*Genetic Resources and Crop Evolution* **44:** 315–326, 1997.

© 1997 *Kluwer Academic Publishers. Printed in the Netherlands.*

# Genetic diversity in cucumber (*Cucumis sativus* L.): III. An evaluation of Indian germplasm †

Jack E. Staub[1,*], Felix C. Serquen[1] & James D. McCreight[2]
[1]*USDA-Agricultural Research Service, Vegetable Crops Unit, Department of Horticulture, 1575 Linden Dr., University of Wisconsin, Madison, WI 53706 USA;* [2]*USDA-Agricultural Research Service, Crop Improvement and Protection Unit, 1636 East Alisal Street, Salinas, CA, USA (*author for correspondence)*

Received 19 July 1996; accepted in revised form 3 February 1997

*Key words:* cluster analysis, ethnobotany, isozymes, plant collection, plant exploration, plant germplasm, principal component analysis

## Summary

Genetic variation in cucumber (*Cucumis sativus* L. var. *sativus*) accessions from India was assessed by examining variation at 21 polymorphic isozyme loci. Forty-six accessions acquired by the U.S. National Plant Germplasm System (NPGS) before 1972 were compared with 146 accessions collected during a 1992 U.S.–India expedition to the states of Rajasthan, Madhya Pradesh and Uttar Pradesh, India. Isozymic profiles of these Indian accessions were also compared with 707 previously examined U.S. NPGS cucumber accessions. Two distinct groups (Group 1 and Group 2) were identified within accessions collected in 1992 ($0.025 < P < 0.01$). Variation at *Ak-2*, *Fdp-2*, *Gr*, *Mdh-2*, *Mpi-1*, *Per*, *Pgm* and *Skdh* was important in the detection of this difference. A previously unreported *Pgm* allele [*Pgm (3) – 105*] was detected in accessions collected in a market in Pali, Rajasthan. Group 1 contained 37 (27 Madhya Pradesh + 10 Uttar Pradesh) accessions and Group 2 contains 102 (84 Rajasthan + 18 Madhya Pradesh) accessions. Seven accessions (5 Madhya Pradesh + 2 Rajasthan) were not associated with either group. The accessions 20664 (Tonk, Rahjasthan), 20666 (Jaipur, Rahjasthan), 20872 (Sehore, Madhya Pradesh), 20881 (Ashtok, Sehore, Madhya Pradesh) and 21026 (Bhatta, Dehra Dun, Uttar Pradesh) were heterozygous for at least nine loci and represent the genetic diversity within this collection. Isozymic variation in U.S. NPGS accessions acquired before 1972 differed significantly ($P < 0.005$) from those collected during 1992. All loci were important in the detection of this difference, except *Ak-2*, *Pep-pap*, and *Pgd-2*. When Indian accessions taken collectively (i.e., those acquired before 1972 and during 1992) were compared with an array of 707 *C. sativus* accessions examined previously, relationships between accessions grouped by country or subcontinent differed from those found in previous work.

## Introduction

The genus *Cucumis* of the Cucurbitaceae contains two major commercially grown vegetables, cucumber (*Cucumis sativus* L. var. *sativus*; $2n = 2x = 14$; hereafter referred to as *C. s.* var.*sativus*) and melon (*Cucumis melo* L.; $2n = 2x = 24$) (Jeffrey, 1980; Kirkbride, 1993). Cucumber is indigenous to India, and has been domesticated for at least 3,000 years (Whitaker & Davis, 1962). *Cucumis sativus* var. *hardwickii* (Royle) Alef. (hereafter referred to as *C. s.* var. *hardwickii*; $2n = 2x = 14$) is a wild, sympatric botanical variety of *C. s.* var. *sativus* that grows in the foothills of the Himalayan mountains (Deakin et al., 1971). Because *C. s.* var. *hardwickii* possesses a sequential fruiting and multiple branching habit not present in *C. s.* var.

† The cost of publishing was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked as advertisement solely to indicate this fact. Mention of a trade name, proprietary product, or specific equipment does not constitute a guarantee or warranty by the USDA and does not imply is approval to the exclusion of other products that may be suitable.

*sativus*, it has potential for increasing fruit yield in commercial cucumber (Horst & Lower, 1978; Staub et al., 1993).

Genetic markers (morphological and biochemical) have been employed to characterize the genetic diversity present in the cucumber collection (Knerr et al., 1989; Meglic et al., 1996). The degree of genetic diversity in *C. s.* var. *sativus* is relatively low compared with other cross-fertilized species of *Cucumis* (Dane, 1976, 1983; Esquinas-Alcazar, 1977; Knerr et al., 1989). A recent assessment of 753 accessions in the NPGS cucumber collection using 21 isozyme loci revealed that an average of 1.4 loci were polymorphic per enzyme system and 2.2 alleles were present per polymorphic locus (Meglic et al., 1996). The isozymic variation present in these accessions allowed discrimination among *C. s.* var. *sativus* accessions, and between *C. s.* var. *hardwickii* and *C. s.* var.*sativus* accessions. All accessions of *C. s.* var. *sativus* were isozymically distinct from four *C. s.* var. *hardwickii* accessions, which were themselves dissimilar from each other. Isozyme profiles allowed for discrimination among accessions when accessions were grouped by continent or sub-continent.

Genetic erosion has resulted in a depletion of both the number of crop species and the genetic diversity expressed by the amount of genetic variation within a species (Frankel, 1972; Harlan, 1975). Rapid development of elite cultivars has hastened the displacement of old varieties and landraces. Thus, in many species the broad genetic base needed for crop improvement continues to shrink.

Plant collection expeditions were undertaken by the U.S. government in the early part of this century to acquire vegetable germplasm from Asia. These included expeditions led by D. Fairchild in 1901, B. Lathrop and A. Armour in 1902, J.F. Rock in 1920/1921, W. Koelz in 1939/1948, and H.S. Gentry in 1953 (U.S. National Agricultural Archives, Geil and Giel files). Since 1953, cucumber (*C. s.* var. *sativus* and *C. s.* var. *hardwickii*) germplasm donated by international public and private sources has resulted in the formation of a U.S. National Plant Germplasm System (NPGS) cucumber collection consisting of approximately 1,360 accessions (personal communication, K. Reitsma, NPGS, Ames, Iowa, 1994). Approximately 194 ($\sim$14%) were obtained directly from India. In addition, the NPGS *Cucumis* collection contains about 291 accessions of 17 wild African diploid *Cucumis* species ($2n = 2x = 24$).

In 1992, the U.S. and Indian governments sponsored an expedition to collect *Cucumis* spp. in the states of Rajasthan, Madhya Pradesh, and Uttar Pradesh, India (Figure 1). Collection sites were chosen at 15 to 20 km intervals based on the availability of specific *Cucumis* spp. Collections were made from cultivated and non-cultivated areas, vegetable markets (subji mundi), and from seed dealers. One hundred and three cucumber collections were made in the northwestern (Jodhpur, Pali, Nagaur, Bikaner and Sri Gangagnangar districts), eastern/central (Sikar, Jaipur, and Tonk districts), southeast (Bhilwara and Chittaurgarh districts) and southwestern (Udaipur, Sirohi, and Jodhpur districts) regions of Rajasthan. In Madhya Pradesh (Chhatarpur, Bhopal, Hoshangabad, Indore, Dhar, Dewas, and Sehore districts) 63 collections were made. In Uttar Pradesh, 18 accessions were collected at low (Dehra Dun; $\sim$450m) and moderately high (Mussoorie; $\sim$2,300m) elevations in the foothills of the Himalayas.

