GENETIC BASE BROADENING IN TOMATO
(*Lycopersicon esculentum* Mill.)

Gabriel Saavedra Del Real

PhD Thesis

The University of Edinburgh

August 2001

Abstract

Genetic variability in many modern crops is very limited because of bottlenecks during domestication and past selection pressures. This narrow genetic base has resulted in a lack of genetic variability in some crops, and increasing the susceptibility to biotic and abiotic stresses, and may limit responsiveness to market needs. Tomato (*Lycopersicon esculentum* Mill.) is one of many autogamous crop species in which the exploited germplasm has been severely reduced as a result of the process of domestication, and particularly because the initial germplasm used to generate much of the material exploited in current varieties, represented a very small fraction of the initial variability available. The concept of genetic base broadening has been suggested as a means of mitigating this lack of diversity in modern crops, with the aim to utilise the rich genetic resources available in wild relatives, vintage varieties, and landraces. Genetic base broadening programmes involve the systematic utilisation of an arrangement of genetic variability in such a way as to generate a mass of newly adapted gene stocks available as parents in breeding programmes.

This research examines options available within a genetic base broadening programme, limited by space and time. Different populations were created by hybridisation in order to examine options and feasibility within a base broadening programme. These included a study of the genetic diversity of the genus *Lycopersicon,* using 43 accessions of different taxa to examine the level of genetic variability in tomato, and the richness of diversity available in wild relatives and vintage/landrace tomato cultivars. Hybridisation was conducted as part of genetic base broadening programme to create inter-taxon and intra-taxon crosses between selected tomato cultivars and wild relatives. As part of possible strategies, double crosses between inter-taxon populations were tested and analysed. The created populations were selfed and examined using morphological and molecular markers for polymorphism, genetic distances and heterozygosity indices from genetic population analysis computational program packages Popgene and NTSYS.

Results are presented for these populations over a number of generations and reviewed against possible strategies for conservation and utilisation of this sample of

populations for future breeding programmes. Results showed that there is large genetic diversity at morphological and molecular level between and within *Lycopersicon* taxa. *L. esculentum* presented limited genetic diversity within the accessions examined, and a narrow genetic base. However, substantial sources of genetic diversity are available to incorporate into the cultivated tomato from both wild relatives and old varieties and landraces of the cultivated species.

After hybridisation, the created populations did not follow the expectation of autogamous crops, and revealed only a tendency toward decreasing genetic variability in further generations. The F_1 generation behaved as expected, for both morphological and molecular markers, but in F_2 and F_3 generations, the results fluctuated from increasing to decreasing values for all indices examined. However, from the data obtained it was possible to theorise about the number of parents to be involved, and the created population size that should be used in genetic base broadening programmes, along with strategies for the conservation of the created genetic variability.

The methods utilised in this project, morphological and molecular markers, gave valuable information about the genetic diversity in self-pollinating generations. However, morphological characters were more limited than molecular markers in respect to information accuracy, because of the number and type of traits selected. The sample size affected both type of markers. From the genetic indices utilised, average gene diversity (H_s), total gene diversity (H_t), and effective number of alleles (A_e) were more informative than the arbitrary mean proportion of polymorphic loci (P) and number of polymorphic alleles (A). However, all indices had some merit and usefulness in analysing the data obtained in this research.

For the future, it is hoped to use the experience gained with *Lycopersicon* spp utilising morphological and molecular markers in order to answer some more of the questions that will arise in any genetic base broadening programme.

Dedication

This work is dedicated to my wife Rossana, my children Ignacio and Maria Jose, and my parents.

Declaration

I declare that all the work in the thesis was done purely by myself except where identified.

Edinburgh, January 2002

Acknowledgements

I wish to express my sincere gratitude to many people who helped me in completing this thesis. In particular, I would like to thank the following people whose assistance was invaluable to me.

My supervisors Dr. W. Spoor and Dr. L. Harrier for their encouragement, help, criticism during the experimental work, analysis of data, and writing up this thesis.

Dr. M. Möller of Royal Botanic Garden Edinburgh for his friendship, interest, help, support, and for expending hours discussing and analysing this work.

Dr. H. Parzies of University of Hohenheim for his friendship, support and constant discussions about subjects involved in the thesis.

The technical staff of Crop Science Department of the Scottish Agricultural College especially Jeanette Taylor, Pamela Muntch, Linda Patterson, Robert Redpath, Adrian Thompson, and Helen Kearney.

Especial acknowledge to Instituto de Investigaciones Agropecuarias and CRI La Platina for the financial assistance through a studentship during these years. Some words of thankful to the former CRI La Platina Director Mr. Daniel Claro and former Head of the Horticulture Department Mr. Moises Escaff, because they believed in me and made me possible to accomplish this task.

Finally, I would like to thank my family for their constant support, encouragement and everything they did for me.

Table of contents

Table of tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Table of figures

Chapter 3

Table of plates

Chapter 1

Chapter 3

Chapter 4

List of Abbreviations

A	Adenine
A	Number of polymorphic alleles per locus
A_e	Effective number of alleles
AMOVA	Analysis of molecular variance
ANOVA	Analysis of variance
C	Cytosine
CGN	Centre for Genetic Resources
CpDNA	Chloroplast DNA
CPRO-DLO	Centre for Plant Breeding and reproduction Research
CTAB	Hexadecyltrimethylammonium bromide
Cvr.	Cultivar
DNA	Deoxyribonucleic acid
DNTP	Deoxynucleotide triphosphate (Adenine, Cytosine, Guanine and, Thyamine)
DTT	Dithiothreitol
EDTA	Ethylen diaminetetraacetic acid
FAO	Food and Agriculture Organisation
F_{st}	Fixation index
G	Guanine
GEM	Germplasm enhancement of maize
HCl	Hydrochloric acid
HOPE	Hierarchical open-ended population enrichment
H_s	Average gene diversity
H_t	Total gene diversity
I	Shannon's information index
INIA	Instituto de Investigaciones Agropecuarias
IPGRI	International Plant Genetic Resources Institute

IPK	Institute für Pflanzengenetik und Kulturpflanzenforschung
KCl	Potassium chloride
MDS	Multidimensional scale
$MgCl_2$	Magnesium chloride
MtDNA	Mitochondrial DNA
NTSYS	Numerical Taxonomical System
OP	Open pollinated
P	Probability
P	Proportion of polymorphic loci
PAUP	Phylogenetic Analysis Using Parsimony
PCA	Principal Components Analysis
PCR	Polymerase Chain Reaction
PGU	Plant Growth Unit
PVP	Polyvinylpyrrolidone
QTL	Quantitative Trait Loci
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
SAC	Scottish Agricultural College
SE	Standard Error
SSR	Simple Sequence Repeat
STR	Short Tandem Repeats
T	Thymine
TBE	Tris-borate EDTA
TGRC	Tomato Genetics Resources Center
UPGMA	Unweighted Pair-group Method with Arithmetic Average
USDA-ARS	United States Department of Agriculture-Agricultural Research Service
UV	Ultra-violet

Chapter 1

Introduction, background and rationale

1.1 General introduction

In most crop species, as a result of domestication and subsequent breeding/selection processes exploited populations represent a small fraction of the variability available. The lower overall genetic diversity of modern cultivars of autogamous species may also reflect genetic "bottlenecks" to which these species have been subjected. This may be because of natural phenomena such as polyploidy, mating and dispersion systems, and geographical barriers, or during their introduction to new regions away from the centre of origin. In some cases only a limited number of seeds/propagules (or accessions) were carried back by explorers, and this has served as the essential base of modern cultivars today. Over the intervening years many genotypes have been lost as a result of the disappearance of old varieties and landraces, and their replacement by new more productive varieties, apparently more adapted to biotic and abiotic stresses of these localities. More recently, commonly used breeding techniques such as backcross, pedigree selection, or hybrid production have been effective, in terms of producing new varieties with highly prescribed characteristics, but this has been obtained at the further expense of genetic diversity (Sneep, 1979; Rick, 1987; Miller and Tanksley, 1990).

During the period of scientific breeding, utilisation of the available genetic diversity has been poor; for years plant breeders have confined their programmes to a relatively small part of the overall genetic resources (Gepts, 1993; Kannenberg and Falk, 1995) and plant breeders have depended to a large extent on the recycling of a limited gene-pool (Berg, 1997). However, over a similar time frame, a great amount of germplasm such as wild relatives, old varieties, landraces, and other breeding material has been stored in genebanks; this is a potentially valuable, but relatively under exploited, source of genetic variability. Only a small amount of this variability has been introgressed into crop species, and then usually aiming to solve a specific problem (most frequently disease resistance) involving a single or a few genes. Most plant breeders are very constrained and cannot afford to work with germplasm that even temporarily dilutes agronomic performance or quality (Kannenberg and Falk, 1995), since working "elite" germplasm is easily disrupted by crosses with

unimproved "exotic" germplasm from landraces and wild relatives (Tanksley and Nelson, 1996). Even though there are good reasons to diversify breeding sources, the fact remains that breeding progress continues in most crops, albeit at variable and sometimes slowing rate, and breeders must develop cultivars that meet the standards of highly competitive markets of today. In addition, exploitation of heterogeneity and crop evolution in farmers' fields are outside the scope of most, particularly commercial plant breeding research (Berg, 1997).

Concern about this perceived lack of genetic diversity and the resulting genetic vulnerability of our food plants has led researchers and policy makers to assess the situation in different crops (Simmonds, 1993; Van Beuningen and Busch, 1997). Genetic base broadening is one approach which has been suggested, in the Leipzig Agreement (FAO, 1996), as a means of providing a viable sustainable genetic base from which varieties can be selected either directly or following hybridisation with existing and currently exploited genetic base of a crop species.

1.2 Genetic base broadening

1.2.1 Definitions

Genetic base broadening has been defined as 'composite crosses' (Suneson, 1956), 'incorporation' (Simmonds, 1993) and 're-synthesis' (Becker *et al.*, 1995). However, regardless of the term used, the definition *per se* has been the same. Genetic base broadening is the incorporation and re-synthesis of populations from wild relatives, landraces and/or old varieties into relatively new varieties or accessions, with the aim of enhancing the ability to respond to biotic and abiotic stresses in future breeding generations. Genetic base broadening involves the systematic utilisation of an arrangement of genetic variability in a manner likely to generate a mass of newly adapted genotypes to be made available as source material in breeding programmes (Simmonds, 1993).

Strictly, base broadening should be without preconceived aims, partly because it is difficult to predict future requirements, and partly because such preconceptions may influence the construction of the initial populations. The intention must be to create

populations that have an enhanced ability to respond to any local need. These populations, selected for local adaptation, would contribute to the sustainability of agricultural systems and be an immediate source of variability in the case of unexpected local environmental changes (biotic or abiotic).

Garanko (1991) commented that broad-based genetic materials are essential to meet a number of breeding objectives. Breeders can no longer be dependent on the basic stock of cultivars which they inherited from their predecessors. Furthermore, previously unexploited germplasm can perhaps lead to the discovery of new developmental pathways and ecological adaptations that may be important to meet the needs of changing agronomic practices (Kannenberg and Falk, 1995). Wild species and primitive cultivars present valuable initial material to turn into the breeding programs, and the number and diversity of original ancestors can provide insight into the relative genetic diversity within and among populations (Van Beuningen and Busch, 1997). Kannenberg and Falk (1995) argue that diversification of a crop breeding base must be through introgression of new germplasm via meritorious cultivars or lines that are from different genetic backgrounds but competitive with commercial germplasm. Both the potential for long term genetic gain and the reduction of genetic vulnerability may depend, in part, on the initial genetic diversity present in the genetic base of the crop (Van Beuningen and Busch, 1997).

Introgression has been usually described as backcrossing new genes controlling desired characters into adapted accessions (Cooper *et al.*, 1998; Ortiz, 1998). However, the methodology only allows a few foreign genes to be introduced at one time (Simmonds, 1993). Single major genes are usually the objectives of 'Introgression' programmes, but transfer of undesired genes is likely to be greater if the transfer is from more distantly related wild species, owing to the high diversity of the alleles (Carver and Taliaferro, 1992). This approach has been used mostly for major disease resistant genes, but some polygenic inherited stress responses and quality traits have also been transferred (Cooper *et al.*, 1998).

'Incorporation' creates new genetic stocks, where a variety of new allelic combinations can be expected. It does not emphasise specific gene transfer, as in

introgression, but requires rigorous characterisation and evaluation of the phenotypes generated. In general, phenotypes of exotic materials may not provide any useful or usable guide to breeding directly, but probabilities of broader gene combinations are higher (Rick and Chetelat, 1995; Ortiz, 1998). Incorporation is more powerful than introgression for broadening the genetic base, but demands the assurance of a long-term programme utilising approaches that are more population-oriented than gene- or character-oriented (Cooper *et al.*, 1998).

1.2.2 Genetic base broadening efforts in crops

The narrowness of the genetic base has been established in many crops (Hawkes, 1979; Sneep, 1979; McClean and Hanson, 1986; Miller and Tanksley, 1990; Simmonds, 1993; Dubreuil and Charcosset, 1998). However, only rarely has recognition of this problem resulted in deliberate breeding effort to broaden the genetic base available to the breeders. Frequently, techniques such as traditional introgression and even more sophisticated approaches involving gene transfer from other species, along with the technologically demanding approaches for improving selection efficiency within a narrow gene pool (marker assisted selection) have been proposed as means to maintain or increase the productivity of new cultivars. These 'quick fix' solutions may continue for sometime to create more productive cultivars, but it is yet to be proven that these are sustainable approaches in the long term.

The Global Plan of Action and Leipzig Declaration by FAO (FAO, 1996a) recognised that there was currently a need for genetic enhancement for many crops now and this group could only be expected to increase in the future. As a result of Leipzig Agreement two expert workshops were supported by FAO and IPGRI to discuss broadening the genetic base of crops, the first Rome in 1997 and the second in Edinburgh (UK) in 1999. This was followed by a special forum on base broadening at the 3rd International Crop Science Congress 2000 in Hamburg. Plans have been established to develop a web-based forum, both to encourage general developments, but also to alert interested groups in progress in crop species, or countries for linkages and funders. Therefore, FAO has given considerable support to these actions, because this methodology has the potential to be one of the most

environmentally benign of agricultural technologies. The search for genetic resistance or tolerance to biotic and abiotic stresses could decrease the use of the many contaminant and pollutant products used in modern agriculture. The result of this search could be a more sustainable and environmentally friendly agriculture.

Examples of deliberate base broadening activities, although relatively rare, can be found in a range of crops. In outbreeding species such as maize (*Zea mays* L.), tropical germplasm was adapted to conditions in southern maize-growing regions in the USA (Goodman, 1985); Salhuana *et al.* (1993) reported a national project for Germplasm Enhancement of Maize (GEM) in the USA; Kannenberg and Falk (1995) designed a breeding system for maize called HOPE - Hierarchical Open-ended Population Enrichment. In sorghum (*Sorghum bicolour* (L.) Moench), Ethiopian and Sudanese landrace germplasm was successfully incorporated into adapted Indian cultivars (Mengesha and Rao, 1982). In clonal crops, usually outbreeders, the narrow gene pool of potato (*Solanum tuberosum* L.) has been enhanced by the creation of *neotuberosum* populations from wild relatives from the Andigena Group (Gledinning, 1979; Plaisted, 1982; Mendoza, 1989; Simmonds, 1993). In Cassava (*Manihot esculenta* Crantz) (Nassar, 1989) and sugarcane (*Saccharum officinarum* L.) (Chave, 1991; Simmonds, 1993) there have also been researches enhancing the genetic base. In the case of inbreeding crop species such as barley (*Hordeum vulgare* L.) there have been approaches for broadening the genetic base, such as the incorporation of exotic germplasm into barley breeding pools in the Nordic Region (Lehmann, 1991; Vetelainen *et al.*, 1996), and the recurrent introgressive population enrichment (RIPE) in Canada (Kannenberg and Falk, 1995). Becker *et al.* (1995) reported re-synthesis research in oil seed rape (*Brassica napus* L.). Other autogamous crops have also been studied for genetic base broadening such as oat (*Avena sativa* L.) (Frey, 1994), soyabean (*Glycine max* (L.) Merr.) (Wang, 1994), rice (*Oryza sativa* L.)(Li *et al.*, 1994), and Arabica coffee (*Coffea arabica* L.) (Agwanda and Awuor, 1989). Several other researchers have suggested the use of incorporation in other crops using wild relatives and landrace populations as sources of genetic variability. Thus, Ahmad *et al.* (1996) proposed an incorporation programme in cultivated lentils (*Lens culinaris* ssp *culinaris*) because of the very limited variability

found in this crop. Such a programme has also been suggested in tomato (*Lycopersicon esculentum* Mill.) by Garanko (1991) and Rick and Chetelat (1995); and in common bean (*Phaseolus vulgaris* L.) by Welsh *et al.* (1995).

For the future, the introduction of novel, favourable alleles and gene combinations from wild relatives and other sources into gene pools and the efficient international exchange of germplasm may both contribute to broadening the genetic base to maximise genetic gains and reduce genetic vulnerability (Van Beuningen and Busch, 1997).

1.3 Tomato as a model for genetic base broadening

Tomato has been selected as a model within autogamous crops because it is an amenable crop, easy to cross between species; it has a narrow genetic base; it has a large number of wild relatives; landraces exist and finally it is possible to obtain at least two generations per year under artificial experimental conditions.

1.3.1 Tomato as crop

Tomato (*Lycopersicon esculentum* Mill.) is one of the most important vegetable crops in the world, representing 20.66% of the vegetable production in the world in 1998 with 89 million tonnes (FAO, 1999). The main producer countries are China and the USA; however in respect to yield, the Netherlands and UK are the highest with 466,667 and 283,333 kg per hectare respectively, though in these countries production is under protection (plastic tunnels and glasshouses). In comparison, the world average yield reaches just 28,343 kg per hectare; and countries such as the USA (65,063 kg/ha), Chile (63,430 kg/ha), and Australia (44,944 kg/ha) are able to produce higher yields from very substantial areas of field grown crops. In area, China and India are the largest tomato growing countries with 539,000 and 350,000 ha respectively; and globally in 1998 there were 3.1 million hectares under tomato cultivation (FAO, 1999).

The main uses for tomato are as fresh fruit and for processing. For fresh consumption, tomato is grown either in the open field or under protected conditions. The use of greenhouses and other protection systems are common practice in

northern latitudes, while field growing is the most common approach in developing countries. Processing tomatoes are all grown under field conditions; to be dried or canned, or made into juice or paste.

1.3.2 Botanic and taxonomic classification

Lycopersicon is a relatively small genus within the large and diverse family *Solanaceae*, which consists of around 90 genera (Table 1.1). Within the sub-family *Solanoideae*, *Lycopersicon* belongs to the largest tribe, *Solaneae*. This tribe consists of around 18 genera containing the genus *Lycopersicon* and the closely related genus *Solanum* (Hogenboom, 1979). Rick (1979a) considered that there are profound differences between *Lycopersicon* and *Solanum* in terms of cytogenetic evolution, and morphological/physiological differentiation, therefore separation of the two genera is justified. However, latest systematic research, using molecular sequencing techniques, indicated that *Lycopersicon* has evolved from within a paraphyletic genus *Solanum* (Spooner *et al.*, 1993; Olmstead and Palmer, 1997; Knapp and Spooner, 1999). Thus, *Lycopersicon* species could be regarded as belonging to the genus *Solanum*, but the systematic treatment of *Solanum sensu stricto* will be applied for the present study, and the genus *Lycopersicon* utilised.

Table 1.1 Taxonomic classification of *Lycopersicon* (NCBI, 2001)

Kingdom	*Plantae*
Division	*Tracheophyta*
Sub-division	*Spermatophytina*
Class	*Angiospermae*
Sub-class	*Dicotyledoneae*
Order	*Solanales*
Sub-order	*Solanineae*
Family	*Solanaceae*
Sub-family	*Solanoideae*
Tribe	*Solaneae*
Genus	*Lycopersicon*

All species in the genus *Lycopersicon* are typical of the *Solanoideae* sub-family, each having an identical chromosome number (2n=2x=24), regular flowers, compressed seeds and curved embryo (Taylor, 1986).

Müller (1940) quoted by Taylor (1986) subdivided the genus *Lycopersicon* based on fruit colour: *Eulycopersicon* (red-fruited species such as *L. esculentum*, *L. esculentum* var. *cerasiforme*, *L. cheesmanii*, and *L. pimpinellifolium*) and *Eriolycopersicon* (green-fruited species such as *L. hirsutum*, *L. pennellii*, *L. parviflorum*, *L. chmielewskii*, *L. peruvianum*, and *L. chilense*), but this division is arbitrary and does not correspond to more fundamental differences between species. Rick (1976) divided the genus into species that can be easily crossed with the edible tomato, *esculentum-complex* and those which cannot, *peruvianum-complex* (Rick, 1979a; Warnock, 1988; Miller and Tanksley, 1990) (Table 1.2).

Table 1.2 *Lycopersicon* classification according to Rick (1976)

Esculentum-complex	***Peruvianum-complex***
L. esculentum Mill.	*L. chilense* Dun.
L. esculentum var. *cerasiforme*(Dun.) Gray	*L. peruvianum* (L.) Mill.
L. pimpinellifolium (Jusl.) Mill.	*L. peruvianum* var. *humifusum* Mull.
L. cheesmanii Riley	
L. cheesmanii var. *minor* (Hook. F.) Porter	
L. parviflorum Rick, Kes., Fob. & Holle	
L. chmielewskii Rick, Kes., Fob. & Holle	
L. hirsutum Humb. and Bonpl.	
L. hirsutum var.*glabratum* (Mull.) Luckwill	
L. pennellii (Corr.) D'Arcy	
L. pennellii var. *puberulum* (Corr.) D'Arcy	

1.3.3 Morphology and pollination biology

Lycopersicon esculentum is a tropical or sub-tropical plant that has been adapted to a wide range of environments far different from its original home. The species is perennial by nature, but it is almost universally cultivated as an annual.

Tomato plants are characterised by a herbaceous growth habit, odd pinnate leaves, yellow coloured corolla and anthers, and soft edible berry fruits. Tomato flowers develop in cymes and flowering is centrifugal; flowers are abracteate and hermaphrodite; the pistil is enveloped by stamens forming the characteristic of this genus, a flask-shaped anther cone.

Genus *Lycopersicon* has a very unusual pattern of anther dehiscence: anthers split laterally and the split occurs soon after or during the corolla opening. Pollen is released inside the anther cone and emerges through the communal channel formed by the joining of each elongated anther.

Tomato crop plants are usually autogamous, but in regions with excessive activity of pollinating insects, about 10 to 15% natural cross pollination occurs (Taylor, 1986). Rate of success of effective pollination is influenced by temperature and relative humidity, the optimum ranging from 22 to 28°C and 70 to 85% humidity (Stevens and Rick, 1986).

The fruit is a berry, in wild species it is 2-celled, each cell being regular in shape and with a somewhat dry placenta. In cultivated forms, the cells can be numerous, irregular in size and outline, with the placenta markedly succulent. Each fruit contains many seeds, approximately disc-shaped and covered with hairs. (Hector, 1936; Taylor, 1986; Kaul, 1991).

1.3.4 Tomato crop: origin and evolution

The tomato is a relatively new crop. The oldest records date back less than 400 years, a brief time when compared with the oldest available records of many other crops (Boswell, 1937; Smartt and Simmonds, 1996). It has been suggested by several publications that the original site of domestication was Mexico (Jenkins, 1948; Rick, 1958; Hawkes, 1991; Villand *et al.*, 1998), where the ancient people called it

"Tomatl". The cherry tomato (*Lycopersicon esculentum* var. *cerasiforme*) is almost certainly the direct ancestor of the modern cultivated form (Boswell, 1937; Rick and Forbes, 1975; McClean and Hanson, 1986; Taylor, 1986; Hanson and McClean, 1987; Hawkes, 1991; Villand *et al.*, 1998), but Bretó *et al.* (1993) also included *L. pimpinellifolium*, because of its close phylogenetic position to *L. esculentum*. Most *Lycopersicon* spp occur as weeds in South America, where currently all the wild relative species are to be found (Hawkes, 1991). It has been suggested that there was an export of plants or seeds toward the northern part of the continent by native people of South America (Rick, 1991).

Deliberate selection and breeding to adapt tomato to specific growing areas has been in progress for little more than 200 years (Stevens and Rick, 1986). The short history of the tomato crop begins with its introduction into Europe by the Spanish early in the 16th Century. In the 17th Century it was grown in England for ornament only, although it was known to be eaten elsewhere. By the end of the 18th Century it was grown in fields in Italy and used extensively as food, but it was half century or more before people in the USA generally dared to eat it (Boswell, 1937).

Stevens and Rick (1986) comment that prior to 1860 no cultivars had been developed in the USA; the few that were used had been imported, mostly from England, but with a few from France. The efforts at selection by early growers of the crop in Europe, together with natural factors, produced a very interesting and effective assortment of general types (Boswell, 1937).

Research attempts leading to the current popularity of tomatoes date back to 1905 when Halsted *et al.* (1905) reported the occurrence of single gene mutants in tomato. In 1909, Winkler evaluated the tomato cytologically and found it to have $2n = 2x = 24$ chromosomes. Linkage of mutant genes was first noted in 1917 by Jones (De Verna and Paterson, 1991). In the late 1920s rapid progress was made in cultivar development as hybridisation followed by selection in segregating generations was utilised (Stevens and Rick, 1986). Prior to 1925, tomato improvement was largely a result of selection of new genotypes within existing heterogeneous cultivars or selection of chance variance which resulted from spontaneous mutations, natural outcrossing or recombination of pre-existing genetic variation. By the mid-1930s

breeders were developing procedures to improve selection efficiency. Starting in 1940, accelerated introgression of useful exotic traits contributed to significant improvement, manifested in a 4x/5x fold increase yield (Rick and Chetelat, 1995). Much of the later work on tomato breeding has been carried out through commercial companies. Necessarily, therefore, this has been subject to secrecy, particularly in the case of greenhouse cultivars (Stevens and Rick, 1986). However, there are still breeding programmes in public funded companies in some countries, but without the commercial impact of private multinationals, which invest in research and marketing at the same level.

1.3.5 Genetic variability in *L. esculentum*

The tomato is one of the many self-pollinated crop species in which genetic variability of the exploited germplasm has been severely reduced by the processes of domestication and the breeding of new cultivars outside the native region. Results reported by Miller and Tanksley (1990), Van der Beek *et al.* (1992) and Rus-Kortekaas *et al.* (1994) show the relatively low amount of genetic variation detected with RFLP and RAPD markers among *L. esculentum* cultivars.

The lower overall genetic diversity of modern cultivars may in part reflect the genetic 'bottleneck' to which modern tomato cultivars were subjected during their introduction from Latin America to Europe (and later to the USA) (Boswell, 1937). For instance, only a limited number of seeds (accessions) were carried back and which served as the basis of modern cultivars of today (Garanko, 1991).

Rick and Chetelat (1995) suggested that the initial genetic variability of the ancestral form may have already been at low level, and was further diminished by the combination of autogamy and repeated reproductive bottlenecks (Rick, 1976; Miller and Tanksley, 1990; Williams and St. Clair, 1993). Perhaps, domestication from wild relatives to *Lycopersicon esculentum* was accompanied by a transition from exserted to inserted stigmas with consequent change from facultative outcrossing to enforced autogamy (Rick, 1979b), such that all representatives of *Lycopersicon esculentum* are self-compatible and exclusively inbreeding (Taylor, 1986). Within *L. esculentum*, apart from induced variation and variability resulting from the occasional

introgression of traits from wild species (Rick, 1979b), little genetic diversity can be found (Miller and Tanksley, 1990; De Verna and Paterson, 1991; Bretó *et al.*, 1993; Williams and St. Clair, 1993; Rick and Chetelat, 1995). Williams and St. Clair (1993) reported that one striking feature of the Unweighted Pair Group Method Arithmetic Average (UPGMA) dendograms and Phylogenetic Analysis Using Parsimony (PAUP) cladograms, utilising RAPD and RFLP markers, was that nine of the modern cultivars analysed grouped on the same branch, even though the samples contained introgressed germplasm from different wild species. Miller and Tanksley (1990) proposed that the lower diversity observed in the modern cultivars may reflect popular breeding methods. The basic breeding methodology used, following hybridisation, includes pedigree and backcross methods. Both methods produce homozygous lines of which only a very limited number become a cultivar.

As far back as 1937, Boswell gave warning about limited differences between cultivars, initially because several commercial firms and seed growers had given special attention to the isolation of superior stocks and strains of a number of leading commercial varieties. Therefore, if the differences between cultivars are still decreasing, methods to identify cultivars unambiguously need to be developed. Molecular markers techniques have been indicated as tools for fingerprinting cultivars (Lindhout *et al.*, 1991).

One approach to create genetic variability looking for desirable traits in tomato has been deliberate mutagenesis. Methods used to induce mutants have been highly varied and include treatments with radium, x-rays, Uv-light, induced osmotic stress, neutrons, and chemical mutagens (De Verna and Paterson, 1991). The tomato has been a classical species for mutational studies. An increased number of mutants, as a result of induced mutagenesis and the discovery of isozyme variants, has enhanced the repertoire of stocks available for mapping and other purposes (De Verna and Paterson, 1991). Rick (1984) listed 688 monogenic mutants. A few notable examples of mutants are: ripening mutants *rin* (Clayberg *et al.*,1970), *nor* (Clayberg *et al.*,1973) and *Nr* (Clayberg *et al.*,1960); male sterile mutants *ms* series (Rick and Butler, 1956), jointless *j* (Rick and Butler, 1956) and *j-2* (Clayberg *et al.*,1960); anthocyanin-deficient and hairless (*hl ini*) which can be used as a marker (Rick and

Butler, 1956); self-pruning (*sp*)for machine harvest (Rick and Butler, 1956); and *vgc* for fruit colour development (Bohm *et al.*, 1966). Some mutants have been of great value for studies in physiological processes, mineral transport and metabolism (Rick, 1987).

Another way to induce variability is via tissue culture to exploit somaclonal variation (Hostika and Hanson, 1984; Evans, 1987; O'Connell and Hanson, 1987), but De Verna and Paterson (1991) reported that there had been not much success with this method.

Another source of genetic variation has been the utilisation of tomato wild relatives developing sesquidiploid hybrids of *Lycopersicon esculentum* and *Solanum lycopersicoides* (Rick *et al.*, 1986). Rick *et al.* (1988) reported that this hybrid has served as a vehicle to develop monosomic alien addition lines and diploid individuals carrying traits derived from the wild species. Unfortunately, the extra chromosomes always carried many undesirable genes along with the useful and the plants were usually weedy and low producers (Griffiths *et al.*, 1996).

1.3.6 Genetic diversity in *Lycopersicon* spp

Tomato wild relatives are mostly distributed from northern Chile to southern Colombia and from the Pacific Ocean coast to the eastern foothills of the Andes. Curiously, the close relative *Lycopersicon esculentum* var. *cerasiforme* is the only wild tomato species found outside South America. It is widely distributed in Peru, Ecuador, but also in Mexico (Garanko, 1991). Also *L. cheesmanii* is the only wild taxon endemic in the Galapagos Island (Ecuador).

Traditionally, variability has been measured by morphological characteristics, but recently biochemical/molecular methods have become more popular. These methods (isozymes, Random Amplified Polymorphic DNA or RAPD, Amplified Fragment Length Polymorphism or AFLP, Single Sequence Repeat or SSR, Restriction Fragment Length Polymorphism or RFLP) permit simple and more accurate estimates of genetic variability within populations and at other levels (Garanko, 1991). Miller and Tanksley (1990) found that the level of DNA polymorphism, detected through RFLP, within accessions and species of *Lycopersicon* was highly

correlated with the mating system, self-incompatible species harboured, on average, more than ten-fold variation within accessions compared with self-compatible species (Stevens and Rick, 1986; Miller and Tanksley, 1990).

The amount of genetic variation found within accessions as opposed to between accessions (estimated as the average genetic distance between individuals within accessions as opposed to between accessions) differed greatly among the *Lycopersicon* species (Miller and Tanksley, 1990). More genetic variation could be found within a single accession of the self-incompatible species (e.g. *L. peruvianum*) than among all accessions of any one of the self-compatible species (e.g. *L. esculentum* or *L. pimpinellifolium*) (Bretó *et al.*, 1993; Rick and Chetelat, 1995). Miller and Tanksley (1990) estimated that the three self-incompatible species (*L. hirsutum*, *L. pennellii* and *L. peruvianum*), in terms of total variation, together contained nearly three times as much genetic variation as the four self-compatible species combined (*L. esculentum*, *L. pimpinellifolium*, *L. cheesmanii*, and *L. parviflorum*).

Williams and St. Clair (1993) suggested that the low diversity observed in *L. cheesmanii* may be due to the use of two accessions, but Miller and Tanksley (1990) also found this species to have less diversity than *L. esculentum*. However, one form of *L. cheesmanii*, characterised by its highly ornate and elaborately subdivided leaflets, has been given subspecific status. This group is more common than the typical form and is known as botanic variety "minor" (Taylor, 1986).

L. peruvianum is a remarkably polymorphic species (Rick, 1979a; Bretó, *et al.*, 1993). A high level of genetic variability is evident between individuals of the same population, between populations of a given race, and between races. The variation in this species, expressed in morphological as well as biochemical and genetical characters, is so extreme that one seldom faces two plants of identical genotype (Rick, 1979a).

1.3.6.1 Phylogenetic relationship and genetic distances among *Lycopersicon* spp

There are common barriers to effective hybridisation of the different *Lycopersicon* species, including hybrid sterility, which leads to incompatibility and incongruity

(Hogenboom, 1979). The importance of these barriers varies considerably but generally is proportional to the phylogenetic distance between parents (Rick and Chetelat, 1995).

McClean and Hanson (1986) found that mitochondrial DNA (mtDNA) phylogeny placed *L. chmielewskii* closest to *L. pennellii*, also places *L. hirsutum* and *L. esculentum* as close relatives, and the two other red-fruited species (*L. pimpinellifolium* and *L. cheesmanii*) were found to be closer to the green-fruited species than to the cultivated tomato. The study of Palmer and Zamir (1982), based on chloroplast DNA (cpDNA), placed *L. chmielewskii* close to *L. peruvianum*, while cross-compatibility data (Rick, 1979a) suggested a more distant relationship between these two species.

In the case of closely related species *L. parviflorum* and *L. chmielewskii*, studies of allozymes have confirmed the complete uniformity within populations of '*parviflorum*' and the extensive heterozygosity shown by the outbreeder '*chmielewskii*'. *L. parviflorum* is assumed to have evolved from '*chmielewskii*' and to have become genetically isolated from the parent species by virtue of inbreeding (Taylor, 1986). Rick (1983) also reported that *L. chmielewskii* and *L. parviflorum* are sibling species.

In spite of many unique characteristics, such as very short internodes, there appears to be no doubt that *L. cheesmanii* is closely related to *L. esculentum* and *L. pimpinellifolium* (Rick and Fobes, 1975; Palmer and Zamir, 1982; Rick, 1983; Hanson and McClean, 1987), because these three species have coloured fruited (Taylor, 1986); probably *L. pimpinellifolium* gave rise to *L. cheesmanii* (Bretó *et al.*, 1993).

Similarly, the presence of crossing barriers has been taken to strengthen the case for regarding *L. chilense* as a true species and not simply a form or variety of *L. peruvianum* (Rick and Lamm, 1955).

1.3.6.2 Crossability

In respect to cross-compatibility, Rick (1979a) commented that comparative chromosomal morphology and pairing displayed a remarkable degree of coherence in tomato species. However, in the *Solanaceae* family, self-incompatibility is gametophytically controlled. The self-incompatible nature of pollen is conditioned by its own haploid genome, including self-incompatible gene (*s*). This gene prevents pollen tube growth in styles expressing the same allele (Hogenboom, 1979). Several considerations lead to the conclusion that such barriers (incompatibility) were acquired secondarily to geographic isolation and other kinds of genetic differentiation (Rick, 1979a). So, evolution of the mating system and adaptation to specific habitats must have played major roles in the speciation processes within the *Lycopersicon* species (Bretó *et al.*, 1993).

