Theor Appl Genet (1984) 69:39–45

Multivariate analysis applied to tomato hybrid production

S. Balasch [1], F. Nuez [1], G. Palomares [1] and J. Cuartero [2]

[1] Departamento de Genética, Universidad Politécnica, Camino de Vera, 14, Valencia-22, Spain
[2] Estación Experimental La Mayora, Algarrobo, Málaga, Spain

Received March 7, 1984; Accepted April 15, 1984
Communicated by R. Hagemann

Summary. Twenty characters were measured on 60 tomato varieties cultivated in the open-air and in polyethylene plastic-house. Data were analyzed by means of principal components, factorial discriminant methods, Mahalanobis D^2 distances and principal coordinate techniques. Factorial discriminant and Mahalanobis D^2 distances methods, both of which require collecting data plant by plant, lead to similar conclusions as the principal components method that only requires taking data by plots. Characters that make up the principal components in both environments studied are the same, although the relative importance of each one of them varies within the principal components. By combining information supplied by multivariate analysis with the inheritance mode of characters, crossings among cultivars can be experimented with that will produce heterotic hybrids showing characters within previously established limits.

Key words: *Lycopersicon esculentum* – Tomato hybrids – Polycross – Genetic distance – Multivariate analysis

Introduction

Genetic diversity among parentals is considered as an important factor for obtaining heterotic hybrids (Moll et al. 1962; Hawkins et al. 1965; Khanna and Chaudhary 1974; Chandra 1977). Geographic diversity has been used as a not-too-adequate index of genetic diversity (Timothy 1963; Murty and Arunachalam 1966). More recently, electrophoretic analysis, which distinguishes the different biochemical characteristics of the varieties, has been employed to determine the level of genetic differences among them (Tanksley and Rick 1980). Although these analyses determine the possible genic variability with accuracy, they are scarcely practical for the breeder, who is mainly interested in quantitative polygenic characters which depend greatly on genic interactions. Another complementary way of approaching the problem is by means of multivariate analysis techniques on a sufficient number of morphologic characters.

The principal aim of this work is the use and comparison of different multivariate techniques in classifying an important number of tomato varieties. A simple multivariate technique for variety classification is principal components analysis, which has been greatly discussed because it acts on average character values, and therefore gives the same weight to each character regardless of inter- and intra-variety distributions of each character. In this work, this technique is compared with other more precise ones which take the different variability of characters into account. They are: factorial discriminant analysis and calculation of Mahalanobis D^2 distances.

Material and methods

Sixty tomato varieties from the USA, Canada, Holland, Germany, Spain, Russia, Nigeria, England, Poland and Argentine were used (Table 1). Cultivation took place in the open-air and in polyethylene plastic-houses. The interest of trying out two environmental conditions lies in their possible influence on the experiment conclusions through interactions genotype-environment.

Characters evaluated were:

– Number of fruit/plant – Total production/plant – Average fruit weight – Earliness: estimated according to two types of indexes:

a) Temporal earliness index (TEI), defined as the harvest, or a fraction of it, over which 30% of the total production in weight

Table 1. Tomato varieties studied and abbreviations used

Americano A (AA) [a]	Harold 12088 (H12088) [a]	Muchamiel (Mu) [a]	Stamm 1280 (S1280)
Atkinson (Atk)	Healani (H)	Novedad (N)	Sub Artic Cherry (SAC)
Atom (At)	Heinz 1360 (H1360)	Oxheart (O)	Sub Artic Maxi (SAM) [a]
Aurora (Au)	Hellfrucht Frühstamm (He) [a]	Pakmor V. F. (P) [a]	Sub Artic Plenty (SAP)
Busch (B) [a]	Kalaohi (K)	Pearson Al (Pe)	Super Roma (SR) [a]
Cal Ace (CA)	L-7 (L-7)	Piervil (Pi) [a]	Supersonic (Su) [a]
Cal J (CJ)	L-16 (L-16)	Platense (Pl)	Tiny Tim (TT)
Campbell 28 (C28)	L-20 (L-20)	Ponderosa Pink (PP)	Valenciano (V) [a]
Cliaton (Cl)	Madrigal (M)	Porter (Po)	Vesset (Ve) [a]
Cuarenteno (C)	Manalucie (Mn) [a]	Red Top (RT) [a]	V.F. 105-2 (VF)
Early Pak 7 (EP7) [a]	Marglobe (Mrg) [a]	Resaplús N.V.F. (R) [a]	V.F. 145-Gus (VFG)
Floradel (F) [a]	Marmande (Mrm)	Royal Ace (RA) [a]	W. Virginia 36 (WV36) [a]
Fortuna c (Fc) [a]	Marroqui (Mr)	Rutgers (Ru)	W. Virginia 63 (WV63)
Futuro Bird (FB)	Money Maker (MM) [a]	Salomé (S)	W. Virginia 700 (WV700)
Grosse Fleisch (GF)	Moss (Mo)	Severianín (Se) [a]	75/79 (75/79)