Seed of these accessions were subsequently increased in 1993 and 1994 at the North Central Regional Plant Introduction Station (NCRPIS), Ames, Iowa, and accessions were evaluated for morphological characteristics (principally fruit). Complete passport data and descriptions of morphological traits of all accessions are available through NCRPIS and the U.S. Department of Agriculture Germplasm Resources Information Network (GRIN), Beltsville, Maryland.

An assessment of the U.S. NPGS cucumber collection by Meglic et al. (1996) included 46 accessions of direct Indian origin acquired before 1972. There were no accessions *C. s.* var. *sativus* accessions received from India between 1972 and 1992. The 1992 U.S./India expedition increased the cucumber collection $\sim$11% and represents a three-fold increase in accessions from India. This study was designed to: (1) determine the genetic relationships among cucumber accessions collected during the 1992 U.S.–India expedition using 21 polymorphic isozyme loci; (2) identify accessions possessing unique isozymic profiles within this collection; (3) characterize genetic relationships among all Indian accessions in the U.S. NPGS collection, and; (4) determine the relative genetic relationship of Indian accessions to other cucumber accessions in the U.S. NPGS. The latter was accomplished by comparing the genetic variation detected in samples collected in 1992 with that found in a previous study (Meglic et al., 1996), and then interpreting this variation in light of available ethnobotanical information.

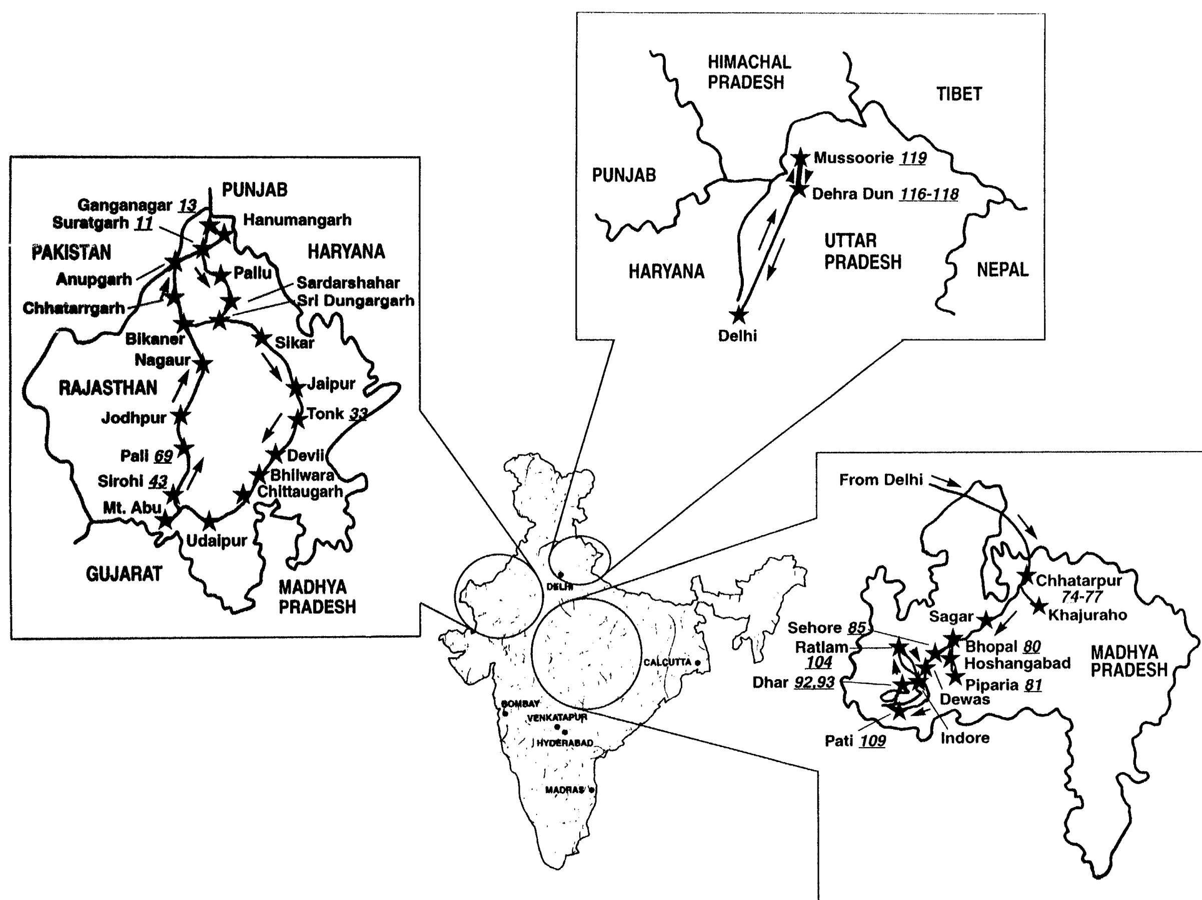



*Figure 1.* Collection site locations and routes of the 1992 U.S.–India *Cucumis* collection expedition to Rajasthan, Madhya Pradesh, and Uttar Pradesh, India (numbers identify collection sites in consecutive order of acquisition).

## Materials and methods

### *Plant material*

Poor germination and/or unique day-length requirements prevented the increase of 38 Indian cucumber accessions collected in 1992. Seed of 146 of the 184 *C. s.* var. *sativus* accessions collected during the 1992 U.S./Indian germplasm expedition were increased in 1993 and 1994. The 146 Indian accessions collected in 1992 and 46 accessions acquired directly from India prior to 1971 were obtained from NCRPIS in 1995. NCRPIS provided seed of 86, 50, and 10 accessions from Rajasthan, Madhya Pradesh, and Uttar Pradesh collected in 1992, respectively.

Since the group of 146 have not been given plant introduction numbers, they will be referred to herein by their NCRPIS number. The 46 accessions acquired before 1972 were examined by Meglic et al. (1996). Data obtained from all these Indian accessions were compared to the 707 additional accessions [*C. s.* var. *hardwickii* (4) and *C. s.* var. *sativus* (703)] previously examined by Meglic et al. (1996) and Knerr et al. (1989).