It is difficult to generalise in respect to fertility of inter-taxon hybrids and comportment of later generations. The situation varies from complete fertility with no cytogenetic irregularities in later generations, as in *L. esculentum* x *L. pimpinellifolium* which can be reciprocally hybridised (Taylor, 1986), to combinations with appreciable F_1 (genic) sterility and inviability, reduced recombination, modified segregation ratios, and other problems in F_2 generations (Rick, 1979a).

L. parviflorum, *L. cheesmanii* and *L. cheesmanii* var. *minor* can be reciprocally hybridised with cultivated tomato without any major interspecies barrier. In the case of *L. pennellii*, this species freely hybridises with members of the '*esculentum*-complex' giving fertile hybrids showing no sign of disturbed chromosome pairing. The inter-taxon hybrid can be easily backcrossed to *L. esculentum*, provided that the tomato is used as the female parent. *L. pennellii* also hybridises unilaterally with *L. pimpinellifolium*, *L. cheesmanii*, *L. parviflorum*, and *L. hirsutum*. *L. pennellii* cannot be crossed with either *L. chilense* or *L. peruvianum*, and is therefore behaving as a classic member of the '*esculentum*-complex' (Table 1.2) (Taylor, 1986).

L. hirsutum var. *typicum* shows unilateral incompatibility with the cultivated tomato. Normal seed and hybrid plants can easily be obtained using *L. esculentum* as the female, but the reciprocal cross does not result in fruit set. *L. hirsutum* var. *glabratum*

is more tolerant of foreign pollen than the var. *typicum* forms in this species, and it is reciprocally compatible with the crop species and its close relatives (Taylor, 1986). When a self-compatible tomato species *L. esculentum* is crossed as female with either of the self-incompatible species, *L. peruvianum* or *L. hirsutum* var. *typicum* (both species have exserted styles), pollen tube growth is not inhibited in the style, whereas in the reciprocal crosses, the pollen tube growth is arrested and the cross fails (Kaul, 1991).

L. chilense is separated from cultivated tomato by severe barriers to prevent intercrossing. The stigma of the wild species will not accept *L. esculentum* pollen and the flowers rapidly abscise. Although the reciprocal cross results in fruit development, viable seeds are produced only rarely. Ayusa *et al.* (1986) reported that F_1 plants from the cross *L. esculentum* x *L. peruvianum* were strongly self-incompatible, but cross-compatible with *L. esculentum* as staminate parent.

There are several approaches to overcome these interspecific barriers. In tomato crosses, Gradziel and Robinson (1985) found that bud pollination 2 to 4 days before flowering followed by 4h of 95 to 99% relative humidity led to the avoidance of self-incompatibility in some genetic lines of *L. hirsutum*. Generally, the related *Lycopersicon* species are inter-compatible only when a self-compatible species is used as female parent (Kaul, 1991). Also the barrier between '*esculentum*-complex' and '*peruvianum*-complex' can be broken down by the application of embryo culture, which succeeds only when the member of the '*esculentum*-complex' is used as the female parent (Rick, 1979b). However, techniques such as protoplast fusion have also proved to be successful in overcoming incompatibility barriers (Adams and Quiros, 1985; O'Connell and Hanson, 1985; Handley *et al.*, 1986).

To summarise, self-fertility with various degrees of facultative outcrossing is found in *Lycopersicon chmielewskii, L. esculentum, L. pimpinellifolium* and the self-compatible biotypes of *L. hirsutum* and *L. pennellii*. Obligate outcrossing are self-incompatible biotypes of *L. chilense, L. hirsutum, L. peruvianum, and L. pennellii. L. cheesmanii and L. parviflorum* are completely autogamous.

1.3.7 Utilisation of tomato genetic resources

Among cultivated species, tomato is in a highly favourable position with respect to available germplasm of related wild species. Nearly every wild taxon is represented by an ample number of accessions which represents the range of genetic variation, geographic distribution, varied ecological sites, etc. Since 1940, resistance genes for at least 42 major diseases have been discovered in exotic germplasm of which 20 have been used in horticultural tomatoes (Rick and Chetelat, 1995). Doolittle (1954) reported a great amount of tolerance and resistance sources to different diseases and nematodes in wild tomato available to introgress into the tomato crop.

Utilisation of these exotic genetic resources has been assisted recently by the application of various molecular genetic methodologies. Intensive mapping of the tomato genome by Tanksley *et al.* (1992) via *L. esculentum* x *L. pennellii* hybrids paved the way for these and many other important investigations (Rick and Chetelat, 1995). Wide hybridization has played an invaluable role in providing desirable variation for those interested in increasing the diversity of *L. esculentum* (Table 1.3). The most significant application of wide hybridisation in the improvement of the tomato has been in providing novel sources of disease and pest resistance (De Verna and Paterson, 1991). Although this resistance has been derived from all known wild relatives of the tomato, certain species such as *L. chilense*, *L. peruvianum*, *L. hirsutum*, and *L. pimpinellifolium* appear to be the richest sources (Rick and Chetelat, 1995). However, Williams and St. Clair (1993) comparing old and modern cultivars suggested that relatively few new alleles have been introgressed by interspecific crosses that have introduced economically important traits into the modern cultivars

Table 1.3 Some characters introgressed by inter-taxon crosses from wild relatives into tomato.

Species	Character	Reference
L. pimpinellifolium	Resistance to *Fusarium* wilt	Taylor, 1986
L. pimpinellifolium	Resistance to bacterial speck (*Pseudomonas tomato*)	Taylor, 1986
L. pimpinellifolium	Long shelf life	Lobo et al., 1990

L. cheesmanii	Jointless pedicel gene (*j2*)	Stevens and Rick, 1986
L. cheesmanii	Tolerance to salinity	Taylor, 1986
L. parviflorum	Solid soluble content	Taylor, 1986
L. chmielewskii	Solid soluble content	Taylor, 1986
L. hirsutum	Resistance to tomato fruit worm (*Heliothis zea* Boddie)	Taylor, 1986
L. hirsutum	Resistance to sugar beet army worm (*Spodoptera exigua* Hübner)	Taylor, 1986
L. hirsutum	High content of A-tomatine	Rick and Chetelat, 1995
L. hirsutum	Resistance to bacterial speck (*Pseudomonas tomato*)	Taylor, 1986
L. hirsutum	Resistance to root knot nematode (*Meloidogyne spp*),	Taylor, 1986
L. hirsutum	Resistance to *Septoria lycopersici*	Taylor, 1986
L. hirsutum	Resistance to TMV (Tomato mosaic virus)	Taylor, 1986
L. hirsutum	Cold tolerance genes	Stevens, 1980
L. chilense	TMV resistance gene $Tm2^2$	Rick and Chetelat, 1995
L. pennellii	Resistance to *Fusarium* race 3	Mc Grath *et al.*, 1987
L. pennellii	Resistance to *Fusarium* race 2	Scott and Jones, 1991
Solanum lycopersicoides	White anthers (*Wa*),	De Verna *et al.*, 1987a
Solanum lycopersicoides	Day length sensitivity (*Dls*)	Rick *et al.*, 1988
Solanum lycopersicoides	Bifurcate inflorescence (*Bif*)	Chetelat *et al.*, 1989
Solanum lycopersicoides	Fimbriate leaves (*Fmb*)	Rick *et al.*, 1988
Solanum lycopersicoides	Frilly leaves (*Frl*)	Rick *et al.*, 1988
Solanum lycopersicoides	Lacinate leaves (*Lac*)	Rick *et al.*, 1988
Solanum lycopersicoides	Rugose leaf surface (*Rug*)	Rick *et al.*, 1988

1.4 Aims and objectives of the study

The aims of this research are related to aspects of the creation and subsequent management of populations, to increase the genetic base of autogamous crops; and will include different types of populations created from various sources. Although it is not possible to undertake an entire base broadening approach for any crop within three years, this project has chosen a tomato crop as a model to help designing a strategy for genetic base broadening in other autogamous crops.

This project intends to examine:

1) How much variability exists in genetic material from different sources?

- Levels of variability within populations
- Levels of variability between populations

2) How do created populations behave after hybridisation and selfing/outcrossing?

- Comparison of these populations with parent profiles and *L. esculentum* populations

3) Decline of variability in succeding generations. Is it possible to maintain this variability in later generations?

- To what extent do DNA polymorphism and diversity increase/decrease?
- How is maintenance of variability affected by self-pollination?

At the outset of a project aimed at broadening the genetic base of an autogamous crop species, a number of other questions have to be considered:

- How large should be the scale of operation?
- How many parental lines should be utilised?
- What should the range of parents be, or how wide should be the choice of parental material?
- Can the population be large enough to generate variability for many years even at low rates of natural outcrossing?
- Will there be a need for continued hybridisation?
- How will selection be minimised particularly during the initial phase?

- Should the material be exploited at one or many different sites?
- Does the selection of different sites lead to maintenance of variability?

In the light of studies undertaken, it is intended to examine some of these questions in a review of strategy for base broadening.

Chapter 2

Materials and methods

2.1 Materials

2.1.1 Germplasm

Seeds from 43 accessions, listed in Table 2.1, were obtained from the Centre of Genetic Resources (CGN, part of CPRO-DLO, Wageningen, The Netherlands), Tomato Genetics Resources Center (TGRC, Department of Vegetable Crops, University of California, Davis, California, U.S.A.), United States Department of Agriculture – Agricultural Research Service (USDA-ARS, Plant Genetic Resources Unit, Cornell University, Geneva, NY, U.S.A.), Institute für Pflanzengenetik und Kulturpflanzenforschung (IPK, Gatersleben, Germany), and Instituto de Investigaciones Agropecuarias (INIA, CRI-La Platina, Santiago, Chile). The six groups of species in this study included 12 open-pollinated (OP) vintage cultivars and landraces, 4 modern OP cultivars; 7 modern F_1 hybrids within *Lycopersicon esculentum* and 1 wild type accession of *Lycopersicon esculentum* var. *cerasiforme*; 9 *Lycopersicon cheesmanii*; 8 *Lycopersicon hirsutum*; 4 *Lycopersicon parviflorum*; 4 *Lycopersicon pennellii*; and 6 *Lycopersicon pimpinellifolium*.

2.1.2 Plant growth conditions

The experiments were carried out in the Plant Growth Unit (PGU) and laboratories of the Department of Biotechnology in the Scottish Agricultural College (SAC), Edinburgh, Scotland, and within the experimental station of CRI-La Platina in the Instituto de Investigaciones Agropecuarias (INIA), Santiago, Chile.

From each accession 5 tomato plants were grown in the PGU greenhouses at 22 ± 2°C and under 14/10 hours light/dark cycles. Light source was 400 watt, high-pressure sodium lamps. Plants under field conditions were grown in Santiago at 33°34`S latitude and 70°38`W longitude, altitude 625 meters above sea level., in a clay loam soil during the southern hemisphere spring-summer season 1999-2000.

For greenhouse growth of plants, seeds were sown in seedling trays (9 x 6 holes) containing a mixture of compost (Irish moss peat) : perlite (3:1). Seedlings were transplanted to 13 cm pots containing the same substrate at the stage of 3 to 4 true leaves. Plants were transplanted again to 18 cm pots at 20 to 25 cm height stage and

tied to a cane as support. Plants were watered every day and fed twice a week with Tomato Feed Premier (pbi Home & Garden Ltd.) fertiliser containing N:P:K (5:5:10) and trace elements.

2.2 Methods

2.2.1 Molecular markers

2.2.1.1 Plant DNA extraction

Chemicals utilised in DNA extraction and evaluation are listed and described in Table 2.2, while reagents utilised are presented in Table 2.3.

Young leaves from individual plants in each accession were sampled and total genomic DNA was isolated according to a modified CTAB method of Hachizume *et al.* (1996).

Fresh leaf tissue with main veins removed (0.3g) was frozen in liquid nitrogen and ground to a fine powder using a mortar and pestle. One ml of DNA extraction buffer (Table 2.4) was added to the homogenate of leaf tissue, mixed and poured into a 2 ml Eppendorf microcentrifuge tube. The cationic detergent CTAB facilitated DNA extraction because it solubilized cell membranes, while Tris-HCl (pH 8.0) preserved the DNA against degradation by native enzymes, such as lipolytic lipoxygenases, DNases, and/or secondary products released from the cells upon disruption. The extraction buffer also included EDTA as metal-dependent enzyme inhibitor, because it chelated divalent cations as Mg^{2+} and Ca^{2+}. Reducing agents DTT and PVP were included to protect the DNA against tannins, quinones, disulphides, peroxidases, and polyphenoloxidases action.

The sample tube was incubated for 30 minutes at 65°C in a heat block (Techna, DB-3) and allowed to cool at room temperature. The incubation at 65°C results in the formation of a CTAB-DNA complex, denaturation of many proteins, and dissociation from other DNA contaminants (Milligan, 1992; Taylor *et al.*, 1993).

Table 2.1 ***Lycopersicon*** **spp wild relatives, landraces, and varieties used in the** ***Lycopersicon esculentum*** **genetic base broadening project.**

Specie	Accession	Cultivar	Cultivar Type	Origin		Donator
1 Lycopersicon cheesmanii	CGN 15820			Galapagos	Ecuador	CGN - Holland
2 Lycopersicon cheesmanii	LA 0166			Galapagos	Ecuador	TGRC-Davis-USA
3 Lycopersicon cheesmanii	PI 379035	LA 1404		Galapagos	Ecuador	USDA-ARS-USA
4 Lycopersicon cheesmanii var. *minor*	LA 0317			Galapagos	Ecuador	TGRC-Davis-USA
5 Lycopersicon cheesmanii var. *minor*	PI 379040	LA 1411		Galapagos	Ecuador	USDA-ARS-USA
6 Lycopersicon esculentum		Limachino	Landrace			INIA-Chile
7 Lycopersicon esculentum	LA 2338	Ailsa Craig	Vintage			TGRC-Davis-USA
8 Lycopersicon esculentum	LA 0516	Ace	Vintage			TGRC-Davis-USA
9 Lycopersicon esculentum	LA 2414	Cal Ace	Modern			TGRC-Davis-USA
10 Lycopersicon esculentum	LA 3238	Earliana	Vintage			TGRC-Davis-USA
11 Lycopersicon esculentum	LA 2711	Edkawy	Landrace			TGRC-Davis-USA
12 Lycopersicon esculentum	LA 0534	Lukullus	Vintage			TGRC-Davis-USA
13 Lycopersicon esculentum	LA 2706	Moneymaker	Modern			TGRC-Davis-USA
14 Lycopersicon esculentum	LA 0502	Marglobe	Vintage			TGRC-Davis-USA
15 Lycopersicon esculentum	LA 0180	San Marzano	Vintage			TGRC-Davis-USA
16 Lycopersicon esculentum	LA 3008	San Marzano	Vintage			TGRC-Davis-USA
17 Lycopersicon esculentum	LA 0012	Pearson	Vintage			TGRC-Davis-USA
18 Lycopersicon esculentum	LA 1506	Stone	Vintage			TGRC-Davis-USA
19 Lycopersicon esculentum	PI 303782	Red Top	Vintage			USDA-ARS-USA
20 Lycopersicon esculentum	PI 270243	Roma	Modern			USDA-ARS-USA
21 Lycopersicon esculentum		Super Roma	Modern			INIA-Chile

Specie	Accession	Cultivar	Cultivar Type	Origin		Donator
22 *Lycopersicon esculentum*		1702-FA-144	F_1 hybrid			INIA-Chile
23 *Lycopersicon esculentum*		Boa F_1	F_1 hybrid			INIA-Chile
24 *Lycopersicon esculentum*		Cobra F_1	F_1 hybrid			INIA-Chile
25 *Lycopersicon esculentum*		Grandeur F_1	F_1 hybrid			INIA-Chile
26 *Lycopersicon esculentum*		Presto F_1	F_1 hybrid			INIA-Chile
27 *Lycopersicon esculentum*		Super Max F_1	F_1 hybrid			INIA-Chile
28 *Lycopersicon esculentum*		Vita F_1	F_1 hybrid			INIA-Chile
29 *Lycopersicon esculentum* var. *cerasiformes*	LA 1673			Lima	Peru	TGRC-Davis-USA
30 *Lycopersicon hirsutum*	LA 1353	PI 365934		Cajamarca	Peru	TGRC-Davis-USA
31 *Lycopersicon hirsutum*	LYC 4/88					IPK-Germany
32 *Lycopersicon hirsutum* var. *glabratum*	G 29255	VIR 3950			Ecuador	USDA-ARS-USA
33 *Lycopersicon hirsutum* var. *glabratum*	LA 1223	PI 365903		Chimborazo	Ecuador	TGRC-Davis-USA
34 *Lycopersicon hirsutum* var. *glabratum*	PI 199381			Lima	Peru	USDA-ARS-USA
35 *Lycopersicon parviflorum*	LA 1322	PI 379032		Cusco	Peru	TGRC-Davis-USA
36 *Lycopersicon parviflorum*	LA 1326	PI 379033		Apurimac	Peru	TGRC-Davis-USA
37 *Lycopersicon parviflorum*	T 1264/94					IPK-Germany
38 *Lycopersicon pennellii*	LA 0716	PI 246502		Arequipa	Peru	TGRC-Davis-USA
39 *Lycopersicon pennellii*	PI 473422			Lima	Peru	USDA-ARS-USA
40 *Lycopersicon pennellii* var. *puberulum*	LA 1926	CGN 15819		Ica	Peru	TGRC-Davis-USA
41 *Lycopersicon pimpinellifolium*	PI 230327				Ecuador	USDA-ARS-USA
42 *Lycopersicon pimpinellifolium*	PI 270449				Mexico	USDA-ARS-USA
43 *Lycopersicon pimpinellifolium*	PI 390739				Peru	USDA-ARS-USA

Then 1 ml of chloroform : isoamyl alcohol (24:1) was added to the tube. The tube was gently shaken, and then centrifuged (12,000 x g) at room temperature (22°C) for 4 minutes. The aqueous phase (approximately 750 µl) was transferred to a new 1.5 ml microcentrifuge tube and an equal volume of chloroform was added, mixed and centrifuged as above. The extraction processes with chloroform:isoamyl alcohol and chloroform remove problematic complex carbohydrates and denatured proteins (Taylor *et al.*, 1993).

The aqueous phase (approximately 600 µl) was transferred to a new microcentrifuge tube and an equal volume of precipitation buffer (Table 2.4) containing 1% of 2-mercaptoethanol was added. The addition of 2-mercaptoethanol inhibits any oxidation reaction occurring (Taylor *et al.*, 1993). The sample was gently mixed and the suspension was allowed to precipitate at room temperature for 30 minutes. The CTAB and nucleic acids form an insoluble complex under reduced salt conditions and which precipitates, leaving the remaining carbohydrates dissolved in the supernatant.

Centrifugation at room temperature for 3 minutes (12,000 x *g*) was followed by pouring off the supernatant and dissolving the pellet in 100 µl of 1M NaCl : TE (Table 2.4) and incubating at 65°C for 15 minutes in dry block. Resuspending the pellet in 1M NaCl:TE increases the concentration of salt and reprecipitates the DNA. The polysaccharides remain soluble in high salt concentration and do not co-precipitate with the DNA (Milligan, 1992).

Then 2.5 volumes (250 µl) of cold ethanol (-20°C) was added, mixed and centrifuged as above. The precipitate was washed in 250 µl of 70% ethanol, dried and redissolved in 100 µl of sterile distilled water. Co-precipitated RNA was eliminated by adding 1µl of RNase (500 µg/ml) and incubated for 15 minutes at 37°C. The DNA stock was stored at -20°C in a freezer.

Table 2.2 Chemicals name, use, formulae, molecular weight, and source utilised in the experiments.

Common Name	Use*	Chemical Name	Formula	Molecular Weight	Source
2-mercaptoethanol	1	2-hydroxyethylmercaptan	C_2H_6OS	78.13	Sigma, M-7154
Agarose	2-3				Amresco, 0710
Agarose Ultrapure-1000	3				GibcoBRL, 10975
Bleach	5	Sodium hypochloride			
Boric acid	3		H_3BO_3	61.83	Sigma, B-6768
Bromophenol blue	2-4	3',3",5',5"-tetrabromophenolsulfonephthalein	$C_{19}H_9Br_4O_5SNa$	691.90	Sigma, B-6131
Chloroform	1		$CHCl_3$	119.40	Sigma, C-2432
Chloroform : Isoamyl alcohol	1	24 : 1			Sigma, C-0549
CTAB	1	Hexadecyltrimethylammonium bromide	$C_{19}H_{42}NBr$	364.50	Sigma, H-6269
DTT	1	Dithiothreitol	$C_4H_{10}O_2S_2$	154.20	Sigma, D-9779
EDTA	1-3	Ethylen diaminetetraacetic acid	$C_{10}H_{14}N_2O_8Na_2$ $2H_2O$	373.20	Amresco, 0105
Ethanol	1		C_2H_6O	46.06	Amresco, E 193
Ethidium bromide	1-2	2,7-diamino-10-ethyl-9-phenylphenanthridinium bromide	$C_{21}H_{20}N_3Br$	394.30	Sigma, E-8751
Isoamyl alcohol	1	3-methyl-1-butanol	$C_5H_{12}O$	88.15	Sigma, I-9392
Metaphor agarose	3				Flowgen, 50180
PVP – 40	1	Polyvinylpyrrolidone			Sigma, PVP-40
Sodium chloride	1		NaCl	58.44	Sigma, S-3014
Sucrose	4	a-D-glucopyranosyl b-D-fructofuranoside	$C_{12}H_{22}O_{11}$	342.30	Sigma, S-7903
TBE – buffer	2-3	Tris - borate - EDTA buffer (5X)			Sigma, T-4415
TE – buffer	1	Tris - EDTA buffer (100x)			Sigma, T-9285
TRIZMA base	1-3	(tris[hidroxymethyl]amino methane)	$C_4H_{11}NO_3$	121.10	Sigma, T-8524

* 1) DNA extraction; 2) DNA evaluation; 3) Electrophoresis; 4) Loading buffer; 5) Seeds extraction.

Table 2.3 Reagents, concentration and source utilised in the experiments.

Reagent	Use*	Concentration	Source
Amplitaq DNA Polymerase, Stoffel Fragment	1	10 U/ml	Perkin Elmer, F0717
10x Stoffel buffer:	1		Perkin Elmer, H1240
Tris-HCl (pH 8.3)		100 mM	
KCl		100 mM	
Magnesium Chloride ($MgCl_2$)	1	25 mM	Perkin Elmer, H0994
dNTPs-mix (dATP, dCTP, dGTP, and TTP)	1	5 mM	Bio Gene Ltd., 300-113
Low DNA Mass Ladder	2	470 ng/4 ml	Gibco BRL, 10068-013
100bp DNA Ladder	1	1 mg/ml	Gibco BRL, 15628-019
RNAse, DNAse-free from bovine pancreas	3	500 mg/ml	Boehringer, 1119 915

* 1) RAPD and microsatellite PCR reaction; 2) DNA evaluation; 3) DNA extraction.

Table 2.4 Buffers reagents and concentrations.

Reagent	Stock	Extraction Buffer		Precipitation Buffer		1M NaCl : TE	
		in 100 ml	Final Concent.	in 100 ml	Final Concent.	in 100 ml	Final Concent.
CTAB		2 g	2%	1 g	1%		
Tris-HCl (pH 8.0)	1M	10 ml	100mM	5 ml	50mM	1 ml	10 mM
EDTA	0.5M	4 ml	20mM	2 ml	10mM	0.2 ml	1mM
Na Cl	5M	28 ml	1.4mM			20 ml	1M
PVP - 40		3 g	3%				
DTT		0.1 g	0.10%				
Distilled water		58 ml		93 ml		78.8 ml	

2.2.1.2 DNA evaluation, quantification and pooling

Qualitative and quantitative assessment of DNA was carried out by agarose gel electrophoresis in a horizontal submerged Flowgen medium size chamber (24.5 cm between electrodes). The agarose gel 3 mm thick was made by dissolving 1.2% agarose in 1xTBE buffer (89mM Tris-HCl, 89mM boric acid, 5mM EDTA) in a beaker and heating it in a microwave oven for 2 minutes, swirling once the solution started boiling. The solution was left cooling at room temperature, when it reached 60°C the gel was poured into a frame and a 20 well comb put into the warm liquid. The gel frame was allowed to cool and solidify at room temperature, and once set it was placed in a cold chamber at 4°C for 30 minutes.

The DNA sample was diluted 10 fold (2µl in 18µl of sterile distilled water), then 2 µl of the dilution mixed with 2 µl of gel loading buffer (sucrose 40% (w/v), bromophenol blue 0.25% (w/v), stored at 4°C) and 4µl of mixture loaded in each wells. As a standard comparison a low DNA mass ladder (Gibco BRL) in a similar proportion was used.

The gel was run for 1 hour in 0.5xTBE buffer at 94V, then stained in ethidium bromide (0.5 µl/ml) for 30 minutes.

Gel visualisation and evaluation was completed under UV illuminator at 302 nm, the image was captured by Flowgen IS-500 gel documentation system and analysed by densitometry in a Flowgen IS-1000 gel analysis system, comparing the total area of the bands produced with the standard bands. From F_2 generation forward a spectrophotometric analysis was included using the DNA analyser Eppendorf BIO Photometer.

Once DNA quality and quantity (ng/µl) was known, an aliquot of the stock was diluted to a concentration of approximately 10 ng/µl as working sample. A similar aliquot from the working sample of 25 µl was taken from each of the 5 plants per accession, mixed pooling the DNA before starting the PCR experiments, and kept frozen at -20°C for the duration of the experimentation.

2.2.1.3 Polymerase Chain Reaction (PCR)

Reagents utilised in PCR reactions are presented and described in Table 2.3.

2.2.1.3.1 Microsatellites or Short Tandem Repeat (STRs)

A set of 18 microsatellites primers of 18 – 20 oligonucleotides in length were tested (Table 2.5). These primers were extracted from literature available on tomato microsatellites (Smulders *et al*, 1994; Broun & Tanksley, 1996; Provan *et al.*, 1996; Arens *et al.*, 1995). Fifteen primers were selected based on whether they demonstrated polymorphism between and within *Lycopersicon* spp.

After testing several amplification reaction protocols (Vosman & Arens, 1997; Broun & Tanksley, 1996; Provan *et al*, 1996; Arens *et al*, 1995; Morgante & Olivieri, 1993) it was found that 10mM Tris-HCl (pH 8.3), 10mM KCl, 0.2mM deoxynucleotide triphosphates (dNTPs), 0.2μM each primer, 0.05U Stoffel fragment DNA polymerase, 2.5mM $MgCl_2$ and 20 ng genomic DNA per 15μl reaction volume represented the best results. The reaction components were mixed in a 0.5ml Eppendorf microcentrifuge tube in the order described above. Amplifications were performed using a TRIO-Thermoblock (Biometra) or Gene E (Techna) thermal cycler, both devices have hot lid, therefore mineral oil to avoid evaporation was not required. Amplification conditions were 1 cycle at 94°C for 30 seconds (hot starting), 30 cycles at 94°C for 30 seconds (denaturation), 2 minutes at specific annealing temperature for every primer pair, and 3 minutes at 72°C (extension), final extension of 7 minutes at 72°C was followed by soaking at 4°C. PCR products were stored in a fridge at 4°C for no more than 24 hours before electrophoresis. The annealing temperature was calculated as the sum of 4°C for each C and G bases, and 2°C for each A and T bases, then subtracted 5°C from the sum.

Fifteen μl of PCR products were electrophoresed in 3.5% Metaphor Agarose (Flowgen):Ultrapure 1000 (Gibco BLR) (2:1). The agarose was prepared by dissolving in chilled 1xTBE and incubating for 30 min at room temperature, with continuous stirring until the agarose was completely hydrated. A conical flask

containing the mixed agarose was wrapped in plastic film, pierced for ventilation and weighed. The suspension was heated for 2 min in a microwave oven, stirred and heated till boiling, and boiled for 1 min. The flask was weighed again and the difference was made up with distilled water. The boiled agarose was cooled, and when it reached 60°C, it was poured into the frame and the comb set. Once the gels had set, they were kept wrapped in plastic film at 4°C in cold room overnight to strengthen the resolution capacity of the gel.

As a standard comparison 4 µl of 100bp ladder (Gibco BRL) solution (8µl 100 bp ladder in 108 µl loading buffer) was eletrophoresed in addition to the PCR samples. Electrophoresis was performed in a horizontal Flowgen chamber using a double gel (40 cm between electrodes), at 3.9 V/cm for 2 hours in 0.5X TBE buffer.

The gel was stained as before in ethidium bromide (100 µl EtBr in 200 ml distilled water), visualised under UV and photographed.

2.2.1.3.2 Random Amplified DNA Polymorphism (RAPD)

The random-sequence primers for the polymerase chain reaction (PCR) were 10-base oligonucleotides in length, at least 50% G/C in content and lacked internal inverted repeats (Waugh & Powell, 1992). The series OPA, OPH, OPI, and OPL from Operon Inc. were tested and additional primers from other series included (Table 2.6). Amplification reaction conditions were similar to those reported by Hachizume *et al* (1996) consisting of 10 mM Tris-HCl (pH 8.3), 10 mM KCl, 0.2 mM each dNTPs, 40 pM primer, 3.5 mM $MgCl_2$, 10 ng genomic DNA, and 0.05 U Stoffel fragment DNA polymerase per 10µl reaction volume in a 0.5 ml microcentrifuge tube. DNA amplification was performed in a thermal cycler from Biometra (TRIO-Thermoblock) or Techna (Gene E), programmed for 1 cycle at 94°C for 30 seconds (hot starting), then 45 cycles at 94°C for 30 seconds (denaturation), 40°C for 2 minutes (annealing) and 72°C for 3 minutes (extension). One cycle of 7 minutes at 72°C as final extension was completed and followed by soaking at 4°C.

Ten µl of amplification products were separated by electrophoresis in a 2% agarose gel as detailed in 2.2.1.2. As a standard comparison 4 µl of a 100 bp ladder (8µl 100 bp ladder and 108 µl loading buffer) was utilised. The gel was run at 156V for 2

hours in 1xTBE buffer, stained in ethidium bromide (0.5 μl/ml) for 30 minutes and visualised under UV illuminator at 302 nm and recorded.

Table 2.5 Microsatellite oligonucleotide primers utilised in the experiments.

Locus	Repeat	Forward Primer	Reverse Primer	Number of alleles	Fragment Size (bp)
LE20592	$(TAT)_{15}$ $(TGT)_4$	5'- CTG TTT ACT TCA AGA AGG CTG-3'	5'-ACT TTA ACT TTA TTA TTG CCA CG-3'	7	166
LE21085	$(TA)_2$ $(TAT)_9$	5'- CAT TTT ATC ATT TAT TTG TGT CTT G-3'	5'-ACA AAA AAA GGT GAC GAT ACA-3'	5	104
LECHI3	$(TA)_6$ $(GA)_4$	5'- TAA CAA TCA AAA GAA CTT CGC-3'	5'-ATC CCC TTA TTG ATT ACA TCC-3'	2	128
LEEF1A	$(TA)_8$ $(ATA)_9$	5'- AAA TAA TTA GCT TGC CAA TTG -3'	5'-CTG AAA GCA GCA ACA GTA TTT -3'	8	131
LEGAST1	$(TA)_{12}$ and $(TG)_4$ GT $(TG)_5$	5'- GTT CTT TTG GTG GTT TTC CT-3'	5'-TTA TTT CTC TGT TGT TGC TG-3'	4	204
LEGTOM5	$(TA)_{10}$	5'- AAA GAT AAA GCA TGA AAT GAA -3'	5'-GGA GTT GAG ATA AAG TGA AGA -3'	2	181
LEILV1B	$(T)_8$ $(TA)_{10}$ $(T)_5$	5'- GAT CGA CAC ATT TGA ATT TGT -3'	5'-GGT CAC TAA TTA ATT GAT TCC -3'	3	143
LELE25	$(TA)_{11}$	5'- TTC GAT AAA GCA TGA AAT GAA -3'	5'-CTC TAT TAC TTA TTA TTA TCG -3'	3	225
LELEUZIP	$(AAG)_6$ TT $(GAT)_7$	5'- GGT GAT AAT TTG GGA GGT TAC-3'	5'-CGT AAC AGG ATG TGC TAT AGG-3'	4	105
LEMDDNa	$(TA)_9$	5'- ATT CAA GGA ACT TTT AGC TCC-3'	5'-TGC ATT AAG GTT CAT AAA TGA-3'	4	211
LEPRP4	$(TAT)_3$ $(TGT)_5$	5'- TTC ATT TCT TGC AAC TAC GAT -3'	5'-CAT ACT AGC AAC ATC AAA GGG -3'	4	200
LERBCO	$(TG)_4$ $(TA)_6$ TGTA	5'- CAT TAC CTC CAT TGC TA-3'	5'-GTC TCG TAC TTC TTC AT-3'	4	172
LESSRPSPGa	$(TATT)_5$	5'- GAA TAT ATC GGG GAC AAT CTC -3'	5'-AAC GAA ATC TTT GTT CAG TGA -3'	2	219
LESSRPSPGb	$(C)_{16}$	5'- AAC ATT AGT TTG ATT GGA TGG-3'	5'-TTA AAC TTG CTT GAC TTT CC-3'	5	332
TG51	GATA/GACA	5'- CCT ACG TAC CTA CCC ATG T -3'	5'-ACA TAC AAA CAG AGA GAC AAA -3'	8	140

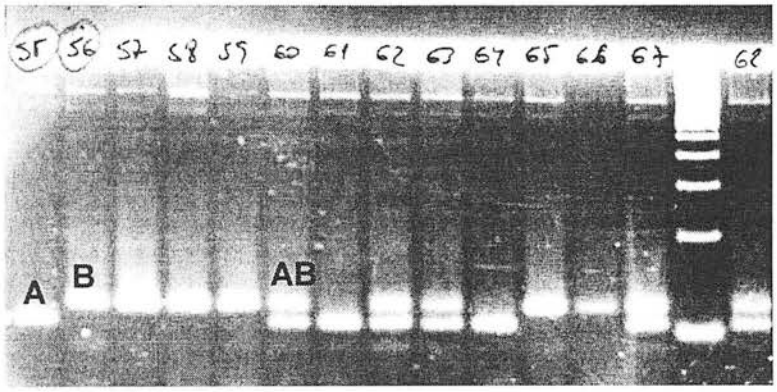
Table 2.6 Random amplified polymorphic DNA oligonucleotide primers utilised in this experiment.

