *Capsicum.* Int Board Plant Genet Res, Rome, Italy

Cherry JP (1975) Comparative studies of seed proteins and enzymes of species and collections of *Arachis* by gel electrophoresis. Peanut Sci 2:57–65

Davis RJ (1964) Disc electrophoresis. 2. Method and application to human serum proteins. Ann NY Acad Sci 121: 404–427

Duke EJ, Glassman E (1968) Evolution of xanthine dehydrogenase in *Drosophila.* Genetics 58:101–112

Ellison WL, Alston RE, Turner BL (1962) Methods of presentation of crude biochemical data for systematic purposes, with particular reference to the genus *Bahia* (Compositae). Am J Bot 49:599–604

Eshbaugh WH (1970) A biosystematic and evolutionary study of *Capsicum baccatum* (Solanaceae). Brittonia 22:31–43

Gottlieb LD (1977) Electrophoretic evidence and plant systematics. Ann Mo Bot Gard 64:161–180

Harlan JR, deWet JMJ (1971) Toward a rational classification of cultivated plants. Taxon 20:509–517

Jensen RJ, McLeod MJ, Eshbaugh WH, Guttman SJ (1979) Numerical taxonomic analysis of allozymic variation in *Capsicum* (Solanaceae). Taxon 28:315–327

Johnson BL, Barnhart D, Hall O (1967) Analysis of genome and species relationship in the polyploid wheats by protein electrophoresis. Am J Bot 54:1089–1098

Johnson BL, Thein MM (1970) Assessment of evolutionary affinities in *Gossypium* by protein electrophoresis. Am J Bot 57:1081–1092

Ladizinsky G (1975) Seed protein electrophoresis of the wild and cultivated species of the section *Faba* of *Vicia*. Euphytica 24:785–788

Ladizinsky G, Alder A (1975a) The origin of chickpea as indicated by seed protein electrophoresis. Isr J Bot 24: 183–187

Ladizinsky G, Alder A (1975b) The origin of chickpea *Cicer aeritinum* L. Euphytica 25:211–217

Ladizinsky G, Hymowitz T (1979) Seed protein electrophoresis in taxonomic and evolutionary studies. Theor Appl Genet 54:145–151

Margoliash E, Fitch WM (1968) Evolutionary variability of cytochrome C primary structures. Ann NY Acad Sci 151:359–381

McDaniel RG (1970) Electrophoratic characterization of proteins in *Hardeum*. J Hered 61:243–247

McLeod MJ, Eshbaugh WH, Guttman SI (1979) A preliminary biochemical systematic study of the genus *Capsicum*-Solanaceae. In: Hawkes JG, Lester RN, Skelding AD (eds) The biology and taxonomy of the Solanaceae. Academic Press, New York, pp 701–713

McLeod MJ, Guttman SI, Eshbaugh WH, Rayle RE (1983) An electrophoretic study of evolution in *Capsicum* (Solanaceae). Evolution 37:562–574

Mies DW, Hymowitz T (1973) Comparative electrophoretic studies of trypsin inhibitors in seeds of the genus *Glycine*. Bot Gaz 134:121–125

Ornstein L (1964) Disc Electrophoresis. 1. Background and theory. Ann NY Acad Sci 121:321–349

Paulis JW, Wall JS (1977) Comparison of the protein compositions of selected corns and their wild relatives teosirte and *Tripascum*. J Agric Food Chem 25:265–270

Pickersgill B (1971) Relationships between weedy and cultivated forms in some species of chili peppers (Genus *Capsicum*). Evolution 25:683–691

Pickersgill B, Heiser CB Jr, McNeill J (1979) Numerical taxonomic studies on variation and domestication in some species of *Capsicum*. In: Hawkes JG, Lester RN, Skelding AD (eds) The biology and taxonomy of the Solanaceae. Academic Press, New York, pp 679–700

Siddiq EA, Nerkar YS, Mehta SL (1972) Intra and interspecific variation in soluble proteins of *Oryza sativa* L. Theor Appl Genet 42:351–356

Vaidyanath K (1981) Electrophoretic analysis of soluble seed proteins in the study of phylogeny and species differentiation in *Oryza* I. *Sativa* complex. Indian J Bot 4:38–44

Zohary D, Hopf M (1973) Domestication of pulses in the old world. Science 182:887–894