### *Sample preparation and electrophoresis*

All accessions were surveyed for isozymic variation by horizontal starch gel electrophoresis (Knerr & Staub, 1992) using a random sample of 15 to 20 plants from each accession (Meglic et al., 1996). Accessions can be identified by their sites of collection which were designated by consecutive numbers in order of the states explored (Rajasthan, Madhya Pradesh, and then Uttar Pradesh). Sites are referred to herein by village, district and state (e.g., Kalapata, Dhar, Madhya Pradesh)

*Table 1*. Description of cucumber (*Cucumis sativus* L. var. *sativus*) Group 1 accessions collected during 1992 in India. Accessions grouped by association after principal component analysis using 21 isozyme loci as framing criterion.

| Accession number[1] | Site no.[2] | Collection site descriptors | | | Comments from passport data[4] |
|---|---|---|---|---|---|
| | | Region[3] | City | District | |
| Subgroup A | | | | | |
| 20901 | 88 | M | Khandawa | Khandawa | O.P. variety Poona, Mahendra Hybrid Seed Co., Jalna, India |
| 20912 | 90 | M | Kalapata | Dhar | Landrace, 'Balam Khiri', fruit white immature and light yellowish mature, ∼ L/D = 4.3 |
| 20919 | 91 | M | Golana | Dhar | Landrace, ∼9 km north of Sodpur |
| 20922 | 92 | M | Sodpur | Dhar | Landrace, fruit green mature, ∼ L/D = 1.8 |
| 20949 | 98 | M | Kuhad | Dhar | Landrace, fruit green immature and yellowish-green mature, ∼ L/D = 2.0 |
| 20958 | 101 | M | Bagri | Dhar | Landrace |
| 20983 | 108 | M | Barwani | Khargon | Landrace, fruit green immature and red-brown netted mature, ∼ L/D = 1.6 |
| 21022 | 117 | U | Brahmanwala | Dehra Dun | Landrace, fruit green immature and yellow mature, ∼ L/D = 2.4 |
| Subgroup B | | | | | |
| 20880 | 85 | M | Ashtok | Sehore | O.P. unnamed variety, Prandey Seed Co., Faizabad, India |
| 20881 | 85 | M | Ashtok | Sehore | O.P. unnamed variety, Prandey Seed Co., Faizabad, India |
| 20902 | 88 | M | Khandawa | Khandawa | O.P. variety Barsati, Pahuja Seeds Ltd., New Dehli, India |
| 20907 | 89 | M | Gopalpura | East Nimar | Landrace, fruit green immature and white mature, ∼ L/D = 3.0 |
| 20937 | 95 | M | Dhamnod | Dhar | O.P. variety Poona Kheera, Pocha Seeds Ltd., Pune, India |
| 20945 | 96 | M | Manawar | Dhar | O.P. variety Long Khiri, source from local market |
| 20956 | 100 | M | Kunda | Dhar | Landrace, fruit green immature and greenish-yellow mature, ∼ L/D = 7.5 |
| 20964 | 101 | M | Bagri | Dhar | Landrace, fruit green immature and yellowish-green mature, ∼ L/D = 3.0 |
| 20980 | 107 | M | Rajpur | Khargon | Landrace, fruit green immature and netted brown mature, ∼ L/D = 5.0 |
| 20994 | 112 | M | Kshipra | Dewas | Landrace, fruit green immature and yellow mature, ∼ L/D = 6.0 |
| 20999 | 114 | M | Dewas | Dewas | O.P. variety Long Green, source from local seed dealer |
| 21017 | 115 | M | Londiya | Dewas | Landrace, fruit green immature and red-brown netted mature, ∼ L/D = 3.4 |
| 21022 | 117 | U | Brahmanwala | Dehra Dun | Landrace, fruit dark green immature and yellow mature, ∼ L/D = 5.0 |
| 21025 | 118 | U | Mehunwala | Dehra Dun | Landrace, fruit green, ∼ L/D = 8.0 |
| 21029 | 119 | U | Bhatta | Dehra Dun | Landrace, fruit green immature and orange netted mature, ∼ L/D = 2.3 |
| 21471 | 123 | U | Dehra Dun | Dehra Dun | Landrace, fruit dark green immature and orange mature, ∼ L/D = 4.0 |
| Subgroup C | | | | | |
| 20882 | 85 | M | Ashtok | Sehore | O.P. variety unnamed, Kisan Seed Co., Faizabad, India |
| 20898 | 86 | M | Kothi | Khandawa | O.P. variety Balam Kakri, epidermis green, mesocarp yellow to orange, ∼ L/D = 4.0 |
| 20903 | 88 | M | Khandawa | Khandawa | O.P. variety Summer Best, Pahuja Seeds Ltd., New Delhi, India |
| 20938 | 95 | M | Dhamnod | Dhar | O.P. variety unnamed, local seed dealer |
| 20965 | 102 | M | Dhar | Dhar | O.P. variety Long Green, P.N. Verma and Co., Makrandnagar, Uttar Pradesh, India |
| 20968 | 103 | M | Ratlam | Ratlam | O.P. variety Green long, Shira Seed Co., New Delhi, India |
| 20985 | 109 | M | Anjad | Khargon | Landrace, 'Bhusat', fruit white immature and white to yellow mature, ∼ L/D = 3.0 |
| 21020 | 116 | U | Dehra Dun | Dehra Dun | O.P. unnamed variety, Gagan Seed Co., New Delhi |
| 21021 | 116 | U | Dehra Dun | Dehra Dun | O.P. unnamed varitety, Parshad Seed Co., Rura, Kanpur, Uttra Pradesh |
| 21026 | 119 | U | Bhatta | Dehra Dun | Landrace, fruit green immature and orange mature, ∼ L/D = 10.0 |
| 21027 | 119 | U | Bhatta | Dehra Dun | Landrace, fruit green immature and mature, ∼ L/D = 3.5 |
| 21028 | 119 | U | Bhatta | Dehra Dun | Landrace, fruit green immature and yellow mature, ∼ L/D = 2.8 |

[1] Number given at Regional Plant Introduction Station, Ames, IA.
[2] See Figure 1.
[3] States of Madhya Pradesh = M and Uttar Pradesh = U.
[4] O.P. = open pollinated, L/D = length/diameter ratio of mature fruit.

(Table 1). Where a village bears the same name as the district, only the village is used to identify the collection site (e.g., Tonk, Rajasthan). Specific collection site information regarding sites not mentioned herein can be acquired from the expedition's trip report (Staub & McCreight, 1993).

Cotyledons were harvested from 7-day-old seedlings germinated and grown in vermiculite. Samples were bulked for analysis such that 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 TRIS-HCl at pH 7.1 (Knerr et al., 1989). Ground samples were stored at 5°C ($< 2$ h) before horizontal starch gel electrophoresis according to Knerr and Staub (1992). Modified staining solutions of Allendorf et al. (1977), Brewer (1970),

*Table 2*. Isozymic comparisons among cucumber (*Cucumis sativus* L. var. *sativus*) accessions in the U.S. NGPS collection acquired during a 1992 U.S.–India collection expedition to India, and between these accessions and those acquired by the NGPS system before 1972 using combined probability independence tests (Sokal & Rohlf, 1981).

