Theor Appl Genet (1989) 78:119–128

Genetic diversity in *Cucumis sativus* L. assessed by variation at 18 allozyme coding loci

L. D. Knerr [1], J. E. Staub [1], D. J. Holder [1] and B. P. May [2]

[1] USDA/ARS, Department of Horticulture, University of Wisconsin, Madison, WI 53706, USA
[2] Department of Natural Resources, Cornell University, Ithaca, NY 14853, USA

Received December 28, 1988; Accepted March 17, 1988
Communicated by A. R. Hallauer

Summary. The genetic diversity of the U.S. *Cucumis sativus* L. germplasm collection [757 plant introductions (PI) representing 45 countries] was assessed using 40 enzymes which represented 74 biochemical loci. Polymorphisms were observed at 18 loci (*G2dh-1, Gpi-1, Gpi-2, Gr-1, Gr-2, Idh, Mdh-1, Mdh-2, Mdh-3, Mpi-2, Pepla-2, Peppap-2, Per-4, Pgd-1, Pgd-2, Pgm-1, Pgm-3,* and *Skdh*). Two PIs (285606 and 215589) contained alleles [*G2dh-1(1)* and *Per-4(2)*, respectively] which did not occur in any other PI. Other alleles which occurred in low frequencies (in <1% of the PIs) included *Gpi-1(3), Gpi-2(3), Gr-1(3), Gr-2(1), Idh(1), Mdh-1(2), Mdh-2(1), Peppap-2(1),* and *Pgd-1(1)*. Individual loci containing more than one allele in greater than 20% of the PIs included *Mpi-2, Pepla-2, Pgd-2,* and *Pgm-1*. Multivariate analyses aided in the reduction of data (principle components), depicted relationships among PIs (cluster), and identified the most discriminating enzyme loci (*Pgm-1, Pepla-2, Gr-1, Pgd-2, Mpi-2,* and *Skdh*) (classification and regression tree).

Key words: Biochemical loci – Cucumber – Isozymes – Multivariate analysis – Starch-gel electrophoresis

Introduction

Although more is known about the genetics of cucumber (*Cucumis sativus* L.; 2n = 14) than any other member of the cucurbitaceae (Robinson et al. 1976), limited information exists regarding genetic diversity (Peterson 1975) and linkage relationships (Fanourakis and Simon 1987) in this species. Of 88 catalogued, simply inherited traits in cucumber (Robinson et al. 1976; Cucurbit Genetics Cooperative 1985), many are difficult to assess and require analysis of mature plants and/or progeny testing. These circumstances have hindered construction of a genetic map for this horticulturally important genus. It would be valuable to characterize simple genetic markers which could be used to assess genetic diversity, to permit the efficient identification of linkage groups and to facilitate construction of a genomic map.

Allozymes have been used as genetic markers and tools for linkage studies in a variety of agronomic and horticultural crops (Goodman et al. 1980; Weeden and Lamb 1987). The analytic and predictive power of this methodology is based on the simple genetic control of allozymic variation and its lack of modification by environmental influences (Tanksley and Orton 1983). Allozyme polymorphisms reflect levels of genetic variability (Ayala 1970), and thus have been used for the identification of cultivars (Decker 1985; Fedak 1974) and the examination of taxonomic relationships (Crawford 1985). Such genetic control has led to the identification of marker genes associated with single genes of commercial importance such as disease resistance and morphological traits (Pollack et al. 1984; Zamir and Ladizinsky 1983).

Electrophoretic variation has been used to characterize interspecific biochemical relationships and genetic variation in *Cucumis* (Dane 1976; Esquinas-Alcazar 1977; Perl-Treves et al. 1985; Staub et al. 1987). Dane (1976) found no differences in peroxidase banding patterns among 28 accessions of *C. sativus* and one accession of a botanical variety *C. sativus* var. *hardwickii* (R.) Alef. (hereafter referred to as *C. hardwickii*). Similarly, Esquinas-Alcazar (1977) reported little or no genetic variation present for six enzyme systems among two *C. sativus* populations.

Research partially supported by Asgrow, DeRuiter, Nickerson-Zwaan, Nunhems, and Sun Seed Companies; and the Graduate School, University of Wisconsin, Madison

Information on enzyme variation in *C. sativus* from our laboratory has previously been limited to seven polymorphic loci and a small proportion of the U.S. germplasm collection (Staub et al. 1985). Therefore, a comprehensive survey was initiated to identify other polymorphisms and to determine the extent of enzyme variability in *C. sativus*. These data can form the basis for subsequent inheritance and linkage studies among polymorphic loci.

Materials and methods

Plant material and sample preparation. Seven hundred and fifty-three *C. sativus* and four *C. hardwickii* plant introductions (PI) were obtained from the United States Department of Agriculture. Regional Plant Introduction Station, Ames, Iowa in March, 1987. This represented the entire germplasm collection available at that time.

Initially, a random sample of cotyledons was taken from the collection to determine the most appropriate buffer system for electrophoresis of 49 enzymes. Subsequently, a random array of 16% of the accessions was evaluated for variation using these enzymes. The remainder of the collection was then surveyed for potential polymorphisms based on information obtained in this initial evaluation.

Electrophoresis. Cotyledons of 15 individuals from each PI were harvested from 7- to 14-day-old seedlings germinated and grown in vermiculite. Approximately 0.01 g of cotyledonary tissue from each seedling was ground in 0.1 ml of a buffer solution containing 0.67 g/l TRIS Base and 7.02 g/l of TRIS-HCl, pH 7.1, and centrifuged at $100 \times g$ for 5 min. Individuals were sampled and either used immediately or stored at $-70\,^{\circ}C$ before horizontal starch gel electrophoresis using the techniques of May et al. (1979) and May (1980).

Gels consisted of either 42 g or 56 g of a 1:1 mixture of hydrolyzed potato starch (Sigma Co., St. Louis/Mo.) and Connaught starch (Connaught Laboratories, Willowdale, Ontario, Canada) dissolved in either 300 ml or 400 ml of buffer, respectively. Gel and electrode buffers described by Clayton and Tretiak (1972), Ridgway et al. (1970), and two by Selander et al. (1971) were used and are referred to as C (pH 6.1 gel and electrode), R (pH 8.5 gel, 8.1 electrode), S-4 (pH 6.7 gel, 6.3 electrode), and S-9 (pH 8.0 gel, 7.8 electrode), respectively (Table 1).

Filter paper wicks (3×8 mm, Schleider and Schuell No. 470) were dipped into the homogenized cotyledon tissue and loaded on the gels. A wick dipped in red food coloring (SCM Corporation, Westlake, Ohio; in water, propylene glycol, and artificial colors) was placed at the end of each gel to monitor the rate of movement of the front. To standardize the relative mobilities of the electromorphs (bands), extracts of the *C. sativus* inbred processing cucumber line GY-14a were loaded on each gel and band mobilities were recorded in relation to GY-14a bands. An electric potential of 125–250 V (≤ 75 mA) was applied to the gels until the marker dye migrated to within 30 mm of the end of the gel. A gel was sliced horizontally into four or more sections by drawing monofilament thread through it. Modified staining solutions of Allendorf et al. (1977), Brewer (1970), and Shaw and Prasad (1970) were used to visualize banding patterns of 40 enzyme systems (Table 1).