Code	Sequence	Molecular Weight
OPA - 01	5'- CAG GCC CTT C -3'	2955
OPA - 02	5'- TGC CGA GCT G -3'	3035
OPA - 03	5'- AGT CAG CCA C -3'	2988
OPA - 04	5'- AAT CGG GCT G -3'	3059
OPA - 05	5'- AGG GGT CTT G -3'	3090
OPA - 06	5'- GGT CCC TGA C -3'	2995
OPA - 07	5'- GAA ACG GGT G -3'	3108
OPA - 08	5'- GTG ACG TAG G -3'	3099
OPA - 09	5'- GGG TAA CGC C -3'	3044
OPA - 10	5'- GTG ATC GCA G -3'	3059
OPA - 11	5'- CAA TCG CCG T -3'	2979
OPA - 12	5'- TCG GCG ATA G -3'	3059
OPA - 13	5'- CAG CAC CCA C -3'	2933
OPA - 14	5'- TCT GTG CTG G -3'	3041
OPA - 15	5'- TTC CGA ACC C -3'	2939
OPA - 16	5'- AGC CAG CGA A -3'	3037
OPA - 17	5'- GAC CGC TTG T -3'	3010
OPA - 18	5'- AGG TGA CCG T -3'	3059
OPA - 19	5'- CAA ACG TCG G -3'	3028
OPA - 20	5'- GTT GCG ATC C -3'	3010
OPL - 12	5'- GGG CGG TAC T -3'	3075
OPL - 16	5'- AGG TTG CAG G -3'	3099
OPL - 18	5'- ACC ACC CAC C -3'	2893
OPH - 01	5'- GGT CGG AGA A -3'	3108
OPH - 11	5'- CTT CCG CAG T -3'	2970
OPH - 14	5'- ACC AGG TTG G -3'	3059
OPH - 16	5'- TCT CAG CTG G -3'	3010
OPI - 15	5'- TCA TCC GAG G -3'	3019

2.2.1.4 Image analysis

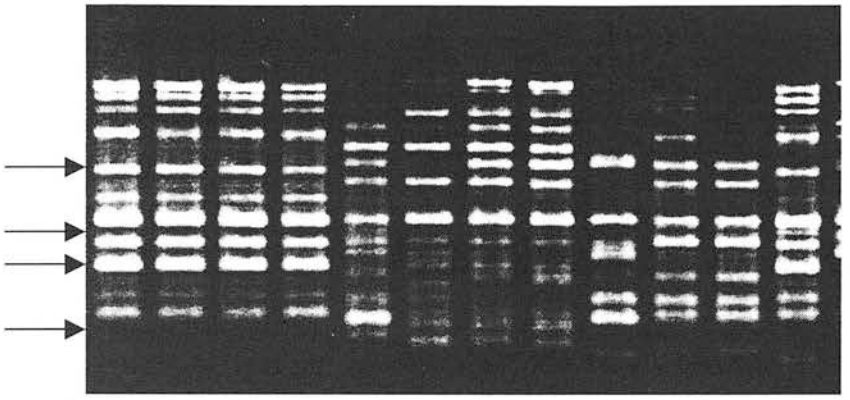
Microsatellites bands were analysed by treating the shared bands present in each lane as co-dominant markers. In Plate 2.1 the arrows are showing the parental bands A and B (in lanes 55 and 56), which were evaluated as alleles AA and BB; all the other lanes represent F_2 segregating individuals from the cross A x B and were evaluated as AA if only band A was present, BB if band B was present , or AB if both bands were present. All microsatellite analyses in this studied were carried out utilising this method, unless stated in the respective section

Plate 2.1. Agarose gel electrophoresis of PCR amplification products of *Lycopersicon* microsatellite locus LE21085 consisting in parental *accessions L. esculentum* var. *cerasiforme* (55) and *L. esculentum* cv. Limachino (56), and from 57 to 68 F_2 segregating individuals from the cross 55 x 56.



In the case of RAPD, the image was analysed evaluating polymorphism by absence (0) or presence (1) of bands, as it is showed in Plate 2.2. The arrows display the polymorphic bands in the electrophoresis agarose gel.

Plate 2.2 Agarose gel electrophoresis of PCR amplification products from 12 *Lycopersicon* taxa RAPD markers.



Arrows show polymorphic positions with primer OPA-18; 31 = *L. esculentum* cv. 1702-F_1-144; 33 = *L. esculentum* cv. Boa F_1; 34 = *L. esculentum* cv. Cobra; 40 = *L. esculentum* var. *cerasiforme*, LA-1673; 41 = *L. hirsutum*,, LA-1353; 42 = *L. hirsutum*, G29255; 44 = *L. hirsutum* var. *glabratum*, LYC 4/88; 45 = *L. hirsutum* var. *glabratum*, LA-1223; 47 = *L. hirsutum* var. *glabratum*, PI-1993181; 49 = *L. parviflorum*, LA-1322; 50 = *L. parviflorum*, LA-1326; 52 = *L. parviflorum*, T-1264/94.

2.2.2 Morphological characters

2.2.2.1 Pollination and crosses

Most plants were allowed to self-pollinate. Only selected accessions as parents were hybridised. After the second truss had formed, three to four flowers per truss were emasculated before opening and pollen release. Stamens from mature flowers were left in a plastic container and dried overnight in a desiccator with silica gel. When the emasculated flowers opened the stigma was receptive, and were put in contact with the pollen. The flower was then tagged and the fruit allowed to develop.

After hybridisation, one fruit from each individual growing in each population was harvested and mixed all together. During the following generations the same procedure was followed.

2.2.2.2 Plant morphology evaluation

There were 16 discontinuous characters evaluated. These were selected from a list published in "Descriptors for Tomato" (IPGRI, 1996). The characters, evaluation stages, and scores are detailed in Appendix 1.

2.2.2.3 Fruit harvest and seed extraction

Fruits were harvested at a ripe stage, each accession and/or hybridisation were bulked in sealed plastic bag, crushed, and left to ferment for week. Then the fruits were washed in a 1% sodium hypochloride solution for 20 minutes, rinsed in water, and the seeds separated from the pulp and skin. Seeds were dried overnight in a petri dish containing a layer of filter paper, then packed in a paper bag and stored in a desicator containing silica gel to keep drying the seeds.

2.3 Population size

Initially 5 plants per accession were grown in the greenhouse. After hybridisation, 8 plants were grown for most populations created from inter- and intra-taxon crosses in all generations studied. However, 40 plants were grown in one population (I-1939) and reciprocal (I-3919), which was randomly selected by a draw from inter-taxon crosses and 20 plants in the case of intra-taxon crosses E-2219 and reciprocal E-1922. The same number of 40 and 20 plants from all these populations quoted before were grown during the development of F_1 to F_3 generations.

2.4 Statistical analysis

Descriptive statistics, analysis of variance (ANOVA) (Snedecor, 1934) for normally distributed characters, Kruskal-Wallis test (Kruskal and Wallis, 1952) for non-normally distributed characters, and Tukey test (Tukey, 1953) for multiple comparison analysis were carried out utilising the statistical program Minitab 11.1. Graphs, tables and figures were produced using the programs MS-Word 97 and MS-Excel 97.

Analysis of molecular variance (AMOVA) was utilised to measure the genetic structure of the populations from which the samples were drawn. It works on binary data (0 and 1) creating a distance matrix between samples. The analysis treats genetic

distances as deviations from a group mean position, and uses the generated deviations as variances. The total sum of squares of genetic distances may then be partitioned representing the within-group and the between-group mean squares. Φ_{st} represents the correlation between random genetic accessions within a group relative to random accessions from the population at large. Φ_{st} statistic is analogous to Wright's F_{st} (Wright, 1965). This multilocus approach, originally, was developed for haplotype data, but it has recently become much applied for RAPD- and binary-data to estimate between populations variability. The data with a hierarchical structure allows an analysis of variance-like approach that can be extended to evaluate molecular marker data even with absence of replicated values for sample.

Genetic similarity values between pairs of genotypes were calculated using Jaccard's coefficient (Sneath and Sokal, 1973). This did not include 0-0 matches as indicator of similarity. Using genetic similarity matrices, dendrograms were constructed according to the unweighted pair-group method with arithmetic average (UPGMA) (Sneath and Sokal, 1973). Data were analysed through the program NTSYS-pc version 1.80 (Numerical Taxonomy System, Applied Biostatistics, NY).

Genetic diversity, in this research, was analysed and quantified in terms of diversity indices. They are mathematical measures of species diversity in a community. Diversity indices provide more information about community composition than simply species richness (i.e. number of species present). The aim was to examine the genetic diversity in *Lycopersicon* taxa and the possible changes when hybridising accessions inter- and intra-taxon. The following diversity indices were used in this research:

- Mean proportion of polymorphic loci (*P*). A locus is defined as polymorphic when the frequency of the most common allele is less than 1 and represent the percentage of all loci that are polymorphic regardless of allele frequencies.
- Mean number of alleles per locus (*A*), which represents the arithmetic mean of the number of alleles per locus across all loci or allele richness.

- Effective number of alleles (A_e). This equals the actual number of alleles only when all alleles have the same frequency, estimates the reciprocal of homozygosity, and it is calculated as:

$$A_e = \frac{1}{\sum x_i^2}$$

where x_i is the population frequency of the ith allele at a locus.

- Gene diversity (H), which is the probability that two alleles randomly chosen from a population will be different (Nei, 1987) and it is calculated as:

$$H = 1 - \sum x_i^2$$

where x_i is the population frequency of the ith allele at a locus.

Average gene diversity (H_S) was calculated as the average of sub-populations, in this case accessions and populations created by hybridisation, it represents the diversity within a population; and total gene diversity (H_t) utilised all populations as a meta-population, it represents the diversity between all populations. The upper bound of gene diversity is 1.0 when calculated utilising co-dominant marker and 0.5 with dominant markers.

- Shannon's information index (I) enables analogous comparisons between co-dominant and dominant markers because it is not bounded by 1.0 and is calculated as:

$$I = \sum x_i \ln x_i$$

where x_i is the population frequency of the ith allele at a locus.

Population differentiation was calculated according to Hartl and Clark (1997) as:

$$F_{st} = \frac{(Ht - Hs)}{Ht}$$

And it represents the partitioning of the diversity between and within present in the populations analysed.

All calculations and statistical analysis of genetic diversity and population differentiation were carried out utilising the program Popgene (Yeh, 1997). During the calculations the data were treated as populations not in Hardy-Weinberg equilibrium and indices calculated as dominant or co-dominant depending the morphological or molecular marker utilised. Allele frequencies were estimated from the information obtained in the gel electrophoresis and the score of the bands present. Significance levels were represented by asterisks, being significant differences ($P \leq 0.05$) '*' and highly significant differences ($P \leq 0.01$) '**'. Non-significant differences ($P > 0.05$) were denoted by 'ns'.

Chapter 3

Genetic diversity of *Lycopersicon* spp germplasm.

3.1 Introduction

Deliberate selection and breeding to adapt tomato to specific growing areas have been in progress for little more than 200 years (Stevens and Rick, 1986). However, a common problem has always been the lack of genetic variability among tomato germplasm. Boswell in 1937 already reported that problems and new requirements had arisen so quickly that tomato breeders could not find naturally occurring chance variants with the desired characteristics fast enough to met these requirements.

The low overall genetic diversity of modern cultivars reflect genetic "bottlenecks" to which modern tomato cultivars were subjected to during their domestication in Latin America and later introduction to Europe. Rick (1976) supposed that only limited number of seeds, and therefore probably accessions, were brought back by explorers and became the base of worldwide tomato breeding. However, natural bottlenecks during species evolution, such as autogamous plants that had hermaphrodite flowers with pistil enveloped by joined stamens and inserted stigma, suggest that the initial genetic variability of the ancestral form may have already been at low level (Rick, 1976; Miller and Tanksley, 1990; Williams and St. Clair, 1993; Rick and Chetelat, 1995).

Moreover, breeders have been selecting material mainly with inserted stigmas with the aim of enforcing autogamy, but this low diversity has been further reduced by the use of breeding methods that promote genetic uniformity, such as pedigree selection or single-seed descent. The number of cultivars released per year have been increasing, but the genetic and morphological differences between them decreasing. Within commercial breeding, relatively few dominant cultivars have come to be used as suitable parental material, usually in newly released cultivars showing only slight or 'cosmetic' differences, and which justifies a new name.

Among cultivated species, tomato is in highly favourable position with respect to germplasm availability in related wild species and landraces or old varieties. Nearly every taxon is characterised by a large number of accessions representing a range of genetic variation, geographic distribution and ecological niches. However, very few breeders are willing to use wild relatives because of the difficulty and time it takes to remove unwanted 'wild' characters. Nevertheless, resistance to at least 42 major

diseases has been discovered in exotics since 1940 and over 20 such resistances have been bred into horticultural tomatoes, a number that is continually rising (De Verna and Patterson, 1991; Rick and Chetelat, 1995). Landraces and old varieties have a greater useful diversity than modern cultivars, but their utilisation requires larger screening programmes and more expense. Therefore breeders prefer to restrict their programmes to the small amount of genetic diversity present in a few advanced lines and the introgression of individual specific traits from wild relatives (Cooper *et al.*, 1998)

Modern tomato varieties are closely related to the wild species *L. esculentum* var. *cerasiforme* and the two taxa can be freely crossed, in agreement with the subdivision *esculentum-complex* of the *Lycopersicon* genus by Rick (1976). Although unilateral relationships are common, hybrids can be obtained from nearly all combinations without need of special techniques, such as embryo rescue (Rick, 1979a). Of the diverse difficulties of crossing between *Lycopersicon* species, the most influential are blocks to hybridisation and hybrid sterility. It is difficult to generalise in respect to the fertility of inter-taxon hybrids and behaviour in later generations. This can range from complete fertility of reciprocal hybrids to combinations with strong F_1 sterility and inviability (Taylor, 1986; Rick, 1979a).

The objectives of this chapter are to demonstrate the levels of morphological and molecular diversity within *L. esculentum* accessions and the amount of genetic diversity available within the genus *Lycopersicon*. Based on these observations, the creation of a range of different types of populations will be studied in relation to base broadening objectives.

3.2 Morphological diversity in *Lycopersicon* germplasm

The phenotypic expression of morphological characters is usually divided into discontinuous (qualitative) and continuous (quantitative) variation. Most characters in nature are continuous or metric characters, such as yield, fruit size, tolerance/resistance to biotic or abiotic stresses, etc. However in a breeding programme both kind of traits can be useful to characterise individuals, populations, or species, and to analyse the diversity present between and within them.

In general phenotypic discontinuous variation is associated to one or two genes controlling the trait in a Mendelian manner. These qualitative characters are usually not strongly affected by environmental factors, unlike quantitative characters. Continuous variation in phenotype is associated to the collective action of many genes lying at different quantitative trait loci (QTL) and is usually highly influenced by environmental conditions. Thus data obtained in one site are unique. However, there are certain patterns of behaviour that can be studied and utilised in breeding programmes through statistical methods of quantitative genetics. These methods allow the calculation of indices of genetic variation and analysis of the quantitative genetic variation in QTLs.

In the following section, morphological evaluation of 38 accessions of 7 species of *Lycopersicon*, were assessed for 16 qualitative characters with the aim of examining the variation and genetic distances between accessions and species.

3.2.1 Morphological characteristics

Lycopersicon species form a cohesive group in respect to the following characteristics: herbaceous growth; sprawling or prostrate habit; stem organisation in sequences of 2- or 3-leaved sympodia; odd pinnate segmented leaves; cymose inflorescence; ebracteate, bright yellow, chasmogamous, pentamerous, hermaphrodite flowers with pistils enveloped by the connate or connivent anthers; and the fruit is a soft berry (Kaul, 1991; Taylor, 1986; Rick, 1979a).

In this research, of the 18 qualitative morphological characters analysed, 3 presented common characteristics for all accessions under study. All had pubescent hypocotyls, trusses with multiple flowers, and yellow corollas.

There was a considerably diversity observed in leaf shapes between and sometimes within species, as shown in Plate 3.1. Within *L. esculentum* there were some slight differences between accessions, generally within the called "potato leaf type" and "tomato leaf type". These leaf types were very similar to *L. esculentum* var. *cerasiforme* and *L. pimpinellifolium*, but the shape was very different in comparison to *L. pennellii*, *L. hirsutum*, *L. cheesmanii*, and *L. parviflorum*.

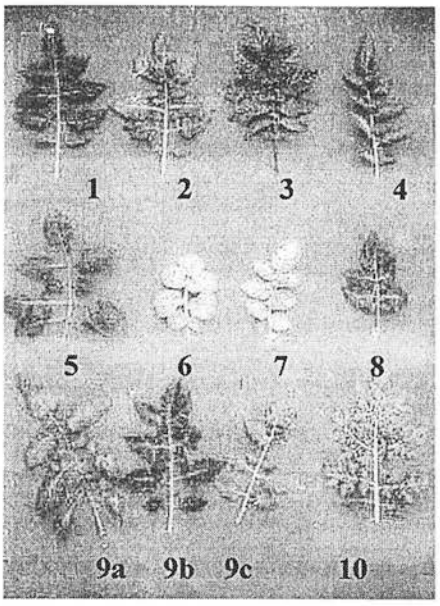


Plate 3.1 Leaf diversity of 10 samples of *Lycopersicon* spp.

1.- *Lycopersicon esculentum*
2.- *Lycopersicon esculentum* var. *cerasiforme*
3.- *Lycopersicon hirsutum*
4.- *Lycopersicon hirsutum* var. *glabratum*
5.- *Lycopersicon parviflorum*
6.- *Lycopersicon pennellii*
7.- *Lycopersicon pennellii* var. *puberulum*
8.- *Lycopersicon cheesmanii*
9.- *Lycopersicon pimpinellifolium* (a,b and c)
10.- *Lycopersicon cheesmanii* var. *minor*

Table 3.2.1 presents information in relation to flower characteristics and fruit morphology in accessions of the genus *Lycopersicon* examined in this research. In respect to style position, self-incompatible species, such as *L. hirsutum* and *L. pennellii*, possess highly exserted style. The character of green fruit at maturity, in these species, is strongly associated to characters such as highly exserted style and self-incompatibility. All representative cultivars and accessions of *L. esculentum* were self-compatible and exclusively inbreeding (Taylor, 1986), since domestication was accompanied by a transition from exserted to inserted stigmas and consequent change from facultative outcrossing to enforced autogamy (Rick, 1979b). Most of the species with red fruits, such as *L. pimpinellifolium*, *L. cheesmanii* and *L. esculentum* var. *cerasiforme*, presented styles at the same level as the anthers or slightly exserted, a characteristic correlated with autogamy.

The exterior colour of the immature fruit did not show much variation between species. In respect to fruit pubescence, *L. esculentum* and *L. esculentum* var. *cerasiforme*, two taxa very closely related, showed only sparse hairiness (few hairs covering the fruit). In the other more distantly related species, the hairiness increased from intermediate levels to dense (fruits completely cover with hairs). Fruit size in tomato wild relatives was always small to very small (less than 2 cm diameter) in comparison to *L. esculentum* cultivars which have been selected for bigger fruit size.

3.2.2 Genetic distances between *Lycopersicon* species

3.2.2.1 Distribution of phenetic similarities between *Lycopersicon* species and accessions

In order to analyse statistically phenetic distances between *Lycopersicon* spp, a dissimilarity matrix was created using the method described by Gower (1985) for a multivariate analysis of morphological traits. For each categorical character, the distance between two accessions was scored as zero if the character matched, and one if they did not. To create a morphological distance matrix, the individual trait distances for each pair of lines were added, then divided by the number of traits scored in both lines.

This matrix was transformed into similarities utilising the additive inverse (Appendix 2, part 1). Morphological data for 13 traits were available for 35 out of 38 accessions of 6 species of *Lycopersicon*. The 3 remaining accessions (*L. esculentum* cv. Cal Ace, *L. parviflorum* LA-1326, and *L. pennellii* PI-473422) did not complete the growing cycle and they were ommitted from the matrix.

To visualise the relationships between species, the similarity matrix was converted to a two dimension coordinate plot with the multidimensional scaling (MDS) procedure (Schiffman et al., 1981), using the program NTSYS-pc version 1.80. The stress parameter for this MDS procedure was 0.465, defined by Kruskal (1964) as "poor". Figure 3.2.1 shows that there are two main groups, one including all *L. esculentum* accessions and the close relative *L. esculentum* var. *cerasiforme*. The other group includes all remaining wild relatives. The *esculentum* group formed a close group, except for two cultivars, Limachino (Chilean landrace) and Super Roma that were located at the bottom of the plot. In the other extreme, cv. Edkawi (Egyptian landrace) was aligned at the same horizontal level with wild cherry tomato (*L. esculentum* var. *cerasiforme*).

Table 3.2.1 Reproductive morphological characters in *Lycopersicon* spp. (Details in Appendix 1)

Species	Style Position	Exterior colour of immature fruit	Fruit pubescence	Fruit size	Exterior colour of mature fruit
L. cheesmanii	Same level as stamen	Greenish-white to light green	Sparse to intermediate	Very small	Orange to red
L. cheesmanii var. *minor*	Slightly exserted	Green	Dense	Very small	Orange
L. esculentum	Inserted to same level as stamen	Light green to green	Sparse	Intermediate to very large	Orange to red
L. esculentum var. *cerasiforme*	Same level as stamen	Light green	Sparse	Small	Red
L. hirsutum	Highly exserted	Light green	Dense	Small	Green
L. hirsutum var. *glabratum*	Highly exserted	Green	Dense	Small	Green
L. parviflorum	Inserted to same level as stamen	Light green to green	Intermediate to dense	Very small	Green to red
L. pennellii	Highly exserted	Greenish-white	Intermediate	Small	Green
L. pennellii var. *puberulum*	Highly exserted	Greenish-white	Intermediate	Small	Green
L. pimpinellifolium	Same level as stamen to slightly exerted	Greenish-white to green	Intermediate	Very small	Red

Within the group of wild relatives, each species clustered with its own kind, but it was noticeable that the green fruited species (*L. pennellii*, *L. hirsutum and L. parviflorum*) grouped together. *Lycopersicon hirsutum* accessions were very closely grouped, and showed little differences in the traits analysed. This is reflected in the grouping of 4 out of 5 accessions at just one point. *L. pimpinellifolium* accessions had more variation between them than the other species which presented relatively close individual groups. The two accessions of *L. cheesmanii* var. *minor* were located apart from *L. cheesmanii* entries. Accessions of *L. parviflorum* and *L. pennellii* showed some differences within them, reflected in the position of every one in the plot.

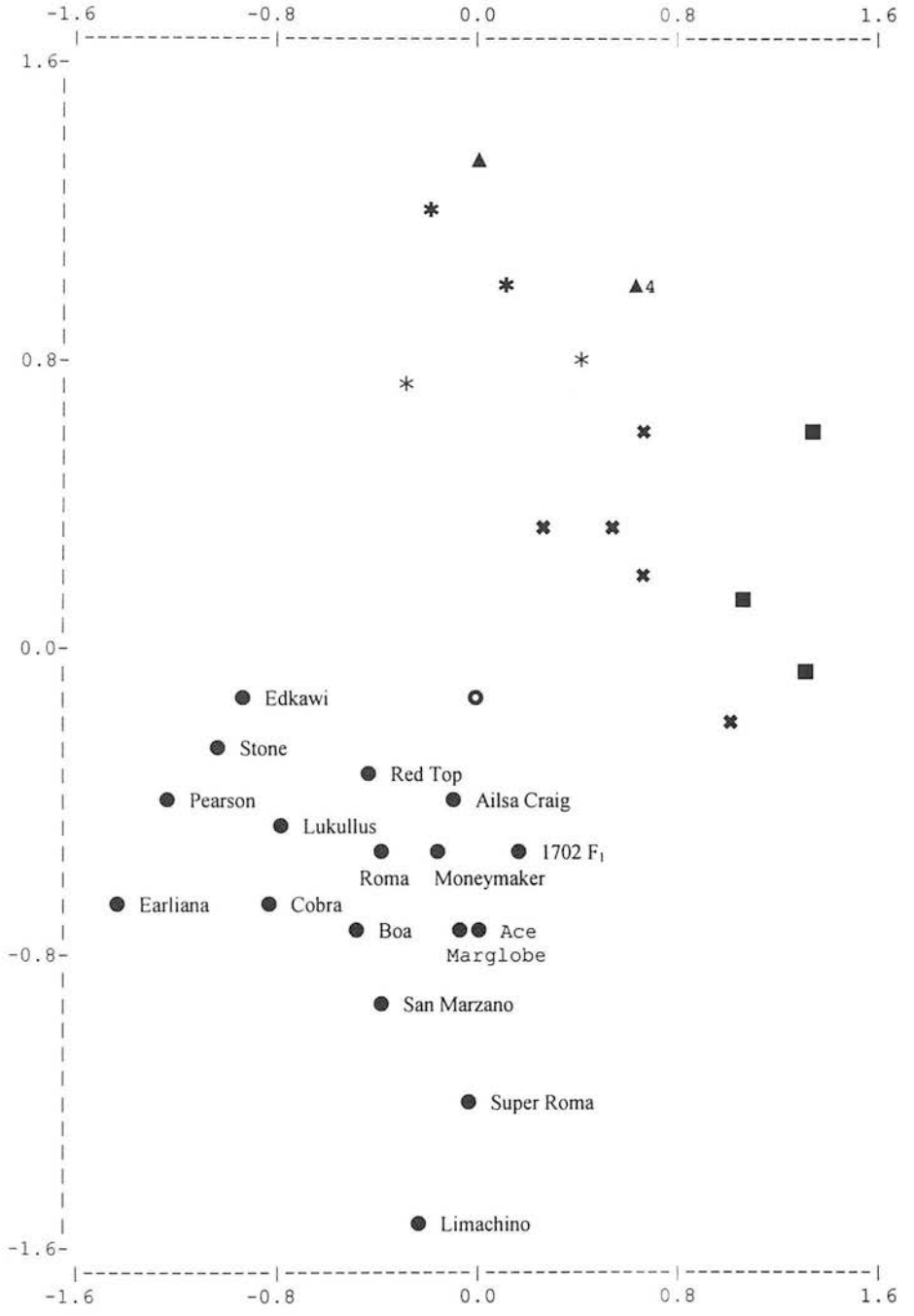
3.2.2.2 Phenetic similarity analysis within and between *Lycopersicon* spp

Morphological data from 13 traits were used to generate a phenetic similarity matrix for 35 accessions of *Lycopersicon*. Similarities were analysed grouping all the combinations of accessions for each species in the matrix, excluding self-combinations giving value 1, and treated as a whole utilising descriptive statistics (Table 3.2.2). Some species such as *L. parviflorum* and *L. pennelli* were considered in this analysis though they were represented by only two accessions, but *L. esculentum* var. *cerasiforme* was discarded because no comparison was possible with just one entry.

The highest genetic similarity mean was presented by *L. hirsutum* with 0.88, ranging from 1. 00 to 0.31, while the lowest means were found in *L. parviflorum*.

Within *L. esculentum* accessions there was high variation of distances, reflecting also the high number of accessions analysed.

Figure 3.2.1 Multidimensional scale (MDS) presentation of data of 13 morphological traits for 35 *Lycopersicon* accessions.



▲ *Lycopersicon hirsutum*
✱ *Lycopersicon pennellii*
∗ *Lycopersicon parviflorum*
✖ *Lycopersicon cheesmanii*
■ *Lycopersicon pimpinellifolium*
✖ *Lycopersicon cheesmanii* f. sp. *minor*
● *Lycopersicon esculentum*
o *Lycopersicon esculentum* var. *cerasiforme*

Table 3.2.2 Descriptive statistical analysis of a genetic similarity matrix based on 13 morphological traits within *Lycopersicon* accessions.

	N	Mean	SE	Max	Min
L. esculentum	136 (17)	0.51	± 0.01	0.85	0.23
L. cheesmanii	10 (5)	0.66	± 0.05	0.85	0.39
L. pimpinellifolium	3 (3)	0.72	± 0.11	0.92	0.54
L. hirsutum	10 (5)	0.88	± 0.08	1.00	0.31
L. parviflorum	1 (2)	0.46	-	-	-
L. pennellii	1 (2)	0.77	-	-	-

Numbers between brackets correspond to the number of accessions analysed per species; N = number of observations; SE = standard error; Max = maximum value; Min = minimum value.

3.2.3 Analysis of molecular variance (AMOVA) of morphological characters in *Lycopersicon* spp

The 35 accessions were grouped according to taxa, then the genetic distance matrix built from the morphological categorical data transformed to binary as explained in section 3.2.2.1, was analysed by Analysis of Molecular Variance (AMOVA) procedure (Excoffier *et al.*, 1992) calculating the variance between and within taxa for the morphological characters (Table 3.2.3). The results indicated highly morphological differentiation ($P<0.01$) between *Lycopersicon* taxa analysed, where 37.1% of the total variation found was attributable to morphological differences between and 62.9% within taxa.

Table 3.2.3 Analysis of molecular variance of 13 morphological characters in 34 accessions of *Lycopersicon* spp.

Source of variation	Sum of squares	df	Mean square	Percentage
Within taxa	412,487.99	28	14,731.71	62.9%
Between taxa	278,095.62	5	55,619.12	37.1%
TOTAL	690,583.61	33		

Variance within taxa	14,731.71
Variance between taxa	8,688.57
Φ_{st}	0.371

3.3 Molecular diversity in *Lycopersicon* germplasm

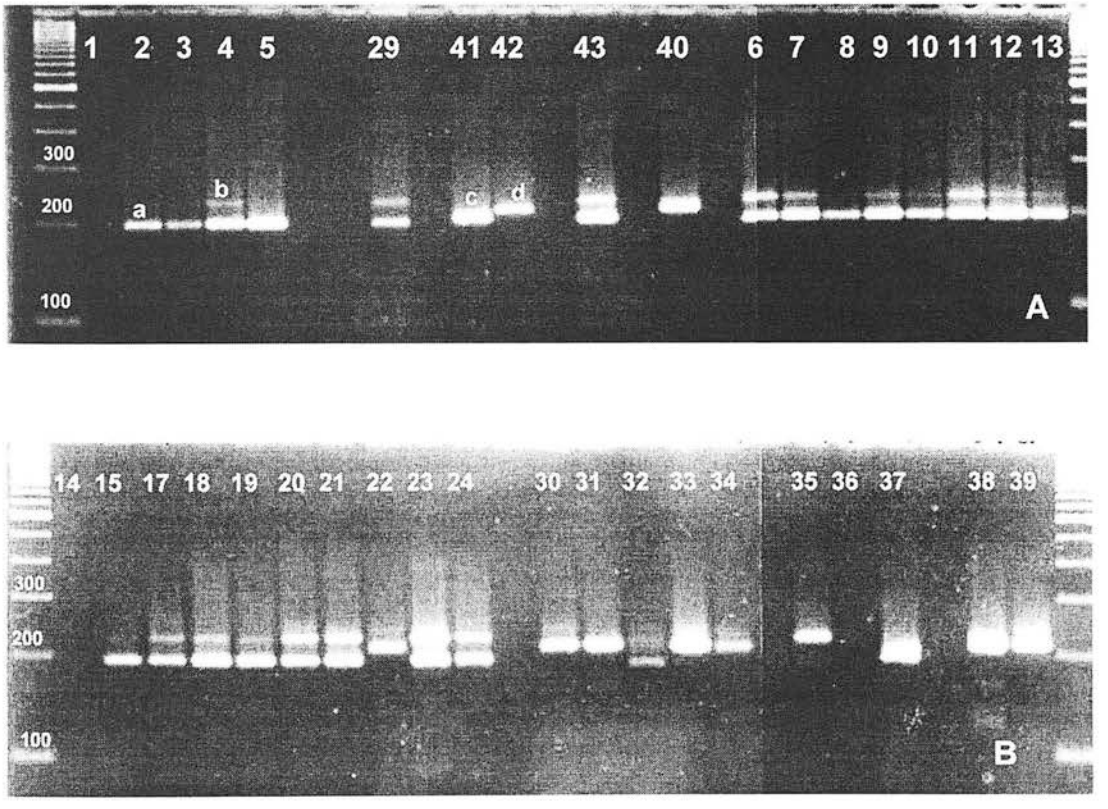
3.3.1 Microsatellite markers

Microsatellites, or short tandem repeats (STRs), or simple sequence repeats (SSRs) are a common feature in the eukaryote genome. They contain a basic repeat motif of 2 – 8 base pairs (Hamada et al, 1982, 1984; Tautz & Rentz, 1984). Such STRs/SSRs can be found in large numbers and are relatively evenly distributed throughout the genome. It has been suggested that the variability of microsatellites is due to variations in the number of copies of the basic repeat unit, likely caused by slippage of the polymerase during replication (Schlötterer & Tautz, 1992) or unequal crossing-over (Schlötterer, 1998). Microsatellite analysis has shown high variability even in populations which showed low levels of variation in allozymes and mitochondrial DNA (Schlötterer, 1998). Amplifying these regions through polymerase chain reaction (PCR) using a unique pair of flanking oligonucleotides as primers, almost regularly presents comprehensive polymorphisms because of different number of repeats (Morgante & Olivieri, 1993). Most microsatellite loci are selectively neutral and as they are embedded in single copy DNA, this facilitates the unambiguous scoring of alleles (Schlötterer, 1998).

3.3.1.1 Population structure and diversity

In this analysis microsatellite bands were treated as dominant markers. This approach was adopted because there was no information available from segregating populations to determine ranking order of alleles. Therefore the alleles could not be scored unequivocally, as it is showed in Plate 3.2. In lane 2 band 'a' is clearly one allele, also lanes 41 and 42 show bands 'c' and 'd' as one allele. However, in lane 4 are present bands 'a' and 'b', and in lane 32 bands 'c' and 'd', that are repeated in several other lanes. The microsatellite locus LEPRP4 is described with a size of about 200 bp, but considering the variability of microsatellites and the closeness of the bands in the gel, it is difficult to score them. Similar is the case showed in Plate 3.3 for locus LEGAST1, which is described with an expected size of 204 bp and marked with an arrow.

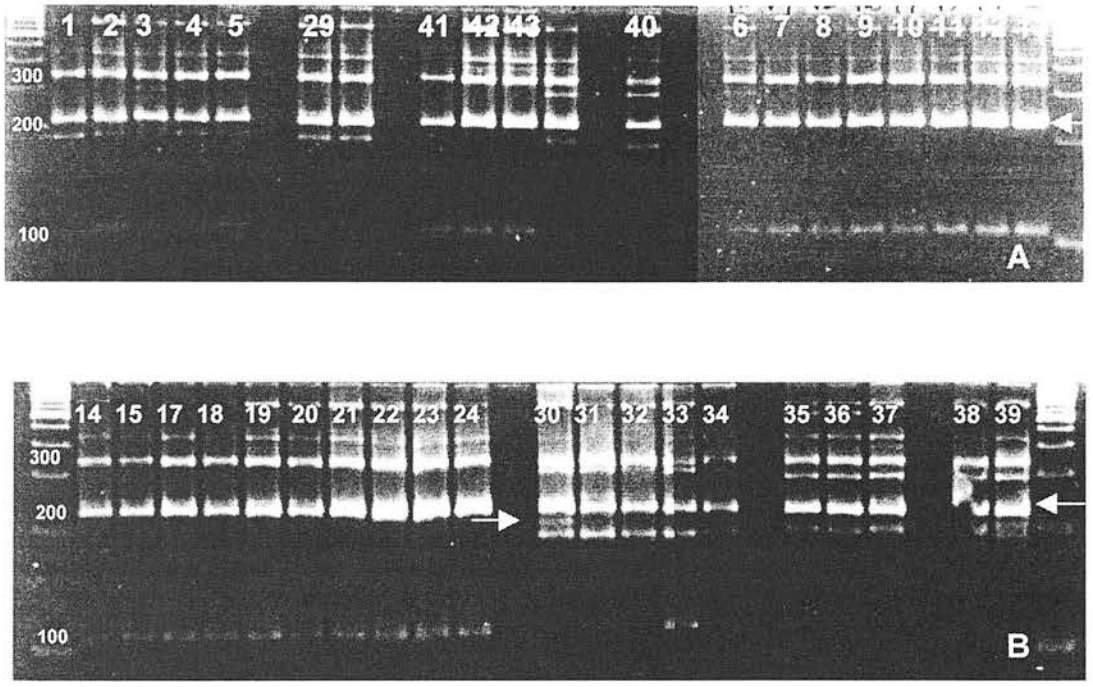
Plate 3.2 Agarose gel electrophoresis of PCR amplification products of *Lycopersicon* spp microsatellite locus LEPRP4.



Left lane = molecular marker in base pairs; numbers on the other lanes indicate *Lycopersicon* spp accessions in Table 2.1

In order to score the alleles unequivocally the dominant marker approach was adopted, scoring the bands of the same size as 1 if the band was present in each lane across the gel and 0 if the band was absent.

Plate 3.3 Agarose gel electrophoresis of PCR amplification products of *Lycopersicon* spp microsatellite locus LEGAST1.