| Enzyme | Acronym | Locus[1] | 1992 India expedition[2] | | 1992 vs. before 1972[4] | |
|---|---|---|---|---|---|---|
| | | | ln $P^3$ | Combined $P^3$ | ln *P* | Combined *P* |
| Adenylate kinase | AK | *Ak-2* | −6.91 | <0.001 | −1.51 | 0.22 |
| | | *Ak-3* | −0.02 | 0.98 | −5.12 | 0.006 |
| Fructose diphosphatase | FDP | *Fdp-1* | −3.91 | 0.02 | −6.91 | <0.001 |
| | | *Fdp-2* | −0.16 | 0.85 | −6.91 | <0.001 |
| Glucosephosphate isomerase | GPI | *Gpi* | −2.30 | 0.10 | −6.91 | <0.001 |
| Glutathione reductase | GR | *Gr* | −6.91 | <0.001 | −6.91 | <0.001 |
| Glycerate dehydrogenase | G2DH | *G2dh* | Monomorphic | | Monomorphic | |
| Isocitrate dehydrogenase | IDH | *Idh* | −2.53 | 0.08 | −6.91 | <0.001 |
| Malate dehydrogenase | MDH | *Mdh-1* | −0.92 | 0.40 | −3.61 | 0.03 |
| | | *Mdh-2* | −3.51 | 0.03 | −6.97 | <0.001 |
| | | *Mdh-3* | −1.31 | 0.27 | −6.97 | <0.001 |
| Manosephosphate isomerase | MPI | *Mpi-1* | −4.27 | 0.01 | −6.97 | <0.001 |
| | | *Mpi-2* | −1.05 | 0.35 | −6.97 | <0.001 |
| Peptidase/glycyl-leucine | PEP-GL | *Pep-gl* | −2.66 | 0.07 | −5.30 | 0.005 |
| Peptidase/leucyl-alanine | PEP-LA | *Pep-la* | −1.39 | 0.25 | −6.97 | <0.001 |
| Peptidase/phenylalanyl-proline | PEP-PAP | *Pep-pap* | −0.92 | 0.40 | −0.31 | 0.74 |
| Peroxidase | PER | *Per* | −6.81 | <0.001 | −6.97 | <0.001 |
| Phosphoglucomutase | PGM | *Pgm* | −6.91 | 0.001 | −6.97 | <0.001 |
| 6-phosphogluconate dehydrogenase | PGD | *Pgd-1* | −1.20 | 0.30 | −6.97 | <0.001 |
| | | *Pgd-2* | −0.16 | 0.85 | −1.90 | 0.15 |
| Shikimate dehydrogenase | SKDH | *Skdh* | −6.91 | 0.001 | −6.97 | <0.001 |
| Overall probability[5] | | | | | | |
| $X^2$ | | | −60.8 | | −114.4 | |
| *P* value | | | $0.025 < P < 0.01$ | | $P < 0.005$ | |

[1] Loci designated by previous examination (Knerr & Staub, 1992; Meglic, 1994) using standard criteria and nomenclature (Richmond, 1972).
[2] Comparison among 147 accessions collected from the states of Rajasthan, Madhya Pradesh, and Uttar Pradesh, India.
[3] $X^2$ test with 1 df.
[4] Comparisons between 147 accessions collected from India in 1992 and 46 accessions acquired from India before 1972.
[5] $X^2$ test with 40 df.

and Shaw and Prasad (1970) resolved reproducible isozyme banding patterns in 15 of the 51 enzyme systems evaluated (Table 2).

Gels were formulated of either 42 g or 56 g of a 1:1:1 mixture of hydrolyzed potato starch (Sigma Co., St. Louis, MO, U.S.A.), Connaught starch (Connaught Laboratories, Willowdale, Ontario, Canada), and Starch Art hydrolyzed potato starch (Starch Art, Smithsville, TX, U.S.A.) dissolved in either 300 ml or 400 ml of buffer, respectively. Gel and electrode buffers described by Allendrof et al. (1977), Clayton and Tretiak (1972), Markert and Faulhaber (1965), Ridgway et al. (1970), and Selander et al. (1971) were used (Table 2). These will be referred to as gel systems A (pH 7.1 gel buffer, 7.0 electrode buffer), C (pH 6.1 gel buffer and electrode buffer), R (pH 8.5 gel buffer, 8.1 electrode buffer), S-4 (pH 6.7 gel buffer, 6.3 electrode buffer), and S-9 (pH 8.0 gel buffer, 7.8 electrode buffer), A (pH 7.1 gel buffer, 7.0 electrode buffer), and M (pH 8.7 gel buffer, 8.7 electrode buffer). To visualize isozymic variation in malate dehydrogenase (MDH), S-4 gel buffer was adjusted to pH 6.2 and electrode buffer to pH 5.8 (Knerr et al., 1995).

### *Evaluation of germplasm*

Isozyme banding patterns observed in 15 enzyme systems and 21 mapped loci were recorded (Table 2;

Meglic & Staub, 1996). Since the *C. s.* var. *sativus* line GY14a was used as a control in this and previous studies, alleles at loci were precisely located by visual observation. Genetic nomenclature for describing allozymic variation followed a modified form (Staub et al., 1985; Knerr & Staub, 1992) previously described by Richmond (1972). Loci coding for enzymes are written with the first letter in uppercase and the rest lowercase. Enzymes coded by multiple loci 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 (Meglic and Staub, 1996). For example, an allele of *Mdh-3* with a band that migrated 2 mm faster than the most common allele was assigned the designation *Mdh-3 (2)–102*.

Allele frequency was assumed to be $p = 0.5$ and $q = 0.5$ in bulked samples of each accession. In linkage studies involving isozymes used in this study, allelic frequencies were not significantly ($P = 0.05$) different from $p = 0.5$ and $q = 0.5$ (Knerr & Staub, 1992, Meglic & Staub, 1996b). In addition, a chi-square analysis of a random sample of 10 heterozygous plant introductions (PIs) and breeding lines indicated that allelic frequencies at 10 loci were not significantly different ($P = 0.05$) than $p = q = 0.5 \pm 0.04$ (unpublished data). Thus, estimates of allelic frequencies were calculated according to the protocol of Widrlechner et al. (1992). Allelic frequency data matrices are available through GRIN.

*Analytical procedures*

Several principal component analyses (PCA) were utilized to depict affinities among accessions (objectives 1, 3 and 4) (Harris, 1975). Chi-square analyses were used to determine which isozymes were critical in the discrimination of Indian accessions (SAS Institute, 1992). Distinctive accessions were identified by their unique isozymic profiles and their relationship to other accessions after PCA (objective 2).

The correlation matrix of allele frequencies was subjected to PCA (SAS/SAT, 1992), which generates linear combinations (principal components) of the original variables (loci) which maximally discriminate among the accessions (Harris, 1975). Eigen values served to measure the cumulative portions of the total variance accounted for by each principal component. To depict genetic relationships, the first three principal components were plotted in three-dimensions with STATISTICA software (1994). The relative spatial relationship of Indian accessions collected in 1992 to those acquired before 1972 were compared. In addition, 1992 accessions and accessions collected before 1972 were compared to 707 Indian accessions grouped by continent and sub-continent reported by Meglic et al. (1996).

Isozymes useful for discrimination were identified by two-sample chi-square goodness-of-fit tests (Gibbons, 1976). The presence or absence of allozymes were compared for each accessions, and chi-square analyses were performed on allelic differences. To make an overall comparison of two populations, the results of chi-square tests were combined using the 'Combined Probabilities from Independence tests' procedure of Sokal and Rohlf (1981). This procedure allowed for the test of difference between groups of accessions collected in 1992 (objective 1) and between accessions acquired before 1972 and in 1992 (objective 3). This non-parameteric procedure combines the outcome of all frequency chi-square tests of each isozyme locus to allow for an overall significance test between populations.