Evaluation of germplasm. Isozyme banding patterns were recorded for analysis. The inheritance of the allozymes used in this study are not known. However, based on information from other organisms (Markert 1975; Neale et al. 1984; Chiang et al. 1987), it is possible to assign provisionary genetic bases to the observed electromorphs. Nomenclature follows a modified form (Staub et al. 1985, 1987) previously described by Richmond (1972), such that loci coding for enzymes (uppercase) are designated by the first letter being uppercase and the rest lowercase. If an enzyme is coded by multiple loci, these are designated by hyphenated numerals and are numbered from most cathodal to most anodal. Alleles of a given locus are numbered from most cathodal to most anodal and enclosed in parentheses. The most common allele of a locus was designated 100, and all other alleles were assigned a value based on the mobility (mm) of their homomeric protein product relative to that of allele 100. For example, an allele of *MDH-3* which had a mobility 2 mm greater than the most common allele was assigned the designation *MDH-3(2)-102.*

Analytical procedures. Three types of multivariate analyses were utilized to depict affinities and dissimilarities among individual PIs and PIs grouped by geographic region. PIs were placed into 45 groups (regions) according to the source country indicated by the U.S. Plant Introduction Station in Ames/IA.

Principle component analysis (PCA) generated linear combinations (principle components) of the original variables (loci) which maximally discriminated among the PIs (Harris 1975). This was the initial step in which the complex data set involving 18 enyzme coding loci and 757 observations (PIs) was simplified to render it more amenable to interpretation. Eigenvalues were used to compute the cumulative portions of the total multiple variance that was accounted for by each component. The percentage of the variation explained by principle components varied with the data set used in analysis. PCA was also used to identify PIs with identical isozyme phenotypes. Duplicate individuals were eliminated from the data set used for final analyses.

Another hierarchial procedure, compact linkage analysis (Ray 1982), was used to group PIs individually and by geographic region (Note: "compact linkage analysis" does not infer a genetic test). Clusters were merged sequentially based on their Euclidean distance using an algorithm which initially used each accession or region as a cluster. PIs or regions with similar isozyme phenotypes were placed in close proximity on the resulting dendrogram.

A third procedure, classification and regression tree (CART) analysis (Brieman et al. 1984), identified enzyme loci which were most discriminate in the analyses. This was accomplished with a portion of the data provided by the cluster analysis. The cluster dendrogram was pruned at a maximum linkage distance of 0.95 to yield 19 groups of countries to be used in the subsequent classification analysis. CART examined the data set in unity and sequential binary splits were made using enzyme differences in conjunction with a rule-based system. Each split produced a classification node. An optimal set of classification rules was used to discriminate between the groups of countries as quantified by an impurity function. A classification tree was produced which provided information on enzyme differences which best discriminated the 19 groups. Loci were also ranked according to their relative importance in this discriminatory procedure.

Results

Eighteen polymorphic loci were identified among the 757 accessions surveyed. *G2DH* variation is coded by 1 locus

Table 1. Enzymes assayed, their abbreviations, Enzyme Commission (EC) designations, buffer system upon which they best resolved, and number of loci scored

Enzyme	Abbreviation	E.C. designation	Buffer system	No. of loci scored
acid phosphatase	ACP	3.1.3.2	C	1
adenosine deaminase	ADA	3.5.4.4	not resolveable	
adenylate kinase	AK	2.7.4.3	S-4	1
alanine aminotransferase	ALAT	2.6.1.2	S-9	2
alcohol dehydrogenase	ADH	1.1.1.1	S-9	1
aldolase	ALD	4.1.2.13	C	1
alkaline phosphatase	AKP	3.1.3.1	C	1
aspartate aminotransferase	AAT	2.6.1.1	R	3
catalase	CAT	1.11.1.6	R	1
diaphorase	DIA	1.6.4.1	C	3
esterase	EST	–	R	3
fumarase	FUM	4.2.1.2	not resolveable	
β-galactosidase	β-GAL	3.2.1.23	C	2
galactosaminidase	GAM	–	S-9	2
general protein	PRO	–	R	2
glucose-6-phosphate dehydrogenase	G6PDH	1.1.1.49	C	2
gluosephosphate isomerase	GPI	5.3.1.9	R	2
α-glucosidase	α-GLU	3.2.1.20	not resolveable	
β-glucosidase	β-GLU	3.2.1.21	C	1
glutamate dehydrogenase	GDH	1.4.1.2	R	1
glutathione reductase	GR	1.6.4.2	S-4	3
glucoronidase	GUS	–	not resolveable	
glyceraldehyde-3-phosphate dehydrogenase	GAPDH	1.2.1.12	S-4	1
glycerate dehydrogenase	G2DH	1.1.1.29	R	1
glycerol-3-phosphate dehydrogenase	G3P	1.1.1.8	not resolveable	
guanine deaminase	GDA	3.5.4.3	not resolveable	
hexoseaminidase	HA	3.2.1.52	R	1
inorganic pyrophosphatase	PP	–	S-9	1
isocitrate dehydrogenase	IDH	1.1.1.42	S-9	1
lactate dehydrogenase	LDH	1.1.1.27	R	1
leucine aminopeptidase	LAP	3.4.11.1	C	1
malate dehydrogenase	MDH	1.1.1.37	S-4	3
malic enzyme	ME	1.1.1.40	C	1
α-mannosidase	α-MAN	3.2.1.24	R	1
manosephosphate isomerase	MPI	5.3.1.8	S-4	2
methylumbelliferyl phosphatase	MUP	–	S-4	3
nucleoside phosphorylase	NP	–	not resolveable	
peptidase with glycyl-leucine	PEPGL	3.4.13.11	not resolveable	
peptidase with leucyl-alanine	PEPLA	3.4.13.11	S-4	2
peptidase with leucyl-glycyl-glycine	PEPLGG	3.4.13.11	R	3
peptidase with phenyl-alanyl-proline	PEPPAP	3.4.13.11	R	2
peroxidase	PER	1.11.1.7	S-4	6
phosphoglucomutase	PGM	5.4.2.2	R	3
6-phosphogluconate dehydrogenase	PGD	1.1.1.43	S-4	2
3-phosphoglycerate kinase	PGK	2.7.1.10	C	2
shikimate dehydrogenase	SKDH	1.1.1.25	S-4	1
superoxide dismutase	SOD	1.15.1.1	R	2
triosephosphate isomerase	TPI	5.3.1.1	R	2
xanthine dehydrogenase	XDH	1.1.1.204	not resolveable	

with 2 alleles, *GPI* by 2 loci with 3 alleles each, *GR* by 2 loci with 3 and 2 alleles, *IDH* by 1 locus with 2 alleles, *MDH* by 3 loci with 2 alleles each, *MPI* by 1 locus with 4 alleles, *PEP-LA* by 1 locus with 5 alleles, *PEP-PAP* by 1 locus with 2 alleles, *PER* by 1 locus with 2 alleles, *PGM* and *PGD* by 2 loci with 2 alleles each, and *SKDH* by 1 locus with 2 alleles. Monomorphic and non-resolvable loci are included in Table 1.