[a] Varieties with available data per plant

(TEI_Y) or in fruit number (TEI_N) would theoretically be reached.

b) Earliness production index (EPI), corresponding to production achieved either in weight (EPI_Y) or in fruit number (EPI_N), up to a previously established harvest. In this experiment the harvest chosen was the one over which approximately 30% of total production for most varieties was achieved.

Both index types have different significances. The TEI index shows the relative production distribution over the productive cycle of the plant independent that such production is high or low. EPI index estimates early production, therefore depending on production relative distribution and on total production.

– Vegetative characters: number of leaves up to first cluster, length between clusters (in the intervals 1st to 2nd, 2nd to 3rd and 3rd to 4th) and number of leaves between clusters (in the same three above mentioned intervals).

– Characteristics of the fruit, corresponding to a series of measurements on the third fruit of the second cluster, which is considered as representative of the plant and variety according to studies carried out by UPOV (1975). These characteristics are: fruit weight, polar and equatorial diameter, number of locules, ribbing, pointed, hollowness and cracking. The latter four mentioned characteristics, which are qualitative, were controlled by numeric punctuation for the smaller or greater incidence of the character.

– Per one plant showing green-back fruit on the total amount of plants controlled per variety.

– Abnormal fruit. In the open air three measurements were considered: per one plant with one or two, three or more, and with at least one, abnormal fruit. This fruit is produced by flowers with an abnormal number of petals and stamens, and its characteristic is to lenthen and curve transversally, providing an easily recognizable shape.

All these characters were measured on a maximum of 20 plants for each variety.

In polyethylene plastic-houses, data were collected plant by plant for 23 varieties in order to be able to carry out factorial D^2 distances. This allowed the comparing of principial components analysis with the other two techniques. Principal coordinates analysis is used as multidimensional scaling technique for graphic representation of Mahalanobis distances.

Results and discussion

Principal components analysis under either environment produces only small differences regarding the components formed. In plastic-houses, the three principal components explain 72% of the total variability, while in the open air four components are necessary to explain 69%. Such components are not equivalent one by one, as their interpretation suffers slight changes according to the environment. The most remarkable differences lie in the interchange between vegetative characters and number of fruit as principal integrants in components 1 and 2. Similarly, the third axis or component, in plastic-houses summarizes in the 3rd and 4th, ones in the open air, mainly influenced by TEI and EPI respectively. Nevertheless, the principal explicative characters are the same under both environments, and moreover, graphic representation of the characters on the planes formed by different axis points out some character associations, which shows their similar discriminant capacity. Therewith, by trying to choose the most easily measurable characters among such associations, the following ones can be catalogued as most important for a tomato variety definition: one measurement of fruit size (e.g. weight of the third fruit of the second cluster), one vegetative characteristic of the plant (e.g. average number of leaves between clusters) and number of fruit per plant, TEI_N and EPI_N. Cavicchi and Giorig (1976) draw similar conclusions except for non-inclusion of the character fruit number, although it must be pointed out that these authors worked on a much more reduced number of varieties.

Varieties representation on the axis formed by such principal components leads to grouping varieties similarly under both environments. Under plastic-house conditions 15 variety groups were differentiated (Fig. 1). In the open-air (Fig. 2), although extreme situations preserve perfectly, some proximate varieties intercross. Logically, among more proximate varieties, sample effects and interactions genotype environment will produce a slightly different type of associations for each environment. In the mentioned representations