Initially, groups of accessions among collections made in 1992 were identified by PCA. Differences in allelic frequencies for each isozyme were tested in each population for deviations from expected values (i.e., $p = q = 0.5$; accuracy to two decimal places). A probability of difference was determined from deviations for each isozyme in each population by consulting a standard chi-square table. The probabilities from these independent tests were then compared for each population, and a combined probability of the difference between these populations were obtained from a chi-square test (accuracy to three decimal places). Each of these combined probabilities ($P$) were converted to a natural logarithm ($\ln P$), and subsequently all $\ln P$'s within population were summed to yield an overall probability. This within population overall probability was used to accept or reject the hypothesis that groups identified among accessions collected in 1992 (objective 1) were different. This procedure was also used to determine whether genetic differences existed between accessions acquired before 1972 and those collected in 1992.

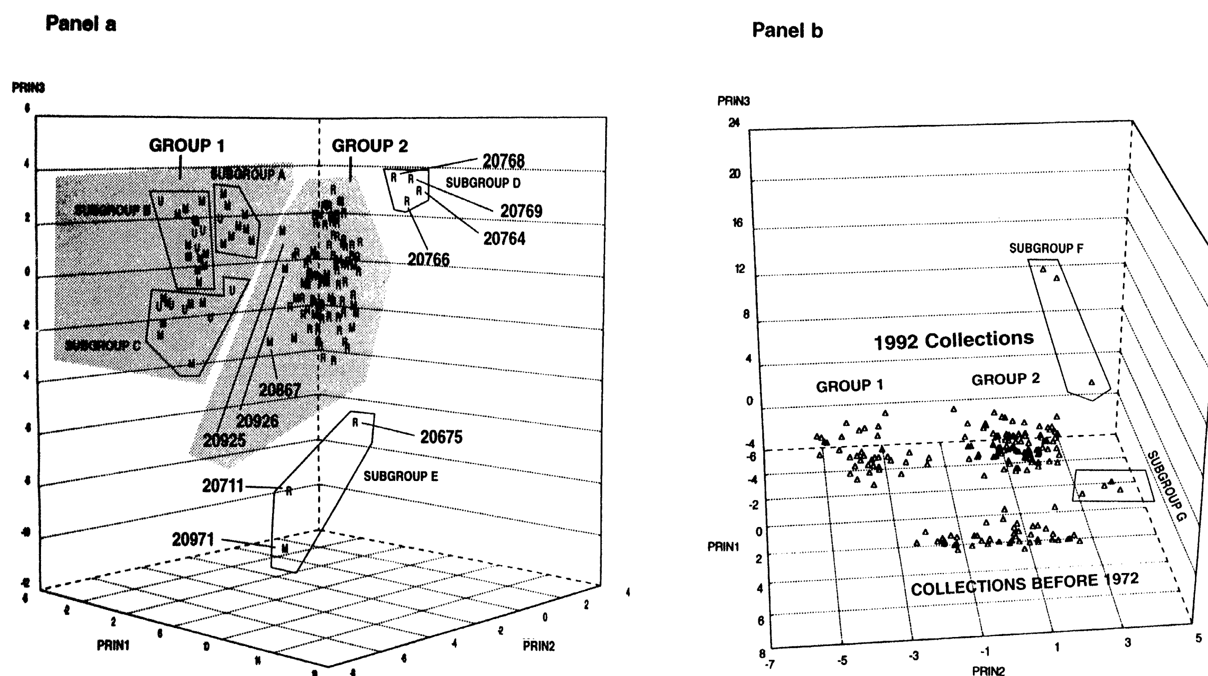



*Figure 2.* Groupings after principal component analyses of accessions of Indian origin in the U.S. National Plant Germplasm System (NPGS) cucumber collection (*Cucumis sativus* L.). Panel a depicts the proximity of 146 accessions acquired in 1992 (R = Rajasthan, M = Madhya Pradesh and U = Uttar Pradesh, India), and panel b depicts relationships of those accessions and 46 Indian accessions in the U.S. NPGS collected before 1972.

## Results and discussion

Differences among accessions resulted from presence of polymorphisms and relative differences in allelic frequencies (Table 2). Principal component analysis of allelic variation allowed for the depiction of distinct groupings of accessions collected in 1992 (Figure 2). These groupings were different from a single grouping of accessions acquired before 1972 (Meglic et al., 1996). Allelic variation in Indian accessions allowed for comparisons with other accessions in the U.S. NPGS collection grouped by continent and sub-continent.

### *Comparisons among Indian accessions collected in 1992*

Principal component analysis of Indian accessions collected in 1992 resulted in two distinct groupings of accessions (Figure 2, panel a; Group 1 and Group 2). The first three principal components accounted for approximately 46% of the observed variation, and were used to examine genetic relationships. Eight isozyme loci (*Ak-2*, *Fdp-1*, *Gr*, *Mdh-2*, *Mpi-1*, *Per*, *Pgm*, and *Skdh*) were important in elucidating these major groups (Table 2). Variation at *G2dh* was not detected. The independence tests at each locus (combined *P* values) taken collectively indicated that these groups were different (overall probability of difference = $0.025 < P < 0.01$).

Group 1 contains 37 (27 Madhya Pradesh + 10 Uttar Pradesh) accessions and Group 2 contains 102 (84 Rajasthan + 18 Madhya Pradesh) accessions (Figure 2, panels a and b). Seven accessions were not associated with either group. The relative dispersion among accessions within Group 1 was greater than that observed in Group 2. There were no accessions from Rahjasthan present in Group 1, and Group 1 contained all the accessions from Uttar Pradesh.

A previously unreported *Pgm* allele [*Pgm (3)–105*] was observed in four heterozygous [*Pgm (23)*] accessions (20764–67; Kira Ki Dhadi, Pali, site 69) from Rajasthan. Five accessions collected in another market in Pali and two accessions collected from the same market as accessions 20764–67 did not possess *Pgm (3)*. Moreover, all other Rajasthan accessions were homozygous *Pgm (22)*, and individuals possess-

ing *Pgm (3)* were not recorded in any other accession examined.

Although there were differences among accessions within Group 2, accessions from Madhya Pradesh and Rahjasthan were random and evenly dispersed within the group (Figure 2, panel a). By rotation of the graphic created by PCA, three spatially distinct subgroups (A–C) were identified within Group 1. All subgroups contained accessions from Madhya and Uttra Pradesh (Table 1). Subgroup A contained seven accessions from Madhya Pradesh and one accession from Uttar Pradesh. Twelve accessions from Madhya Pradesh and four accessions from Uttar Pradesh formed subgroup B. Subgroup C contained seven accessions from Madhya Pradesh and five from Uttar Pradesh.

All Madhya Pradesh accessions in Group 1 were collected south of Bhopal from Dewas to Pati, the southern-most collection site (Figure 1). Madhya Pradesh accessions in Group 2 were collected between New Delhi and Bhopal, from Sehor north to Bhopal and between Bhopal and Piparia (Group 2). In contrast to Madhya Pradesh accessions in Group 1, those in Group 2 were uniformly fixed at *Idh (2)*, *Mpi-1 (2)*, *Pep-gl (2)*, *Pgm (2)* and *Skdh (1)* (data not presented). In contrast, Madhya Pradesh and Uttar Pradesh accessions in Group 1 were fixed at *Pgm (1)* and *Skdh (2)*. These alleles might be associated with traits which received differing selection pressure. For instance, *Pgm (1)* is loosely linked ($\sim$14 cM) to the gene (*dm*) which conditions resistance to downy mildew [causal agent: *Pseudoperonospora cubensis* (Berk. & Curt) Wei] (Meglic and Staub, 1996).