Table 2 lists (by the source in the PI collection) the frequencies of enzyme polymorphisms (frequencies of accessions possessing multiple alleles) and Table 3 contains allelic frequencies and relative mobilities of the ho-

Table 2. Frequencies of 18 enzyme polymorphisms in *Cucumis sativus* L. resulting from a survey of the U.S. germplasm collection (755 accessions)[a]

Country-number represented	Locus																	
	Gpi-1	*Gpi-2*	*Gr-1*	*Gr-2*	*G2dh*	*Idh*	*Mdh-1*	*Mdh-2*	*Mdh-3*	*Mpi*	*Pep-la*	*Pep-pap*	*Per*	*Pgm-1*	*Pgm-3*	*Pgd-1*	*Pgd-2*	*Skdh*
Afghanistan-16	0.00	0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.56	0.06	0.00	0.00	0.81	0.00	0.00	0.38	0.00
Australia-1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
Brazil-1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
Burma-2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.50
Canada-4	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.25	0.00	0.00	0.50	0.25
China-20	0.05	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.38	0.50	0.00	0.00	0.25	0.00	0.00	0.50	0.10
Czechoslovakia-29	0.07	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.66	0.55	0.00	0.00	0.48	0.00	0.00	0.35	0.00
Denmark-3	0.00	0.00	0.67	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.67	0.00	0.00	0.67	0.00	0.00	1.00	0.00
Egypt-7	0.00	0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.43	0.14	0.14	0.00	0.71	0.14	0.00	0.14	0.14
England-2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ethiopia-2	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	1.00	0.00	0.00	0.50	0.50
France-6	0.17	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.17	0.00	0.00	0.50	0.17	0.00	0.50	0.00
Greece-1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
Hong Kong-1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
Hungary-12	0.00	0.00	0.17	0.00	0.00	0.00	0.08	0.00	0.00	0.92	0.42	0.00	0.00	0.67	0.08	0.00	0.25	0.08
India[b]-4	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.25	0.50	0.00	0.50	0.25	0.00	0.00	0.25	0.00
India[c]-49	0.02	0.00	0.08	0.06	0.00	0.00	0.00	0.00	0.00	0.16	0.27	0.06	0.00	0.33	0.00	0.00	0.37	0.04
Indonesia-1	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Iran-63	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.02	0.65	0.21	0.00	0.00	0.71	0.00	0.02	0.29	0.06
Iraq-1	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00
Israel-7	0.00	0.00	0.28	0.00	0.00	0.00	0.00	0.00	0.00	0.57	0.14	0.00	0.00	0.14	0.00	0.00	0.43	0.00
Japan-45	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.16	0.24	0.02	0.00	0.00	0.04	0.04	0.00	0.11	0.04
Kenya-2	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00
Korea-8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	1.00	0.00	0.00	0.00	0.13	0.00	0.00	0.38	0.13
Lebanon-4	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.50	0.00	0.00	0.75	0.00	0.00	1.00	0.25
Malaysia-2	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.50
Netherlands-20	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.05	0.85	0.00	0.00	0.00	0.20	0.00	0.00	0.20	0.05
New Zealand-2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Pakistan-2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.33	0.33
Philippines-4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.50	0.25	0.00	0.00	0.00
Poland-11	0.00	0.00	0.18	0.00	0.01	0.00	0.00	0.00	0.00	0.91	0.18	0.00	0.00	0.46	0.00	0.00	0.55	0.09
P.R.C.-88	0.05	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.11	0.07	0.01	0.02	0.00	0.01	0.00	0.00	0.22	0.02
Puerto Rico-1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
Spain-3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.33	0.00	0.00	0.33	0.00	0.00	1.00	0.00
Sweden-4	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.50	0.00	0.00	0.75	0.00	0.00	0.75	0.00
Syria-8	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.88	0.13	0.00	0.00	0.63	0.00	0.00	0.50	0.00
Taiwan-6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.17	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Thailand-2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.50	0.00	0.00	1.00	0.00
Turkey-178	0.01	0.02	0.09	0.00	0.00	0.00	0.01	0.02	0.00	0.55	0.41	0.00	0.00	0.61	0.05	0.01	0.40	0.14
U.A.E.-1	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00
U.S.A.-16	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.06	0.06	0.50	0.06	0.00	0.00	0.13	0.06	0.00	0.25	0.13
U.S.S.R.-39	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.90	0.41	0.00	0.00	0.46	0.00	0.00	0.49	0.36
West Germany-5	0.00	0.00	0.80	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.40	0.00	0.00	0.80	0.00
West Pakistan-4	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.50	0.00	0.00	0.75	0.00	0.00	0.75	0.25
Yugoslavia-67	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.58	0.08	0.00	0.00	0.69	0.00	0.00	0.13	0.42
Zimbabwe-1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total-755	0.01	0.05	0.10	0.01	0.01	0.01	0.01	0.01	0.03	0.52	0.23	0.01	0.01	0.43	0.02	0.01	0.32	0.13

[a] Frequencies calculated by: (no. accessions polymorphic/total). [b] *Cucumis sativus* var. *hardwickii* (R.) Alef. [c] *Cucumis sativus* L.

momeric products for each allele for the 18 polymorphic loci. Through electrophoretic and morphological observation two PIs, 164433 (India) and 426629 (Pakistan), were determined to be a species other than *C. sativus* and were not included in the analyses. This reduced the data set to 755 observations. Over half of the loci contained multiple alleles in less than 1% of the accessions and 4 loci (*Mpi-2, Pepla-2, Pgd-2,* and *Pgm-1*) contained more than 1 allele in more than 20% of the accessions (Table 2). Seventy-eight percent of the PIs contained more than 1 allele for at least 1 locus. The frequency of the most common allele (allele 100) for a given locus varied from 0.52 to 1.00. Allelic frequencies varied among loci and among geographic regions (Table 3).

Several unique PIs were observed which possessed rare alleles or multiple alleles at several loci. Only PI 285606 (Poland) possessed three distinct *G2dh-1* electromorphs. Multiple alleles for *IDH* and *PER* were found only in *C. hardwickii* (PI 183967, PI 215589). Allozymic variation was recorded in PI 183967 (India) for *IDH* and *PER* and in PI 215589 (India) for *PER*. Other PIs possessing rare alleles were:

1) 164734 (India) – *Gpi-1(3);*
2) 176525, 176952 (Turkey) – *Gpi-2(3);*
3) 193496 (Ethiopia), 109275 (Turkey) – *Gr-1(3);*
4) 179676, 183056, 183127 (India), 470254 (Indonesia) – *Gr-2(1);*
5) 171613 (Turkey), 209064 (U.S.), 326594 (Hungary) – *Mdh-1(2);*
6) 172849, 174164, 185690 (Turkey), 357835 (Yugoslavia), 419214 (Hong Kong) – *Mdh-2(1);*
7) 163213, 163216 (India), 188749 (Egypt), 432861, 432897 (People's Republic of China) – *Peppap-2(1);*
8) 169380, 175692, 185697 (Turkey), 222782 (Iran) – *Pgd-1(1).*

PI 288992 (Hungary) was polymorphic for 6 loci (*Gr-1, Mpi-2, Pepla-2, Pgd-2, Pgm-1,* and *Pgm-3*), 6 PIs [174177 (Turkey), 176524 (Turkey), 209064 (USA), 263047 (USSR), 264229 (France), and 267741 (Japan)] showed variation at 5 loci, and 31 PIs were poylmorphic for 4 loci (Table 4). Rare alleles are important in their impact on analysis, use in linkage studies, and possible use in germplasm identification. Accessions with multiple polymorphisms will prove useful in linkage studies as a means of maximizing the number of paired allelic combinations while minimizing the number of crosses.