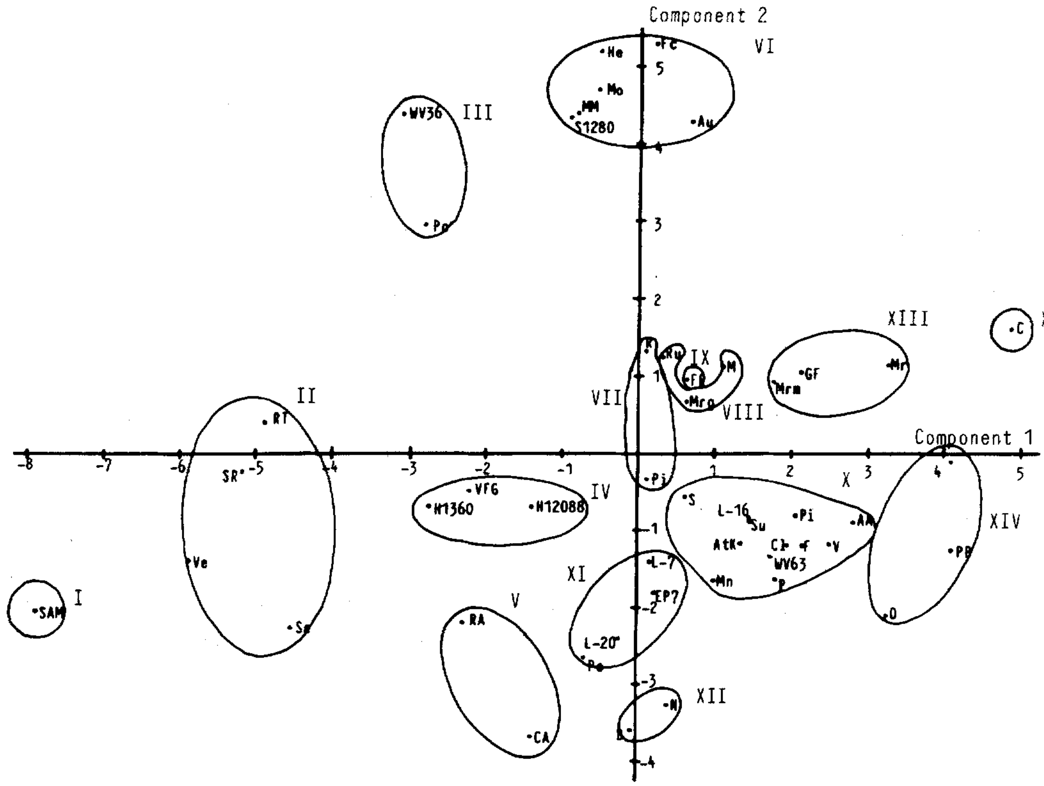

Fig. 1. Representation of varieties on the first two components of principal components analysis. Cultivation under plastic-house

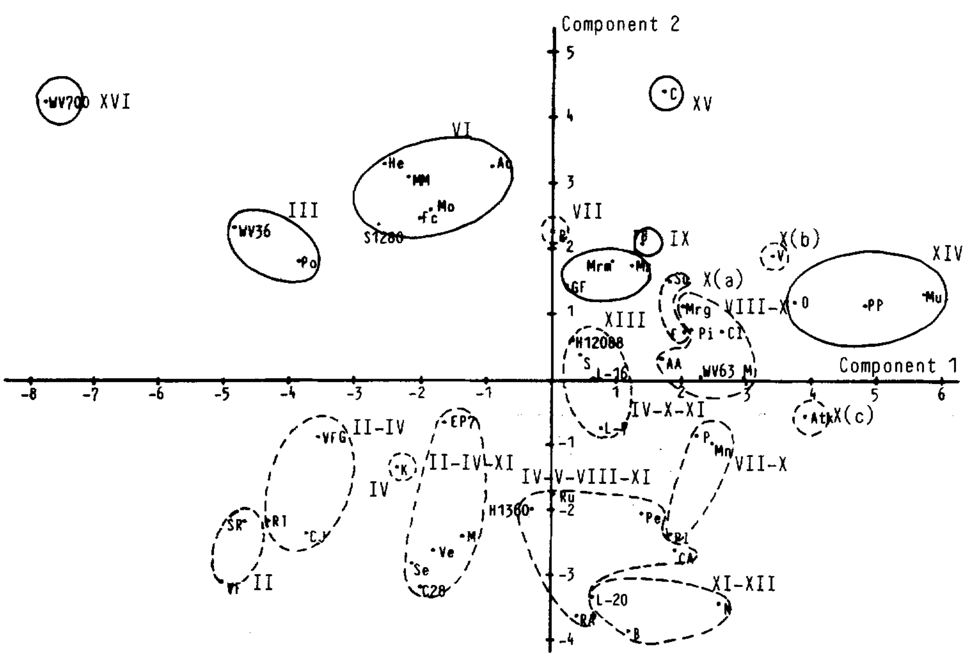

Fig. 2. Representation of varieties on the first two components of principal components analysis. Open-air cultivation