Left lane = molecular marker in base pairs; numbers on the other lanes indicate *Lycopersicon* spp accessions in Table 2.1; arrows indicate locus position.

In this section (3.3.1.1) *L. esculentum* var. *cerasiforme* was not included in the analysis, because there was only one accession.

3.3.1.1.1 Polymorphic loci

In the results presented in Table 3.3.1, out of the 55 microsatellite loci assessed, 53 (96.36%) were polymorphic. *L. esculentum* had the largest number of polymorphic loci (52%), followed by *L. hirsutum* var. *glabratum* and *L. hirsutum* with 33% and 30%, respectively. *L. pimpinellifolium* (20%) and *L. pennellii* (18%) showed the lowest number.

Table 3.3.1 Number and proportion of polymorphic loci (*P*) in 6 species and 2 infraspecific categories belonging to the genus *Lycopersicon* based on 55 microsatellite marker data.

Species	Number of polymorphic loci	Proportion of polymorphic loci
L. esculentum	29	0.52
L. cheesmanii	14	0.25
L. cheesmanii var. *minor*	14	0.25
L. pimpinellifolium	11	0.20
L. parviflorum	16	0.29
L. hirsutum	17	0.30
L. hirsutum var. *glabratum*	18	0.33
L. pennellii	10	0.18

A non-parametrical statistical analysis utilising the Kruskal-Wallis test showed no significant differences between species (Details in Appendix 3, part 1). These differences could be considered as an index for the variability within species. However the proportion of polymorphic loci (*P*) does not reflect the real genetic variation in a population, because is very sensitive to the number of samples analysed.

3.3.1.1.2 Diversity indices

Diversity indices for the 38 accessions of 6 *Lycopersicon* spp and 2 infraspecific categories were calculated. Each accession was regarded as a sample, each taxon was considered as a population, with several accessions as samples. The species as a whole were treated as metapopulations allowing the calculation of each index as an overall. The statistical analyses of the indices were carried out using the ANOVA procedure for genetic indices, which were tested for normal distribution. These indices were the average gene diversity (H_S) and the Shannon's information index (I). Indices whose values are not distributed normally, such as the number of polymorphic alleles per locus (A) and the effective number of alleles (A_e), were

analysed with the non-parametric Kruskal-Wallis test. Results are given in Table 3.3.2, and details of the statistical analysis are given in Appendix 3, part 1.

ANOVA analyses produced significant differences between species for some genetic indices such as the number of polymorphic alleles per locus (A), the effective number of alleles (A_e) and the Shannon's information index (I). The results showed a significant difference of A, A_e and I only for *L. pennelli* in comparison with all other taxa. The mean number of alleles per locus (A) differed very significantly ($P<0.01$) between taxa and varied from 1.18 in *L. pennelli* to the higher value 1.53 presented by *L. esculentum* and *L. hirsutum* var. *glabratum*. Considering all the taxa together 1.96 (SE: ±0.03) alleles per locus are found in average, so there are more polymorphic alleles per locus between than within taxa of *Lycopersicon* spp. In the case of effective number of alleles (A_e), the means between taxa differed significantly ($P<0.05$) with a range from 1.12 in *L. pennellii* to 1.30, the highest value corresponding to *L. esculentum*. The average for all 6 species and 2 infraspecific categories was 1.48 (SE: ±0.04). The mean effective number of alleles, according to Hartl and Clark (1997), estimates the reciprocal of homozygosity. Based on this estimate, homozygosity of the *Lycopersicon* spp accessions was about 67%.

Table 3.3.2 Mean of 4 diversity indices for 6 *Lycopersicon* spp and 2 infraspecific categories based on 55 microsatellite markers.

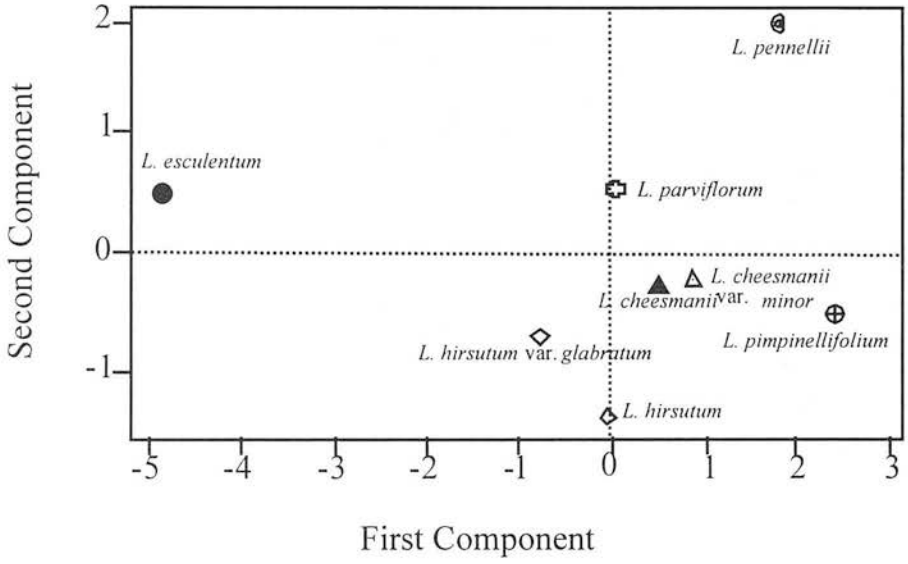
Species	N	A	A_e	H_s	I
		**	*	ns	*
L. esculentum	18	**1.53 a** (±0.07)	**1.30 a** (±0.05)	**0.18** (±0.03)	**0.27 a** (±0.04)
L. cheesmanii	3	**1.26 a** (±0.06)	**1.17 a** (±0.04)	**0.10** (±0.02)	**0.14 a** (±0.03)
L. cheesmanii var. *minor*	2	**1.26 a** (±0.06)	**1.18 a** (±0.04)	**0.11** (±0.03)	**0.18 a** (±0.04)
L. pimpinellifolium	3	**1.21 a** (±0.06)	**1.15 a** (±0.04)	**0.08** (±0.02)	**0.12 a** (±0.03)
L. parviflorum	3	**1.29 a** (±0.06)	**1.18 a** (±0.04)	**0.11** (±0.02)	**0.16 a** (±0.04)
L. hirsutum	2	**1.31 a** (±0.06)	**1.22 a** (±0.04)	**0.13** (±0.03)	**0.19 a** (±0.04)
L. hirsutum var. *glabratum*	3	**1.53 a** (±0.06)	**1.23 a** (±0.05)	**0.13** (±0.03)	**0.20 a** (±0.04)
L. pennellii	3	**1.18 b** (±0.05)	**1.12 b** (±0.04)	**0.07** (±0.02)	**0.10 b** (±0.03)

N = number of accessions; A = Number of polymorphic alleles per locus; A_e = Effective number of alleles; H_S = Average gene diversity; I = Shannon's information index; numbers between brackets correspond to standard error; significance * = $P<0.05$; ** = $P<0.01$; ns = no significance; same letters show no statistical differences.

The average gene diversity (H_S) did not differ significantly between species ranging from 0.07 to 0.18. The total gene diversity (H_t) of the taxa was 0.30 (SE: ±0.02), therefore the probability that 2 randomly sampled alleles in the whole sample are different is higher than 30%. However, Shannon's information index (I) showed significance between means ranging from 0.10 to 0.27 and a total value of 0.45 (SE: ±0.03) for all *Lycopersicon* spp.

A principal components analysis (PCA) was carried out based on genetic diversity indices (Figure 3.3.1). The first component clearly separates *L. esculentum* from the wild relatives, and the second component separates *L. pennellii* and *L. parviflorum* from the main group, but also from *L. esculentum*. These components explained 82.2% and 16.6% of the total variation at diversity indices level.

Figure 3.3.1 Principal components analysis of 5 genetic diversity indices in 6 *Lycopersicon* spp and 2 infraspecific categories based on 55 microsatellite markers.



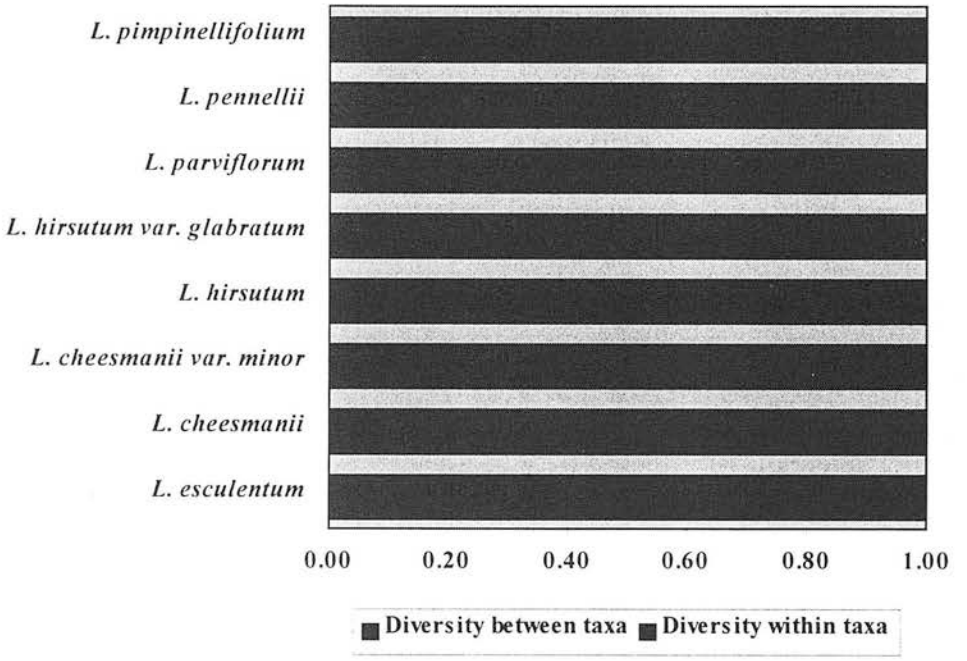
3.3.1.1.3 Genetic diversity within and between species

The total diversity (H_t) of species analysed can be divided into two fractions: diversity found between and within species. The fixation index (F_{st}) gives the relative amount of the total diversity that is found between species and can be expressed as a percentage. Data extracted from H_s (Table3.3.2) and H_t in diversity indices was used to obtain the F_{st} values for each species.

The results given in Figure 3.3.2 show that diversity between taxa ranged from 40% to 76%. Most of diversity present in tomato wild relatives was found between taxa. In contrast, in *L. esculentum* ocurred mostly within the taxa. Wild relatives, *L.*

pennellii and *L. pimpinellifolium* presented little diversity within taxa (24% and 28%), closely followed by *L. cheesmanii* (33%), but all other taxa were in the range from 36 % to 45%. Overall, the mean F_{ST} for all species was 61.74%.

Figure 3.3.2 Diversity partitioning within and between taxa in *Lycopersicon* based in 55 microsatellite markers.

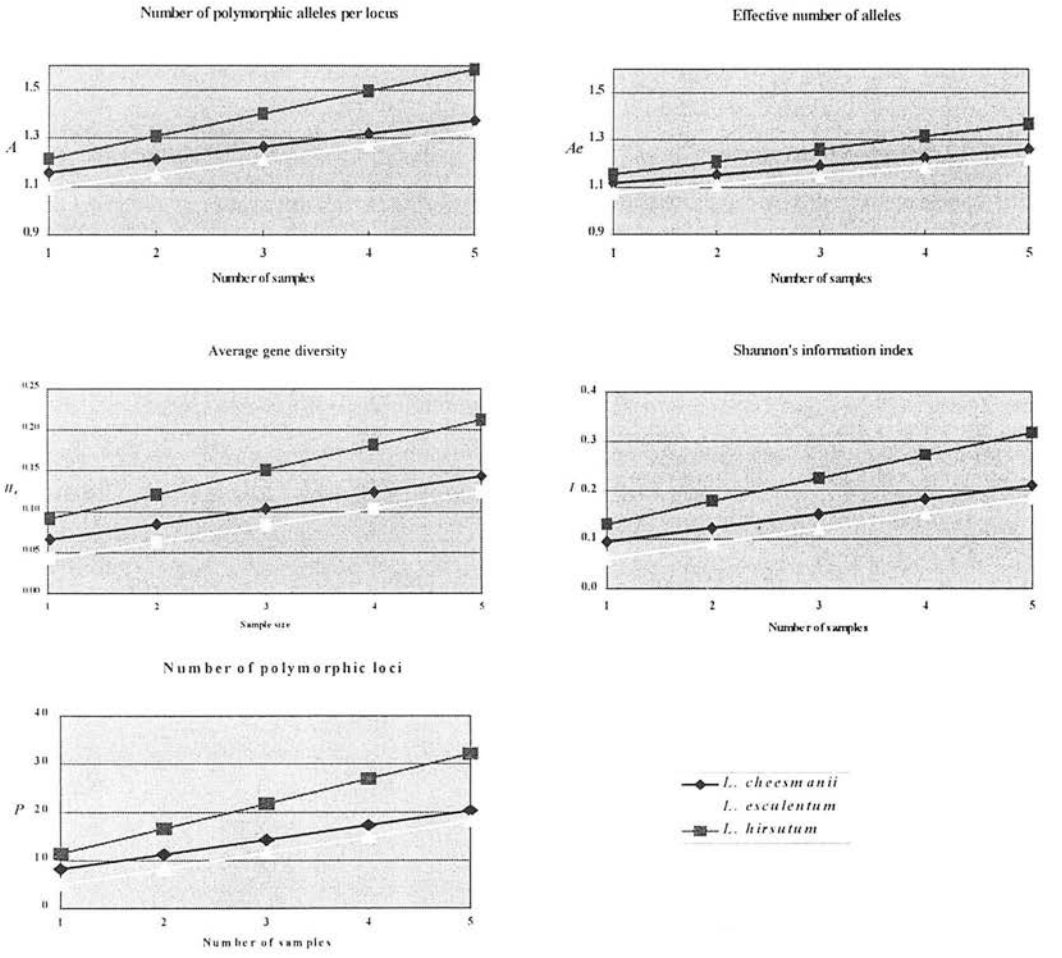


3.3.1.1.4 Genetic diversity relationships in respect to the number of accessions sampled

In order to determine how the genetic indices behaved in respect to the number of samples analysed, a regression analysis was carried out. As there were not enough samples to perform the regression in all species, only the three presenting more populations were selected: *L. hirsutum*, *L. cheesmanii*, and *L. esculentum*. The results of these relationships are presented in Figure 3.3.3 and the statistical analysis in Appendix 3, part 2. The five indices analysed (A, A_e, H_S, I, and P) showed a tendency for *L. hirsutum* to increase when the number of samples rose higher than any other species; *L. cheesmanii* is located at a lower level very close to *L. esculentum*. Slopes between *L. esculentum* and *L. cheesmanii* were similar in all

indices, but different from *L. hirsutum*. These results suggest that increasing the number of samples in the different species to be analysed, more genetic diversity could be found in accessions of *L. hirsutum* than *L. esculentum* and *L. cheesmanii*. It is probably that the inbreeders *L. esculentum* and *L. cheesmanii* have close levels of genetic diversity because of the bottlenecks they have undergone during their evolution.

Figure 3.3.3 Plot of the average number of polymorphic alleles per locus (*A*), effective number of alleles (A_e), gene diversity (H_s), Shannon's information index (*I*), and number of polymorphic loci (*P*), versus number of samples in *Lycopersicon* spp accessions.



3.3.1.2 Genetic similarity analysis associated to *Lycopersicon* spp accessions

3.3.1.2.1 Distribution of genetic similarities

The distribution of genetic similarities between *Lycopersicon* accessions is presented in Figure 3.3.4 and 3.3.5 that shows a histogram based on genetic similarities matrix (Appendix 2, part 2) for 38 accessions of *Lycopersicon* spp and for 18 accessions of *L. esculentum*, respectively. The purpose of this section is to examine and compare the genetic similarity distribution of *Lycopersicon* spp and *L. esculentum* accessions, with the aim to observe where these differences lay.

The mean genetic similarities among species utilised in this study was 0.38 (SE: ±0.01), while the distribution ranged from nearly 0.00 (distant) to a maximum of 1.00 (similar). The higher concentration of observations was located between 0.20 and 0.50. There was in the frequencies a tendency to skew toward the end with less genetic similarities.

In the case of *L. esculentum* accessions, the mean genetic similarities were 0.64 (SE: ±0.01), ranging from 0.34 to 1.00, and most observations were concentrated between 0.60 and 0.80, but skewed toward the end with most similarities.

Figure 3.3.4 Histogram of a genetic similarity matrix for 38 accessions of *Lycopersicon* spp based on 55 microsatellite markers.

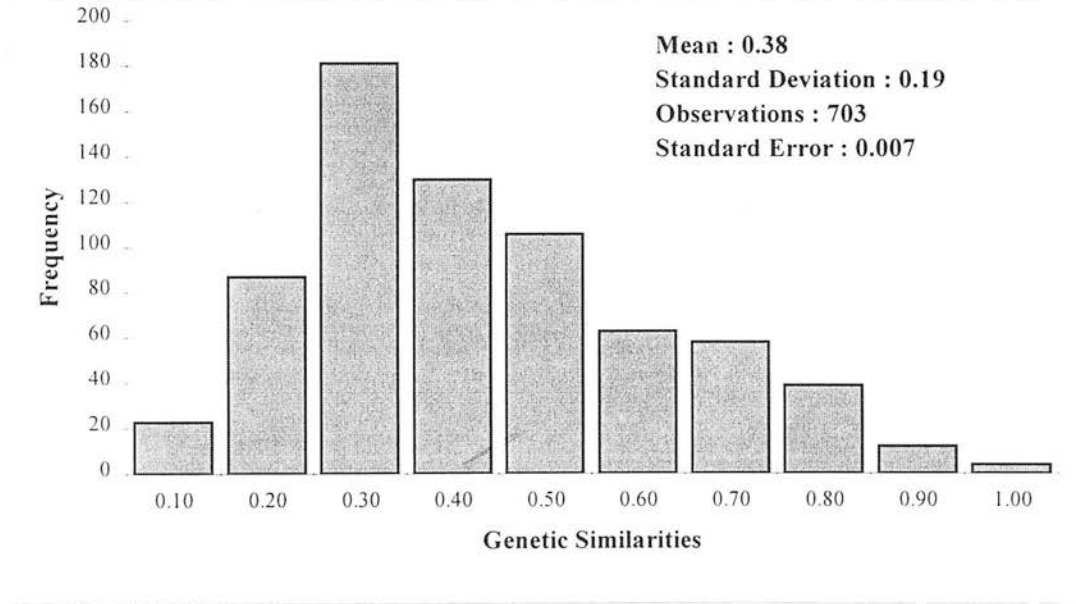
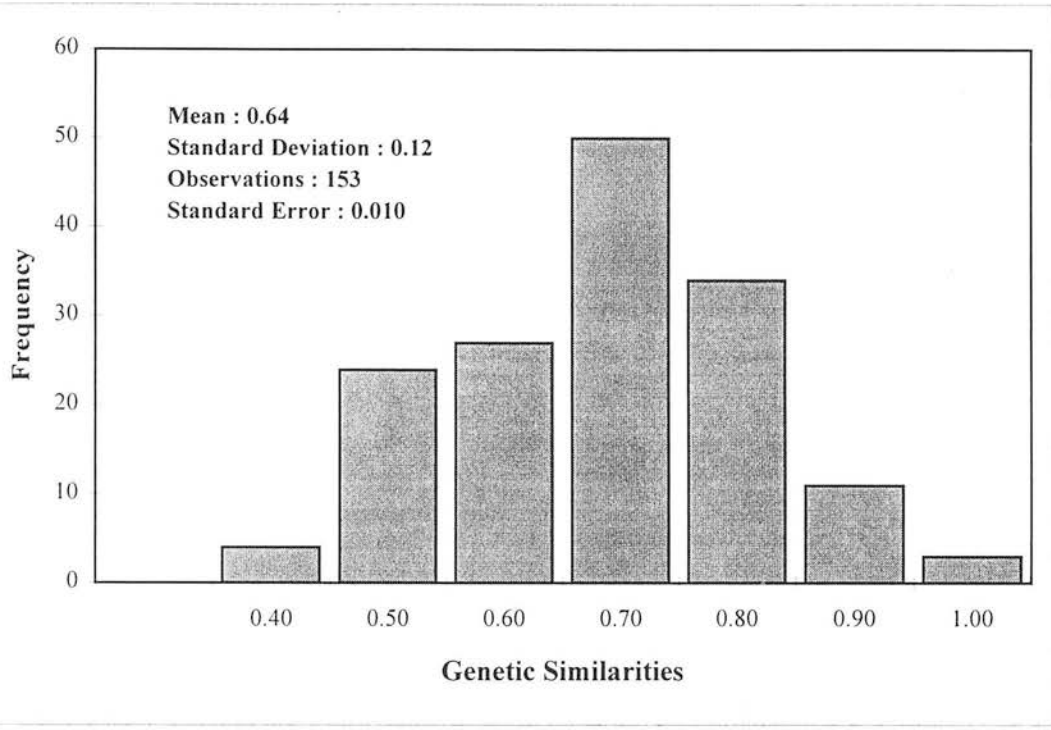


Figure 3.3.5 Histogram of a genetic similarity matrix for 18 accessions of *L. esculentum* based on 55 microsatellite markers.



3.3.1.2.2 Genetic similarity analysis within *Lycopersicon* spp accessions

The genetic similarity matrix (Appendix 2, part 2) generated from microsatellite data for the 38 accessions of *Lycopersicon* spp was analysed statistically within each species; *L. esculentum* var. *cerasiforme* was not included in this analysis because there was only one accession available. The results shown in Table 3.3.3 were analysed for descriptive statistics. Genetic similarities within species showed that *L. pennellii* and *L. pimpinellifolium* revealed accessions with similar values for microsatellite markers, and most dissimilar accessions were located within *L. hirsutum*. The closest distance between maximum and minimum was displayed by *L. pennellii* and *L. pimpinellifolium*. The higher mean corresponded to *L. esculentum* and the lower to *L. hirsutum*.

Table 3.3.3 Descriptive statistical analysis of a genetic similarity matrix for 55 microsatellite markers data within 6 *Lycopersicon* spp.

Species	N	Mean	SE	Max	Min
L. cheesmanii	10	0.56	±0.05	0.83	0.33
L. esculentum	153	0.64	±0.01	0.75	0.34
L. hirsutum	10	0.46	±0.03	0.65	0.31
L. parviflorum	3	0.62	±0.19	1.000	0.43
L. pennellii	3	0.63	±0.16	0.67	0.42
L. pimpinellifolium	3	0.57	±0.06	0.67	0.47

N = number of observations; SE = Standard Error; Max = Maximum value; Min = Minimum value.

3.3.1.3 Relationship between *Lycopersicon* spp accessions

A dendrogram based on the cluster analysis of a similarity coefficient matrix was constructed for all accessions of *Lycopersicon* spp utilised in this study. The cluster analysis was carried out based on these values by the unweighted pair group method with arithmetic average (UPGMA). As shown in Figure 3.3.6, most accessions of wild green-fruited taxa, species such as *L. hirsutum*, *L parviflorum* and *L. pennellii*, grouped together and are clearly separated from the red-fruited. However, one red-fruited accession of *L. parviflorum* (T1264/94 from IPK, Germany) and one green-fruited *L. hirsutum* var. *glabratum* (PI-199381 from USDA-ARS, USA) clustered out of their groups closer to red-fruited accessions and species. The only accession of *L. esculentum* var. *cerasiforme* grouped within the *L. esculentum* accessions; and *L. pimpinellifolium* entries were located adjacent to the *esculentum* group; next came the accessions of *L. cheesmanii*, showing their isolated evolution on the Galapagos Islands.

In order to obtain further information about the grouping of the wild relatives and cultivated accessions, a principal component analysis (PCA) was carried out utilising the similarity matrix. The PCA presented in Figure 3.3.7 reflected the relationship within and between wild species and *L. esculentum* entries that were also obtained in

the dendrogram. The first and second component could explain 18% and 6% of the variation, respectively. The first axis obviously classified wild species apart from the cultivated types (including *L. esculentum* var. *cerasiforme* and *L. parviflorum* entry T1264/94). The second axis separated cultivated tomato into F_1 hybrids and some modern open pollinated (OP) cultivars from old cultivars, both landraces were split one in each sector.

Figure 3.3.6 Dendrogram for *Lycopersicon* spp obtained using UPGMA based in similarity matrix from microsatellite data.

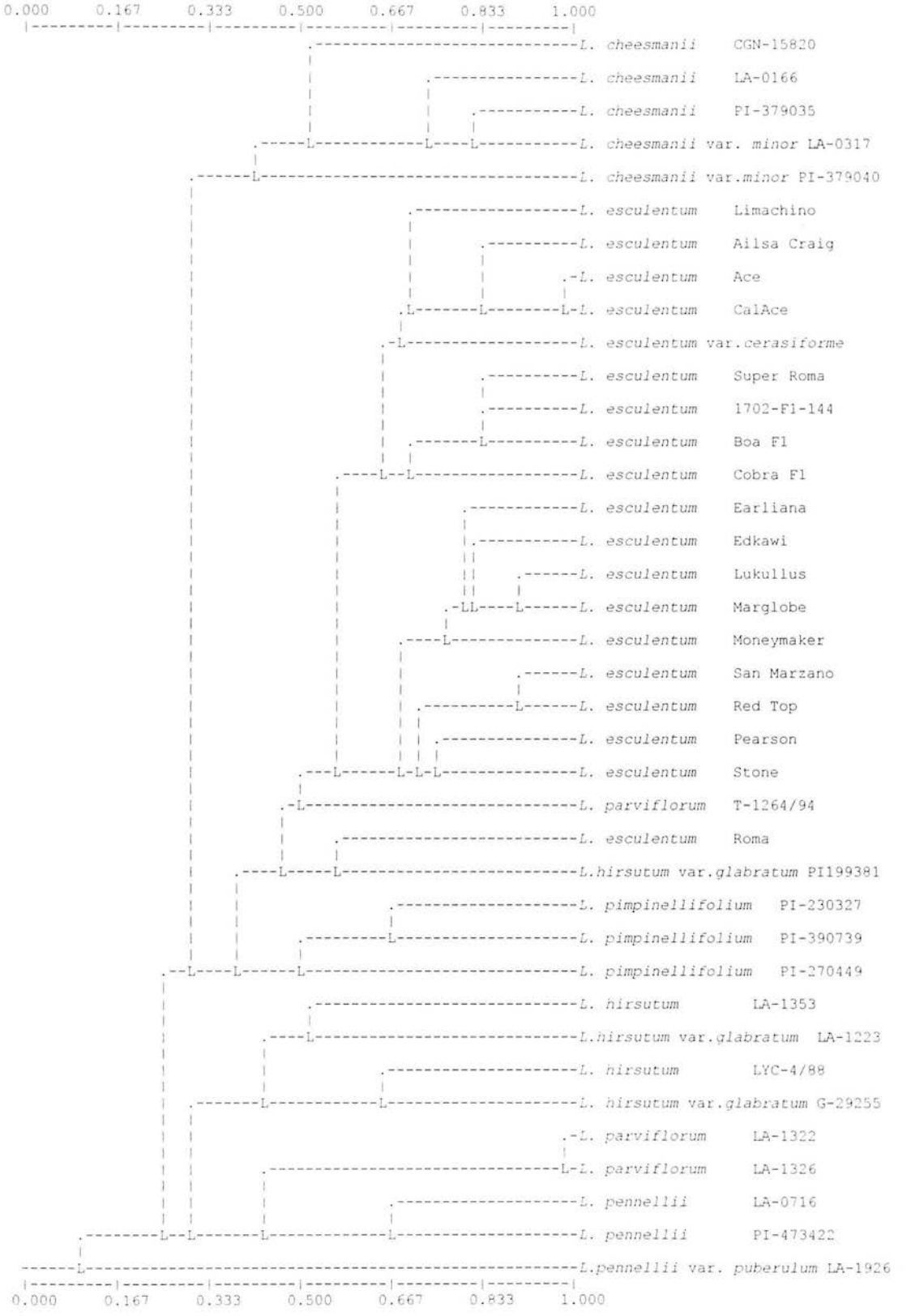
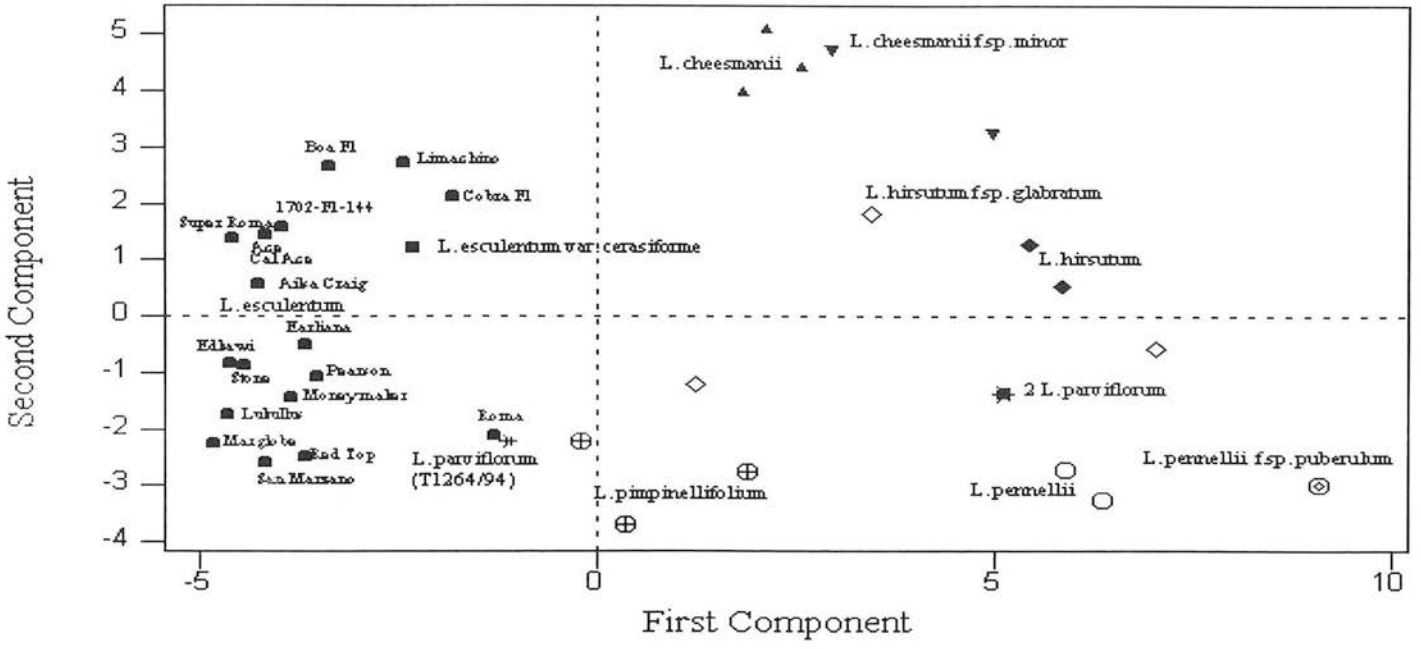


Figure 3.3.7 Principal components analysis (PCA) of 38 accessions of *Lycopersicon* based on a similarity matrix.



▲ *L. cheesmanii* ▼ *L. cheesmanii* var. *minor* ✱ *L. parvifloru* ♦ *L. hirsutum*
◇ *L. hirsutum var. glabratum* ⊕ *L. pimpinelifolium* ◉ *L. pennelli* var. *puberulum* ○ *L. pennellii*
⬟ *L. esculentum* ■ *L. esculentum* var. *cerasiforme*

Analysis of molecular variance (AMOVA) was conducted for 37 accessions of *Lycopersicon* spp utilising the similarity matrix (Excoffier *et al*, 1992). The results (Table 3.3.4) showed a highly significant value for Φ_{st} ($P<0.01$) and that only a 40.1% of the genetic variation was accounted for between species. The remaining 59.9% of the variation can be found within species.

Table 3.3.4 Analysis of molecular variance (AMOVA) of a genetic similarity matrix based in 55 microsatellite markers of 37 accessions of *Lycopersicon* spp.

	Sum of squares	df	Mean square	Percentage
Total within taxa	1,274.46	31	41.11	59.9%
Total between taxa	926.49	5	185.30	40.1%
TOTAL	2,201.95	36		

Variance within taxa	41.11
Variance between taxa	27.55
Φ_{st}	0.401

3.3.1.4 Genetic indices for red- and green-fruited species in *Lycopersicon* spp

A further population analysis, grouping the species in green- and red-fruited, was carried out to determine whether there were statistical differences between both groups. The results displayed in Table 3.3.5 did not provide statistical significance between both groups in any of the parameters. Details are presented in Appendix 3, part 3. Few differences were observed between groups, but most of the values were very close. In the case of the fixation index (F_{ST}), red-fruited species showed that 66% of the diversity lies between and 34% within species of the group, while in green-fruited 57% was between and 43% within species.

Table 3.3.5 Genetic diversity statistics and population partitioning parameters for 2 groups of *Lycopersicon* spp based on the fruit colour.

	P	A	A_e	H_s	I	F_{st}
	ns	ns	ns	ns	Ns	
Red-fruited species	**0.81**	**1.82** (±0.05)	**1.40** (±0.05)	**0.09** (±0.01)	**0.37** (±0.03)	**0.66**
Green-fruited species	**0.85**	**1.86** (±0.05)	**1.42** (±0.05)	**0.11** (±0.01)	**0.39** (±0.03)	**0.57**

P = proportion of polymorphic loci; A = number of polymorphic alleles per locus; A_e = effective number of alleles; H_S = average gene diversity; I = Shannon's information index; F_{St} = fixation index; numbers between brackets correspond to standard error; ns = no statistical significance.

3.3.2 Random amplified polymorphic DNA (RAPD) markers

Advances in the application of polymerase chain reaction (PCR) has made possible to score individuals at a large number of loci. A method developed simultaneously by Welsh and McClelland (1990), Williams *et al.* (1990), and Caetano-Anollés *et al.* (1991), later called random amplified polymorphic DNA or RAPD has been utilised for different purposes because of its simple and fast methodology, small amount of DNA required; each primer used has the potential to detect multiple bands, and the costs of utilising this technique are low. RAPD is one of the main techniques utilised for characterisation of germplasm (Hu and Quiros, 1991; Kresovich *et al.*, 1992; Wilkie *et al.*, 1993) and analysis of genetic diversity (Pejic *et al.*, 1998; Villand *et al.*, 1998; Mengistu *et al.*, 2000).

Polymorphisms in RAPD are the result of variations in the sequence of the primer-binding sites (e.g. point mutation), which impede stable linkage with the primer, or from indels (insertions/deletions) that change the band size. In respect to inheritance, they are transmitted mainly as dominant markers (Waugh and Powell, 1992), but also co-dominantly (Kawchuk *et al.*, 1994). However, most often they are treated as dominant marker, because if one allele at a RAPD site is unamplifiable, then the marker/marker homozygote cannot be distinguished from the marker/null heterozygote. Provided there is only a single amplifiable allele per locus, this does

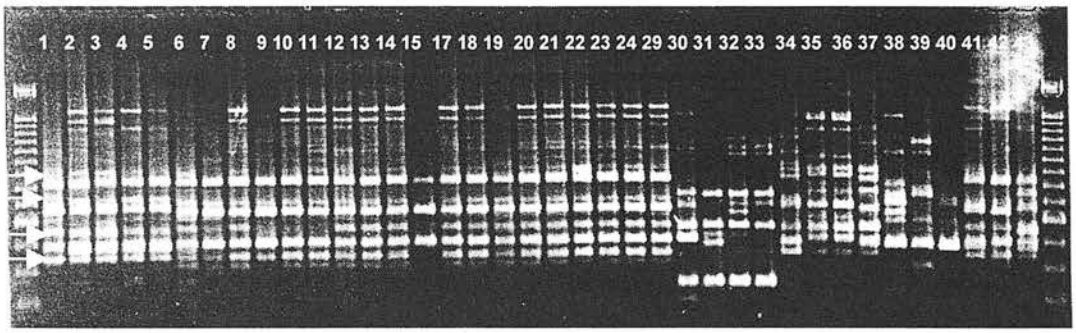
not bias the estimation of allele frequencies necessary for population genetic analysis, but it does reduce the accuracy of such estimation relative to analysis with co-dominant markers (Lynch and Milligan, 1994).