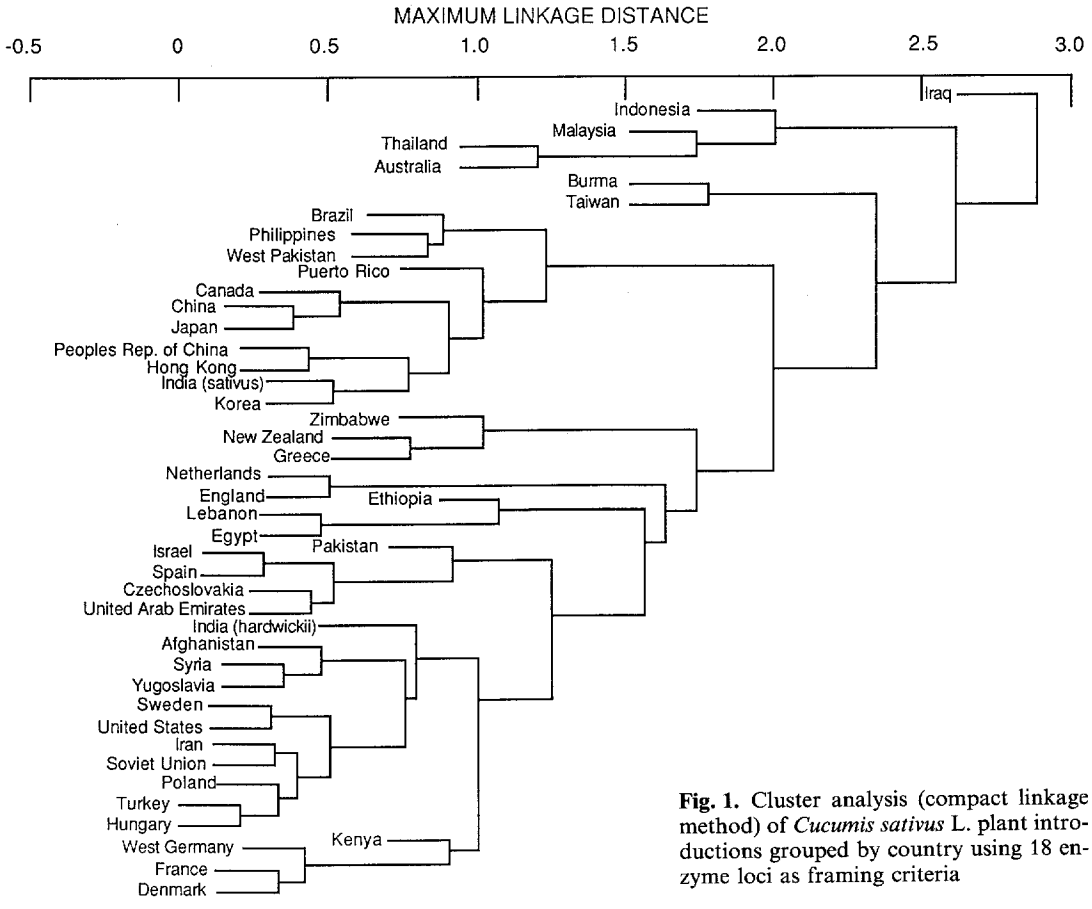

Fig. 1. Cluster analysis (compact linkage method) of *Cucumis sativus* L. plant introductions grouped by country using 18 enzyme loci as framing criteria

Table 3. Allelic frequencies of 18 polymorphic loci in *Cucumis sativus* L. resulting from a survey of the U.S. germplasm collection (755 accessions)

Country locus: allele	Assigned allelic nomenclature for electrophoretic phenotypes of enzymes*																					
	Gpi-1			*Gpi-2*			*Gr-1*			*Gr-2*		*G2dh*		*Idh*		*Mdh-1*		*Mdh-2*		*Mdh-3*	
	1	2	3	1	2	3	1	2	3	1	2	1	2	1	2	1	2	1	2	1	2
Afghanistan	0.00	1.00	0.00	1.00	0.00	0.00	0.07	0.93	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Australia	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Brazil	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Burma	0.50	0.50	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Canada	0.00	1.00	0.00	1.00	0.00	0.00	0.14	0.86	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
China	0.03	0.97	0.00	0.99	0.01	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0.98	0.02
Czechoslovakia	0.03	0.97	0.00	1.00	0.00	0.00	0.03	0.97	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Denmark	0.00	1.00	0.00	1.00	0.00	0.00	0.17	0.83	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Egypt	0.00	1.00	0.00	1.00	0.00	0.00	0.07	0.93	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
England	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Ethiopia	0.00	1.00	0.00	1.00	0.00	0.00	0.00	0.75	0.25	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
France	0.07	0.03	0.00	1.00	0.00	0.00	0.21	0.79	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Greece	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Hong Kong	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Hungary	0.00	1.00	0.00	1.00	0.00	0.00	0.08	0.92	0.00	0.00	1.00	0.00	1.00	0.00	1.00	0.96	0.04	0.00	1.00	1.00	0.00
India[a]	0.00	0.98	0.02	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.12	0.88	1.00	0.00	0.00	1.00	1.00	0.00
India[b]	0.01	0.99	0.00	1.00	0.00	0.00	0.03	0.97	0.00	0.03	0.97	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Indonesia	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.50	0.50	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Iran	0.00	1.00	0.00	1.00	0.00	0.00	0.02	0.98	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0.99	0.01
Iraq	0.00	1.00	0.00	0.00	1.00	0.00	0.25	0.75	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0.71	0.29
Israel	0.00	1.00	0.00	1.00	0.00	0.00	0.14	0.86	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Japan	0.00	1.00	0.00	1.00	0.00	0.00	0.02	0.98	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0.85	0.15
Kenya	0.00	1.00	0.00	1.00	0.00	0.00	0.25	0.75	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Korea	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0.89	0.19
Lebanon	0.00	1.00	0.00	1.00	0.00	0.00	0.14	0.86	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Malaysia	0.00	1.00	0.00	1.00	0.00	0.00	0.50	0.50	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Netherlands	0.00	1.00	0.00	1.00	0.00	0.00	0.09	0.91	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0.98	0.02
New Zealand	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Pakistan	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Philippines	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Poland	0.00	1.00	0.00	1.00	0.00	0.00	0.09	0.91	0.00	0.00	1.00	0.04	0.96	0.00	1.00	1.00	0.00	0.00	1.00	0.91	0.09
P.R.C.	0.03	0.97	0.00	1.00	0.00	0.00	0.02	0.98	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0.71	0.29
Puerto Rico	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Spain	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Sweden	0.00	1.00	0.00	1.00	0.00	0.00	0.25	0.75	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Syria	0.00	1.00	0.00	0.88	0.12	0.00	0.06	0.94	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Taiwan	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0.25	0.75
Thailand	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Turkey	0.01	0.99	0.00	0.98	0.01	0.01	0.01	0.98	0.01	0.00	1.00	0.00	1.00	0.99	0.01	0.08	0.01	0.08	0.92	1.00	0.00
U.A.E.	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
U.S.A.	0.00	1.00	0.00	1.00	0.00	0.00	0.11	0.89	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.03	0.97	0.97	0.03
U.S.S.R.	0.00	1.00	0.00	1.00	0.00	0.00	0.05	0.95	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0.97	0.03
West Germany	0.00	1.00	0.00	1.00	0.00	0.00	0.30	0.70	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
West Pakistan	0.00	1.00	0.00	1.00	0.00	0.00	0.25	0.75	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Yugoslavia	0.00	1.00	0.00	1.00	0.00	0.00	0.01	0.99	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00
Zimbabwe	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00

[a] *Cucumis sativus* var. *hardwickii* (R.) Alef.