the absence of the varieties 'Cal J,' 'Healani', 'Atom', 'Sub Arctic Plenty', 'Kalaohi', 'West Virginia 700', 'Sub Arctic Cherry', '75/59', 'Tiny Tim', 'Campbell 28' and 'V.F. 105-2', under plastichouse conditions and 'Sub Arctic Maxi', 'Atom', 'Sub Arctic Plenty', 'Sub Arctic Cherry', '75/59' and 'Tiny Tim' in the open air, can be noted. The reason is not the disposition of vegetative characters data of the same. By carrying out another principal components analysis on all varieties after eliminating vegetative characters, it is possible to obtain a relative localization of such new varieties with regards to the rest. Furthermore, bearing in mind that all such new varieties have a rather low and strong determinated habitat, except for 'West Virginia 700', they can be situated approximately in the following areas:

– Plastic-house: 'Sub Arctic Cherry', 'Atom', '75/59', 'Sub Arctic Plenty' and 'Tiny Tim', in group I. 'Cal J', 'Healani', 'Kalaohi', 'Campbell 28' and 'V.F. 105-2' between groups II and IV, and 'West Virginia 700', near group III.

– Open-air: 'Sub Arctic Maxi', 'Atom', 'Sub Arctic Plenty', 'Sub Arctic Cherry', '75/59' and 'Tiny Tim', near the groups II and II–IV.

To carry out an homogeneous comparison among multivariate techniques for variety classification, factorial discriminant analysis, calculation of Mahalanobis distances and principal component analysis were carried out under plastic-house conditions with the 23 varieties marked by an asterix on Table 1.

Regarding factorial discriminant analysis, although the discriminant canonical functions are not comparable one by one to factors obtained by means of principal components analysis, the characters with greater explicative importance coincide. These are the same as the ones previously enumerated for the principal components analysis on the 60 varieties, except that for discriminant analysis production also seems to be important, mainly for the third discriminant function.

Graphic representations of varieties according to both methods (Figs. 3 and 4) show the great similarity of associations between both analysis methods. These associations moreover correspond with the ones obtained by principal components analysis on all varieties found plastic-house. It can therefore be asserted that the minor precision which principal components analysis evidently offers, with regards to factorial discrimi-

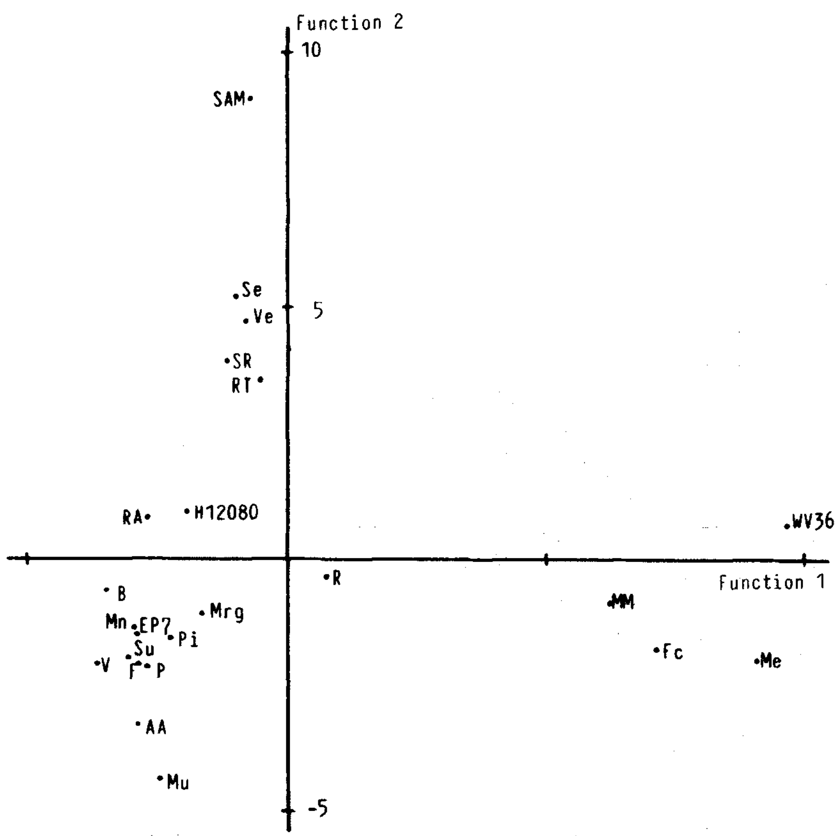

Fig. 3. Representation of 23 varieties on the first two canonical functions of discriminant factorial analysis. Cultivation under plastic-house