In this part of the study RAPD are treated as dominant markers to establish relationships and genetic similarities between and within 38 accessions of *Lycopersicon*.

3.3.2.1 Population structure and diversity in *Lycopersicon* spp accessions

In this analysis RAPD bands were treated as dominant markers being evaluated as presence (1) or absence (0). Plate 3.4 shows the results of an agarose gel electrophoresis of PCR products from RAPD primer OPA-16 and the arrows mark some of the polymorphic loci. Bands of the same size were scored as 1 if the band was present in each lane across the gel and 0 if the band was absent.

Plate 3.4 Agarose gel electrophoresis of PCR products from RAPD markers primer OPA-16 in *Lycopersicon* spp accessions.



Left and right lanes correspond to molecular markers; arrows indicate examples of polymorphic loci; numbers indicate the accessions in Table 2.1; 1 to 5 = *L. cheesmanii*; 6 to 24 = *L. esculentum*; 29 = *L. esculentum* var. *cerasiforme*; 30 to 34 = *L. hirsutum*; 35 to 37 = *L. parviflorum*; 38 to 40 = *L. pennellii*; 41 to 43 = *L. pimpinellifolium*.

3.3.2.1.1 Polymorphic loci

The 28 primers used to screen the 38 *Lycopersicon* accessions produced a total of 268 amplified DNA fragments. The number and proportion of polymorphic loci (*P*) within species are displayed in Table 3.3.6 (Details in Appendix 3, part 4). *L.*

esculentum presented the highest, with 162 out of 268 (62%), and *L. cheesmanii* var. *minor* the lowest, with 40 out of 268 (15%), polymorphic bands. The non-parametric Kruskal-Wallis test showed no statistical significance for this index. In general, most species showed fairly low proportions, between 20% and 30%, of polymorphic loci. However, overall 262 out of 268 bands (97.76%) were polymorphic.

Table 3.3.6 Number and proportion of polymorphic loci (*P*) in 6 species and 2 infraspecific categories belonging to the genus *Lycopersicon* based on 268 RAPD markers.

Species	Number of Polymorphic loci	Proportion of polymorphic loci
L. esculentum	167	0.62
L. cheesmanii	90	0.33
L. cheesmanii var. *minor*	40	0.15
L. pimpinellifolium	75	0.28
L. parviflorum	79	0.29
L. hirsutum	56	0.21
L. hirsutum var. *glabratum*	92	0.34
L. pennellii	64	0.24

3.3.2.1.2 Diversity indices

Four genetic diversity indices (number of polymorphic alleles per locus (A), effective number of alleles (A_e), average gene diversity (H_S), and Shannon's information index (I)) are presented in Table 3.3.7. The structure of the analysis was similar to that utilised for microsatellites (3.3.1.1.2), ANOVA was carried out on normally distributed indices such as H_S and I, and the non-parametric Kruskal-Wallis procedure for non-normally distributed indices such as A and A_e. Details of the statistical analysis are presented in Appendix 3, part 4. All indices analysed differed statistically between taxa ($P<0.01$). A range between 1.16 to 1.62 of average polymorphic alleles per locus was found. Considering all taxa as a metapopulation A

was 1.98 (±0.01), confirming the results obtained with microsatellites that there are more polymorphic alleles per locus in average between than within taxa. *L. esculentum* and *L. cheesmanii* var. *minor* were statistically different from all other taxa for this index.

The mean effective number of alleles (A_e) for all 8 taxa was 1.55 (±0.02), with a range from 1.11 to 1.32. Considering A_e as the inverse of homozygosity, then the most homozygous species was *L. cheesmanii* var. *minor* with 90%. The mean Nei (Nei, 1987) average gene diversity (H_S) for 8 taxa was 0.12 (±0.01), the average diversity of species varying from 0.07 (*L. cheesmanii* var. *minor*) to 0.19 (*L. esculentum*). The total gene diversity (H_t) in the entire sample was 0.32 (±0.01). The species *L. hirsutum* and *L. pennellii* were very close for H_S. The Tukey's test indicated that *L. esculentum* was statistically different from all the other taxa, as it also was *L. cheesmanii* var. *minor*. The other taxa showed no differences. Shannon's information index (I) was low across the species, the mean for 8 species was 0.18 (±0.02) with values fluctuating from 0.10 to 0.29, and I as an overall of species was 0.14 (±0.01).

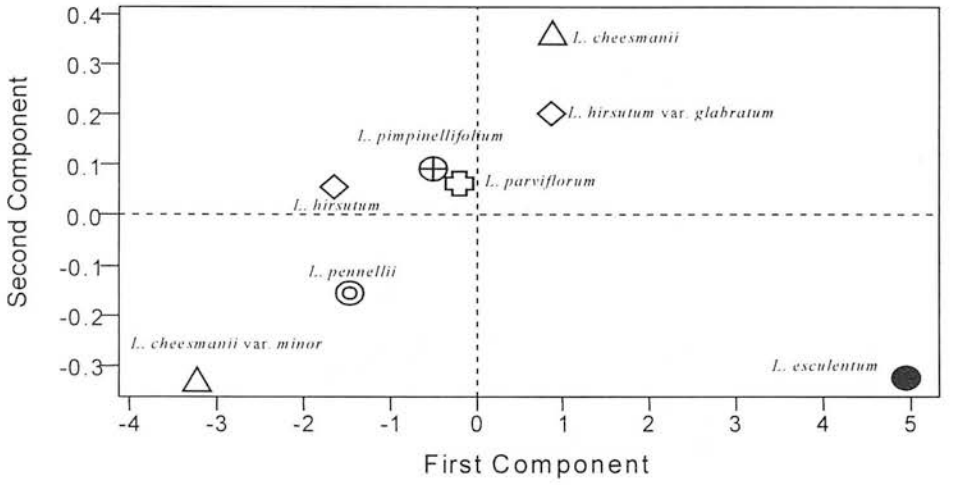
A principal component analysis was performed for 6 genetic indices for 8 taxa (Figure 3.3.8). The first component separated *L. esculentum* from the other *Lycopersicon* spp, but also *L. cheesmanii* var. *minor* in the other side of the plot. Most of the taxa grouped close together, but *L. cheesmanii* was located separately in the upper part from all other taxa. The first component explains 99% of the total variation.

Table 3.3.7 Mean of 4 diversity indices for 6 *Lycopersicon* spp and 2 infraspecific categories based in 268 RAPD markers.

Species	n	A	A_e	H_s	I
		**	**	**	**
L. esculentum	18	**1.62 a** (±0.03)	**1.32 a** (±0.02)	**0.19 a** (±0.02)	**0.29 a** (±0.02)
L. cheesmanii	3	**1.35 bc** (±0.03)	**1.24 bc** (±0.02)	**0.14 bc** (±0.02)	**0.20 bc** (±0.02)
L. cheesmanii var. *minor*	2	**1.16 bd** (±0.02)	**1.11 bd** (±0.02)	**0.07 bd** (±0.01)	**0.10 bd** (±0.01)
L. pimpinellifolium	3	**1.29 bc** (±0.03)	**1.20 bc** (±0.02)	**0.11 bc** (±0.01)	**0.17 bc** (±0.02)
L. parviflorum	3	**1.30 bc** (±0.03)	**1.20 bc** (±0.02)	**0.12 bc** (±0.01)	**0.17 bc** (±0.02)
L. hirsutum	2	**1.23 bc** (±0.03)	**1.16 bc** (±0.02)	**0.09 bc** (±0.01)	**0.14 bc** (±0.02)
L. hirsutum var. *glabratum*	3	**1.36 bc** (±0.03)	**1.23 bc** (±0.03)	**0.14 bc** (±0.01)	**0.20 bc** (±0.02)
L. pennellii	3	**1.25 bc** (±0.03)	**1.16 bc** (±0.02)	**0.09 bc** (±0.01)	**0.14 bc** (±0.02)

n = number of accessions; A = Number of polymorphic alleles per locus; A_e = Effective number of alleles; H_S = Average gene diversity; I = Shannon's information index; numbers between brackets correspond to standard error; significance * = $P<0.05$; ** = $P<0.01$; ns = no significance; same letters show no statistical differences.

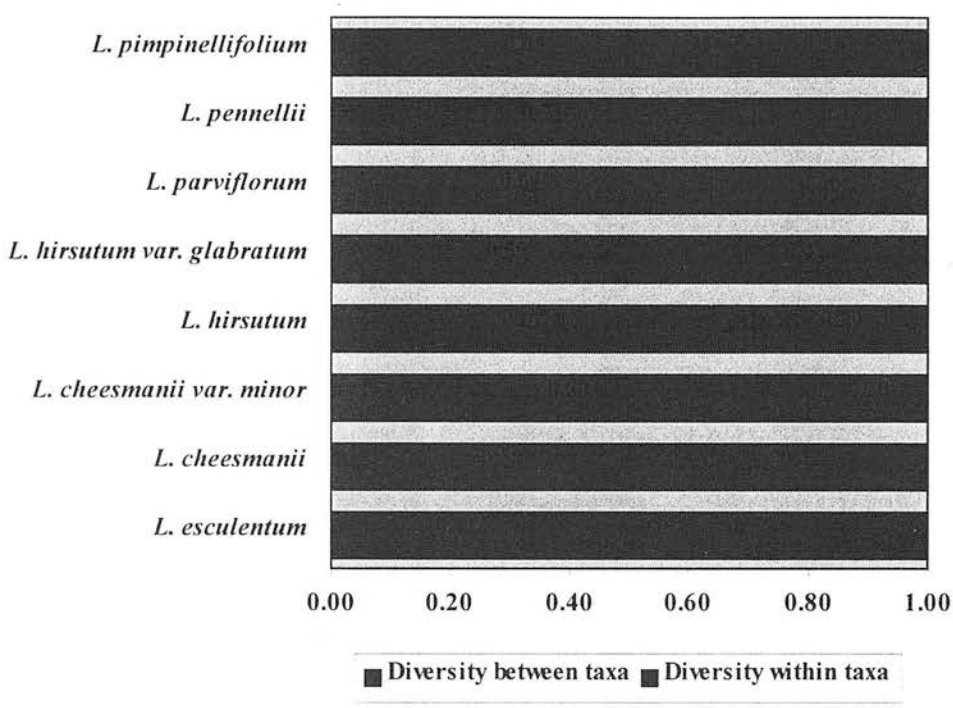
Figure 3.3.8 Principal components analysis for 6 genetic diversity indices in 6 *Lycopersicon* spp and 2 infraspecific categories based in 268 RAPD markers.



3.3.2.1.3 Genetic diversity partitioning

F-statistics for the 6 species and 2 infraspecific categories are given in Figure 3.3.9. The F_{st} explains that most of the diversity found in all species was due to differences between species, except in *L. esculentum* where 58.2% of the diversity lay within species. In contrast the F_{st} value for *L. cheesmanii* var. *minor* showed that 80% of the diversity was found between species.

Figure 3.3.9 Diversity partitioning within and between taxa in *Lycopersicon* based on 268 RAPD markers.

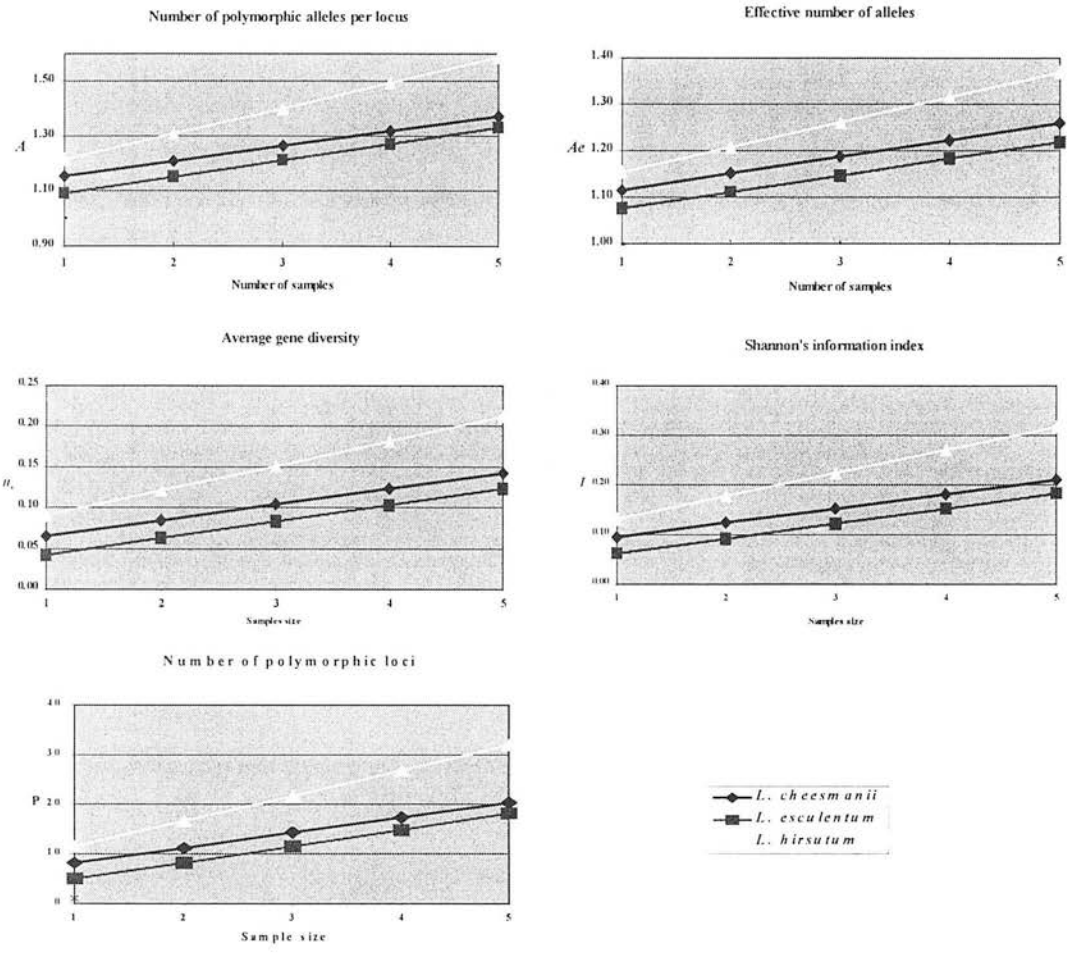


3.3.2.1.4 Genetic diversity relationship in respect to the number of accessions sampled

The relationship between genetic diversity indices, based in 268 RAPD markers, and number of accessions sampled was studied using analysis of regression for each index and species. In this part of the research only three species were analysed, the other species were disallowed because of the low number of samples which would make any extrapolation of the results and statistical analysis difficult.

The results are displayed in Figure 3.3.10 and the statistical analysis in Appendix 3, part 5. Most regressions displayed in the plots were statistically very significant ($P<0.01$), but *L. cheesmanii* in H_S and *L. esculentum* in A_e were not statistically significant. For the five indices, *L. hirsutum* presented very steep slope in comparison to a shallow *L. esculentum* and *L. cheesmanii.* These results were very similar to those obtained with microsatellites and confirm the suggestion that there are more genetic diversity in *L. hirsutum* than in the other species examined, when the number of samples is increased.

Figure 3.3.10 Plot of average number of polymorphic alleles per locus (A), effective number of alleles (A_e), average gene diversity (H_s), Shannon's information index (I), and proportion of polymorphic loci (P) based in 268 RAPD markers versus number of samples in *Lycopersicon* accessions.



3.3.2.2 Genetic similarities based on RAPD data associated to *Lycopersicon* spp.

3.3.2.2.1 Distribution of genetic similarities

Figure 3.3.11 is a histogram of the distribution of genetic similarities based on the matrix presented in Appendix 2 (part 3), between 38 accessions belonging to 6 species of *Lycopersicon*. Mean genetic similarity between species was 0.41 (SE: ±0.005), the distribution ranged from 0.18 to 0.88, and a concentration of similarities was found in the region of 0.30 and 0.50.

Figure 3.3.11 Histogram of genetic similarities between 38 accessions of *Lycopersicon* spp based on 268 RAPD markers.

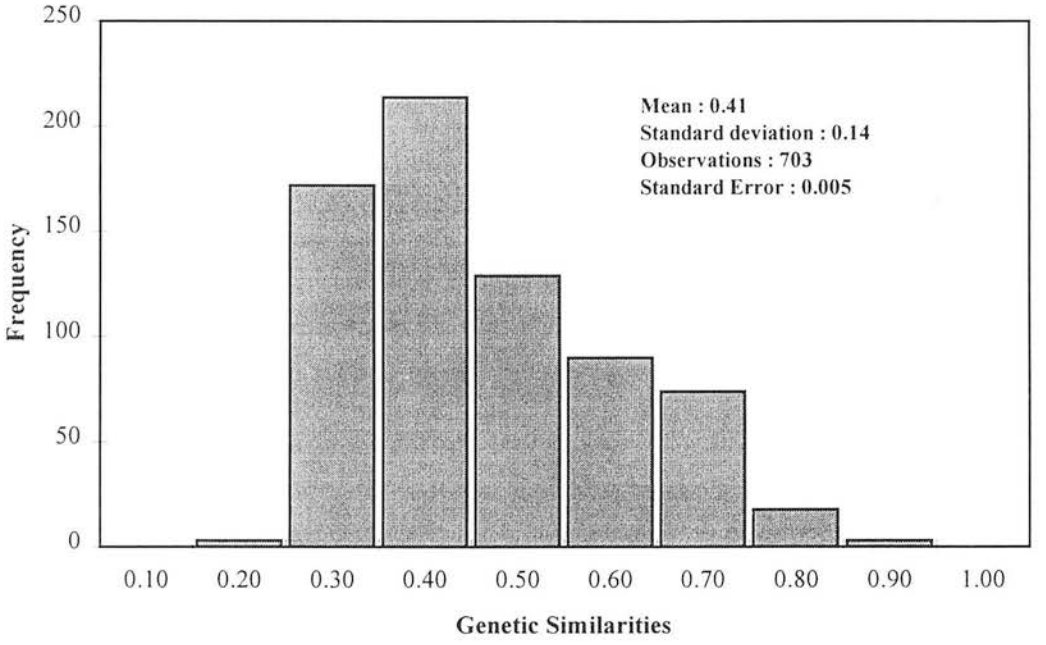
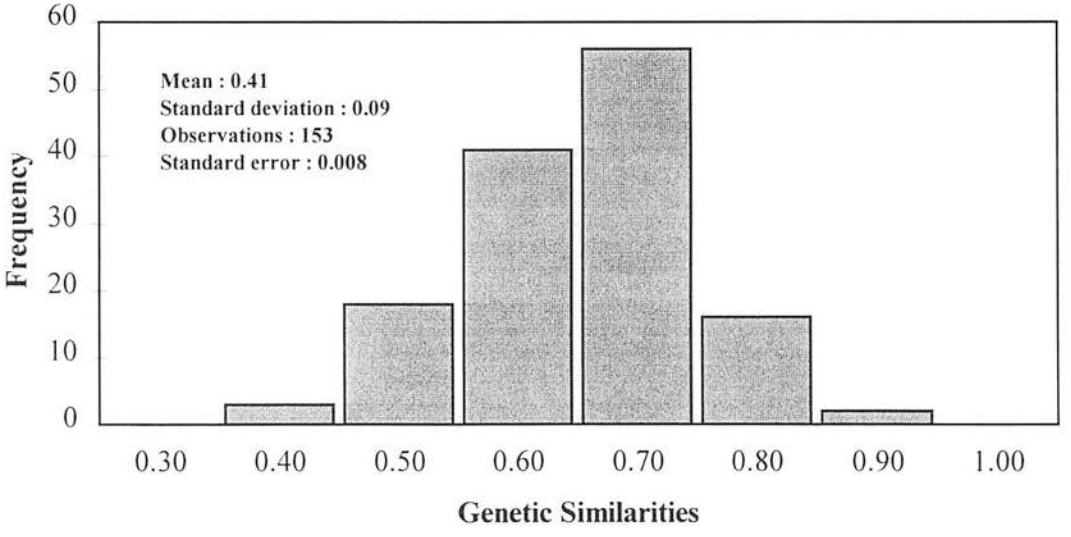


Figure 3.3.12 Histogram of genetic similarities between 18 accessions of *L. esculentum* based on 268 RAPD markers.



A second histogram (Figure 3.3.12) shows the genetic similarities between 18 *L. esculentum* accessions based on the same matrix as quoted before. The mean between accessions was 0.59 (SE: ±0.008), the range fluctuation was from 0.33 to 0.88, and most of the observations were between 0.60 and 0.70.

3.3.2.2.2 Genetic similarity analysis within *Lycopersicon* spp accessions.

Each species combination of similarities was grouped and analysed using descriptive statistics. In this analysis *L. esculentum* var. *cerasiforme* was not included because there was only one accession, and so it was not possible to perform any combinations. Results are displayed in Table 3.3.8, it is not possible statistically to compare these means between species because they are the result of genetic similarities within each species. There was not much variation for genetic similarities within species, but *L. hirsutum* and *L. pennelli* showed lower genetic similarities within each species. The greatest difference between maximum and minimum value was presented by *L. esculentum*.

Table 3.3.8 Descriptive statistical analysis of genetic similarity matrix of 268 RAPD markers within 6 *Lycopersicon* spp.

Species	N	Mean	SE	Max	Min
L. cheesmanii	10	0.58	±0.02	0.69	0.50
L. esculentum	153	0.59	±0.01	0.88	0.33
L. hirsutum	10	0.44	±0.04	0.64	0.29
L. parviflorum	3	0.58	±0.12	0.82	0.43
L. pennellii	3	0.54	±0.04	0.61	0.49
L. pimpinellifolium	3	0.61	±0.03	0.66	0.57

N = number of observations; SE = Standard Error; Max = Maximum value; Min = Minimum value.

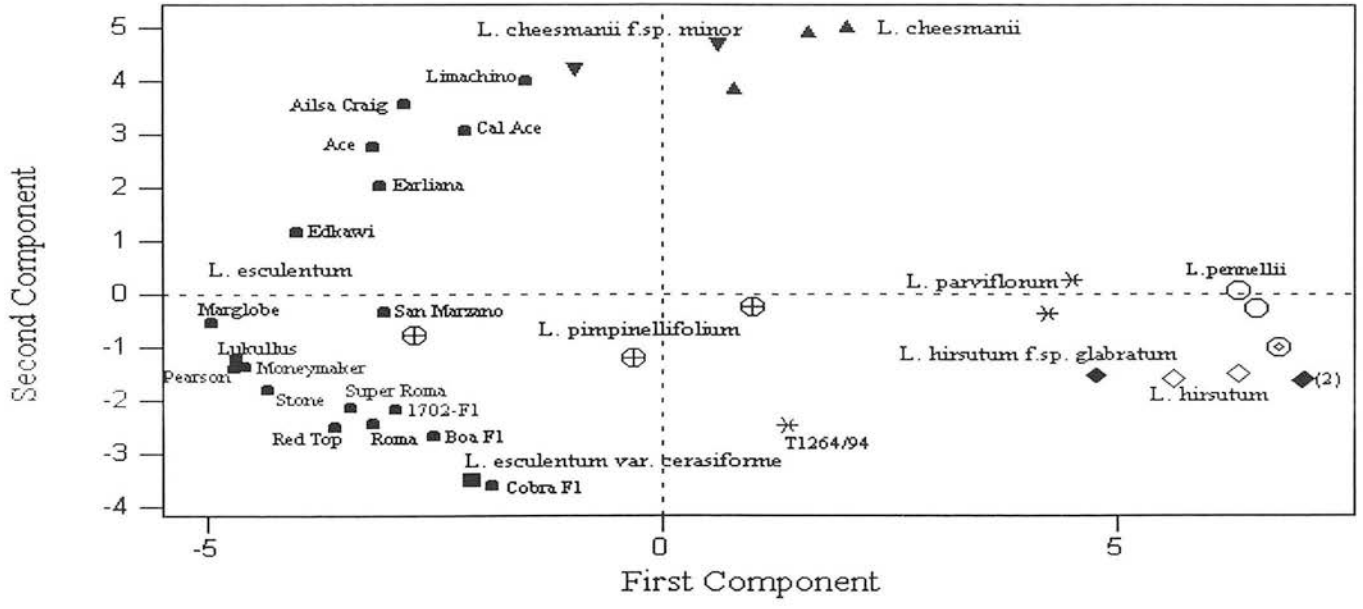
3.3.2.3 Relationship between *Lycopersicon* spp accessions based in RAPD markers.

Using genetic similarity values, a dendrogram was constructed to visualise the relative relatedness among *Lycopersicon* accessions. The cluster analysis was performed utilising the UPGMA method. Figure 3.3.13 shows all accessions clustering within their respective species, except *L. hirsutum* fo. *glabratum* (PI-199381) which is closer to the *L. parviflorum* cluster. Of the eight taxa, *L. esculentum* var. *cerasiforme* was closely linked to *L. esculentum*, and all other red-fruited species formed a major cluster separated from green-fruited species (*L. pennelli*, *L. hirsutum*, and *L. parviflorum*). This cluster of green-fruited taxa included the red-fruited *L. parviflorum* T1264/94.

In order to discover more relationships in the grouping of the different *Lycopersicon* spp accessions, a PCA analysis was carried out using the similarity matrix as raw data (Appendix 2, part 3) (Figure 3.3.14). The PCA shows similar relationships between accessions as inferred from the dendrogram. The first and second components explained 16.0% and 6.5% of the variation between entries, respectively. The first axis divided *L. esculentum* from most of the wild relatives, except *L. esculentum* var. *cerasiforme*, two entries of *L. pimpinellifolium*, and one from *L. cheesmanii* var. *minor*. The remaining accession of *L. pimpinellifolium*, one belonging to *L. cheesmanii*, and one *L. cheesmanii* var. *minor* were located closer to the limit between both groups.

Figure 3.3.13 Dendrogram for *Lycopersicon* spp obtained using UPGMA based in similarity matrix from 268 RAPD markers.

Figure 3.3.14 Principal components analysis (PCA) plot of 38 accessions of *Lycopersicon* based on similarity matrix of 268 RAPD markers.



▲ *L. cheesmanii*
▼ *L. cheesmanii* var. *minor*
✱ *L. parviflorum*
♦ *L. hirsutum*
◇ *L. hirsutum* var. *glabratum*
⊕ *L. pimpinelifolium*
⊙ *L. pennelli* var. *puberulum*
O *L. pennellii*
⬟ *L. esculentum*
■ *L. esculentum* var. *cerasiforme*

The second axis divided tomato cultivars into two groups, one containing two landraces, some old cultivars and the very closely related cultivars Ace and Cal Ace, and the other group containing all the other cultivars, including *L. esculentum* var. *cerasiforme*. Similarly, *L. cheesmanii* accessions were isolated in the upper part of the plot, clearly apart from the other wild types and close to *L. esculentum* accessions, mostly old cultivars and landraces.

An analysis of molecular variance (AMOVA) was carried out over 37 accessions of *Lycopersicon* spp utilising the similarity matrix produced from RAPD markers (Table 3.3.9). In this analysis *L. esculentum* var. *cerasiforme* was not considered because there was just one accession growing. The results that show a highly statistical significance for Φ_{st} ($P<0.01$) and only 40.1% of the variation was due to differences found between species, while 59.9% of the variation was found within species.

Table 3.3.9 Analysis of molecular variance (AMOVA) of genetic similarities matrix based on 268 RAPD markers of 37 *Lycopersicon* spp accessions.

	Sum of squares	df	Mean square	Percentage
Total within taxa	33,515.74	31	1,081.15	59.9%
Total between taxa	26,629.86	5	5,325.97	40.1%
TOTAL	60,145.60	36		

Variance within taxa	1,081.15
Variance between taxa	811.25
Φ_{st}	0.401

3.3.2.4 Genetic indices for red- and green-fruited species in *Lycopersicon* spp.

An analysis was carried out on different species grouped by their fruit colour (red – and green-fruited) based on the similarity matrix presented in Appendix 2, part 3. The statistical analysis for significance between both groups was divided between normally distributed indices (H_S and I) utilising ANOVA procedure and non-normally distributed (P, A, A_e) utilising the non-parametric Kruskal-Wallis test. Results are presented in Table 3.3.10 and details of the statistical analysis in Appendix 3, part 6. In this analysis no index displayed statistical significances between groups, most values between indices were very close, and few differences were observed between both red- and green-fruited species. The F_{ST} shows that 19% of the diversity in red-fruited species lay between species and 81% within species, while in green-fruited species 9% of the variation was found between and 91% within species.

Table 3.3.10 Genetic diversity indices and partitioning parameters for two groups of *Lycopersicon* spp based in fruit colour utilising 268 RAPD markers.

	P	A	A_e	H_s	I	F_{st}
	ns	**ns**	**Ns**	**ns**	**ns**	
Red-fruited species	**0.86**	**1.86** (±0.02)	**1.44** (±0.02)	**0.26** (±0.01)	**0.40** (±0.01)	**0.19**
Green-fruited species	**0.85**	**1.85** (±0.02)	**1.51** (±0.02)	**0.29** (±0.01)	**0.44** (±0.02)	**0.09**

P = proportion of polymorphic loci; A = number of polymorphic alleles per locus; A_e = effective number of alleles; H_S = average gene diversity; I = Shannon's information index; F_{ST} = fixation index; numbers between brackets correspond to standard error; ns = no statistical significance.

3.4 Genetic variability within *L. esculentum* accessions

In order to analyse the genetic diversity present in *L. esculentum* accessions and to investigate contribution to the *esculentum* gene-pool, the accessions were grouped in landraces, old varieties (vintage), modern varieties open-pollinated (OP), and modern F_1 hybrids.

3.4.1 Genetic diversity present in accessions of *L. esculentum*

Means of diversity indices for the four groups of *L. esculentum* accessions are given separately for both genetic markers, microsatellites and RAPD (Table 3.4.1). Statistical analysis of each genetic index, per molecular marker group, was carried out between the four groups of tomato. The statistical analysis of the results was performed utilising the ANOVA procedure for genetic indices which were predominantly normally distributed as H_S and I. Non-normally distributed indices such as A and A_e, were analysed with the non-parametric Kruskal-Wallis test. In the case of F_{St} and P no statistical analysis could be performed because there were no replications. Details of the statistical analysis are presented in Appendix 3, part 7.

The results show no statistical significance for any indices in microsatellite markers. In respect to A, there was not a big difference between means, but F_1 hybrids presented less polymorphic alleles per locus than the other groups. A_e showed that there was a high homozygosity for all groups of *L. esculentum*, between 79% to 90%. Old and modern OP varieties presented more polymorphic loci than landraces and F_1 hybrids. F_{st} means show F1 hybrids with a higher value and modern varieties OP with a lower.

However, for RAPD marker analysis, all genetic indices means analysed presented high significance ($P<0.01$). The number of polymorphic alleles per locus (A) showed slightly higher values for RAPD than for microsatellites, and also for A_e a decrease in homozygosity values (ranging from 88% to 73%). In respect to H_S and I, there were no great differences between RAPD and microsatellites data. RAPD gave consistently higher F_{St} and P values than microsatellites did.

Table 3.4.1 Mean genetic diversity indices (A, A_e, H_s, I, F_{st}, and P) of grouped *L. esculentum* accessions in landraces, old varieties, modern varieties OP, and F_1 hybrids for two different genetic markers.

	A	A_e	H_s	I	F_{st}	P
Microsatellites	ns	ns	ns	ns		
Landraces	**1.26** (±0.06)	**1.18** (±0.04)	**0.11** (±0.02)	**0.15** (±0.04)	**0.38**	**0.26**
Old varieties	**1.35** (±0.06)	**1.20** (±0.05)	**0.12** (±0.02)	**0.18** (±0.04)	**0.32**	**0.35**
Modern varieties OP	**1.40** (±0.07)	**1.27** (±0.05)	**0.15** (±0.03)	**0.23** (±0.04)	**0.11**	**0.40**
F_1 hybrids	**1.16** (±0.05)	**1.11** (±0.04)	**0.06** (±0.02)	**0.09** (±0.03)	**0.63**	**0.16**
RAPD	**	**	**	**		
Landraces	**1.19 bd** (±0.03)	**1.13 b** (±0.02)	**0.08 b** (±0.01)	**0.11 b** (±0.02)	**0.60**	**0.18**
Old varieties	**1.47 a** (±0.04)	**1.26 a** (±0.03)	**0.15 a** (±0.02)	**0.23 a** (±0.02)	**0.20**	**0.48**
Modern varieties OP	**1.39 ac** (±0.04)	**1.26 a** (±0.03)	**0.15 a** (±0.02)	**0.22 a** (±0.02)	**0.23**	**0.41**
F_1 hybrids	**1.24 bc** (±0.03)	**1.37 b** (±0.02)	**0.09 b** (±0.01)	**0.13 b** (±0.02)	**0.53**	**0.24**

A = number of polymorphic alleles per locus; A_e = effective number of alleles; H_S = average gene diversity; I = Shannon's information index; F_{st} = fixation index; P = proportion of polymorphic loci; ** = high statistical significance (P<0.01); ns = no significance; same letters show no statistical differences.

3.4.2 Genetic distances between *L. esculentum* accessions

Genetic similarity between groups within *L. esculentum* obtained from microsatellite data was high, as shown in Table 3.4.2. The most similar groups were old varieties and modern varieties OP, while the most distant relationship was F_1 hybrids and old varieties. Table 3.4.3 shows the genetic distance/similarity of the four tomato groups,

but based in RAPD markers data. Again the closest link was between old varieties and modern varieties OP, and the most distant, F_1 hybrids from landraces.

Table 3.4.2 Genetic similarity and genetic distance between landraces, old varieties, modern varieties OP, and F_1 hybrids of *L. esculentum* based on 55 microsatellite markers.

populations	1	2	3	4
1 Landraces	****	0.93	0.95	0.88
2 Old varieties	0.07	****	0.97	0.83
3 Modern varieties OP	0.05	0.03	****	0.87
4 F_1 hybrids	0.13	0.19	0.14	****

Genetic identity (above diagonal) and genetic distance (below diagonal).

Table 3.4.3 Genetic similarity and genetic distance between landraces, old varieties, modern varieties OP, and F_1 hybrids of *L. esculentum* based on 168 RAPD markers.

populations	1	2	3	4
1 Landraces	****	0.89	0.86	0.75
2 Old varieties	0.11	****	0.95	0.88
3 Modern varieties OP	0.14	0.06	****	0.87
4 F_1 hybrids	0.25	0.13	0.14	****

Genetic identity (above diagonal) and genetic distance (below diagonal).

3.4.3 Relationship between grouped *L. esculentum* accessions

Utilising the genetic distance values a dendrogram was constructed in order to visualise relative genetic relatedness among *L. esculentum* groups. For each molecular marker and utilising the UPGMA method dendrograms were constructed (Figure 3.4.1 and 3.4.2), which show similar links for both markers and reflects clearly the relationship found in genetic distance/similarity matrix.

Figure 3.4.1 Dendrogram of genetic distance utilising UPGMA method of grouping *L. esculentum* accessions based on 55 microsatellite markers.