[b] *Cucumis sativus* L.

* Allozymes which occurred in highest frequency were given the mobility designation 100. All other alleles produced protein products with relative mobilities to allozyme 100 as follows: *Gpi-1(1)*, 98; *Gpi-1(3)*, 102; *Gpi-2(2)*, 102; *Gpi-2(3)*, 104; *G2dh* to be determined; *Gr-1(1)*, 97; *Gr-1(3)*, 103; *Gr-2(1)*, 96; *Idh(1)*, 94; *Mdh-1(2)*, 101.5; *Mdh-2(1)*, 98; *Mdh-3(2)*, 102; *Mpi(2)*, 103; *Mpi(3)*, 106; *Mpi(4)*, 109; *Pep-la(1)*, 92; *Pep-la(2)*, 94; *Pep-la(3)*, 96; *Pep-la(4)*, 98; *Pep-pap(1)*, 95; *Per(2)*, 105; *Pgm-1(2)*, 102.5; *Pgm-3(2)*, 102.5; *Pgd-1(1)*, 98; *Pgd-2(2)*, 102.5; *Skdh(2)*, 102.5

Although PCA allowed for a sizable reduction in the data set, none of the variables could be eliminated since the first ten principle components explained only 49% of the total variability. Among these ten components, none had a sufficiently low eigenvector (<0.2) to warrant elimination. The original 755 accessions could be reduced to 238 through amalgamation of PIs with identical eigenvectors.

While the dendrogram produced after cluster analysis using individual PI groupings (238) did not lend itself to clear interpretation, results of analysis by geographic region are noteworthy (Fig. 1) CART revealed that *Gr-1*, *Mpi-2*, *Pepla-2*, *Pgd-2*, *Pgm-1*, and *Skdh* were most discriminating. Predictably, these most discriminating loci were those possessing polymorphisms at the highest frequencies.

Table 3 (continued)

Mpi				*Pep-la*					*Pep-pap*		*Per*		*Pgm-1*		*Pgm-3*		*Pgd-1*		*Pgd-2*		*Skdh*	
1	2	3	4	1	2	3	4	5	1	2	1	2	1	2	1	2	1	2	1	2	1	2
0.63	0.37	0.00	0.00	0.00	0.00	0.00	0.03	0.97	0.00	1.00	1.00	0.00	0.41	0.59	0.80	0.20	0.00	1.00	0.33	0.67	0.80	0.20
0.50	0.50	0.00	0.00	0.00	0.00	0.50	0.50	0.00	0.00	1.00	1.00	0.00	0.50	0.50	1.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00
1.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.75	0.00	1.00	1.00	0.00	0.25	0.75	1.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00
0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	0.25	0.75	1.00	0.00	0.00	1.00	0.38	0.62	1.00	0.00
0.75	0.25	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	0.38	0.62	1.00	0.00	0.00	1.00	0.75	0.25	0.87	0.13
0.61	0.34	0.05	0.00	0.00	0.00	0.03	0.03	0.94	0.00	1.00	1.00	0.00	0.13	0.87	1.00	0.00	0.00	1.00	0.74	0.26	0.89	0.11
0.39	0.60	0.01	0.00	0.00	0.00	0.00	0.28	0.72	0.00	1.00	1.00	0.00	0.69	0.31	1.00	0.00	0.00	1.00	0.59	0.41	1.00	0.00
0.66	0.17	0.17	0.00	0.00	0.00	0.00	0.33	0.67	0.00	1.00	1.00	0.00	0.67	0.33	1.00	0.00	0.00	1.00	0.50	0.50	1.00	0.00
0.50	0.50	0.00	0.00	0.00	0.00	0.00	0.36	0.64	0.07	0.93	1.00	0.00	0.43	0.57	0.93	0.07	0.00	1.00	0.07	0.93	0.93	0.07
0.50	0.50	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	1.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00
0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.01	0.99	1.00	0.00	0.50	0.50	1.00	0.00	0.00	1.00	0.50	0.50	0.75	0.25
0.64	0.36	0.00	0.00	0.00	0.00	0.00	0.08	0.92	0.00	1.00	1.00	0.00	0.75	0.25	0.93	0.07	0.00	1.00	0.57	0.43	1.00	0.00
0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	0.50	0.50	1.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00
1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	0.50	0.50	1.00	0.00
0.46	0.54	0.00	0.00	0.00	0.00	0.00	0.21	0.79	0.00	1.00	1.00	0.00	0.67	0.33	0.96	0.04	0.00	1.00	0.21	0.79	0.96	0.04
0.38	0.38	0.24	0.00	0.13	0.06	0.00	0.06	0.75	0.00	1.00	0.75	0.25	0.38	0.62	1.00	0.00	0.00	1.00	0.13	0.87	1.00	0.00
0.61	0.39	0.00	0.00	0.01	0.02	0.05	0.19	0.73	0.03	0.97	1.00	0.00	0.22	0.78	1.00	0.00	0.00	1.00	0.46	0.54	0.93	0.07
0.50	0.50	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	1.00	0.00
0.52	0.48	0.00	0.00	0.00	0.00	0.01	0.25	0.74	0.00	1.00	1.00	0.00	0.41	0.59	1.00	0.00	0.01	0.99	0.19	0.81	0.94	0.06
0.91	0.09	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	0.50	0.50	1.00	0.00	0.00	1.00	0.50	0.50	0.00	1.00
0.25	0.75	0.00	0.00	0.00	0.00	0.00	0.44	0.56	0.00	1.00	1.00	0.00	0.43	0.57	1.00	0.00	0.00	1.00	0.44	0.56	1.00	0.00
0.83	0.17	0.00	0.00	0.00	0.00	0.00	0.01	0.99	0.00	1.00	1.00	0.00	0.04	0.96	0.98	0.02	0.00	1.00	0.81	0.19	0.99	0.01
0.25	0.25	0.50	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	0.25	0.75	1.00	0.00	0.00	1.00	0.50	0.50	1.00	0.00
0.50	0.50	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	0.06	0.94	1.00	0.00	0.00	1.00	0.44	0.56	0.94	0.06
0.50	0.50	0.00	0.00	0.00	0.00	0.00	0.25	0.75	0.00	1.00	1.00	0.00	0.38	0.62	1.00	0.00	0.00	1.00	0.38	0.62	0.87	0.13
0.50	0.50	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	0.50	0.50	0.75	0.25
0.48	0.52	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	0.86	0.14	1.00	0.00	0.00	1.00	0.33	0.67	0.98	0.02
0.25	0.75	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00
0.50	0.50	0.00	0.00	0.00	0.00	0.00	0.50	0.50	0.00	1.00	1.00	0.00	0.50	0.50	1.00	0.00	0.00	1.00	0.75	0.25	0.50	0.50
1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.38	0.62	0.00	1.00	1.00	0.00	0.25	0.75	0.75	0.25	0.00	1.00	0.50	0.50	1.00	0.00
0.46	0.54	0.00	0.00	0.00	0.00	0.00	0.09	0.91	0.00	1.00	1.00	0.00	0.59	0.41	1.00	0.00	0.00	1.00	0.36	0.64	0.95	0.05
0.91	0.08	0.00	0.00	0.00	0.00	1.00	0.00	0.99	0.01	0.99	1.00	0.00	0.01	0.99	1.00	0.00	0.00	1.00	0.42	0.58	0.99	0.01
1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	0.50	0.50	0.50	0.50
0.17	0.83	0.00	0.00	0.00	0.00	0.00	0.50	0.50	0.00	1.00	1.00	0.00	0.50	0.50	1.00	0.00	0.00	1.00	0.50	0.50	1.00	0.00
0.38	0.62	0.00	0.00	0.00	0.00	0.00	0.25	0.75	0.00	1.00	1.00	0.00	0.38	0.62	1.00	0.00	0.00	1.00	0.38	0.62	1.00	0.00
0.44	0.56	0.00	0.00	0.00	0.00	0.00	0.06	0.94	0.00	1.00	1.00	0.00	0.56	0.44	0.87	0.13	0.00	1.00	0.25	0.75	0.81	0.19
0.67	0.33	0.00	0.00	0.00	0.00	0.00	0.08	0.92	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	1.00	0.00
0.50	0.50	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	1.00	0.00	0.50	0.50	1.00	0.00	0.00	1.00	0.50	0.50	1.00	0.00
0.55	0.44	0.01	0.00	0.00	0.00	0.00	0.26	0.74	0.00	1.00	1.00	0.00	0.62	0.38	0.96	0.04	0.01	0.99	0.23	0.77	0.91	0.09
0.50	0.50	0.00	0.00	0.00	0.00	0.00	0.50	0.50	0.00	1.00	1.00	0.00	0.50	0.50	1.00	0.00	0.00	1.00	0.50	0.50	1.00	0.00
0.41	0.53	0.03	0.03	0.00	0.00	0.03	0.12	0.85	0.00	1.00	1.00	0.00	0.36	0.64	0.97	0.03	0.00	1.00	0.47	0.53	0.94	0.06
0.54	0.46	0.00	0.00	0.00	0.00	0.00	0.33	0.67	0.00	1.00	1.00	0.00	0.57	0.43	1.00	0.00	0.00	1.00	0.35	0.65	0.83	0.17
0.50	0.20	0.30	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	0.75	0.25	1.00	0.00	0.00	1.00	0.40	0.60	1.00	0.00
0.76	0.12	0.12	0.00	0.00	0.12	0.38	0.00	0.50	0.00	1.00	1.00	0.00	0.38	0.62	1.00	0.00	0.00	1.00	0.38	0.62	0.87	0.13
0.50	0.50	0.00	0.00	0.00	0.00	0.00	0.04	0.96	0.00	1.00	1.00	0.00	0.55	0.45	1.00	0.00	0.00	1.00	0.10	0.90	0.78	0.22
0.00	1.00	0.00	0.00	0.00	0.50	0.00	0.00	0.50	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	0.13	0.87	1.00	0.00