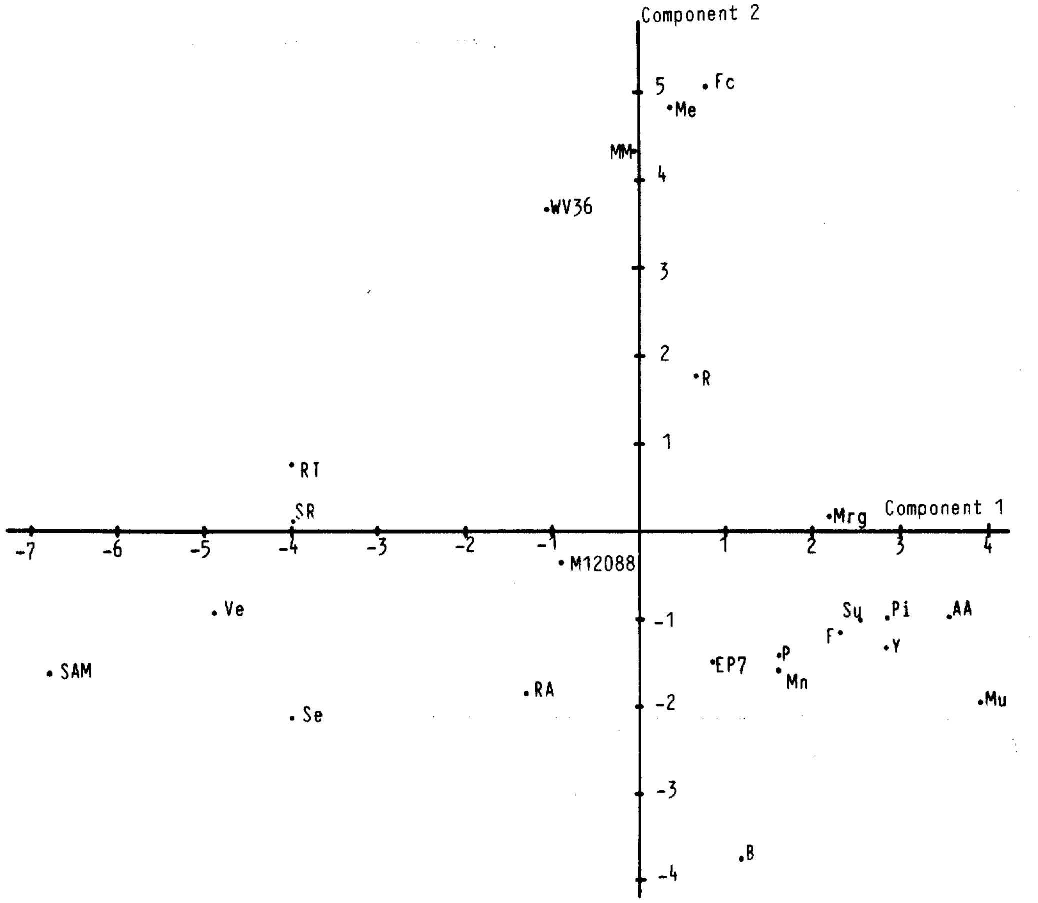

Fig. 4. Representation of 23 varieties on the first two components of principal components analysis. Cultivation under plastic-house

Table 2. Mahalanobis distances among parental lines belonging to polycross group 1. Cultivation under plastic-house