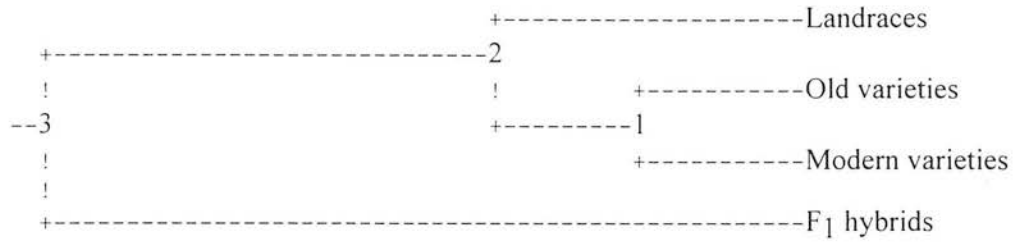


Figure 3.4.2 Dendrogram of genetic distance utilising UPGMA method of grouping *L. esculentum* accessions based on 168 RAPD markers.

3.5 Relationship among genetic distance/similarity matrices

To analyse the relationship among the genetic distance/similarity matrices from morphological, microsatellite and RAPD data, in *Lycopersicon* spp accessions a Mantel test (Mantel, 1967) was carried out. The results showed high correlation between microsatellite and RAPD data ($Z = 89.22$; $r = 0.921$; $P<0.01$), but neither morphological and microsatellite data ($Z = 24.82$; $r = -0.110$; $P<0.01$), nor morphological and RAPD data ($Z = 26.45$; $r = -0.1110$) showed correlation.

3.6 Discussion

Morphological diversity in Lycopersicon spp.

Analysis of 16 morphological characters in *Lycopersicon* spp made discrimination between species possible, although three characters were common to all accessions.

Diversity in the other characters analysed presented great variation, especially in leaf shape, size and colour. Closely related taxa to *L. esculentum* such as *L. esculentum* var. *cerasiforme* and *L. pimpinellifolium* presented very similar leaves, with some slight morphological differences, while at the other extreme *L. pennellii* and *L. hirsutum* were completely different. Other taxa such as *L. parviflorum* and *L. cheesmanii* presented great variation within each taxon, but as expected no similarity to *L. esculentum* was found in leaf shape.

A highly exserted style is a common characteristic for green-fruited species such as *L. pennellii* and *L. hirsutum*, which are self-incompatible and exclusively outbreeders. It acts as a primary physical barrier to avoid direct contact of self pollen and style within a flower, and also facilitates cross-pollination, mainly by insects (Rick, 1979b). Other species such *L. parviflorum*, that contained green- and red-fruited accessions, or red-fruited *L. cheesmanii* and *L. pimpinellifolium* have styles at the same level of stamens or only slightly exserted, characteristic of facultative outbreeding species. All representatives of *L. esculentum* and *L. esculentum* var. *cerasiforme* showed stigmas inserted at the level as the anthers. The accessions of *L. esculentum* are all self-compatible and exclusively inbreeding (Taylor, 1986), because domestication was accompanied by a transition from exserted to inserted styles and consequent change from facultative outcrossing to enforced autogamy (Rick, 1979b). This was probably as an effect of non-deliberate selection, as a result of selection of other features associated with homozygosity.

With respect to fruit characteristics, there were few differences between species for exterior colour of immature fruit. However, green-fruited species have very densely pubescent fruit in comparison with *L. esculentum* and its closely related species. These presented medium to sparse pubescence, again probably selected together with fruit size during the domestication of tomato.

Multidimensional scaling (MDS) showed all red-fruited species clustering in one sector and green-fruited in another. However, there was a clear separation between tomato cultivars and wild relatives, although *L. esculentum* var. *cerasiforme*, as expected because of its relatedness, grouped almost within the *L. esculentum* cluster.

Little morphological diversity was found within *L. hirsutum*; four of the accessions clustered in one point. It is possible that this reflected limited morphological variability within the species, but also it could be due to close similarities between the accessions received from the genebanks. However, the dispersion of points observed in *L. cheesmanii* show the diversity found within this species. The separate position of *L. cheesmanii* var. *minor* accessions shows the differences that exist between both these taxa and also the original sites within the Galapagos Islands, all from different islands and some from the coast and the other from inland. The scattered distribution of *L. pimpinellifolium* can also be expected due to their different regional origins (see Table 2.1), and therefore isolation from each other, limiting gene flow between them. Accessions of *L. parviflorum* and *L. pennellii* also presented a great diversity within them, especially *L. parviflorum* which is usually described as a green-fruited species, but also included a red-fruited accession.

Genetic similarity within species showed a high variation in *L. esculentum*, but this was expected since most of the selected traits analysed were related to fruit characteristics. These traits are some of the most variable in the accessions utilised in this study, and controlled just by few genes. The results of comparing genetic similarity of wild relatives and *L. esculentum* show a closer relationship with *L. esculentum* var. *cerasiforme*, as expected, but a greater distance with all other species involved. This study only examined a very limited number of morphologic characteristics. The potential value of these characters for current breeding is questionable, but the variation found in these characters is indicative of the variation that may exist within related species and which may have value in breeding programmes of tomato. Similarly the AMOVA analysis also demonstrated the high variability present in *Lycopersicon* spp, but also the morphological diversity present in old varieties, landraces and modern cultivars, including F_1 hybrids of *L. esculentum*.

Molecular diversity in Lycopersicon spp.

Different indices can be used to assess genetic diversity between and within populations and species. Some are more sensitive than others, but the information is

valuable in the interpretation of results. This is the case of the proportion of polymorphic loci (*P*), where differences can be considered as an index for variability within species. However, it does not reflect the real genetic variation in a population, because a slightly polymorphic locus is counted as much as a very polymorphic one and is very sensitive to the number of samples analysed. For both markers used, microsatellites and RAPD, *L. esculentum* presented the higher level of polymorphisms (0.52 and 0.62, respectively). This is expected because of the number of samples analysed in comparison to the wild relatives. The level of polymorphic loci, in the other species and for both markers, was 0.34 and 0.15, relatively low. These values are similar to other inbreeding species, such as diploid wheat relatives (between 0.30 and 0.17) (Hedge *et al.* 2000), or for cowpea (*Vigna unguiculata* (L) Walp) (between 0.39 and 0.14) (Pasquet, 2000), both with allozymes.

Average gene diversity (H_S) in cultivated tomato based on microsatellite markers (H_S=0.18) and RAPD markers (H_S=0.19) were within the range reported by Villand *et al.* (1998) for primary and secondary centres of diversity (0.22 and 0.14, respectively) utilising RAPD. Most of the wild relatives of tomato presented lower values for H_S than *L. esculentum,* though with no statistical significance. For both markers, *L. hirsutum* var. *glabratum* showed a higher mean than the other species, but the values were very similar. These indices in general do not show much information because of the unbalanced number of samples per species and the limitations of using scoring system of presence or absence. However, by plotting a regression of each index and balancing the number of samples per species with more than five accessions for both molecular markers, it is possible to indicate that wild species show a clear tendency for steeper slopes than *L. esculentum*, especially the self-incompatible outbreeder *L. hirsutum.* This indicates that there is more genetic diversity in natural populations than in domesticated *L. esculentum*. This is supported by results of previous work from Miller and Tanksley (1990) and Hamrich and Godt (1997) who suggested that self-incompatible species contains more genetic variation than self-compatible species.

A PCA for each molecular marker was carried out to investigate whether the combination of all genetic indices can be used to explain spatial species structure. It

was possible to show that *L. esculentum* is completely separated from the wild relatives, also in the partitioning of diversity in both markers *L. esculentum* presented almost 60 % of the diversity coming from differences within populations, while all other taxa most diversity was between populations.

From both molecular data sets, genetic similarity matrices were constructed and analysed as histograms. The histograms showed a huge range of similarities, from 0.10 to 1.00; this is expected because of the distribution of frequencies from the matrix of genetic similarities obtained combining many different taxa. However, histograms of *L. esculentum* accessions only showed closer similarities, between 0.40 to 1.00, values similar to those found by Nienhuis and Bosco (1994) and Villand *et al.* (1998) utilising RAPD. The frequencies of *Lycopersicon* spp accessions in the histogram skewed toward the lower end with less genetic similarities, while *L. esculentum* accessions skewed in the other direction. This distribution of genetic similarity frequencies shows the close relationship within the tomato accessions analysed and the dissimilarity with *Lycopersicon* spp accessions, which can possess novel characters to incorporate in *L. esculentum*.

The means of genetic similarity within species showed no great differences, but some species exhibited a large variation between minimum and maximum values. This genetic variability, expressed as a function of the genetic identity within species, showed differences useful for characterising individual accessions and selection of parents for the next steps in this study. It is not proposed to utilise this information as part of a phylogenetic study in *Lycopersicon* spp, because of the limitations of time and objectives of this work. However, the means of genetic similarity for wild species in comparison to *L. esculentum* showed that *L. esculentum* var. *cerasiforme* is genetically the closest relative. The other accessions of wild species presented larger genetic distance in respect to domesticated tomato cultivars indicating that there is genetic diversity available, which can be exploited in a genetic base broadening programme.

Dendrograms (Figures 3.3.6 and 3.3.13) and PCA analysis (Figures 3.3.7 and 3.3.14) of genetic similarity matrices for both markers also showed a clear separation of wild relatives and *L. esculentum* cultivars, and also a separation already noticed in

morphological analysis, between red- and green-fruited species. This was also noted by Miller and Tanksley (1990) utilising RFLPs and Peralta and Spooner (2001) using DNA sequences of the structural gene granule-bound starch synthase (GBSSI or "waxy"). Both kinds of analysis confirmed the close relationship between *L. esculentum* and *L. esculentum* var. *cerasiforme*, as well as to *L. pimpinellifolium*. This agrees with results of Rick and Fobes (1975) who also proposed *L. esculentum* var. *cerasiforme* as a hypothetical ancestor of domesticated *L. esculentum* in Mexico. It also agrees in part with Quiros (1974) who indicated an ancestry from a pre-*Lycopersicon* ancestor to *L. peruvianum* to *L. hirsutum* to *L. esculentum* var. *cerasiforme* and subsequently to the domesticated tomato. The same author suggests that the other species, such as *L. chilense*, *L. parviflorum*, *L, chmielewskii* and *L. cheesmanii*, are diverging types from this major stalk.

Accessions of *L. hirsutum* showed scattered distribution in the plots, suggesting that they contained more genotypic than phenotypic diversity as four accessions were identical morphologically. Genetic similarities between *L. pennellii* accessions were not very close for both molecular markers, in average they had 63% to 54% similarities with microsatellites and RAPD respectively. These results imply that there is genetic diversity available to exploit within accessions in this species. The evolution in isolation (allopatric speciation) of the Galapagos Islands species, *L. cheesmanii* and *L. cheesmanii* var. *minor*, was confirmed in both dendrograms and PCA plots (Figure 3.3.6, 3.3.7, 3.3.13, and 3.3.14). This species is a rich source of genes for breeding programmes, such as jointless pedicel gene (*j2*), high content of solid solubles and ascorbic acid, salt tolerance, and others (Taylor, 1986). However, this species could posses certain weaknesses against the vast range of mainland pests and diseases affecting *Lycopersicon* because of limited exposure to them during the processes of natural selection and evolution. Another interesting case is *L. parviflorum*. Warnock (1988) described it as green-fruited species, but with affinities to the *esculentum* complex. Both PCA plots show green-fruited accessions situated far apart from the red-fruited accession T-1264/94. Furthermore two green-fruited accessions analysed with microsatellites were identical, and similarly with RAPD they were very closely located, implying a very close relationship. The origins of

both accessions in Peru are not far apart, and this could explain in part the genetic similarities. The other red-fruited species, *L. pimpinellifolium*, clustered close to *L. esculentum* indicating close genetic relationship between both species (Rick, 1977; Rick and Fobes, 1975; Rick *et al.*, 1979).

AMOVA analysis of microsatellite and RAPD markers showed, in both cases, that 60% of the total variation found was due to differences within species. This is an important finding for this project because it indicates that there is high genetic variability, at the molecular level, present not only between species, but also within them.

In the genetic analysis between red- and green-fruited species, genetic diversity indices did not present statistical differences, although there were some slightly higher values in green-fruited species, for both markers. These results were not entirely expected, since green-fruited species are mainly outbreeders, and there should be more diversity in this group in comparison to red-fruited species. However, the results are based on a relatively small number of accessions, which may not reflect the real diversity and heterogeneity existing in these species.

Genetic variability within L. esculentum accessions

An analysis of genetic variability within *L. esculentum* accessions was carried out in order to observe and characterise four groups of accessions: landraces, old varieties, modern varieties OP, and F_1 hybrids. It has to be accepted at the outset that the analysis may have been affected by the low number of samples within each group. However, means of genetic diversity indices showed no statistical significance for microsatellite markers, but highly significant results ($P<0.01$) for RAPD markers. For most indices, old and modern varieties presented similar levels of genetic diversity reflecting the narrow genetic base used in their breeding process and also the close parentage that breeding techniques have produced in this crop. Differences between both groups can be taken as marginal for this experiment, considering that most cultivars utilised were selected because they were not introgressed with wild genetic material. Landraces and F_1 hybrids were significantly different and lower than the other groups in RAPD markers. In the case of F_1 hybrids, these lower values

reveal the decreasing genetic base utilised in these modern cultivars as a reflection of the breeding method, which first reduces heterozygosity in each parental line and therefore genetic diversity because of selection of few individuals as parental stock. But also the loci analysed could have not been involved in heterosis and not associated to phenotypic characters in the F_1 hybrids. Landraces showed the statistical effect of a limited number of samples analysed. These facts are clearly demonstrated in the proportion of polymorphic loci (P), where old varieties and modern varieties OP showed almost twice as many polymorphic loci compared to F_1 hybrids and landraces. Analysing A_e as the reciprocal of homozygosity, all groups showed high levels of homozygosity for both markers (between 73% to 90%), a result expected for a highly inbreeding species such as *L. esculentum*. In addition, average gene diversity (H_S) showed low values, in agreement with that found by Villand *et al.* (1998) for secondary centres of diversification (Vavilov, 1926) (H_S = 0.13). A low H_S is expected because of bottlenecks and selection pressure that reduce variability within cultivars, and post-domestication facts such as adaptation to new environments. In respect to partitioning the diversity available in these groups, the results showed that in old varieties and modern varieties OP most diversity lay in differences within populations. On the other hand in F_1 hybrids and landraces most diversity was between populations. Genetic identity and dendrograms based on these two matrices confirm the results obtained in genetic indices, where a close relationship between old and modern varieties OP was observed, and at some distance F_1 hybrids and landraces. However, in both dendrograms landraces clustered with the group of OP varieties and old cultivars indicating a closer relation with them.

In general, the results presented, regardless of the limitations stated, show the low genetic base in *L. esculentum* and the availability in wild species of genetic resources usable in breeding programmes. It also shows that there is still variability available within tomato crop cultivars, especially old varieties and landraces.

Differences between microsatellite and RAPD markers analysis

Microsatellite and RAPD markers did not show large differences in genetic distances between taxa, or partitioning genetic diversity with AMOVA, or grouping red- and green-fruited taxa in the dendrograms. However, there were differences locating some accessions within the dendrograms and PCA analysis, which can be explained by the sampling region of the genome that each marker utilise. Microsatellites are markers locus-specific, therefore they sample only small regions of the genome and the diversity present in that region. Conversly, RAPD markers sample randomly regions on the genome that contains segment sharing sequence similarity to the primer. The loci sampled by RAPD are more representative of the genome, but they are limited by reproducibility and quality of data, because as dominat marker heterozigosity is not detectable.

Chapter 4

Parent Selection and F_1 Characterisation

4.1 Introduction

Any research project involving breeding and hybridisation needs parent selection, but this is particularly important in a genetic base broadening project where the philosophy and aims are to conserve as much as possible of the genetic variation of the species/accessions utilised. There is a real need to balance the number of parents (and their individual characters) with the ability to handle the number of individual crosses and subsequent populations.

Parent selection is one of the most important steps in the present project, because within the correct choice lies the future of the research; therefore, morphological and molecular characterisations of species and accessions are a fundamental prerequisite in this study which aims to investigate the value of different approaches.

The aims of this chapter include:

i) to present and characterise the accessions of *L. esculentum* and its wild relatives selected as potential parents.

ii) to select a set of polymorphic primers for microsatellites and RAPD analysis.

iii) to review hybridisation and F_1 generation populations.

4.2 Parent selection

From the 38 accessions of 8 taxa characterised in Chapter 3, only 10 accessions involving 6 taxa, as genetically and morphologically diverse as possible, were selected as parents for inter- and intra-taxon hybridisation. Table 4.1 gives the genetic material selected, their identification as accession or cultivar and country of origin. All entries, except *L. esculentum* cv. Limachino from INIA-Chile, were obtained from different agencies in USA and Germany, where they have been multiplied and conserved within the gene-banks of TGRC in Davis, California; USDA-ARS in Cornell University, New York; and IPK, Gatersleben, Germany. This germplasm was selected based on information obtained in Chapter 3, details on the rationale for choices are given in this Chapter. In this choice of parents, no *L. cheesmanii* accessions were included because of the inability to produce flowering in this species which coincided with flowering in the other material.

Table 4.1 Species and accessions selected as parents for hybridisation.

Species	Accession	Cultivar	Country of origin
L. esculentum		Limachino	Chile
L. esculentum	LA 0516	Ace	USA
L. esculentum	LA 0534	Lukullus	UK
L. esculentum	LA 0502	Marglobe	USA
L. esculentum	LA 0180	San Marzano	Italy
L. esculentum var. *cerasiforme*	LA 1673		Peru
L. hirsutum var. *glabratum*	PI 199381		Peru
L. parviflorum	T1264/94		Peru
L. pennellii var. *puberulum*	LA 1926		Peru
L. pimpinellifolium	PI 270449		Mexico

4.2.1 Morphological characterisation and differences between parent accessions

Of the 18 morphological characters selected and analysed in this study, three were monomorphic. All parental accessions showed presence of hypocotyl pubescence, horizontal leaf attitude, and multiparous inflorescence type. Tables 4.2.1 and 4.2.2 give a brief resume of the 15 traits that characterise each parental accession and made possible the morphological analysis. Some traits were common in some species and accessions, but the combination of them was unique for each entity. However, there were many reasons to select these accessions, therefore detailed characteristics are described in the following paragraphs.

Within *L. esculentum* accessions, the Chilean cultivar Limachino was selected because from a morphological point of view, it is a cultivar with determinate growth containing within its genome the *self-pruning* gene (*sp*) (Butler, 1952; Rick, 1982; Stevens and Rick, 1986) and *potato type leaf* (less segmented leaf) controlled by the *c* gene (Rick and Butler, 1956). These characters are only found in this entry. The fruits were of intermediate size. They were flattened, *fasciated* shape and *multi locular* controlled by *f* (Rick and Butler, 1956) and *lc* (Fryxell, 1954) genes

respectively. The expression of orange fruit colour was controlled by three genes: *red flesh* colour (*R*) (Rick and Butler, 1956), yellow skin or *pigmented fruit epidermis* (*Y*) (Rick and Butler, 1956), and modifier *tangerine* flesh colour (*t*) (Rick and Butler, 1956); and pointed blossom end shape controlled by gene *nipple-tip* (*nt*) (Butler, 1955).

Cultivar Ace is a modern open pollinated (OP) cultivar and is still in use in some regions for field growing. This cultivar was characterised by indeterminate growth (sp^+) and standard leaf shape (*C*). In general the fruits were large in size, but under the experimental growing conditions did not reach their full potential, although the fruits were bigger than any other accession. As in cv. Limachino, the fruits were flattened, fasciated (*f*) and multilocular (*lc*). The colour was red, suggesting the gene combination red flesh colour (*R*), yellow skin *(Y)* and non-tangerine flesh colour *(T)*. Blossom end shape was flat (*Nt*).

Cultivar Lukullus is a greenhouse OP cultivar, no longer in use. Its growth type is indeterminate (sp^+) and presents standard leaf type (*C*). Fruit shape was slightly flattened (*F*), small in size, and presenting usually two or three locules (*Lc*). Fruit colour and blossom end shape were similar to 'Ace'.

Cultivar Marglobe is a very old field growing OP cultivar, whose history can be traced back to the year 1925. No introgression from wild relatives has been reported in the pedigree of this cultivar (Boswell, 1937). It is a cultivar commonly utilised as a control for morphological traits (Stevens and Rick, 1986) and used to be the base of breeding programmes in the USA during the first decades of the last century, as it was the most successful cultivar in those years (Boswell, 1937). It has indeterminate growth (sp^+) and standard leaf type (*C*). Fruit shape was rounded (*F*), small size and presented two locules (*Lc*). Fruit colour and blossom end shape were similar to 'Ace' and 'Lukullus'.

As representative of processing tomatoes, cultivar San Marzano was chosen. It is also an old OP field growing cultivar, although currently not cultivated, but is extensively used in breeding programmes. The growth type was indeterminate (sp^+) and with standard leaf type (*C*). Fruit shape was cylindrical, character controlled by *ovate* gene

(*o*) (Rick and Butler, 1956), with two locules (*Lc*), and small size. Fruit colour was red, as for cv. Ace.

In the *L. esculentum* cultivars there were some common characters for the species, such as ½ purple/½ green hypocotyl colour controlled by *anthocyaninless* gene *a* (Rick and Butler, 1956) and modified by *anthocyanin loser* gene *al* (Rick and Butler, 1956); *yellow corolla* colour controlled by gene *Wf* (Rick and Butler, 1956); inserted style by *exserted stigma* genes *Ex-1, Ex-2* and *Ex-3* (Tikoo and Anand, 1982), and the sparse *fruit pubescence* controlled by gene *P* (Rick and Butler, 1956).

With *L. esculentum* var. *cerasiforme* or "cherry tomato", a very close relative of tomato, there was no choice as only one accession was obtained and grown. The hypocotyl colour was similar to *L. esculentum* (½ purple/½ green) (*A* and *al*), with indeterminate growth type (sp^+) and standard leaf type (*C*). Flower characteristics were yellow corolla colour (*Wf*) with a stigma position at same level as the tips of the anthers (*Ex-*). Fruits were very small (<2 cm diameter), rounded shape (*F*), red colour (*R*, *Y* and *T*), with two locules (*Lc*), and flat blossom end shape (*Nt*).

L. hirsutum var. *glabratum*, accession PI-199381, had purple hypocotyl colour (*A* and *Al*), indeterminate growth type (sp^+), and *peruvianum* leaf type. In addition, the flowers had a bright yellow corolla colour, larger size, and more open shape. It is likely that due to their self-incompatibility, *L. hirsutum* requires insect pollination for reproduction, and have bright colours, attractive shapes and odours to attract pollinators (Prokopy and Owen, 1983; Schoonhoven *et al.*, 1998). A highly exserted style (*ex-*) is another characteristic typical of these self-incompatible taxa, the exserted style avoids contact with own pollen, and acts as a primary barrier preventing self-pollination and facilitating cross-pollination via insects. These species were characterised by hairiness, the fruit were densely pubescent (*p*), very small and round shaped (*F*). The colour was green because of the *green flesh* (*gf*) gene (Clayberg *et al.*, 1967) that controls the persistence of chlorophyll in the locules, and clear skin (*y*). The fruits had a flat blossom end shape similar to all tomato wild relatives utilised in this study.

The selection of *L. parviflorum* was based on fruit colour, because this species has been described as green-fruited (Taylor, 1986) and of the 3 samples grown two had

green fruits. Regardless of the cross-compatibility of this species with *L. esculentum*, the red-fruited accession (T1264/94) was selected in order to increase the chances of positive hybridisation of these species. This entry also produced great amount of flowers, synchronously with the other species. In respect of hypocotyl colour, it showed purple colour similar to *L. hirsutum* var. *glabratum* and all other tomato wild relatives, except *L. esculentum* var. *cerasiforme*. Growth type was semi-determinate, but corresponding to a vine-type growth, very branching and aggressive. The leaves were a very typical shape for this species, with leaflets smaller than in other species and more widely spaced. Flowers showed a yellow corolla colour (*Y*), and a slightly exserted style position; a common characteristic in species exhibiting facultative self-pollination. The fruits had intermediate pubescence (*p*), round shape (*F*), very small size, red exterior colour but pink flesh colour (*R*, *Y* and *T*).

From the three accessions of *L. pennellii* grown, *L. pennellii* var. *puberulum* (LA-1926) was chosen because it developed fruits earlier than other accessions. In respect to flowering timing, amount of flowers and pollen produced all accessions behaved similarly. Accession LA-1926 presented a determinate growth type (*sp*) and a typical *pennellii* leaf type. Leaflets were smaller than *L. esculentum,* more rounded, brighter and sticky. Flowers showed a bright yellow corolla, but were larger than in *L. hirsutum*, and a highly exserted style (*ex-*), typical for self-incompatible entomophilous species. Fruits were very small, with a dense pubescence (*p*) and green colour in exterior and flesh (*gf*).

The choice of parental accession in *L. pimpinellifolium* was based on geographical distances, PI-270449 originated from Mexico and most of the other accessions were from Peru and Ecuador. This accession had a vigorous indeterminate growth type (sp^+), with a typical *pimpinellifolium* leaf type with smaller and more serrate leaflets than *L. esculentum.* The flowers were yellow (*Wf*) with a slightly exserted style (*ex-*), typical for a facultative self-pollinated species. Fruit pubescence was intermediate (*p*), and the fruits very small in size were of round shape (*F*) and pink colour (*R*, *y* and *T*).

Table 4.2.1 Morphological characteristics of *Lycopersicon* spp accessions utilised as parents.

Species	Cultivar/ Variety	Hypocotyl Colour	Growth Type	Leaf Type	Corolla Colour	Style Position	Inmature Fruit Colour	Fruit Pubescence
L. esculentum	Limachino	½ purple	determinate	Potato type	yellow	inserted	greenish-white	sparse
L. esculentum	Ace	½ purple	indeterminate	Standard	yellow	inserted	light green	sparse
L. esculentum	Lukullus	½ purple	indeterminate	Standard	yellow	inserted	green	sparse
L. esculentum	Marglobe	½ purple	indeterminate	Standard	yellow	inserted	greenish-white	sparse
L. esculentum	San Marzano	½ purple	indeterminate	Standard	yellow	inserted	light green	sparse
L. esculentum	var. *cerasiforme*	½ purple	indeterminate	Standard	yellow	same level of stamens	light green	sparse
L. hirsutum	var. *glabratum*	purple	indeterminate	Peruvianum	bright yellow	highly exserted	light green	dense
L. parviflorum		purple	semi-determinate	Parviflorum	yellow	slightly exserted	light green	intermediate
L. pennellii	var. *puberulum*	purple	determinate	Pennellii	bright yellow	highly exserted	dark green	dense
L. pimpinellifolium		purple	indeterminate	Pimpinellifolium	yellow	slightly exserted	greenish-white	intermediate

Details of each morphological character in Appendix 1.

Table 4.2.2 Morphological characteristics of *Lycopersicon* spp accessions utilised as parents.

Species	Cultivar/ Variety	Fruit Shape	Fruit Size	Mature Fruit Colour	Flesh Colour	Fruit Cross Shape	Shape of Pistil Scar	Number of locules	Blossom End Shape
L. esculentum	Limachino	Flattened	Intermediate	Orange	orange	irregular	irregular	>5	pointed
L. esculentum	Ace	Flattened	Intermediate	Red	red	irregular	dot	>5	flat
L. esculentum	Lukullus	Slightly flattened	Small	Red	red	round	dot	2-3	flat
L. esculentum	Marglobe	Rounded	Small	Red	red	round	dot	2	flat
L. esculentum	San Marzano	Cylindrical	Small	Red	pink	angular	dot	2	indented
L. esculentum	var. *cerasiforme*	Rounded	very small	Red	red	round	dot	2	flat
L. hirsutum	var. *glabratum*	Rounded	very small	Green	green	round	dot	2	flat
L. parviflorum		Rounded	very small	Red	pink	round	dot	2	flat
L. pennellii	var. *puberulum*	Rounded	very small	Green	green	round	dot	2	flat
L. pimpinellifolium		Rounded	very small	Pink	pink	round	dot	2-3	flat

Details of each morphological character in Appendix 1

4.2.2 Molecular markers and differences between parental accessions

Molecular markers used in selecting parents from *Lycopersicon* spp accessions were microsatellites and RAPD. Data analysed came from the general screening described in Chapter 3.3. The similarity matrices utilised in this analysis are given in Appendix 2 part 2 for microsatellites and part 3 for RAPD markers. Accessions selected as parents are highlighted in bold.

Considering only the accessions selected as parents, the UPGMA dendogram based on microsatellite analysis (Figure 3.3.6) showed that accessions belonging to *L. esculentum* and including *L. esculentum* var. *cerasiforme* clustered together; two sub-clusters were formed, one consisting of Limachino, Ace and "cherry" tomato, and other with Lukullus, Marglobe and San Marzano. Other parental accessions belonging to wild relative species, such as *L. pennellii* var. *puberulum*, *L. parviflorum*, *L. hirsutum* var. *glabratum*, and *L. pimpinellifolium*, formed other cluster, but reflected the genetic distances between them and other accessions. The closest accessions were *L. esculentum* cvs. Lukullus and Marglobe, with approximately 90% of similarity. The most distant accession was that belonging to *L. pennellii* var. *puberulum*.

The dendogram generated from RAPD data analysis (Figure 3.3.13) showed close similarity to that produced from microsatellite data, but the *L. esculentum* cluster contained *L. pimpinellifolium*; two sub groupings could be observed: one involving Limachino and Ace, and the second Lukullus, Marglobe, San Marzano. The cluster was completed by *L. esculentum* var. *cerasiforme* accession constructing the main cluster. As in the microsatellite dendogram, remaining wild relatives did form another cluster and again the most and least similar accessions were the same.

Mantel's test (Mantel, 1967) established a statistically significant ($P<0.05$) moderate correlation between microsatellite and RAPD genetic similarity matrices for accessions selected as parent ($r = 0.44$).

4.2.3 Analysis of some continuous morphological characters in parents

Continuous morphological characters are usually the expression of several loci within the genome (polygenic characters). These phenotypes are highly influenced by environmental conditions (genotype x environment interactions) (Srb, *et al.*, 1952; Mayo, 1987; Jensen, 1988; Griffiths *et al.*, 1996). The genotype establishes the aptitude for growth and development, while the environment determines the mode of development, resulting in the phenotype (Simmonds, 1979).

Results of the analyses of the continuous morphological characters are displayed in Tables 4.3.1 and 4.3.2. ANOVA was carried out to observe statistical differences between parent accessions; details are shown in Appendix 3 part 8. Tukey's multiple comparison test was used on variables with statistical significance. No results of the accessions *L. hirsutum* var. *glabratum* (PI-199381) and *L. pennellii* var. *puberulum* (LA-1926) were analysed because of the scarcity of fruits during the growing season.

Fruit diameter, length and ratio (diameter/length) showed statistically highly significant differences ($P<0.01$) between parents; *L. esculentum* accessions presented fruits significantly bigger than wild relatives (Table 4.3.1). Within *esculentum* cultivars, Ace and Limachino were also significantly larger in size than the others. In respect to fruit ratio, this character reflects the shape of each accession and statistical differences were expected to be found. Values below 1 represent elongated fruits, such as San Marzano; values closer to 1 represent rounded fruits, such as Lukullus, Marglobe, and wild relatives; and values greater than 1 represent flattened fruits, such as Limachino and Ace.

Fruit weight, solid soluble content, and weight of 1,000 seeds also presented highly significant differences ($P<0.01$) between parents (Table 4.3.2). Significantly heavier fruits were found in accessions of *L. esculentum*, cultivars Ace and Limachino. Mean fruit weights of Lukullus, Marglobe and San Marzano were not significantly different, but they were greater than those of the wild relatives. However, these were not statistically different.

Table 4.3.1 Means of 3 fruit characters (fruit diameter, length and ratio d/l) of *Lycopersicon* spp accessions selected as parents.

Species/Accession	Fruit diameter (cm)	Fruit length (cm)	Fruit ratio (D/L)
	**	**	**
***L. esculentum* cvs.**			
11 – Limachino	5.86 bc (± 0.21)	4.42 a (± 0.14)	1.33 ac (± 0.03)
15 – Ace	6.76 a (± 0.45)	4.74 a (± 0.26)	1.42 a (± 0.02)
19 – Lukullus	3.65 bde (± 0.11)	3.07 bc (± 0.12)	1.20 bc (± 0.05)
21 – Marglobe	3.81 bde (± 0.14)	3.45 bc (± 0.16)	1.11 bd (± 0.02)
22 – San Marzano	2.82 bde (± 0.10)	4.82 a (± 0.36)	0.61 bde (± 0.05)
39 – *L. esculentum* var. *cerasiforme*	1.80 bdf (± 0.05)	1.67 bd (± 0.05)	1.08 bd (± 0.01)
52 – *L. parviflorum*	1.03 bdf (± 0.02)	0.87 bd (± 0.02)	1.19 bd (± 0.01)
60 – *L. pimpinellifolium*	1.44 bdf (± 0.06)	1.29 bd (± 0.07)	1.13 bd (± 0.02)

Numbers between brackets correspond to standard error; D/L = diameter/length; ** = high statistical significance ($P<0.01$); same letters show no statistical differences between means.

Solid soluble content, mainly glucose and fructose (Hewitt and Gavey, 1987), showed highly significant differences ($P<0.01$). *L. parviflorum* accession had significantly higher values than any other entry, but also the wild types had higher values than the *L. esculentum* accessions. Within *L. esculentum* entries, significant differences in sugar content were found between cultivars. Limachino and Marglobe were statistically different to Lukullus, San Marzano and Ace.

Seed size also was highly significant ($P<0.01$); bigger seeds were found within *L. esculentum* accessions compared to wild relatives. Statistically, *L. pimpinellifolium* showed the smallest seeds and *L. esculentum* cvs. Lukullus and Ace the largest.

Table 4.3.2 Means of 3 fruit characters (fruit weight, solid soluble content, and weight of 1000 seeds) of *Lycopersicon* spp accessions selected as parents.

Cross/Accession	Fruit weight (g)	Solid soluble content (brix°)	Weight 1,000 seeds (g)
	**	**	**
***L. esculentum* cvrs.**			
11- Limachino	86.21 bc (± 9.70)	5.10 bde (± 0.13)	2.97 bc (± 0.02)
15 – Ace	127.07 a (± 22.79)	5.77 bde (± 0.46)	3.00 ac (± 0.01)
19 – Lukullus	24.90 bde (± 1.98)	6.68 bdf (± 0.24)	3.18 a (± 0.05)
21 – Marglobe	30.42 bde (± 3.46)	5.23 bde (± 0.27)	2.45 bd (± 0.03)
22 – San Marzano	22.79 bde (± 2.45)	6.20 bdf (± 0.16)	1.59 bde (± 0.02)
39 – *L. esculentum* var. *cerasiforme*	3.46 bdf (± 0.33)	7.09 bd (± 0.24)	1.26 bde (± 0.05)
52 – *L. parviflorum*	0.68 bdf (± 0.03)	10.76 a (± 0.31)	1.13 bde (± 0.02)
60 – *L. pimpinellifolium*	1.98 bdf (± 0.21)	8.59 bc (± 0.20)	0.86 bdf (± 0.03)

Numbers between brackets correspond to standard error; D/L = diameter/length; ** = high statistical significance ($P<0.01$); same letters show no statistical differences between means.

4.3 Selection of microsatellites and RAPD primers to use in F1 and further generations

To analyse a large number of populations through PCR procedures, ideally there should be few oligonucleotide primers, each giving highly polymorphic results between populations, and which are reliable, consistent, reproducible, and comparable (Hoelzel and Green, 1998).

Of the 18 microsatellite and 80 RAPD primers investigated, 6 and 7 respectively were selected, which fulfilled the characteristics described above. In the case of microsatellite markers a total of 26 loci were amplified of which 24 (92%) were polymorphic and two common to all accessions. In RAPD markers, out of 92 amplified bands, 61 (66%) were polymorphic and 31 were common to all accessions. Selected primers for both microsatellites and RAPD markers are presented in Table

4.4, which also displays total number of bands amplified and average number of bands per accession. On average, microsatellite primers presented 4.3 bands considering all the primers, and 1.95 bands per accession. RAPD primers showed 13.1 bands considering the average of all primers and 8.1 bands per accession.