Discussion

Limited genetic variability affects the ability to incorporate new economically important traits into commercial varieties and to study their inheritance. Moreover, genetic variability can assure success in wide crossing. The parental genotype greatly affects the success of interspecific crosses in *Cucumis* and consequently the incorporation of useful alleles from exotic germplasm (Kroon et al. 1979).

The size of the U.S. *C. sativus* germplasm collection (755) is small compared to germplasm collections such as bean (6034) and potato and its wild relatives (3500). The size of our present germplasm collection is marginally adequate to satisfy present needs, as genetic diversity in *C. sativus* is reported to be low (Dane 1976, 1983; Esquinas-Alcazar 1977; Hutchins 1938; Kupper and Staub 1988; Peterson 1975). For example, the collection has been evaluated for pickleworm (*Diaphania nitidalis* Stoll) and powdery mildew [causal agent (*Sphaerotheca fuliginea* {Schl. Ex Fr} Poll.)] resistance. Data indicate that either no accession had an adequate level of resistance or the level of resistance was low for these traits (Barczynska et al. 1988, T. C. Wehner, personal communication). Other traits with limited variability are resistance to whitefly (*Trialeurodes vaporariurum* Westw.) and

Table 4. *Cucumis sativus* L. plant introductions in the U.S. germplasm collection (755) which are polymorphic for four or more biochemical loci

PI No.	Country	Polymorphic loci
288992	Hungary	*Gr-1, Mpi-2, Pepla-2, Pgd-2, Pgm-1, Pgm-3*
264229	France	*Gr-1, Mpi-2, Pgd-2, Pgm-1, Pgm-3*
267741	Japan	*Gr-1, Mpi-2, Pgd-2, Pgm-1, Pgm-3*
263047	Soviet Union	*Gr-1, Mpi-2, Pepla-2, Pgd-2, Pgm-1*
174177	Turkey	*Gr-1, Pepla-2, Pgd-2, Pgm-3, Skdh*
176524	Turkey	*Gpi-1, Gr-1, Pepla-2, Pgd-2, Pgm-1*
209064	United States	*Gr-1, Mdh-2, Mpi-2, Pgm-3, Skdh*
220169	Afghanistan	*Mpi-2, Pgd-2, Pgm-1, Pgm-3*
283902	Czechoslovakia	*Mpi-2, Pepla-2, Pgd-2, Pgm-1*
354952	Denmark	*Mpi-2, Pepla-2, Pgd-2, Pgm-1*
188749	Egypt	*Pgd-2, Pgm-1, Pgm-3, Skdh*
288237	Egypt	*Gr-1, Mpi-2, Pepla-2, Pgm-1*
264228	France	*Gr-1, Mpi-2, Pgd-1, Pgm-1*
267746	India	*Gr-1, Mpi-2, Pepla-2, Pgm-1*
296387	Iran	*Mpi-2, Pepla-2, Pgd-2, Pgm-1*
293432	Lebanon	*Gr-1, Mpi-2, Pepla-2, Pgd-2*
306180	Poland	*Mpi-2, Pepla-2, Pgd-2, Skdh*
319216	Saudi Arabia	*Mpi-2, Pepla-2, Pgd-2, Pgm-1*
263049	Soviet Union	*Mpi-2, Pepla-2, Pgd-2, Pgm-1*
343452	Soviet Union	*Gr-1, Pgd-2, Pgm-1, Skdh*
392292	Soviet Union	*Mpi-2, Pepla-2, Pgd-2, Pgm-1*
458850	Soviet Union	*Pepla-2, Pgd-2, Pgm-1, Skdh*
284699	Sweden	*Gr-1, Pepla-2, Pgd-2, Pgm-1*
165046	Turkey	*Gr-1, Mpi-2, Pgd-2, Skdh*
171163	Turkey	*Gr-1, Pgd-2, Pgm-1, Skdh*
174167	Turkey	*Gr-1, Pgd-2, Pgm-3, Skdh*
174170	Turkey	*Gr-1, Pgd-2, Pgm-1, Pgm-3*
175692	Turkey	*Pepla-2, Pgd-1, Pgd-2, Pgm-1*
175693	Turkey	*Pepla-2, Pgd-2, Pgm-3, Skdh*
204567	Turkey	*Mpi-2, Pepla-2, Pgd-2, Pgm-1*
204568	Turkey	*Mpi-2, Pepla-2, Pgd-2, Pgm-1*
206953	Turkey	*Mpi-2, Pepla-2, Pgd-2, Pgm-1*
344348	Turkey	*Mpi-2, Pgd-2, Pgm-1, Skdh*
264666	West Germany	*Gr-1, Mpi-2, Pgd-2, Pgm-1*
269481	West Pakistan	*Mpi-2, Pepla-2, Pgd-2, Pgm-1*
269482	West Pakistan	*Gr-1, Pepla-2, Pgm-1, Skdh*