	W. Virginia 36	Vesset	S.A. Maxi	Resaplús	Early Pak 7	Royal Ace	Pakmor	Hellfrucht	Muchamiel	Severianín	Valenciano	Busch	Supersonic	Harold 12088	Red Top	Super Roma	Piervil	Floradel	Manalucie	Marglobe	Fortuna c	Money Maker	Americano A
W. Virginia 36	0	182	241	175	226	209	219	89	244	206	229	213	221	208	164	181	212	234	228	184	99	121	208
Vesset	182	0	30	44	54	36	61	155	116	17	86	57	64	31	28	39	64	61	59	57	118	101	82
S.A. Maxi	241	30	0	107	134	89	143	230	221	33	164	128	140	89	63	71	132	138	127	131	190	168	173
Resaplús	175	44	107	0	44	39	39	92	62	52	70	69	36	24	53	69	40	27	34	33	51	44	54
Early Pak 7	226	54	134	44	0	25	19	157	46	73	39	25	28	29	62	74	27	25	21	34	117	116	31
Royal Ace	209	36	89	39	25	0	25	167	81	43	52	37	21	29	52	54	26	23	11	26	128	108	39
Pakmor	219	61	143	39	19	25	0	156	51	83	34	37	22	32	79	93	31	16	20	38	120	115	26
Hellfrucht	89	155	230	92	157	167	156	0	182	169	202	188	172	153	167	188	148	153	163	155	19	35	173
Muchamiel	244	116	221	62	46	81	51	182	0	119	56	40	56	57	99	113	77	38	60	69	124	139	42
Severianín	206	17	33	52	73	43	83	169	119	0	101	61	81	41	43	49	75	67	68	81	130	111	103
Valenciano	229	86	164	70	39	52	34	202	56	101	0	43	28	59	62	69	38	38	37	47	160	146	31
Busch	213	57	128	69	25	37	37	188	40	61	43	0	37	40	64	72	42	32	38	50	147	148	27
Supersonic	221	64	140	36	28	21	22	172	56	81	28	37	0	39	58	66	17	19	9	13	125	105	13
Harold 12088	208	31	89	24	29	29	32	153	57	41	59	40	39	0	47	53	44	24	29	32	104	99	48
Red Top	164	28	63	53	62	52	79	167	99	43	62	64	58	47	0	10	74	78	64	52	116	100	78
Super Roma	181	39	71	69	74	54	93	188	113	49	69	72	66	53	10	0	82	86	67	57	145	124	88
Piervil	212	64	132	40	27	26	31	148	77	75	38	42	17	44	74	82	0	20	19	28	117	103	24
Floradel	234	61	138	27	25	23	16	153	38	67	38	32	19	24	78	86	20	0	14	33	112	105	23
Manalucie	228	59	127	34	21	11	20	163	60	68	37	38	9	29	64	67	19	14	0	21	120	100	27
Marglobe	184	57	131	33	34	26	38	155	69	81	47	50	13	32	52	57	28	33	21	0	109	93	26
Fortuna c	99	118	190	51	117	128	120	19	124	130	160	147	125	104	116	145	117	112	120	109	0	14	128
Money Maker	121	101	168	44	116	108	115	35	139	111	146	148	105	99	100	124	103	105	100	93	14	0	126
Americano A	208	82	173	54	31	39	26	173	42	103	31	27	13	48	78	88	24	23	27	26	128	126	0