Table 4.4 Microsatellite and RAPD primers selected to analyse parental accessions, hybridisation and further generations in *Lycopersicon* spp.

Primer[1]	Total number of bands	Average number of bands/accession	Approx. band size range (bp)
Microsatellites :			
LE20592	4	1.0	177 – 166
LE21085	4	2.1	210 - 98
LEEF1A	4	1.6	231 – 186
LEGAST1	6	3.0	365 – 194
LELEUZIP	4	2.0	177 – 66
LEPRP4	4	2.0	239 – 192
Average	**4.3**	**1.95**	
RAPD :			
OPA – 01	12	7.9	1350 – 525
OPA – 12	8	6.4	800 – 350
OPA – 19	9	6.4	800 – 275
OPH – 01	15	8.9	1400 – 300
OPH - 11	17	8.6	1200 – 275
OPL – 16	14	8.8	1200 – 240
OPL - 18	17	9.6	1500 – 450
Average	**13.1**	**8.1**	

[1] Primers sequences are presented in Tables 2.5 and 2.6.

In the microsatellites, PCR produced one or two bands (alleles) per loci as expected in this kind of molecular marker. More amplified bands were expected in RAPD markers and they ranged in average from 6.4 to 9.6 bands. Another difference between both markers was the band size range. While microsatellite primers produced amplification products ranging from 365 to 66 bp, RAPD produced sizes between 1400 and 240 bp.

4.4 Parents hybridisation

In order to create segregating populations after selecting the parents, manual hybridisations between all accessions were carried out, as explained in Chapter 2 (2.2.21). However, when working with wild relatives there is a possibility of failure in some crosses owing to genetic incompatibilities between species. Most incompatibilities between species or even between accessions are physical or genetic and can be overcome utilising breeding manipulations.

In this project no manipulations, except emasculation, were carried out, and there was no attempt to overcome any hybridisation failure. The *esculentum*-complex species crossed easily with tomato accessions. It was found that green fruited species such as *L. hirsutum* and *L. pennellii* could be utilised as staminate parent for inter-taxon crosses, but they were not good pollen receptors.

In the case of *L. esculentum* accessions, cv. Marglobe presented the lowest level of fecundity based in the number of hybridisations carried out (Table 4.5), 22% as female and 44% as male. The highest levels were presented by cultivars Limachino, Ace, and Lukullus. Within wild relatives, the red-fruited *L. parviflorum* accession showed the highest fecundity with 77% as female and 66% as male.

Table 4.5 Parental accessions, crosses, reciprocal crosses, and seed number obtained.

	Species	Cultivar/accession	Number	11	15	19	21	22	39	47	52	56	60
11	*L. esculentum*	Limachino			103	167	164	0	41	0	162	275	291
15	*L. esculentum*	Ace	LA – 0516	41		91	9	21	1	0	0	85	59
19	*L. esculentum*	Lukullus	LA-0534	161	131		0	7	15	1	15	0	43
21	*L. esculentum*	Marglobe	LA-0502	76	56	0		0	0	0	1	0	0
22	*L. esculentum*	San Marzano	LA-0180	49	43	11	3		10	1	8	15	4
39	*L. esculentum*	var. *cerasiforme*	LA-1673	20	9	18	0	0		0	13	0	0
47	*L. hirsutum*	var. *glabratum*	PI-199381	0	0	0	0	0	0		0	0	0
52	*L. parviflorum*		T1264/94	4	1	15	0	3	24	0		15	16
56	*L. pennellii*	var. *puberulum*	LA-1926	0	0	0	0	0	0	0	0		0
60	*L. pimpinellifolium*		PI-390739	2	0	0	6	0	1	1	29	174	

Crosses: Female (rows) x Male (columns) below diagonal; reciprocal crosses: Female (rows) x Male (columns) above diagonal.

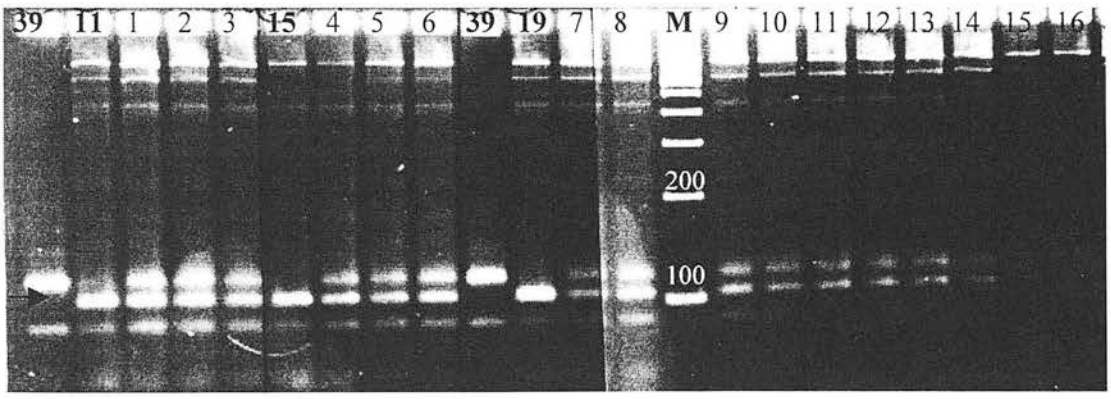
4.5 F_1 generation analysis of morphological and molecular characters

After hybridisation an F_1 generation was grown. In the case of the *Lycopersicon* spp accessions utilised as parents, all morphological and molecular characters studied were genetically homozygous. This was confirmed by growing the parental accessions together with each generation and observing non segregation of characters.

4.5.1 Molecular characteristics

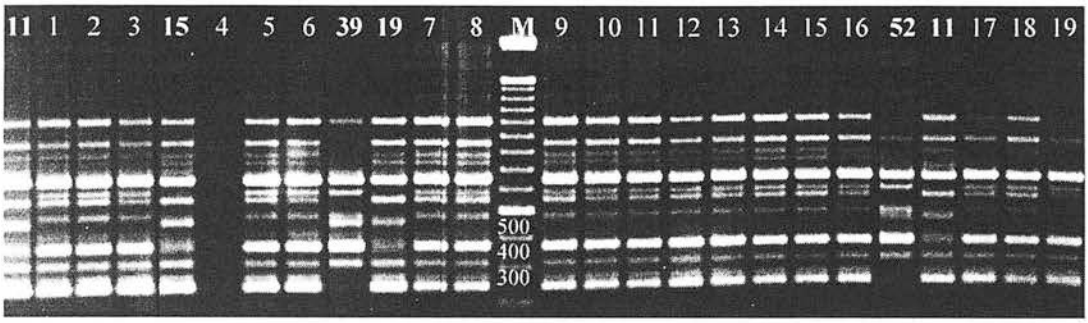
Molecular markers, microsatellites and RAPD, were not examined in F_1 plants because they did not show segregation within individuals, but heterozygosity in respect to parents. Microsatellites presented both alleles sharing the same locus in all individuals studied for all primers, and selected RAPD bands showed always the presence of the dominant allele in all loci, including the heterozygous, making impossible any further genetic analysis.. However, in any genetic population analysis based on molecular markers, the F_1 generation shows the highest level of heterozygosity: for co-dominant markers, such as microsatellites H_S = 1.0 and for dominant markers, such as RAPD H_S = 0.5, because there is only half of the information available. Examples of inheritance of microsatellite and RAPD markers in various F_1s are shown in Plates 4.1 and 4.2.

Plate 4.1 Agarose gel of microsatellite from primer LE-21085 in parents and F_1 populations of inter-taxon crosses in *Lycopersicon* spp.



Arrow shows microsatellite position; **M** = molecular size marker in bp; **39** = *L. esculentum* var. *cerasiforme*; **11** = *L. esculentum* cv. Limachino; **15** = *L. esculentum* cv. Ace; **19** = *L. esculentum* cv. Lukullus; 1 to 3 = F_1 from cross between **39** and **11**; 4 to 6 = F_1 cross between **39** and **15**; 7 to 16 = F_1 cross between **39** and **19**.

Plate 4.2 Agarose gel of RAPD from primer OPA-12 in parents and F_1 populations of inter-taxon crosses in *Lycopersicon* spp.



Arrows show polymorphic positions; **M** = molecular size marker in bp; **39** = *L. esculentum* var. *cerasiforme*; **11** = *L. esculentum* cv. Limachino; **15** = *L. esculentum* cv. Ace; **19** = *L. esculentum* cv. Lukullus; **52** = *L. parviflorum*; 1 to 3 = F_1 from cross between **39** and **11**; 4 to 6 = F_1 cross between **39** and **15**; 7 to 16 = F_1 cross between **39** and **19**; 17 to 19 = F_1 cross between **52** and **11**.

4.5.2 Morphological characteristics

Morphological traits for inter- and intra-taxon crosses were analysed and the results presented in Tables 4.6.1 and 4.6.2. Results of more F_1 accessions are shown in Appendix 4.

One or at most two genes control most morphological traits chosen for characterising accessions and crosses, as described in Section 4.2.1. Such genes behave in Mendelian manner and it is possible to study their segregation ratios in F_2 generation. All F_1 populations presented characteristics such as indeterminate growth type, yellow corolla colour, sparse fruit pubescence, and red mature fruit colour. For other characters, such as hypocotyl colour, presence of ½ purple/½ green hypocotyl was a common feature in most F_1 populations, but in some inter-taxon crosses only purple hypocotyls were also found. Leaf type was a variable characteristic, depending on the parents involved in the cross and, in the case of inter-taxon crosses showed a tendency toward the tomato wild relative types. Within intra-taxon crosses of *L. esculentum*, F_1 populations did not show variation from the standard leaf type. In respect of stigma position, most accessions presented inserted stigmas or stigmas at the same level as the tips of the anthers, but there were two inter-taxon crosses involving *L. parviflorum* as parent that presented slightly exserted stigmas. Immature fruit colour for inter-taxon crosses was mainly greenish-white to light

green colours, except one cross involving *L. esculentum* cv. Limachino and *L. esculentum* var. *cerasiforme* that presented green colour.

The colour of immature fruit in intra-taxon crosses varied from light green to green, and only one accession presented greenish-white colour in all individuals.

The most common fruit shape found in F_1 populations of inter- and intra-taxon *Lycopersicon* spp crosses was rounded, though some crosses involving *L. esculentum* cvs. Ace and Limachino presented slightly flattened fruits, and others containing *L. esculentum* cv. San Marzano presented heart shaped or cylindrical fruits.

There was not much variation for fruit size between all F_1s, most of them showed very small or small fruits; only two F_1s presented intermediate fruit size and both contained as parent *L. esculentum* cv. Ace. Flesh colour in mature fruits ranged between orange and red in most of F_1 populations. From all F_1s grown, only one (E-1522) showed irregular fruit cross sectional shape; all others exhibited a round and angular shape. In shape of pistil scar, a dot was the common character, but intra-taxon crosses including *L. esculentum* cv. Ace presented a stellate shape. The number of locules was very variable between accessions, ranging from 2 to >4, but the most common value was two locules especially within inter-taxon crosses. Blossom end shape was flat for almost all inter-taxon crosses, except one (I-2260) that presented indented shape. In intra-taxon crosses there was a distribution between flat, indented and pointed end shapes.

Table 4.6.1 Morphological characteristics of F_1 plants of *Lycopersicon* spp inter- and intra-taxon crosses.

Crosses	Hypocotyl Colour	Growth Type	Leaf Type	Corolla Colour	Style Position	Immature Fruit Colour	Fruit Pubescence
I – 3919	½ purple	Indeterminate	Standard	Yellow	Same level of stamens	Light green	Sparse
I – 1939	½ purple	Indeterminate	Standard	Yellow	Inserted	Light green	Sparse
I – 5211	Purple	Indeterminate	Parviflorum	Yellow	Same level of stamens	Greenish-white	Sparse
I – 1152	Purple	Indeterminate	Parviflorum	Yellow	Same level of stamens	Light green	Sparse
I – 6021	½ purple	Indeterminate	Pimpinellifolium	Yellow	Inserted	Greenish-white	Sparse
I – 1560	½ purple	Indeterminate	Pimpinellifolium	Yellow	Same level of stamens	Greenish-white	Sparse
E – 1922	½ purple	Indeterminate	Standard	Yellow	Inserted	Light green	Sparse
E – 2219	½ purple	Indeterminate	Standard	Yellow	Inserted	Light green	Sparse

I = inter-taxon cross; E = intra-taxon cross; first two digits correspond to female parent and second two to male parent; details of each cross in Table 4.5; details of morphological character in Appendix 1.

Table 4.6.2 Morphological characteristics of F_1 plants of *Lycopersicon* spp inter- and intra-taxon crosses.

Crosses	Fruit Shape	Fruit Size	Mature Fruit Colour	Flesh Colour	Fruit Cross Shape	Shape of Pistil Scar	Number of locules	Blossom End Shape
I – 3919	Rounded	Very small	Red	Orange	Round	Dot	2	Flat
I – 1939	Rounded	Very small	Red	Red	Round	Dot	2	Flat
I – 5211	Slightly flattened	Very small	Red	Red	Angular	Dot	2	Flat
I – 1152	Slightly flattened	Very small	Red	Red	Round	Dot	2	Flat
I – 6021	Rounded	Very small	Red	Orange	Round	Dot	2	Flat
I – 1560	Slightly flattened	Very small	Red	Red	Angular	Dot	3	Flat
E – 1922	Heart shaped	Small	Red	Orange	Angular	Dot	2	Indented
E – 2219	Heart shaped	Small	Red	Orange	Angular	Dot	2	Indented

I = inter-taxon cross; E = intra-taxon cross; first two digits correspond to female parent and second two to male parent; details of each cross in Table 4.5; details of morphological character in Appendix 1

4.5.3 Analysis of Continuous characters in F_1 generation

Six continuous characters were analysed in F_1 population for each cross, inter- and intra-taxon, and compared statistically with the parents and mean expected value. ANOVA was carried out to determine statistical significance between means and a Tukey's test to establish the differences. Results for selected crosses are displayed in Tables 4.7.1 and 4.7.2. These crosses, and when possible the reciprocal, were selected as a sample from each wild relative x *L. esculentum* and intra-taxon *L. esculentum* group of crosses, including all accessions involved in the hybridisations. Results for the remaining crosses are presented in Appendix 5. Details of ANOVA procedure for both groups are displayed in Appendix 3, part 9.1 and 9.2.

For all crosses, inter- and intra-taxon, diameter, length and ratio of the fruits presented statistically significant differences ($P<0.01$ or $P<0.05$) between the means of cross, reciprocal and parental accessions. Inter-taxon crosses containing *L. esculentum* var. *cerasiforme* showed that mean diameter and length of fruits were smaller than expected value (average of parents) when hybridised with *L. esculentum* accessions normally presenting larger fruits such as cvs. Limachino and Ace. Means were within expected value when crossed with *L. esculentum* cv. Lukullus. Fruit ratio was within expected values in all cases. In these three crosses the F_1 populations were statistically different from *esculentum* cultivars in diameter, length and ratio of fruits. Similar results were found when analysing crosses involving *L. parviflorum and L. pimpinellifolium.*

Within intra-taxon crosses of *L. esculentum*, diameter and length of fruits were within or slightly smaller than expected values. Fruit ratio varied depending on the parents utilised in the cross, but most of the populations presented values close to the expected. Fruit ratios for most crosses were close to 1 (round shape), but those including cvs. Ace and Limachino (slightly flattened to flattened shape) showed a tendency toward these shapes. In the case of crosses involving cv. San Marzano (cylindrical shape), the tendency was towards fruits slightly elongated (values < 1) but not completely cylindrical or pear shaped.

Table 4.7.1 Means of crosses and parental accessions of continuous traits in F_1 generation of inter- and intra-taxon crosses.

Cross/Accession	Fruit diameter (cm)	Fruit length (cm)	Fruit ratio (D/L)
	**	**	*
I – 3919	2.67 bc (± 0.06)	2.48 bc (± 0.06)	1.08 b (± 0.01)
I – 1939	2.70 bc (± 0.10)	2.48 bc (± 0.10)	1.09 a (± 0.02)
Expected value	**2.73**	**2.37**	**1.14**
39 – *L. esculentum* var. *cerasiforme*	1.80 bd (± 0.05)	1.67 bd (± 0.05)	1.08 b (± 0.01)
19 – Lukullus	3.65 a (± 0.11)	3.07 a (± 0.12)	1.20 a (± 0.05)
	**	**	**
I – 5211	1.55 bc (± 0.07)	1.30 bde (± 0.07)	1.20 b (± 0.02)
I – 1152	1.91 bc (± 0.05)	1.68 bc (± 0.05)	1.14 b (± 0.03)
Expected value	**3.45**	**2.65**	**1.26**
52 – *L. parviflorum*	1.03 bd (± 0.02)	0.87 bdf (± 0.02)	1.19 b (± 0.01)
11- Limachino	5.86 a (± 0.21)	4.42 a (± 0.14)	1.33 a (± 0.03)
	**	**	*
I – 6021	2.24 bc (± 0.10)	2.13 bc (± 0.10)	1.06 b (± 0.01)
Expected value	**2.63**	**2.37**	**1.12**
60 – *L. pimpinellifolium*	1.44 bd (± 0.06)	1.29 bd (± 0.07)	1.13 a (± 0.02)
21 – Marglobe	3.81 a (± 0.14)	3.45 a (± 0.16)	1.11 a (± 0.02)
	**	**	**
I – 1560	2.79 bc (± 0.04)	2.38 bc (± 0.05)	1.17 b (± 0.02)
Expected value	**4.10**	**3.02**	**1.28**
60 – *L. pimpinellifolium*	1.44 bd (± 0.06)	1.29 bd (± 0.07)	1.13 b (± 0.02)
15 – Ace	6.76 a (± 0.45)	4.74 a (± 0.26)	1.42 a (± 0.02)
	**	**	**
E – 1922	3.97 a (± 0.09)	3.90 bc (± 0.05)	1.02 bc (± 0.02)
E – 2219	3.30 bc (± 0.16)	3.40 bc (± 0.12)	0.97 bc (± 0.03)
Expected value	**3.24**	**3.95**	**0.91**
19 – Lukullus	3.65 a (± 0.11)	3.07 bd (± 0.12)	1.20 a (± 0.05)
22 – San Marzano	2.82 bd (± 0.10)	4.82 a (± 0.36)	0.61 bd (± 0.05)

Numbers between brackets correspond to standard error; I = inter-taxon crosses; E = intra-taxon crosses; first two digits correspond to female parent; same letters show no statistical differences.

Table 4.7.2 Means of crosses and parental accessions of continuous traits in F_1 generation of inter- and intra-taxon crosses.

Cross/Accession	Fruit weight (g)	Solid soluble content (brix°)	Weight 1,000 seeds (g)
	**	**	**
I – 3919	11.23 bc (± 0.69)	6.86 bc (± 0.44)	2.50 bc (± 0.05)
I – 1939	11.82 bc (± 1.30)	8.75 a (± 0.69)	2.52 bc (± 0.03)
Expected value	**14.18**	**6.89**	**2.22**
39 – *L. esculentum* var. *cerasiforme*	3.46 bd (± 0.33)	7.09 a (± 0.24)	1.26 bd (± 0.05)
19 – Lukullus	24.90 a (± 1.98)	6.68 bd (± 0.24)	3.18 a (± 0.05)
	**	**	**
I – 5211	2.23 b (± 0.34)	7.68 bde (± 0.15)	1.66 bde (± 0.08)
I – 1152	4.08 b (± 0.33)	9.27 bc (± 0.21)	2.43 bc (± 0.03)
Expected value	**43.45**	**8.00**	**2.05**
52 – *L. parviflorum*	0.68 b (± 0.03)	10.76 a (± 0.31)	1.13 bdf (± 0.02)
11- Limachino	86.21 a (± 9.70)	5.10 bdf (± 0.13)	2.97 a (± 0.02)
	**	**	**
I – 6021	7.60 b (± 0.89)	8.16 a (± 0.72)	2.03 bc (± 0.07)
Expected value	**16.20**	**6.91**	**1.66**
60 – *L. pimpinellifolium*	1.98 b (± 0.21)	8.59 a (± 0.20)	0.86 bd (± 0.03)
21 – Marglobe	30.42 a (± 3.46)	5.23 b (± 0.27)	2.45 a (± 0.03)
	**	**	**
I – 1560	12.53 b (± 0.67)	5.25 b (± 0.25)	2.38 bc (± 0.07)
Expected value	**64.53**	**7.18**	**1.93**
60 – *L. pimpinellifolium*	1.98 b (± 0.21)	8.59 a (± 0.20)	0.86 bd (± 0.03)
15 – Ace	127.07 a (± 22.79)	5.77 b (± 0.46)	3.00 a (± 0.01)
	**	ns	**
E – 1922	33.83 a (±1.27)	6.19 (± 0.54)	3.21 a (± 0.08)
E – 2219	23.77 b (± 3.10)	5.71 (± 0.21)	3.20 ac (± 0.04)
Expected value	**23.84**	**6.44**	**2.39**
19 – Lukullus	24.89 b (± 1.98)	6.68 (± 0.24)	3.18 ac (± 0.08)
22 – San Marzano	22.79 b (± 2.45)	6.20 (± 0.16)	1.59 bd (± 0.02)

Numbers between brackets correspond to standard error; I = inter-taxon crosses; E = intra-taxon crosses; first two digits correspond to female parent; same letters show no statistical differences.

Fruit weight for inter-taxon crosses showed the most remarkable diversity between the species involved in the crosses. Fruits of wild species were very small, averaging between 0.68 and 3.46 g. In contrast *L. esculentum* cultivars utilised as parents averaged between 127.07 and 22.79 g, approximately 34 times greater than those from wild species. Statistically significant differences ($P<0.01$) were found in all F_1 populations from inter-taxon crosses and means were further away from the expected values. In the case of intra-taxon crosses, most F_1 populations showed significant differences ($P<0.01$) between means, except the cross of cultivars San Marzano and Marglobe that presented no differences. Fruits tended to be smaller than or very close to expected values.

The mean soluble solid concentration in inter-taxon crosses ranged from 5.25 to 9.27 °Brix, about 1.3 times less than the value of 7.09 – 10.76 °Brix observed in tomato wild relatives. But these values were about 1.3 bigger than the range of 5.10 – 6.68 °Brix in tomato cultivars utilised as parents. Most populations presented means significantly higher than their *L. esculentum* parents, except one cross involving *L. pimpinellifolium* and cultivar Ace (I-1560) that did not differ statistically from the value of *L. esculentum* parent. Most intra-taxon crosses did not differ significantly; the means of F_1 populations were similar to the means of parents and therefore close to the expected value. Only three crosses differed significantly (E-1119 and E-1911, E-1121 and E-2111, and E-2221).

Heavier seeds than the wild parent were found in all inter-taxon crosses (high significance $P<0.01$). The 1000 seeds weight of F_1 populations ranged from 1.66 to 3.48 g, while wild relatives parents were 2 to 2.7 times smaller ranging from 0.86 to 1.26 g, and the seeds were very similar to tomato cultivars (1.59 to 3.18 g). All intra-taxon crosses showed significantly different seed weight ($P<0.01$), most of them had heavier seeds than the expected value and sometimes higher than both parents.

4.6 Discussion

Germplasm selection and characterisation

Morphological differences between some selected *L. esculentum* cultivars, based on the studied traits, are relatively limited. Most of the selected morphological

characters involved fruit characteristics, such as shape, colour, and size, which is the product of years of tomato breeding. Therefore, less similarity was expected in relation to wild relatives that present non-domesticated characteristics. Consequently, it is interesting to include very old and relatively modern cultivars in a genetic base broadening project. Such cultivars can contain genes for characters that have been discarded some years ago, but which today could be part of novel genetic combinations and provide more alternatives for exploitation of genetic variability in a breeding programme.

In respect to tomato wild relatives, selection was confined mainly to plants that produced flowers and fruits during the growing season in the greenhouse. In this study, due to time constraints it was not possible to test large number of species and accessions for response to greenhouse conditions. Although basic knowledge extracted from reports and books provided information on the species, each accession responded in different ways. In addition, there was the need to select species and accessions that had coincidence of flowering.

The self-incompatible *L. hirsutum* var. *glabratum*, accession PI-199381 was selected partly because of coincidence of flowering with other chosen accessions, and also for the large amount of flowers produced under the greenhouse growing conditions. This fact ensured enough pollen for cross-pollination and flowers for emasculation. Furthermore, Martin (1962) reported that this subspecies is more tolerant of foreign pollen than is the typical form of this species. Most importantly, this accession was the only one producing fruits during the growing season. This was also observed with *L. pennellii* var. *puberulum* (LA-1926).

In the case of molecular markers, genetic similarity matrices and dendrograms confirmed the first choice of parents based on morphological characters (Section 3.3). Differences between tomato and wild relatives were sufficient to discriminate between them using molecular markers. Cultivars Lukullus and Marglobe had the closest relationship, with nearly 85% similarity, for both markers. This in part is a reflection of the markers used in the study, but is reinforced by morphological characters such as fruit shape and growth type. However, these cultivars belong to completely different ages and geographical regions of breeding. Marglobe is an old

American OP cultivar for field production and Lukullus a relatively modern OP British cultivar for greenhouse purposes. Cultivar San Marzano grouped in the same cluster with Lukullus and Marglobe, but it was distant from them, almost 60% in both molecular markers. This distance is due to the fact that it has been selected for the different genetic characteristics of processing tomato: high solid soluble content, cylindrical fruit shape, and adaptation to high input production systems. Cultivars Limachino and Ace grouped together, but still there were differences observed between them, they were not identical on a molecular basis. The *esculentum* cluster included *L. esculentum* var. *cerasiforme*, but *L. pimpinellifolium* was included only in RAPD data. This indicates that *L. pimpinellifolium* is also closely related to *L. esculentum*.

Continuous characters showed statistical differences between the accessions selected as parents. For example, diameter, length and weight of fruits showed a clear difference between the cultivars that produce bigger fruits such as Limachino and Ace with all other accessions. There were also statistical differences between *L. esculentum* accessions and wild relatives, because wild tomato accessions tend to produce smaller fruits (usually less than 2 cm in diameter and length). In fruit ratio, wild accessions presented means very close to 1, i.e. means almost round shaped, but *L. esculentum* accessions showed more variability, because they were selected for different shapes (Ku *et al.*, 1999). Important differences were found in soluble solid content. All wild accessions presented statistically higher sugar contents than tomato cultivars, demonstrating again the availability of interesting characteristics in these wild types. This character is the most important for processing tomato, however progress in gene introgression has been hampered because of the linkage of this character with small fruit size, indeterminate growth habit and poor fruit set (MacGillivray and Clemente, 1956; Stevens and Rudich, 1978). Statistical differences were also present between small seeds from wild types and larger seeds from *L. esculentum* cultivars. Related *Lycopersicon* spp produced more and smaller seeds that can be easily disseminated, thus increasing the probability of finding a suitable environment for germination and survival. Conversely, tomato cultivars have been bred for larger seeds with the aim of giving uniform germination and high

vigour under direct field sowing and therefore do not depend upon natural vectors for dispersal, such as wind or birds(Doganlar *et al.*, 2000).

Microsatellites and primer selection

Only 6 microsatellite and 7 RAPD primers were found useful after screening complete sets of oligonucleotide primers. The main requirement for these primers was the display of polymorphisms between the selected parent accessions, allowing them to be characterised and discriminated. This was possible except that two *L. esculentum* cultivars (Marglobe and Lukullus) presented 100% similarity microsatellites, but the same accessions showed differences with RAPD markers.

The differences observed between microsatellite and RAPD markers are the results of the different molecular characteristics of each technique. Polymorphisms in microsatellites are due to slippage of the DNA polymerase during replication (Schlötterer & Tautz, 1992), and also the length of the polymorphism is affected by recombination, insertions and other genetic effects. The primer aims to anneal to a specific locus, usually of nuclear origin. In microsatellites the polymorphisms are due to length variation between alleles. Conversely, RAPD variation is due to mutational effects, with different evolutionary implications, that take place in the annealing site of the primer and between the two adjacent sites responsible for the amplification. Disappearance of bands can be due to base changes (inversions, insertions, or deletions) that change primer targeting sites. RAPD fragments from total extracted DNA are generated from nuclear, chloroplast and mitochondrial genomes. They are not locus specific.

Parents hybridisation

Results of crossing the different *Lycopersicon* spp accessions selected as parents showed the degree of incompatibility expected between red- and green-fruited species (Stevens and Rick, 1986). No cross involving both green-fruited species, *L. pennellii* var. *puberulum* and *L. hirsutum* var. *glabratum* as female set fruits. This is in complete agreement with the subgeneric classification of genus *Lycopersicon* by

Rick (1976). Unilateral incompatibility is the main cause of crossing failure (Chetelat and De Verna, 1991; Foolad, 1996); crosses between *L. esculentum* and *L. pennellii* and *L. hirsutum* are only possible in one direction, green-fruited species as staminate parent. A multigenic and unilateral incompatibility system determines the direction of fertility (Rick, 1969). This incompatibility is manifested by inhibition of *L. esculentum* pollen tube growth in stigma, style or ovaries of *L. pennellii* and *L. hirsutum* (Hardon, 1967).

In the case of intra-taxon *L. esculentum* crosses, low fecundity between some accessions may be explained by a lack of coincidence between pollen maturation and stigma receptiveness. Although pollination was carried out on several occasions on flowers at different stages, stigma receptivity could have varied between the different accessions and fertilisation could not occur.

F_1 generation: morphological and molecular characterisation

The F_1 generation of a cross between two completely homozygous lines has the highest degree of heterozygosity, but they are completely homogeneous. The results of the inter-taxon crosses showed phenotypic uniformity in each F_1 population for the characteristics studied, but most of them expressed dominant characters of tomato wild relatives. For instance, all crosses had indeterminate growth phenotypes, but genotypes of crosses containing as parent cv. Limachino should be sp^+/sp. Therefore determinate growth phenotype (*sp*) is recessive.

Phenotypes for many characters in nature and agriculture show continuous variation. This continuous distribution has been attributed to the collective action of several genes interacting with the environment. Therefore the phenotypic response to environmental stresses change between and sometimes within sites. However, within an F_1 generation it is possible to observe responses like heterosis when crossing two homozygous lines. However, the amount of heterosis is dependent on the genetic difference between the parents (Wricke and Weber, 1986). No heterosis was observed for most of the characters examined in inter- or intra-taxon crosses, except seed weight that presented an increase over the average of both parents or expected value.

All characters involving fruit size (diameter and length) and weight were distant from expected values, dominance of small fruit was evident. In inter-taxon crosses, when large- and small-fruited accessions were crossed, the fruit size of the F_1 hybrids typically resembled that of the smaller fruited parent (MacArthur and Butler, 1938). In intra-taxon crosses, the presence of *oblate* (*o*) gene that elongates the fruit tended to decrease the size, whereas the genes for *fasciation* (*f*) and *tangerine* (*t*) and those increasing locule number increased fruit size (MacArthur and Butler, 1938). For fruit ratio the high level of dominance of the characteristics of the wild accessions was shown in the F_1 generation; the means of F_1s for most characters were almost all different from expected values and similar to those of the wild parent. The wild relative alleles always gave smaller values for the fruit shape ratio, in that way showing a tendency to more round-shaped fruits. In the case of intra-taxon hybridisations, most crosses also showed a tendency toward round to slightly flattened fruits and some hybrids were distant from expected value.

Solid soluble content in most inter-taxon crosses presented the effects of wild relative accessions in the direction expected, increasing the mean concentration; however there were linkages with undesirable characters for the tomato processing industry such as indeterminate growth and small fruit size. Grandillo and Tanksley (1996) found that a major Quantitative Trait Loci (QTL) for this charactyeristic is probably a pleiotropic effect of the gene for the indeterminate growth habit (sp^+). Most intra-taxon crosses showed no statistical variation, as expected in both cases.

The weight of 1000 seeds had opposite effects for inter-taxon crosses to that expected based on the parental means. Heterotic increase of the seed weight was observed in inter- and intra-taxon crosses, where most F_1 accessions presented higher weight of 1000 seeds than expected. Although this type of change is not desirable in tomato breeding, this is an example of what occurs when using diverse gene-pools as parental sources.

Chapter 5

Selfing effect and genetic diversity in created *Lycopersicon* populations

5.1 Introduction

The loss of genetic diversity and, therefore, variability over time in agricultural crops reduces the genetic material available for use by present and future generations. Modern breeding techniques and objectives are leading to crop varieties with potentially dangerous uniformity, in response to market needs and registration (patents) laws.

The incorporation of genes from wild relatives and discarded genetic material into adapted, but genetically depauperate, breeding material is the objective of a genetic base broadening programme to increase genetic diversity. But in autogamous species, inbreeding results in homozygosity, which can again lead to reduced genetic variability in individual populations in a few generations, and therefore a decreasing of the total genetic diversity of the species.

The objective of this chapter is to show the behaviour of F_2 and F_3 generations created from particular inter- and intra-taxon crosses of *Lycopersicon* spp from a morphological and molecular point of view. The material will also be examined as total populations (inter-taxon and intra-taxon) to observe some of the features that one might expect if, as would be expected in a base broadening programme, the products of individual crosses were bulked and treated as one population.

5.2 Morphological characteristics of F_2 and F_3 populations

Most standards established by UPOV for "DUS" (distinctness, uniformity and stability) are based on morphological characteristics of the species. Once a new breed is presented as a cultivar, it is tested by the Official Governmental Organisation and if it satisfies the appropriate requirements, it can be named and commercialisation can start. For breeders, morphological traits continue to be of great importance because it is possible to assess many plants segregating for different characters readily and rapidly.

Segregating populations are the base of any breeding and genetic base broadening programme.

5.2.1 Discrete characters

An assessment of morphological characters from F_2 and F_3 segregating populations was carried out and the results are displayed in Table 5.1. From inter-taxon crosses containing accessions of *L. esculentum* var. *cerasiforme* and *L. esculentum*, 40 plants were analysed for morphological characters; in the case of *L. parviflorum and L. pimpinellifolium* 16 and 20 plants in the intra-taxon *L. esculentum* crosses. Unfortunately it was not possible to assess the segregation ratios in the crosses because of the small number of plants grown per cross for reasons of space in greenhouse. However, a descriptive analysis was carried out. From the 18 characters analysed, 4 were monomorphic for all individuals and the remaining were polymorphic. The monomorphic characters included the absence of hypocotyl pubescence, horizontal leaf attitude, multiparous inflorescence type, and yellow corolla colour. Hypocotyl colour, in all individuals and for all crosses and generations, showed anthocyanin presence, probably due to homozygosity for gene *A*. But the *anthocyanin loser* gene (*al*) did segregate in all of them, displaying phenotypes from ¼ and ½ purple hypocotyl (*alal*) to full purple (*Al-*). In respect to plant growth type controlled by the self-pruning gene *sp*, products had indeterminate growth in crosses involving *L. esculentum* var. *cerasiforme* and *L. pimpinellifolium* with *L. esculentum,* including the crosses with determinate growth accessions. Lines from crosses with *L. parviflorum* showed a different growth character, vine-type growth, but indeterminate. This character description was not found in papers reviewed, or in tomato genes databases. The intra-taxon crosses between *L. esculentum* accessions presented both growth types, indeterminate and determinate, segregating when the crosses involved accessions containing the gene *sp*.

Crosses within *L. esculentum* and with *L. esculentum* var. *cerasiforme* showed standard and potato leaf types, depending on the type of cross. *Potato leaf* type is a recessive character controlled by the gene *c*, which produces leaves less segmented. As was expected, most individuals observed showed the standard leaf type in these accessions. Crosses containing *L. parviflorum* as parents displayed individuals with the standard leaf type, but also segregated to '*parviflorum*' leaf type. The phenotype of this trait is characterised by small, relatively simple leaves carried on slender

stems. Similar results were observed in the case of crosses with *L. pimpinellifolium*, which presented the standard and the '*pimpinellifolium*' leaf type, characterised by a lack of deeply serrated leaf margins. In these last two crosses, wild characters were presented by most individuals in both F_2 and F_3 generations.