tolerance to cucumber green mottle virus. Additionally, economically important forms of some traits in the collection are represented by only one or two accessions. For example, bacterial wilt [*Erwinia tracheiphila* (E. F. Smith) Holland] resistance has only been observed in PI 200815 and PI 200818, both from Burma.

The limited number of accessions in the cucumber collection does not allow for a comprehensive analysis of germplasm diversity within this species. However, within the constraints of analysis it does provide information on diversity within the collection and can provide a basis for further evaluation or collection.

In this study we have identified 18 polymorphic loci in *C. sativus*, suggesting that there is more intraspecific variation among biochemical loci than originally reported (Staub et al. 1985). Although the U.S. *C. sativus* germplasm collection is relatively small, adequate genetic diversity exists among biochemical loci for their use in inheritance and population genetics studies. Albeit a large number of polymorphic biochemical loci have been observed in maize (37) [2n = 20] (Goodman and Stuber 1983), pepper (31) [2n = 24] (McLeod et al. 1983), and tomato (38) [2n = 24] (Bernatzky and Tanksley 1986), linkage and inheritance studies are routinely conducted with fewer genetic markers as in soybean (18) [2n = 40] (Kiang and Gorman 1983) or celery (14) [2n = 22] (Orton 1983). As cucumber has 18 biochemical markers and only seven pairs of chromosomes (2n = 14), the characterization of linkage groups is likely.

Widening the genetic base of cucumber through collection may be needed if diversity in the existing collection is indeed narrow. The International Board for Plant Genetic Resources has not given collection of cucumber a high priority among vegetable crops (1981). Considering the fact that only 7% of the U.S. cucumber collection has its source in India (its center of diversity) and that electrophoretic diversity was found to be relatively low in this region, it seems reasonable to consider developing collection strategies for this area. Since India is the accepted center of diversity for cucumber, further diversity should be found as more land races and wild forms are sampled.

The genetic information obtained from an examination of a germplasm collection is very different from that of natural populations. Accessions are either inbred lines, hybrids, or land race composites. Such accessions are usually adapted to the country of origin, have attributes which make them unique for some economic purpose, and provide a comparatively narrow genetic picture of what exists in that region. For example, in our study, PIs from Japan and the People's Republic of China possessed enzyme polymorphisms at very low frequencies. These PIs are not the result of sampling wild, random mating populations but rather the culmination of artificially directed selection.

Plant introductions developed through generations of selfing during plant breeding could undergo genetic drift, thereby possessing a lower level of heterozygosity when compared to wild populations. One polymorphism, a variant at the *G2dh* locus, was found only in PI 285606 (Polish). This and other observations (rare alleles for *Gpi-1, Gpi-2, Gr-1, Gr-2, Idh, Mdh-1, Mdh-2, Peppap-2, Per-4,* and *Pgd-1*) lends support to the hypothesis that the U.S. germplasm base is narrow and lacks diversity present in other crop germplasm collections.

The variation which exists among alleles, loci, and countries was not uniform. Moreover, alleles of a given locus did not occur in equal frequencies. The occurrence of alleles at a frequency of <0.1 was not uncommon. While frequencies of multiple alleles within loci ranged from 0.001 for *G2dh* to 0.522 for *Mpi*, variation for a given locus among countries ranged between 0 and 1.

These data, together with the fact that there were large differences in the number of PIs available from various countries, indicate that the data and their interpretation are complex.

Multivariate analysis reduced the data for interpretation. Results from these analyses depicted affinities between PIs and geographic regions, and the discriminatory power of particular loci. PCA identified PIs with identical enzyme phenotypes (identical eigenvectors) which consequently resulted in a substantially smaller (68%) data set used in cluster analysis. PCA allowed for characterization of cases in which PIs from different countries were identical for all 18 loci but regionally discrete. For example, PI 188807 (U.S.), PI 211975 (Iran), PI 279469 (Japan), and PI 288238 (Egypt) possessed identical enzyme phenotypes. This type of information provides weak argumentation when hypothesizing geographic affinities among PIs and reflects the complex ancestral structure of the collection.

The cluster analysis dendrogram (Fig. 1) suggested affinities exist among accessions from, for example, West Germany, France, and Denmark. However, in some instances, cluster formation grouped countries (i.e., U.S. and Sweden or New Zealand and Greece) which are geographically distant. In some cases, countries distant on the dendrogram contained PIs with identical enzyme phenotypes. For example, PIs 223841 (Philippines) and 234517 (U.S.) possess identical enzyme phenotypes but the countries are quite distant on the dendrogram (Fig. 1). This occurrence is possible since allelic frequencies of PIs from a given country, and not data from individual PIs, were used in the analysis. These inconsistencies could be due to the number of accessions surveyed or to genetic linkages. The most likely contributor to the similarity of accessions among geographically distant countries is the sharing of accessions among countries. For example, in the USDA cucumber breeding program, several important traits were incorporated into U.S. germplasm via foreign PIs. These include anthracnose (*Colletotrichum orbiculare* Berk. & Mont.) resistance from PIs 197087 (India) and 212233 (Japan), bacterial wilt resistance from PIs 200815 and 200818 (Burma), and gynoecoius sex expression from PI 220860 (Korea). As the parentage of most PIs is unknown, geographic affinities described by clustering procedures require judicious appraisal.

CART provided a list of the variables according to their relative importance in the analysis. *Pgm-1, Pepla-2, Mpi-2, Gr-1, Pgd-2,* and *Skdh* (in rank order of discrimination) were found to be the most important in the discrimination of accessions. These six loci were the same six with the highest frequencies of polymorphisms in the collection. In future electrophoretic studies of *C. sativus* accessions, it may be adequate to initially survey only these six loci as a means of saving time and expense.