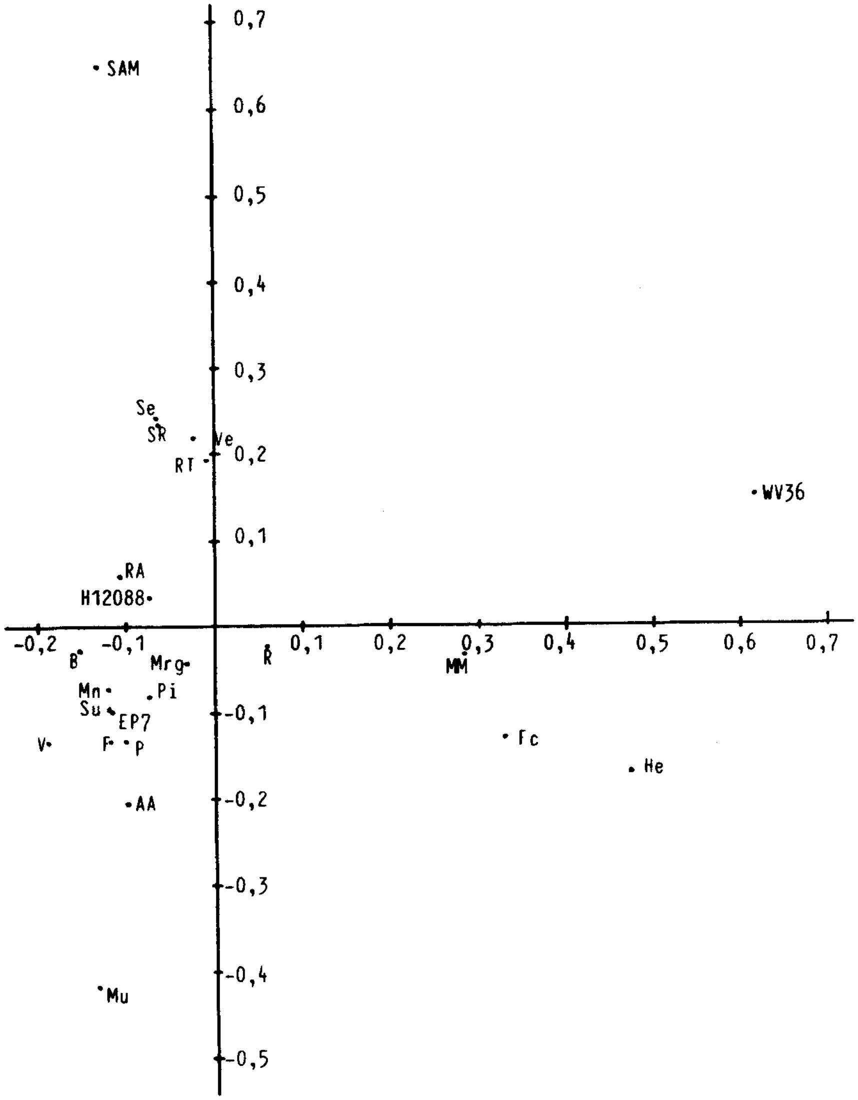

Fig. 5. Representation of Mahalanobis distances of 23 varieties on principal coordinates. Cultivation under plastic-house

nant analysis, does not modify conclusions according to our experience.

Mahalanobis distances (Table 2) are coherent with associations formed according to the other two methods (principal components analysis and factorial discriminant analysis). A principal coordinates analysis on such distances allows a sinthetic and interpretable representation of relationships among varieties to be obtained (Fig. 5).

Summarizing, it seems that variety groups obtained according to the different multivariate techniques are, in a great part, independent from the same. Nevertheless, possibly because there are interactions between genotype and environment, the structure of some groups changes according to whether cultivation takes place in plastic-houses or in the open-air.
It can also be pointed out that the genetic divergence of the varieties is in no way influenced by geographic distribution. In this experiment, there are various examples of different origin varieties classified within the same group: 'Vesset' (Canada) and 'Severianin' (Russia), in group II; 'Hellfrucht' (Germany) and 'Aurora' (USA), in group IV; 'L-16' or 'Valenciano' (Spain) and 'Atkinson', 'West Virginia 63' or 'Pakmor' (USA), in group X; 'L-7' or 'L-20' (Spain) and 'Early Pak 7' or 'Pearson' (USA) in group XI; 'Muchamiel' (Spain) and 'Ponderosa Pink' (USA) in group XIV.

On the other hand, it is convenient to underline the fact that two varieties, such as 'Royal Ace' and 'Cal Ace', where great genetic proximity was known previously, were associated in both cultivation forms.

For many species, heterosis or hybrid vigour is directly related to genetic distance among parentals. The tomato is no exception to this rule (Khanna and Chaudhary 1974; Khanna and Misra 1977; Rajanna et al. 1977; Peter and Rai 1978). Therefore, knowing the relative genetic distances among

Table 3. Possible heterotic crossings conditioned by fruit weight

Average weight level of fruit	Combinations to be experimented
60–100 g	'Harold 12088' × 'Hellfrucht' 'Piervil' × 'Hellfrucht'
100–140 g	'Vesset' × 'Muchamiel' 'Severianín' × 'Muchamiel' 'Hellfrucht' × 'Valenciano' 'Fortuna c' × 'Valenciano' 'Money Maker' × 'Muchamiel' 'Money Maker' × 'Valenciano' 'Money Maker' × 'Busch'
100–180 g	'Piervil' × 'Muchamiel'
>180 g	'Royal Ace' × 'Muchamiel'

varieties, it is possible to plan crossings between parentals in a more scientific way. The mentioned heterosis shows up fundamentally in such complex characters such earliness or total production (Williams 1959; Palenzona and Grillo 1970; Kaul et al. 1972; Dhillon et al. 1975; Lobo and Marin 1975; Babu 1978; Peter and Rai 1978; Singh et al. 1978), owing to which fact it might be convenient to choose genetically very distant varieties for the crossing, but yet conditioned to certain interesting breeding objectives that could be affected by the type of crossing. One of these objectives could be the fruit weight, for which a partially negative dominant inheritance mode seems to be well established (Ibarbia and Lambeth 1969; Butler 1973; Brandolini et al. 1974; Nandpuri and Tyagi 1976; Cuartero and Cubero 1981).