Occasionally several collections were made at the same sampling site. Isozymic profiles of these accessions were similar in some cases (e.g., 21026–21028) and different in others [e.g., 21029 (Uttar Pradesh) vs. 21026–21028] (Table 1). In contrast, accessions collected at the same site in Madhya Pradesh often had differing isozymic profiles (data not presented). For example, at site 85 (Ashtok, Sehor, Madhya Pradesh) five collections were made (20878–20882). Allelic differences among these accessions resulted in their placement into different major groups by PCA [i.e., Group 1 (20880–20882) and Group 2 (20878 and 20879)]. These open-pollinated accessions were produced by different seed companies [20878 (Vanayak Seeds, Kanpur), 20879 (Prandey Seed Co., Faizabad), 20880 (Prandey Seed Co.), 20881 (Prandey Seed Co.), and 20882 (Kisan Seed Co., Faizabad)]. These and other data (not presented) indicate that commercial cucumber varieties possessing detectable genetic variability could be acquired by farmers at some local seed distribution sites.

Genetic diversity was apparent among landraces collected in the same region. For instance, the isozymic profiles of samples of two landraces collected in Sodpur (site 92), Dhar, Madhya Pradesh (20922 and 20925) were different. Accession 20922 in Group 1, subgroup A (Table 1) and 20925 positioned between Groups 1 and 2 (Figure 2, panel a) were acquired from the same farmer. An open-pollinated variety, Green Long (20926; Rama Krishan Seeds, New Delhi), collected at an adjacent site (site 93; Nalcha, Dhar) and a landrace (20867) collected at site 81 (Piparia, Hoshangabad, Madhya Pradesh) were also placed between Groups 1 and 2. These accessions share isozymic similarities with both groups. The isozymic profile of a second landrace collected at site 81 (20868; Group 2) differed considerably from that of 20867. In contrast, the accession 20864 collected at site 80 (Tajpura, Bhopal, Madhya Pradesh) possesses considerable biochemical affinities with 20867.

The collections made in Chhatarpur, Madhya Pradesh provide an additional example of the existence of genetic diversity among landraces within the same geographic area (Figure 1). While the isozymic profiles of 20832 (Atar; site 74), 20837 (Chhatarpur; site 76), and 20840 (Chhatarpur; site 77) were similar, they differed from that of 20835 (Chhatarpur; site 75) (data not presented). Although 20832 was received from a local farmer, 20837 (unnamed variety; Phuja Seeds Ltd., New Delhi) and 20840 ('Jyoti Green Long'; Shiv Seed Co., New Delhi) were acquired from two seed distributors in Chhatarpur. These dealers indicated that seed of some local landraces were increased in other regions of India and then returned for sale in Chhatarpur. The seed distributor which sold accession 20835 indicated that accession 20832 was selected from 20835. This indicates that 20832 and/or similar landraces grown in the Chhatarpur-Atar area could be a source used by seed companies to produce accessions 20837 and 20840.

The isozyme profiles of four landraces (20764, 20766, 20768, 20769) collected in Rajasthan (Figure 1) were similar, but distinct from other accessions from Rajasthan and Madhya Pradesh in Group 2 (Figure 2, panel a). All accessions were collected in a local market in Kira Ki Dhandi, Pali (site 69), and were designated as subgroup D. By comparison, fruit of these accessions were smaller [length: diameter (L/D) ratio = $\sim$2.8] than fruit of 10 landrace collections (20753–20762; L/D = $\sim$6.0) obtained at an adjacent local

market within the same area (Charbjhuja Sadri, Pali, site 69) (Figure 1). The isozymic profiles of accessions 20753–20762 resulted in their placement as a group towards the center of Group 2 (Figure 2, panel a). These and other examples not presented indicate that genetic similarities among cucumber landraces in India can exist across moderate distances (i.e., 15 to 20 km), and that landraces acquired at the same collection site may be different.

The isozyme profiles of two Rahjasthan accessions (20675 and 20711) and one Madhya Pradesh accession (20971) were distinct from either major group (Figure 2, panel a). These were designated as subgroup E. The landrace 20711 bares relatively long fruit ($\sim$30 cm) and was collected in Girwar, Sirohi (site 43) (Figure 1). It possessed biochemical affinities with the landrace 20971 collected in Bidbada, Ratlam (site 104). The landrace 20675 was collected in Banthali, Tonk (site 33), and possessed an isozymic profile which differed from three other landraces (20676–20678; Group 2) collected at the same site. The isozymic profiles of accessions 20676–20678 are similar. All accessions at this site were acquired from different farmers within a village which had been growing these landraces for at least 40 years. Seed of each landrace had been passed through generations within a family. Thus, the case in Banthali represents a condition where landraces were cultivated in close geographic proximity, but are genetically distinct. The fact that the L/D of 20675 ranges between 0.9 to 1.0 (i.e., short and blocky fruit) and the L/Ds of 20676–20678 range between 2.6 to 3.8 supports this contention.

*Accessions possessing unique isozymic profiles*

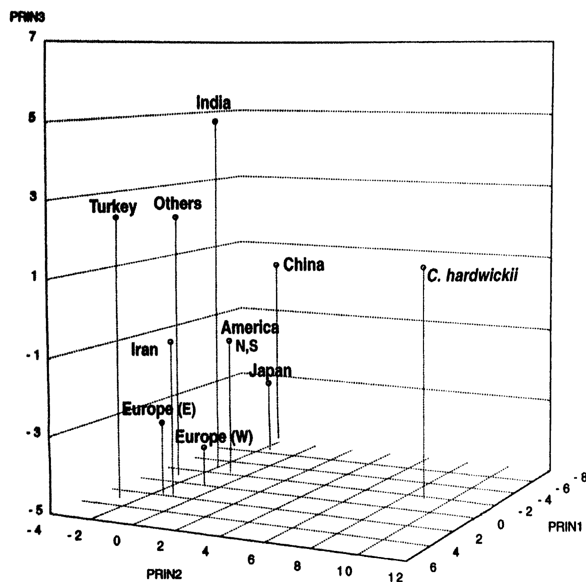
A combination of allelic variation at differing loci resulted in the identification of accessions possessing unique isozyme profiles (Table 2). The isozymic profiles of accessions 20675, 20764, 20766, 20768, 20769, 20711, and 20971 differ from all other new India accessions (data not presented). Their uniqueness is a result of a lack of allelic variation at the loci examined when compared to that of other Indian accessions. Accession 20675 is fixed at all loci, except *Fdp-1*, *Per* and *Mdh-2*, 20764 is fixed at all loci, except *Mdh-3* and *Pgm*, 20766 is fixed at all loci, except *Mdh-1* & *-2*, Pgm, 20768 is fixed at all loci, except at *Ak-2*, and *Mdh-3*, 20769 is fixed at all loci, except for *Ak-3* and *Mdh-3*, 20711 is fixed at all loci, except *Fdp-1* and *Per*, and 20971 is fixed at all loci, except for *Fdp-2*, *Mdh-1*,*-2* & *-3* and *Pep-la*. In contrast, the accessions 20664 (Tonk, Rahjasthan), 20666 (Jaipur, Rahjasthan), 20872 (Sehore, Madhya Pradesh), 20881 (Ashtok, Sehore, Madhya Pradesh), 21026 (Bhatta, Dehra Dun, Uttar Pradesh) are heterozygous for at least nine loci. Thus, it appears that most of the genetic variation for new Indian accessions lies among accessions within Groups 1 and 2.