Studies are currently underway to determine genetic linkages through joint segregation of biochemical loci and morphological single gene traits. Information on biochemical and morphological loci will result in a detailed characterization of linkage groups in cucumber.

Acknowledgements. Mention of a trademark, proprietary product or vendor does not constitute a guarantee or warranty of the product by the U.S. Department of Agriculture and does not imply its approval to the exclusion of other products or vendors that may also be suitable.

References

Allendorf FW, Mitchell N, Ryman N, Stahl G (1977) Isozyme loci in brown trout (*Salmo trutta* L.): Detection and interpretation from population data. Hereditas 86:179–190

Ayala FJ, Mourao CA, Perez-Salas S, Richmond R, Dobzhansky T (1970) Enzyme variability in the *Drosophila willistoni* group. I. Genetic differentiation among sibling species. Proc Natl Acad Sci USA 67:225–232

Barczynska H, Van Klienwee D, Palmer M, Staub JE, Clark R (1988) Evaluation of cucumber germplasm for five pathogens. HortScience 23:778

Bernatzky R, Tanksley SD (1986) Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics 112:887–898

Brewer GW (1970) An introduction to isozyme techniques. Academic Press, New York

Brieman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Wadsworth, Monterey

Chiang YC, Gorman MB, Kiang YT (1987) Inheritance and linkage analysis of phosphoglucose isomerase isozymes in soybeans. Biochem Genet 25:893–900

Clayton JW, Tretiak DN (1972) Amine-citrate buffers for pH control in starch gel electrophoresis. J Fish Res Board Can 29:1169–1172

Crawford DJ (1985) Electrophoretic data and plant speciation. Sys Bot 10:405–416

Cucurbit Genetics Cooperative, Cucurbit Gene List Committee (1985) Gene list for cucumber. Cucurbit Genet Coop Rep 8:86–71

Dane F (1976) Evolutionary studies in the genus *Cucumis*. PhD Diss, Colorado State University, Fort Collins, 202 pp

Dane F (1983) In: Tanksley SD, Orton TJ (eds) Isozymes in plant genetics and breeding Part B. Elsevier, Amsterdam, pp 369–390

Decker DS (1985) Numerical analysis of allozyme variation in *Cucurbita pepo*. Econ Bot 39:300–309

Esquinas-Alcazar JT (1977) Alloenzyme variation and relationships in the genus *Cucumis*. PhD Diss, University of California, Davis, 170 pp

Fanourakis NE, Simon PW (1987) Analysis of genetic linkage in the cucumber. J Hered 78:238–242

Fedak G (1974) Allozymes as aids to Canadian barley cultivar identification. Euphytica 23:166–173

Goodman MM, Stuber CW (1983) In: Tanksley SD, Orton TJ (eds) Isozymes in plant genetics and breeding, part B. Elsevier, Amsterdam, pp 369–390

Goodman MM, Stuber CW, Newton K, Weissinger HH (1980) Linkage relationships of 19 enzyme loci in maize. Genet 96:697–710

Harris, RJ (1975) A primer of multivariate statistics. Academic Press, New York

Hutchins AE (1938) Some examples of heterosis in the cucumber, *Cucumis sativus* L. Proc Am Soc Hortic Sci 36:660–664
International Board for Plant Genetic Resources (1981) Revised priorities among crops and regions. IBPGR Secretariat, Consultative group on international agricultural research, Rome, 18 pp
Kiang YT, Gorman MB (1983) In: Tanksley SD, Orton TJ (eds) Isozymes in plant genetics and breeding, Part B. Elsevier, Amsterdam, pp 369–390
Kroon GH, Custers JBM, Kho YO, Den Nijs APM, Varekamp HQ (1979) Interspecific hybridization in *Cucumis* (L.). I. Need for genetic variation, biosystematic relations and possibilities to overcome crossability barriers. Euphytica 28:723–728
Kupper RS, Staub JE (1988) Combining ability studies between lines of *Cucumis sativus* L. and *Cucumis sativus* var. *hardwickii* (R.) Alef. Euphytica 38:197–216
Markert CL (1975) Isozyme IV. Genetics and evolution. Academic Press, New York
May B (1980) The *Salmonid* genome: evolutionary restructuring following a tetraploid event. PhD Thesis, Pennsylvania State University
May B, Wright JE, Stoneking M (1979) Joint segregation of biochemical loci in Salmonidae: Results from experiments with *Salvelinus* and review of the literature on other species. J Fish Res Board Can 36:1114–1128
McLeod MJ, Guttman SI, Eshbaugh WH (1983) In: Tanksley SD, Orton TJ (eds) Isozymes in plant genetics and breeding, part B. Elsevier, Amsterdam, pp 369–390
Neale DB, Weber JC, Adams WT (1984) Inheritance of needle tissue isozymes in Douglas-fir. Can J Genet Cytol 26:459–468
Orton TJ (1983) In: Tanksley SD, Orton TJ (eds) Isozymes in plant genetics and breeding, part B. Elsevier, Amsterdam, pp 369–390
Perl-Treves R, Zamir D, Navot N, Galun E (1985) Phylogeny of *Cucumis* based on isozyme variability and its comparison with plastome phylogeny. Theor Appl Genet 71:430–436
Peterson CE (1975) Plant introduction in the improvement of vegetable cultivars. HortScience 10:575–579
Pollack LM, Gardner CO, Parkhurst AM (1984) Relationships between enzyme marker loci and morphological traits in two mass selected maize populations. Crop Sci 24:1174–1179
Ray AA (1982) SAS Users guide: statistics. SAS Institute, Cary/NC
Richmond RC (1972) Enzyme variability in the *Drosophila williston* group. 3. Amounts of variability in the superspecies *D. paulistorum*. Genetics 70:87–112
Ridgway GJ, Sherburne SW, Lewis RD (1970) Polymorphism in the esterases of Atlantic herring. Trans Am Fish Soc 99:147–151
Robinson RW, Munger HM, Whitaker TW, Bohn GW (1976) Genes of the Cucurbitaceae. Hortic Sci 11:554–567
Selander RK, Smith MH, Yang SY, Johnson WE, Gentry JB (1971) Biochemical polymorphism and systematics in the genus Peromyseus. I. Variation in the old-field mouse (*Peromyseus polionotus*). In: Studies in genetics, University of Texas Publication, Austin
Shaw CR, Prasad R (1970) Starch gel electrophoresis of enzymes – a compilation of recipes. Biochem Genet 4:297–320
Staub JE, Kupper RS, Schuman D, Wehner TC, May B (1985) Electrophoretic variation and enzyme storage stability in cucumber. J Am Soc Hortic Sci 110:426–431
Staub JE, Fredrick L, Marty TL (1987) Electrophoretic variation in cross-compatible wild diploid species of *Cucumis*. Can J Bot 65:792–798
Tanksley SD, Orton TJ (1983) Isozymes in plant genetics and breeding, parts A & B. Elsevier, Amsterdam
Weeden NF, Lamb RC (1987) Genetics and linkage analysis of 19 isozyme loci in apple. J Am Soc Hortic Sci 112:865–872
Zamir D, Ladizinsky G (1983) Genetics of allozyme variants and linkage groups in lentil. Euphytica 33:329–336