In this way, and parting from average varietal fruit weights, it is possible to have an idea of the most heterotic combinations, according to the level wanted to be obtained for such restrictive character. In this example, considering the four levels for fruit weight, crossings that could be determined in any sense as presumably more heterotic, are presented in Table 3. Some of these crossings were carried out showing a satisfactory agreement with predictions. This same operation could be carried out establishing other characters as improvement aims.

References

Babu RY (1978) Studies on heterosis in tomato (*Lycopersicon esculentum*, Mill.). Mysore J Agric Sci 12:676–677

Brandolini A, Desiderio E, Svanosio A, Vandoni GC (1974) Breeding tomatoes for fresh consumption. Ital Agric 3:95–115

Butler L (1973) Fruit size inheritance in the tomato. Can J Genet Cytol 15:655

Cavicchi S, Giorgi G (1976) Yield in tomato. 2. Multivariate analysis on yield components. Genet Agrar 30:315–326

Chandra S (1977) Comparison of Mahalanobis's method and metroglyph technique in the study of genetic divergence in *Linum usitatissimum*, L. Germ plasm collection. Euphytica 26:141–148

Cuartero J, Cubero JI (1981) Genetics of four economically important characters in tomato (*Lycopersicon esculentum*, Mill.). Z Pflanzenzücht 87:330–338

Dhillon GS, Nandpuri KS, Gupta VP (1975) Heterosis in tomato (*Lycopersicon esculentum*, L.). J Res 12:258–264

Hawkins BS, Pracock HA, Ballard WW (1965) Heterosis and combining ability in upland cotton-Effect on yield. Crop Sci 5:543–546

Ibarbia EA, Lambeth VN (1969) Inheritance of tomato fruit weight. J Am Soc Hortic Sci 94:498–500

Kaul DL, Nandpuri KS, Singh S (1972) Hybrid-vigour studies in tomato. J Res 9:19–26

Khanna KR, Chaudhary RC (1974) The nature of gene action and combining ability for some vegetative characters in tomato. Euphytica 23:159–165

Khanna KR, Misra CH (1977) Divergence and heterosis in tomato. SABRAO J 9:43–50

Lobo AM, Marin VO (1975) Heterosis and combining ability in tomato (*Lycopersicon esculentum*, Mill.). Rev Inst Colomb Agropecu 10:1–10

Moll RH, Salhuan WS, Robinson HF (1962) Heterosis and genetic diversity in variety crosses of maize. Crop Sci 2: 197–209

Murty BR, Arunachalam V (1966) The nature of genetic divergence in relation to plant breeding system in crop plants. Indian J Genet 26A:188–198

Nandpuri KS, Tyagi DD (1976) Inheritance of some characters in tomato. J Res 13:80–84

Palenzona DI, Grillo G (1970) Quantitative characters and heterosis in the tomato. Genet Agrar 24:121–128

Peter KV, Rai B (1978) Heterosis as a function of genetic distance in tomato. Indian J Genet Plant Breed 38:173–178

Rajanna A, Gulshan Lal, Peter KV (1977) Heterozygote advantage as a function of genetic divergence in tomato. Indian J Agric Sci 47:434–437

Singh B, Narendra Kumar, Joshi S (1978) Hybrid vigour in tomato (*Lycopersicon esculentum*, Mill.). Prog Hortic 10: 20–23

Tanksley SD, Rick CM (1980) Isozymic gene linkage map of the tomato: applications in genetics and breeding. Theor Appl Genet 57:161–170

Timothy DH (1963) Genetic diversity, heterosis and use of exotic stocks in Columbia. In: Proc Statist Genet Plant Breed Symp, pp 581–591

Williams W (1959) Heterosis and the genetics of complex characters. Nature 184:527–530