*Relationships among Indian cucumber accessions in the U.S. NPGS*

Principal component analysis identified three distinct groups of Indian accessions; two groups of accessions collected during 1992, and a relatively disperse group containing accessions obtained before 1972 (Figure 2, panel b). The first three principal components explained 47% of the variation present in the data set.

The isozymic profiles of accessions within each group were distinct (Table 2). Significant differences ($p \leq 0.05$) were detected between accessions collected in 1992 and those received before 1972 at most enzyme loci, which was reflected by an overall comparison using all polymorphic loci ($P < 0.005$). All loci were important in the detection of this difference, except *Ak-2*, *Pep-pap*, and *Pgd-2*.

Genetic diversity among Indian accessions acquired before 1972 group was identified as differences depicted in the first principal component (Figure 2, panel b). For instance, the isozymic profiles of PI 163216 (collected by W. Koelz, Belwani province, 1948) and PI 164743 (collected by W. Koelz, 1948; collection site unknown) place these accessions at the extremes of this group (Figure 2, panel b). Although there were changes in the proximity of groups and subgroups, the relative relationships among Indian accessions collected in 1992 (Figure 2, panel a) remained generally the same. For instance, two subgroups were identified (F & G; Figure 2, panel b) which contained accessions possessing isozymic profiles were considered distinct in another analysis (Figure 2, panel a, subgroups D and E). Subgroup F contains the same accessions grouped as subgroup E (Figure 2 panel, a), and were differentiated by the third principal component. Accessions in subgroup G (Figure 2, panel b) included 20675 (site 33; Banthali, Tonk, Rajasthan), 20711 (site 43; Girwar, Sirohi, Rajasthan), and 20971 (site 104; Bidbaba, Ratlam, Madhya Pradesh) (Figure 1). These accessions which were separated from other accessions by the second



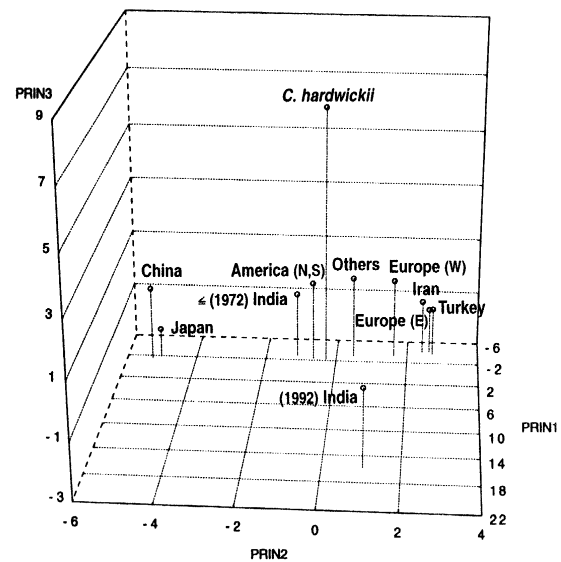




*Figure 3.* Groupings (pooled data by geographic proximity and collection date) after principal component analyses of *Cucumis sativus* L. accessions in the U.S. National Plant Germplasm System (NPGS). Panel a depicts proximity of 707 accessions acquired before 1992 according to Meglic et al., 1996, and panel b depicts relationships of those accessions and 146 Indian accessions obtained in 1992 and 46 accessions acquired before 1972.

principal component and were identical to those accessions in subgroup D (Figure 2, panel a).

### *Relationships of Indian accessions with other cucumber accessions in the U.S. NPGS*

A previous study of variation at 21 isozyme loci (Meglic et al., 1996) determined that genetic differences exist among 753 *Cucumis sativus* accessions in the U.S. NPGS when they were grouped by botanical variety, country or subcontinent (Figure 3, panel a). The four *C. s.* var. *hardwickii* accessions examined were genetically heterogeneous, but were distinct from *C. s. var. sativus* accessions. We used a data set containing the same accessions used by Meglic et al. (1996) plus Indian accessions collected in 1992 to determine whether the previously observed genetic relationships would remain constant. Although differences between the *C. s. var. hardwickii* and *C. s. var. sativus* accessions were evident, the relationships among *C. s. var. sativus* accessions grouped by country or subcontinent differed from Meglic et al. (1996) (Figure 3, panel b).

There is a constant need to conserve genetic resources because plant breeding efforts depend on a continuing and expanding supply of genetic resources. Because cucumber has a relatively narrow genetic base, threats of genetic erosion are of major concern. Genetic erosion of cucumber germplasm was documented during the 1992 U.S./India expedition. For instance, attempts were made to collect in a major cucumber production in northern Rajasthan near Ganganargar (Figure 1). This region was reported to possess cucumber genotypes tolerant to environmental stresses (i.e., water and heat; personal communication, U.C. Srivastava, National Bureau of Plant Genetic Resources, New Delhi, 1992). This region experienced an extended period of drought between 1987 to 1991. Neither seed dealers nor several local farmers were able to provide seed of local landraces because remnant seed had been consumed by farmers during the drought period. Although samples of landraces were obtained from sites at Maraksar, Suratgargh (site 11; 20530), Suratgarh, Ganaganagar (site 13; 20533), and Pilibanga, Ganganagar (site 13; 20539), collections made at some these sites could not be regenerated (e.g., Suratgarh, Ganaganagar), and germplasm once characteristic of the northern-most region of Rajasthan is no longer available.

Such observations suggest specific instances of genetic erosion, but do not provide information regard-

ing the genetic diversity present in India, the center of origin of cucumber (Leppick, 1966). Data from this study provide an estimation of genetic diversity in three Indian states. Indian landraces collected in 1992 from these states were genetically variable, and isozymic profiles differed among accessions within and between collection sites (Figure 2, panels a and b). Although the relative genetic diversity among landraces in some regions of India was limited, isozymic profiles of these accessions were different from Indian accessions received by the U.S. NPGS before 1972. Thus, the 1992 U.S.-Indian expedition captured unique germplasm that contributed significantly to broadening the genetic base of the U.S. NPGS cucumber collection.

Since the release of the first gynoecious line, 'MSU 713-5 (Peterson, 1960), private and public breeders worldwide have concentrated on the production of F1 hybrids for both market and processing cucumbers. Although hybrids provide for a more concentrated uniform fruit set, mass introduction of F1 hybrids can lead to genetic uniformity and thus crop vulnerability. In India, the production and use of F1 varieties has not received wide-spread acceptance. Distribution of commercial F1 cucumber hybrids to remote areas of India is extremely rare. Nevertheless, in some areas local Indian seed distributors indicated that the use of open-pollinated varieties produced by Indian seed companies was increasing. This was particularly evident in Ganganargar where drought had resulted in the elimination of local landraces. Our data indicate that the commercial, open-pollinated accessions examined were genetically diverse when compared to local land races. Thus, it appears that the use of commercially produced, locally-adapted, open-pollinated varieties by Indian growers does not currently contribute substantially to reduced genetic diversity and crop vulnerability in the sites surveyed.

## Conclusion

Genetic variation in cucumber accessions from India was assessed by examining accessions collected during a 1992 U.S.–India expedition, and by comparing this variation with that detected in accessions aquired by the U.S. NPGS before 1972 and an array of U.S. NPGS accessions of diverse origin. Identification of genetically distinct groups of accessions collected in 1992 indicates that variation exists among accessions from different Indian states. Ethnobotanical information obtained from tribal chiefs and local seed dealers was used to explain the origin of unique genetic differences among landraces. Moreover, the genetic differences detected between accessions obtained before 1972 and those collected in 1992 indicates that genetic diversity in the U.S. NPGS collection was significantly enhanced by the addition of the 1992 accessions. The unique genetic nature of the 1992 accessions was confirmed by comparison to a diverse array of U.S. NPGS accessions. This assessment suggests that isozyme analysis is an effective tool for germplasm characterization in cucumber, and that future cucumber collection expeditions to India might be worthwhile. A joint U.S./China government-sponsored *Cucumis* collection expedition was conducted in 1994. Since China is a secondary region of diversity for cucumber and germplasm of Chinese origin has not been freely available since the late 1940's, it would be useful to compare the genetic diversity of accessions recently obtained in China with those now available in the U.S. NPGS.

## References


Allendorf, F.W., N. Mitchell, N. Remain & G. Stahl, 1977. Isozyme loci in brown trout (*Salmon trutta* L.): detection and interpretation from population data. Hereditas 86: 179–190.

Brewer, G.B., 1970. An Introduction to Isozyme Techniques, p. 186. Academic Press, New York.

Clayton, J.W. & D. N. Tretiak, 1972. Amine-citrate buffers for pH control in starch gel electrophoresis. J. Fish. Res. Board Can. 29: 1169–1172.

Dane, F., 1976. Evolutionary studies in the genus *Cucumis.* Ph.D. Diss., 202 pp. Colorado State Univ., Fort Collins.

Dane, F., 1983. Cucurbits. In: S.D. Tanksley & T.J. Orten (Eds). Isozymes in Plant Genetics and Breeding, Part B, 472 pp. Elsevier, Amsterdam.

Deakin, J.R., G.W. Bohn & T.W. Whitaker, 1971. Interspecific hybridization in *Cucumis*. Econ. Bot. 25: 195–211.

Esquinas-Alcazar, J.T., 1977. Alloenzyme variation and relationships in the genus *Cucumis.* Ph.D. Diss., 170 pp. Univ. of California, Davis.

Frankel, O.H., 1972. The significance, utilization and conservation of crop genetic resources. FAO, Rome.

Gibbons, J.D., 1976. Nonparametric Methods for Quantitative Analysis, p. 463. Holt, Rinehart and Winston. New York.

Harlan, J.R., 1975. Our vanishing genetic resources. Science 188: 618–621.

Harris, R.J., 1975. A Primer of Multivariate Statistics, p. 332. Academic Press, New York.

Horst, E.K. & R.L. Lower, 1978. *Cucumis hardwickii*: a source of germplasm for the cucumber breeder. Cucurbit Genet. Coop. Rpt. 1: 5.

Jeffrey, C., 1980. A review of the Cucurbitaceae. Bot. J. Linnean Soc. 81: 233–247.

Kirkbride, Jr. J.H., 1993. Biosystematic Monograph of the Genus *Cucumis* (Cucurbitaceae), p.159. Parkway Publishers, Boone, North Carolina.

Knerr, L.D., J.E. Staub, D.J. Holder & B.P. May, 1989. Genetic diversity in *Cucumis sativus* L. assessed by variation at 18 allozyme coding loci. Theor. Appl. Genet. 78: 119–128.

Knerr, L.D. & J.E. Staub, 1992. Inheritance and linkage relationships of isozyme loci in cucumber (*Cucumis sativus* L.) Theor. Appl. Genet. 84: 217–224.

Knerr, L.D., V. Meglic & J.E. Staub, 1995. A fourth malate dehydrogenase (MDH) locus in cucumber. HortScience 30: 118–119.

Leppick, E.E., 1966. Searching gene centers of the genus *Cucumis* through host parasite relationship. Euphytica 15: 323–328.

Markert, C.L. & L. Faulhaber, 1965. Lactate dehydrogenase isozyme patterns of fish. J. Exp. Zool. 159: 319–332.

Meglic, V., F. Serquen & J.E. Staub, 1996. Genetic diversity in cucumber (*Cucumis sativus* L.): I. A reevaluation of the U.S. germplasm collection. Genet. Res. Crop Evol. 43: 533–546.

Meglic, V. & J.E. Staub, 1996. Inheritance and linkage relationships between allozyme and morphological loci in cucumber (*Cucumis sativus* L.). Theor. Appl. Genet. 92: 865–872.

Peterson, C E., 1960. A gynoecious inbred line of cucumber. Mich. Agri. Expt. Stat. Quart. Bull. 43: 40–42.

Richmond, R.C. 1972. Enzyme variability in the *Drosophila williston* group. 3. Amounts of variability in the superspecies *D. paulistorum.* Genetics 70: 87–112.

Ridgway, G.J., S.W. Sherburne & R.D. Lewis, 1970. Polymorphism in the esterases of Atlantic herring. Trans. Amer. Fish. Soc. 99: 147–151.

SAS Institute, 1992. SAS/STAT user's guide. Release 6.03 edition, p. 1028. SAS Institute, Cary, NC.

Selander, R.K., M.H. Smith, S.Y. Yang, W.E. Johnson & J.B. Gentry 1971, Biochemical polymorphism and systematics in the genus *Peromyseus.* I. Variation in the old-field mouse (*Peromyseus polionotus*). In: Studies in genetics. Univ. of Texas, Austin.

Shaw, C.R. & R. Prasad, 1970. Starch gel electrophoresis of enzymes – a compilation of recipes. Biochem. Genet. 4: 297–320.

Sokal, R.R. & F.J. Rohlf, 1981. Biometry. 2nd edn. W.H. Freeman, New York.

STATISTICA, 1994. StatSoft, Inc., Tulsa, OK. p. 638.

Staub, J.E., C.E. Peterson, L.K. Crubaugh & M.J. Palmer, 1993. Cucumber population WI 6383 and derived inbreds WI 5098 and WI 5551. HortScience 27: 1340–1341.

Staub, J.E. & J.D. McCreight, 1993. *Cucumis* germplasm expedition, India: October–November, 1992. Report to USAID and OICD, Washington, DC.

Staub, J.E., R.S. Kupper, D. Schumann & T.C. Wehner, 1985. Electrophoretic variation and enzyme storage stability in cucumber. J. Amer. Soc. Hort. Sci. 110: 426–431.

United States National Archives. Plant collection itineraries and trip reports. GEIL and GEIL files. U.S. National Archives Library, Washington, DC.

Widrlechner, M.P., L.D. Knerr, J.E. Staub & K.R. Reitsma, 1992. Biochemical evaluation of germplasm regeneration methods for cucumber, *Cucumis sativus* L. FAO/IBPGR Plant Genet. Resources Newsletter 88/89: 1–4.

Whitaker, T.W. & G.N. Davis, 1962. Cucurbits: Botany, cultivation, and utilization. Interscience Publishers, New York.