Style position is a trait controlled by the exserted style gene (*ex*), found only in wild species. The material containing *L. parviflorum* as a parent showed segregation in F_2 and F_3 generations of individuals that presented exserted and inserted styles. The individuals from crosses involving *L. pimpinellifolium* displayed inserted and exserted styles in F_2, but in F_3 only the inserted type. All individuals from the other inter-taxon crosses (*L. esculentum* x *L. esculentum* var. *cerasiforme*) and intra-taxon crosses within *L. esculentum* had styles inserted or at the same level of stamens. This was expected because these two species are closely related and *L. esculentum* var. *cerasiforme* is mainly an autogamous species with inserted styles.

For exterior colour of immature fruits no gene descriptions were found in the literature and *Lycopersicon* genes databases, although it is used as a descriptor character in Descriptors of Tomato (IPGRI, 1996). There were patterns observed in this trait. Intra-taxon crosses and crosses involving *L. esculentum* var. *cerasiforme* showed light green and green immature fruits, but the segregation in crosses involving *L. parviflorum* and *L. pimpinellifolium* produced light green and greenish white fruits.

Fruit pubescence is a trait controlled by the gene *peach* (*p*). All individuals of intra-taxon crosses showed sparse fruit pubescence, but inter-taxon crosses containing *L. esculentum* var. *cerasiforme* and *L. pimpinellifolium* presented segregation in F_2 and F_3 generations for this trait. All fruits and therefore individuals in accessions that include *L. parviflorum* as parent displayed intermediate fruit pubescence in F_2 and F_3 generations. Therefore it is likely that there may be a modifier gene present in the *L. parviflorum* accession.

Table 5.1 General morphological characteristics of F_2 and F_3 generations from inter-taxon crosses involving 3 wild relatives and *L. esculentum*, and intra-taxon crosses within *L. esculentum* accessions.

	L. esculentum var. *cerasiforme*		*L. parviflorum*		*L. pimpinellifolium*		*L. esculentum*	
Character	F_2	F_3	F_2	F_3	F_2	F_3	F_2	F_3
Hypocotyl colour	½ purple to full purple	½ purple to full purple	¼ purple to full purple	¼ purple to ½ purple	½ purple to full purple	½ purple	½ purple to full purple	½ purple to full purple
Hypocotyl pubescence	Absent	Absent	Absent	Absent	Absent	Absent	Absent	Absent
Plant growth type	Indeterminate	Indeterminate	Semi-determinate	Semi-determinate	Indeterminate	Indeterminate	Indeterminate and determinate	Indeterminate and determinate
Leaf attitude	Horizontal	Horizontal	Horizontal	Horizontal	Horizontal	Horizontal	Horizontal	Horizontal
Leaf type	Potato leaf type and standard	Potato leaf type and standard	Parviflorum type and standard	Parviflorum type and standard	Pimpinellifolium type and standard	Pimpinellifolium type and standard	Potato leaf type and standard	Potato leaf type and standard
Inflorescence type	Multiparous	Multiparous	Multiparous	Multiparous	Multiparous	Multiparous	Multiparous	Multiparous
Corolla colour	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow
Style position	Inserted	Inserted and the same level as stamen	Inserted and slightly exserted	Inserted and highly exserted	Inserted and slightly exserted	Inserted and the same level as stamen	Inserted and the same level as stamen	Inserted and the same level as stamen
Exterior colour of immature fruit	Light green	Light green and green	Greenish white and light green	Greenish white and light green	Light green and greenish white	Light green and greenish white	Light green and green	Light green and green
Fruit pubescence	Sparse and intermediate	Sparse and intermediate	Intermediate	Intermediate	Intermediate and sparse	Sparse and intermediate	Sparse	Sparse
Predominant fruit shape	Rounded and slightly flattened	Rounded and slightly flattened	Slightly flattened and rounded	Slightly flattened and rounded	Slightly flattened and rounded	Slightly flattened	All shapes	All shapes
Fruit size	Very small and small	Very small and small	Small	Small	Very small and small	Very small and small	Small and intermediate	Small and intermediate
Exterior colour of mature fruits	Red	Red	Red	Red	Red	Red	Red and orange	Red and orange
Flesh colour of pericarp	Orange, pink and red	Orange, pink and red	Orange and pink	Orange, pink and red	Orange and pink	Pink and red	Red and orange	Red and orange
Fruit cross-sectional shape	Round	Round	Round and angular	Round and angular	Round	Round	Round, angular and irregular	Round, angular and irregular
Number of locules	Two and three	Two and three	Two and three	Two and three	Two	Two	Two to eight	Two to eight
Shape of pistil scar	Dot	Dot	Dot	Dot	Dot	Dot	Dot and irregular	Dot and irregular
Fruit blossom end shape	Flat and indented	Flat	Flat	Flat	Flat	Flat	Flat and indented	Flat and indented

The segregation in inter-taxon crosses showed a tendency toward the wild character, i.e. pubescent fruits.

Fruit shape is a character controlled by several genes with different effects. For example *ovate* (*o*) controls the length/diameter ratio, *fasciated* (*f*) controls the number of locules together with *lc* (*locule number*), and others (Fryxell, 1954; Rick and Butler, 1956). Crosses within *L. esculentum* showed the expected segregation in F_2 and F_3 generations in respect to the fruit shape of the parents. Different shapes could be found, from very flattened to plum types. These characters have been selected and propagated by breeders for many years. However, in crosses involving related tomato species, the fruit tended toward round or slightly flattened shapes, an apparently dominant character present in wild types.

In respect to fruit size, Grandillo *et al.* (1999) reported that there are several QTLs controlling this character. The results of F_2 and F_3 generations for inter-taxon crosses showed that individuals tended to produce larger fruits than the wild relative parent, but smaller than the respective *L. esculentum* accession involved. In the case of intra-taxon crosses, there was segregation, also expected depending on the parents involved, from small to intermediate fruit size. Large fruits were not expected in crosses involving cv. Ace.

Exterior colour of mature fruits was red for most of the individuals in the different generations and crosses. The exception was in intra-taxon '*esculentum*' crosses containing cv. Limachino, where there were individuals segregating to orange colour. Fruit colour is controlled by genes *red* (*R*), *yellow skin* (*Y*) and *tangerine* (*t*). The same genes *red* and *tangerine* are involved in flesh colour of the pericarp that combined with skin colour, gives the fruit colour. There was segregation for this character in both types of crosses, inter- and intra-taxon. Accessions containing *L. esculentum* var. *cerasiforme* showed in F_2 and F_3 generation segregation of orange, red and pink colour. Those containing *L. parviflorum* showed orange and pink fruited individuals in F_2, but orange, pink and red fruits in F_3. In respect to *L. pimpinellifolium*, F_2 individuals presented orange and pink flesh colour, but in F_3 mainly pink and red. In intra-taxon crosses only the populations with cv. Limachino

as parent displayed segregation to red and orange, but all the other combinations were red.

Fruit cross-sectional shape is linked to fruit shape genes. Most individuals in all accessions showed round shape, but in accessions containing *L. parviflorum* there was segregation to angular shape, or more square fruits. In the case of *L. esculentum* crosses, segregation depended on the parents utilised, but individuals showed round, angular and irregular sectional shape.

Number of locules is a trait controlled by gene *few locules* (*Lc*), usually two locules are formed. This character varied in all crosses and generations depending on the parents involved. Segregation was mostly toward two or three locules. In the case of *L. esculentum* crosses, there were individuals presenting two to eight locules. Segregation in both generations was observed.

The shape of the pistil scar was a dot in all inter-taxon crosses accessions, but in intra-taxon crosses there were few individuals showing an irregular shape. Fruit blossom end shape was flat in most individuals in any of the inter- or intra-taxon crosses. However there was segregation in F_2 plants of crosses involving *L. esculentum* var. *cerasiforme* to indented shape only when cv. San Marzano was the other parent. A similar case was observed in F_2 and F_3 generation for intra-taxon crosses.

5.2.2 Continuous characters

Although continuous characters by their nature are variable depending on the environmental conditions, six traits were statistically analysed in F_2 and F_3 generations, to show the differences between crosses (Table 5.2). The number of plants examined was the same as in section 5.2.1 and one fruit was collected per plant for continuous characters analysis. The accessions were grouped in three inter-taxon types of crosses containing one common wild relative as parent and one intra-taxon group with the *L. esculentum* crosses. ANOVA procedure was performed to detect statistical significance between groups and a Tukey's test for multiple comparisons. Details are given in Appendix 3 part 10.

Fruit weight differed significantly ($P<0.01$) in F_2 and F_3 generations. In both cases, as expected, the *L. esculentum* group had larger fruits with means (F_2 with 34.48 g and F_3 with 59.57 g) significantly different from the other groups. The smallest mean was measured for fruits of the *L. parviflorum* group (F_2 with 3.98 g and F_3 with 6.46 g). The groups containing *L. esculentum* var. *cerasiforme* and *L. pimpinellifolium* did not present statistical differences between them in F_2 but they were different from *L. parviflorum*. However, in F_3 generation all three groups were not statistically different.

In F_2 and F_3 generations fruit diameter was differed significantly ($P<0.01$). The means for both generations of *L. esculentum* were larger and different ($P<0.01$) from the other groups. They showed 4.08 cm and 4.95 cm for the F_2 and F_3, respectively. The shortest diameter was displayed by *L. parviflorum* with means of 1.86 cm in F_2 and 2.34 cm in F_3 generation. This group was statistically different from *L. esculentum* var. *cerasiforme* and *L. pimpinellifolium*, however the latter two groups did not show statistical differences.

In respect to fruit length, both generations showed statistically significant differences ($P<0.01$). The group of *L. esculentum* displayed the largest means with 3.51 cm for F_2 and 4.26 cm for F_3 plants. The smallest means in length were observed in the *L. parviflorum* group with 1.61 cm for F_2 and 1.92 cm for the F_3 generation. In the F_2, all groups were statistically different, but in the F_3 the *L. esculentum* var. *cerasiforme* and *L. pimpinellifolium* groups did not show differences.

There was no statistical significance for fruit ratio (diameter/length) in both F_2 and F_3 generations. The ratios fluctuated from 1.10 to 1.17 for F_2 and from 1.09 to 1.23 in F_3 plants.

Table 5.2 Means, statistical analysis and standard error for continuous characters in inter- and intra-taxon crosses between wild relatives and *L. esculentum*.

Crosses between *L. esculentum* and:	N	Fruit Weight (g)		Fruit Diameter (cm)		Fruit Length (cm)		Fruit Ratio (D/L)		Solid Soluble Content (°Brix)		1,000 Seeds Weight (g)	
		F_2**	F_3**	F_2**	F_3**	F_2**	F_3**	F_2 ns	F_3 ns	F_2 ns	F_3**	F_2 **	F_3**
L. esculentum var. *cerasiforme*	40	12.53bc ±0.97	17.82b ±1.87	2.72bc ±0.08	3.13bc ±0.14	2.51bc ±0.07	2.86bc ±0.14	1.10 ±0.04	1.12 ±0.05	6.24 ±0.14	6.27bc ±0.09	2.53b ±0.10	2.29b ±0.18
L. parviflorum	16	3.98bd ±0.20	6.46b ±0.61	1.86bd ±0.02	2.34bd ±0.16	1.61bd ±0.04	1.92bd ±0.07	1.16 ±0.03	1.23 ±0.07	7.06 ±1.11	7.24a ±0.16	2.23b ±0.06	1.92b ±0.13
L. pimpinelifolium	16	8.23bc ±1.35	13.20b ±0.54	2.33bc ±0.17	2.77bc ±0.06	2.12bc ±0.14	2.61bc ±0.19	1.17 ±0.03	1.09 ±0.07	7.58 ±0.58	7.16a ±0.15	2.28b ±0.12	1.90b ±0.06
L. esculentum	20	34.48a ±2.29	59.57a ±5.41	4.08a ±0.14	4.95a ±0.18	3.51a ±0.05	4.26a ±0.13	1.15 ±0.04	1.18 ±0.04	6.32 ±0.13	5.58bd ±0.14	3.15a ±0.18	3.04a ±0.09

N = number of samples; small numbers correspond to standard error; ** = high significance ($P<0.01$); * = significance ($P<0.05$); ns = no statistical significance; same letters show no statistical significance.

Solid soluble content showed no statistical significance in F_2 generation, presenting values from 6.24 °brix for the *L. esculentum* var. *cerasiforme* group to 7.58 °brix for the *L. pimpinellifolium* group. In F_3 there were significant differences ($P<0.01$), with the *L. parviflorum* group having the largest mean with 7.24 °brix and *L. esculentum* group with the lowest, 5.58 °brix. There was no statistical difference between *L. parviflorum* and *L. pimpinellifolium* groups, but *L. esculentum* var. *cerasiforme* and *L. esculentum* groups were different in respect to the other groups.

The 1,000 seed weight showed statistically significant differences ($P<0.01$) in F_2 and F_3 generations. The group of *L. esculentum* presented the heavier seeds (3.15 g in F_2 and 3.04 g in F_3), statistically different from any other group. However, there were no statistical differences between the means of *L. esculentum* var. *cerasiforme*, *L. parviflorum*, and *L. pimpinellifolium*.

5.2.3 Genetic diversity in groups of crosses

A statistical analysis of each genetic diversity index (A, A_e, H_S, and I) of morphological characters, analysed as dominant marker (presence (1) or absence (0) of a character), was carried out for F_2 and F_3 generations. As stated in section 5.2.1, the same number of individuals were sampled for these analyses. Inter-taxon crosses of tomato wild relatives and *L. esculentum* accessions were grouped according to the wild type utilised as one parent, and intra-taxon crosses within *L. esculentum* accessions were treated as a further group. Genetic indices displaying significant differences between groups are indicated by asterisks (Table 5.3). The statistical analysis was performed using the ANOVA procedure for genetic indices that are normally distributed such as H_S and I. Non-normally distributed indices such as A and A_e, were analysed with the non-parametric Kruskal-Wallis test. For the number of polymorphic loci (P) there was no statistical analysis because of the few number of observations available per group. A Tukey's test was carried out and small letters were used to specify significances between different groups (Table 5.3). Details of the statistical analysis are shown in Appendix 3, part 11.

The results (Table 5.3) show a significantly lower number of polymorphic alleles per locus (A) and proportion of polymorphic loci (P) for *L. esculentum* and *L. esculentum*

var. *cerasiforme* groups compared with *L. parviflorum* and *L. pimpinellifolium* in F_2 and F_3 generations. But significantly higher diversity indices (A_e, H_S, and I) were produced in F_2 for *L. parviflorum* and *L. pimpinellifolium* groups than for *L. esculentum* and *L. esculentum* var. *cerasiforme* groups. In the F_3 these three indices, however, showed no statistical difference. It is possible that these indices were higher in F_2 generation because they had less alleles in common between utilised parents than in the case of *L. esculentum* and *L. esculentum* var. *cerasiforme*, therefore the heterozygosity was higher.

5.2.4 Genetic diversity in bulked crosses

Another approach to the analysis of the genetic diversity indices was carried out for the F_2 and F_3 generations: all crosses of tomato wild relatives with *L. esculentum* were grouped in one bulk population of inter-taxon crosses, and crosses involving only *L. esculentum* in intra-taxon crosses. This could be regarded as two approaches to base broadening, i.e. limiting the pool of genetic variability to within a species (i.e. *L. esculentum*) or to deliberately incorporate variability from a much wider source including likely progenitor species (i.e. *L. esculentum* x wild relatives).

The results displayed in Table 5.4 show a tendency to decrease the value from F_2 to F_3 generation in most of the indices of inter-taxon crosses. However, F_{st} was identical in both generations. In the case of intra-taxon crosses, indices such as A, A_e, and P show identical values in both generations, but all other indices showed the same trend as in inter-taxon crosses. Inter-taxon crosses showed higher values than intra-taxon crosses in most indices such as A, A_e, H_S, I, and P, but lower for F_{st}.

Overall, total gene diversity (H_t) in the F_2 was 0.39 (±0.01) and the fixation index (F_{st}) was 0.37 for inter-taxon crosses, while in F_3 plants H_t was 0.35 (±0.02) and F_{st} was identical. In the case of intra-taxon crosses, the F_2 had values of H_t = 0.15 (±0.04) and F_{st} = 0.42, but in the F_3 H_t increased to 0.16 and F_{st} decreased to 0.4.

Table 5.3 Genetic diversity indices for morphological characters in F_2 and F_3 generations of crosses involving *L. esculentum* and wild relatives.

Crosses between *L. esculentum* and:	N	*A*		A_e		H_s		*I*		*P*	
		F_2*	F_3*	F_2**	F_3 ns	F_2**	F_3 ns	F_2**	F_3 ns	F_2	F_3
L. esculentum var. *cerasiforme*	40	1.20b ±0.06	1.16b ±0.05	1.16b ±0.05	1.13 ±0.05	0.13bd ±0.03	0.11 ±0.03	0.19bd ±0.04	0.16 ±0.03	0.20	0.16
L. parviflorum	16	1.55a ±0.12	1.54a ±0.12	1.43a ±0.10	1.38 ±0.10	0.24a ±0.05	0.22 ±0.05	0.34a ±0.08	0.32 ±0.07	0.55	0.54
L. pimpinelifolium	16	1.61a ±0.12	1.52a ±0.13	1.50a ±0.11	1.38 ±0.10	0.27a ±0.06	0.21 ±0.05	0.38a ±0.08	0.31 ±0.07	0.61	0.52
L. esculentum	20	1.33b ±0.10	1.33b ±0.10	1.27b ±0.08	1.27 ±0.09	0.15bc ±0.04	0.16 ±0.05	0.22bc ±0.06	0.23 ±0.06	0.33	0.33

N = number of individuals sampled; *A* = number of polymorphic alleles per locus; A_e = effective number of alleles; H_s = average gene diversity; *I* = Shannon's information index; P = proportion of polymorphic loci; small numbers correspond to standard error; ** = high significance (P<0.01); * = significance (P<0.05); ns = no statistical significance; same letters show no statistical significance.

Table 5.4 Genetic diversity indices for morphological characters in bulks of populations of *Lycopersicon* F_2 and F_3 generations in inter- and intra-taxon crosses.

	N	A		A_e		H_s		H_t		I		F_{st}		P	
Crosses:		**F_2**	**F_3**	**F_2**	**F_3**	**F_2**	**F_3**	**F_2**	**F_3**	**F_2**	**F_3**	**F_2**	**F_3**	**F_2**	**F_3**
Inter-taxon	72	2.15 ±0.03	2.11 ±0.04	1.38 ±0.02	1.32 ±0.06	0.25 ±0.01	0.22 ±0.05	0.39 ±0.01	0.35 ±0.02	0.42 ±0.01	0.38 ±0.06	0.37	0.37	0.90	0.86
Intra-taxon	20	1.33 ±0.10	1.33 ±0.10	1.27 ±0.08	1.27 ±0.09	0.15 ±0.04	0.16 ±0.05	0.26 ±0.02	0.26 ±0.02	0.22 ±0.06	0.23 ±0.06	0.42	0.40	0.33	0.33

N = number of individuals sampled; A = number of polymorphic alleles per locus; A_e = effective number of alleles; H_s = average gene diversity; H_t = total gene diversity; I = Shannon's information index; F_{st} = fixation index; P = proportion of polymorphic loci; small numbers correspond to standard error

5.3 Molecular markers and population characteristics

Molecular markers have been very useful tools to provide information that either confirms previous evidence based on morphological characteristics and/or provides further evidence. In the case of generation analysis after hybridisation and selfing, they can provide valuable information about segregation of alleles at the molecular level.

5.3.1 Microsatellite markers analysis in F_2 and F_3 generations

A statistical analysis of each genetic index was performed between the results of four groups of inter- and intra-taxon crosses, namely *L. esculentum* var. *cerasiforme*, *L. parviflorum*, *L. pimpinellifolium*, and *L. esculentum*. The number of markers and bands examined are presented in Table 4.4 and the data were treated as co-dominant markers.

The statistical analysis of the results was carried out using the ANOVA procedure for genetic indices that are normally distributed such as H_S and I. Non-normally distributed indices (A and A_e) were analysed with the non-parametric Kruskal-Wallis test. In the case of statistical significance Tukey's test was performed for multiple comparisons. Details of the statistical analysis are given in Appendix 3 part 12.

For the number of polymorphic loci (P) there was no statistical analysis because of the few number of observations available per group, because the statistical package Popgene takes all loci in each group as one and gives just one value.

The results (Table 5.5) show a significantly ($P<0.05$) lower number of polymorphic alleles per locus (A) for the *L. esculentum* group compared to three inter-taxon crosses in F_2. The F_3 generation did not give a statistical significance. The effective number of polymorphic alleles (A_e) was statistically significant ($P<0.05$) in F_2 generation. *L. esculentum* var. *cerasiforme* and *L. parviflorum* presented no difference between them, nor did *L. pimpinellifolium* and *L. esculentum* groups. However in F_3 there was no statistical significance between groups.

Average gene diversity (H_S) and Shannon's information index (I) were statistically significant ($P<0.05$) in the F_2 generation, but no significance in the F_3. In the case

Table 5.5 Genetic diversity indices from microsatellite data of F_2 and F_3 generations in 3 inter-taxon crosses and one intra-taxon cross of *Lycopersicon* spp.

Crosses between *L. esculentum* and:	N	*A*		A_e		H_s		*I*		*P*	
		F2*	F3 ns	F2*	F3 ns	F2*	F3 ns	F2*	F3 ns	F2	F3
L. esculentum var. *cerasiforme*	40	1.71a ±0.09	1.57 ±0.09	1.70a ±0.04	1.48 ±0.08	0.35a ±0.06	0.26 ±0.06	0.49a ±0.09	0.37 ±0.09	0.71	0.57
L. parviflorum	16	1.86a ±0.06	1.57 ±0.09	1.82a ±0.06	1.49 ±0.10	0.40a ±0.05	0.26 ±0.06	0.57a ±0.06	0.37 ±0.09	0.86	0.57
L. pimpinellifolium	16	1.71a ±0.08	1.86 ±0.09	1.54b ±0.07	1.73 ±0.08	0.27a ±0.06	0.39 ±0.05	0.39a ±0.08	0.56 ±0.06	0.71	0.86
L. esculentum	20	1.29b ±0.05	1.57 ±0.09	1.28b ±0.05	1.44 ±0.07	0.23b ±0.06	0.24 ±0.04	0.32b ±0.08	0.33 ±0.08	0.29	0.57

N = number of individuals per cross analysed; *A* = number of polymorphic alleles per locus; A_e = effective number of alleles; H_s = average gene diversity; *I* = Shannon's information index; P = proportion of polymorphic loci; small numbers correspond to standard error; ** = high significance ($P<0.01$); * = significance ($P<0.05$); ns = no statistical significance; same letters show no statistical significance.

of F_2 plants, the *L. esculentum* group displayed the lowest H_S value of 0.23, which was different from the other groups. In *I*, the differences were similar, but the lowest value for the *L. esculentum* group was 0.32.

The proportion of polymorphic loci (*P*) showed a low value for the *L. esculentum* group (0.29), but *L. esculentum* var. *cerasiforme* and *L. pimpinellifolium* had similar means (0.71), while *L. parviflorum* had the highest (0.86) for the F_2 generation. In the F_3, only the *L. pimpinellifolium* group showed a different value (0.86), all other groups had identical values (0.57).

Overall, total gene diversity (H_t) in the F_2 generation of inter-taxon crosses was 0.51 (±0.01), but in F_3 was 0.39 (±0.02). In intra-taxon crosses H_t was 0.24 (±0.03) for F_2 and 0.26 (±0.04) for F_3. The results of fixation index (F_{St}) for inter-taxon crosses were 0.20 and 0.13 for F_2 and F_3 respectively, but 0.04 in F_2 and 0.08 in F_3 for intra-taxon crosses.

5.3.2 RAPD marker analysis in F_2 and F_3 generations

With the aim to observe the behaviour of dominant markers (presence or absence of bands) on genetic indices, a statistical analysis of each marker was carried out in the four groups of crosses. The results were analysed utilising the ANOVA procedure for genetic indices for normally distributed (H_S and *I*) and the non-parametrical Kruskal-Wallis test for non-normally distributed (*A* and A_e) indices. In the case of *P*, there was no statistical analysis because of the little data available. In indices statistically significant a Tukey's test was done for multivariate analysis. Details of the analysis are given in Appendix 3 part 13 and the number of markers utilised in RAPD analysis are presented in Table 4.4.

The results are shown in Table 5.6. *A* and A_e had high statistical significant differences ($P<0.01$) between means in both generations, F_2 and F_3. In both indices no differences were found between *L. esculentum* var. *cerasiforme* and *L. parviflorum* groups, either between *L. pimpinellifolium* and *L. esculentum* groups. The lowest values for each index and generation were in *L. pimpinellifolium* group, except for A_e in F_2 showing to *L. esculentum* as the lowest value with 1.13.

The index H_S presented high statistical significance ($P<0.01$) for F_2 and F_3 generations. The lowest H_S were in *L. esculentum* group for F_2 with 0.07 and in *L. pimpinellifolium* group for F_3 with 0.10. The *L. esculentum* group showed to be different to all other groups in F_2, but in F_3 this group was not different with *L. pimpinellifolium*. In both generations, *L. esculentum* var. *cerasiforme* group did not have differences with *L. parviflorum* group, but in F_2 this group showed also no differences with *L. pimpinellifolium*.

Shannon's information index (I) exhibited high statistical significance ($P<0.01$) in F_2 and statistical significance ($P<0.05$) in F_3. The lowest values for I in F_2 were in *L. pimpinellifolium* and *L. esculentum* groups with 0.11, and *L. pimpinellifolium* group in F_3 with 0.15. No differences in means were found between *L. esculentum* var. *cerasiforme* and *L. parviflorum* groups in both generations, either between *L. pimpinellifolium* and *L. esculentum* groups. In respect to proportion of polymorphic loci (P), *L. pimpinellifolium* group showed the lowest proportion in both, F_2 and F_3 generations, with 0.19 and 0.25, respectively. The highest proportion was displayed by *L. esculentum* var. *cerasiforme* with 0.46 in F_2 and 0.45 in F_3.

Overall, total gene diversity (H_t) for inter-taxon crosses showed values of 0.23 (±0.02) in F_2 and 0.12 (±0.03) in F_3. For intra-taxon crosses this index was 0.24 (±0.03) in F_2 and 0.14 (±0.05) in F_3. The F_{st} index in inter-taxon crosses in F_2 was 0.37 and 0.18 in F_3. In the case of intra-taxon crosses F_{st} was 0.40 for F_2 and 0.21 for F_3.

Table 5.6 Genetic diversity indices from RAPD data of F_2 and F_3 generations in 3 inter-taxon crosses and one intra-taxon cross of *Lycopersicon* spp.

Crosses between *L. esculentum* and:	N	*A*		A_e		H_s		*I*		*P*	
		F_2**	F_3 **	F_2**	F_3**	F_2**	F_3**	F_2**	F_3*	F_2	F_3
L. esculentum var. *cerasiforme*	40	1.46a ±0.08	1.46a ±0.11	1.29a ±0.06	1.34a ±0.10	0.17a ±0.01	0.18a ±0.02	0.25a ±0.03	0.26a ±0.03	0.46	0.45
L. parviflorum	16	1.35a ±0.16	1.35a ±0.16	1.24a ±0.13	1.27a ±0.13	0.14a ±0.02	0.15a ±0.02	0.20a ±0.03	0.21a ±0.03	0.35	0.35
L. pimpinellifolium	16	1.20b ±0.12	1.26b ±0.15	1.15b ±0.010	1.18b ±0.11	0.08bc ±0.02	0.10b ±0.02	0.11b ±0.02	0.15b ±0.02	0.19	0.25
L. esculentum	20	1.24b ±0.10	1.27b ±0.15	1.13b ±0.06	1.21b ±0.12	0.07bd ±0.01	0.11b ±0.02	0.11b ±0.02	0.16b ±0.03	0.22	0.26

N = number of individuals analysed per cross; *A* = number of polymorphic alleles per locus; A_e = effective number of alleles; H_s = average gene diversity; *I* = Shannon's information index; P = proportion of polymorphic loci; small numbers correspond to standard error; ** = high significance ($P<0.01$); * = significance ($P<0.05$); ns = no statistical significance; same letters show no statistical significance.

5.3.3 Molecular markers and genetic diversity in bulks of populations

A statistical analysis of genetic diversity indices in bulks of F_2 and F_3 populations was carried out grouping the crosses as in Chapter 5 part 2.4. The number of markers utilised are presented in Table 4.4.

The results presented in Table 5.7 showed that using data from co-dominant microsatellite markers most of the indices in inter-taxon crosses decreased from F_2 to F_3 generation, except for the Shannon`s information index (*I*), which increased. However, intra-taxon crosses displayed an increasing tendency in all indices. The comparison of indices between groups showed higher values for inter-taxon than for intra-taxon crosses, with the exception of *I*, which displayed for F_2 generation a lower value in the inter-taxon group.

In the case of RAPD, indices such as *A*, F_{st} and *P* decreased from F_2 to F_3 generation, while A_e, H_s. H_t, and *I* increased in inter-taxon crosses. However in the intra-taxon group the number of polymorphic alleles per locus (*A*), A_e, H_s, H_t, *I*, and *P* increased, while and F_{st} decreased. Most of the indices decreased from inter- to intra-taxon groups, but in F_2 and F_3 generations F_{st} increased.

Table 5.7 Genetic diversity indices from microsatellite and RAPD data in bulks of populations in F2 and F3 generations of *Lycopersicon* inter- and intra-taxon crosses.

Microsatellites	N	A		A_e		H_s		H_t		I		F_{st}		P	
Crosses:		F2	F3	F2	F3	F2	F3	F2	F3	F2	F3	F2	F3	F2	F3
Inter-taxon	72	2.00 ±0.00	1.86 ±0.03	1.75 ±0.00	1.61 ±0.03	0.41 ±0.01	0.34 ±0.02	0.51 ±0.01	0.39 ±0.02	0.20 ±0.01	0.35 ±0.02	0.20	0.13	1.00	0.86
Intra-taxon	20	1.29 ±0.05	1.57 ±0.08	1.28 ±0.05	1.44 ±0.07	0.23 ±0.06	0.24 ±0.04	0.24 ±0.03	0.26 ±0.04	0.32 ±0.08	0.33 ±0.08	0.04	0.08	0.29	0.57

RAPD	N	A		A_e		H_s		H_t		I		F_{st}		P	
Crosses:		F2	F3	F2	F3	F2	F3	F2	F3	F2	F3	F2	F3	F2	F3
Inter-taxon	72	1.74 ±0.04	1.67 ±0.06	1.38 ±0.03	1.42 ±0.05	0.15 ±0.02	0.20 ±0.03	0.23 ±0.02	0.24 ±0.03	0.35 ±0.03	0.36 ±0.04	0.37	0.18	0.74	0.66
Intra-taxon	20	1.24 ±0.10	1.27 ±0.15	1.13 ±0.06	1.21 ±0.12	0.07 ±0.01	0.11 ±0.02	0.12 ±0.03	0.14 ±0.05	0.11 ±0.02	0.16 ±0.03	0.40	0.21	0.22	0.26

N = number of individuals analysed per cross; A = number of polymorphic alleles per locus; A_e = effective number of alleles; H_s = average gene diversity; H_t = total gene diversity; I = Shannon's information index; F_{st} = fixation index; P = proportion of polymorphic loci; small numbers correspond to standard error.

5.4 Results of Genetic Base Broadening in a field trial

As part of the selfing programme of this project, a complete set of F_1 accessions were grown in CRI La Platina, Santiago, Chile. As a result of this programme, data on insect attack on fruits were taken and statistically analysed. The insect under study was the South American tomato pinworm (SATP) (*Tuta absoluta* Meyrick, Lepidoptera – Gelechiidae). This is an insect that has become a serious tomato pest in countries such as Argentina (Bahamondes and Mallea, 1969), Chile (Povolny, 1975; Larraín, 1986; Estay, 1998), Peru, Bolivia, Ecuador, Colombia, Venezuela (Povolny, 1975), Uruguay (Carvallo *et al.*, 1981), and Brazil (Moreira *et al.*, 1981). Severe SATP attack can cause yield losses of up to 100% (Scardini *et al.*, 1982; Espinoza, 1991). Quality standards for both fresh market and processing tomatoes require the industry to rely heavily on the use of pesticides for SATP control. However, genetic resistance or tolerance to SATP in tomato may provide an alternative method for pest control.

In the segregating F_2 generation, the percentage of damaged fruits at harvest was observed and analysed in groups similar to Section 5.2.2. *L. esculentum* parents utilised in the hybridisation were used as control. The data were statistically analysed as a completely random design using ANOVA for significance and Tukey's test for multiple comparisons. Details are given in Appendix 3 part 14.

The results are presented in Table 5.8. These results show that the means differ significantly ($P<0.01$). The most damaged fruits were in *L. esculentum* parents with a mean of 62.98%, but intra-taxon crosses showed a mean of 54.76%. No statistical differences were found between crosses of *L. esculentum* and the parental group. All inter-taxon crosses displayed statistically ($P<0.01$) lower damage than intra-taxon crosses and tomato parents, but no differences between them. The lowest mean for damaged fruit was observed in *L. parviflorum* group with 14.74%.

Table 5.8 Means and standard error for percentage fruits damaged by the South American tomato pinworm (*Tuta absoluta* Meyrick, Lepidoptera – Gelechiidae)

	Mean**	Standard error
***L. esculentum* parents**	62.98a	±9.28
Crosses of *L. esculentum* with:		
L. esculentum	54.76a	±0.32
L. esculentum* var. *cerasiforme	30.90b	±5.55
L. parviflorum	14.74b	±3.24
L. pimpinellifolium	24.08b	±3.19

** = high statistical significance ($P<0.01$); same small letter show no statistical significance

5.5 Discussion

Morphological characteristics

Most morphological characters analysed in this study are the phenotypical expression of one or two genes. Segregation ratios were not analysed because there were few individuals per segregating generation growing in the greenhouse due to limited space. Therefore, it is not possible to determine the ratio, but observe the tendency expressed in the morphology of the plants. The results of these observations in both F_2 and F_3 generations showed that most individual plants from inter-taxon crosses tend more toward the wild than the domesticated character, such as smaller fruits or vine-type growth.

The problem in the use of wild relatives as source of variation seems to be less in the characterisation and identification of desirable characters, but more in the difficulty of introgressing these traits into domesticated germplasm without introducing undesirable associated characters of the wild relatives (Hawtin *et al.*, 1996). Once an adapted germplasm has been obtained, the introduction of new traits from wild relatives or landraces can present severe difficulties to the breeder. This is even more so for traits under complex genetic control. For these reasons, breeders are reluctant to incorporate massively germplasm from wild relatives into adapted stocks. Modern cultivars are usually the first choice for breeders looking for better characters.

Conversely, a genetic base broadening programme envisages to incorporate novel characters to adapted germplasm and maintain the genetic variability at the highest level.

The only main morphological difference in the intra-taxon crosses of tomato cultivars in the present project was in the type of vegetative growth (determinate vs indeterminate), fruit shape and size. Tomato cultivars produce fruits with extreme variation in both shape and size. The diverse fruit types have been selected for particular purposes either for their utility or for their novelty (Ku *et al.*, 1999). It is possible that humans initially selected for mutations associated with larger fruits and variable shapes, and gradually sufficient mutations accumulated to produce the present day cultivars (Grandillo *et al.*, 1999).

The difference in fruit weight between inter- and intra-taxon crosses was expected, but no comparison between F_2 and F_3 generations was possible as they grew in different seasons. Fruit ratio confirmed the visual estimation of shape in most plants, corresponding to a slightly flattened to round shape in both group of crosses.

In respect to solid soluble content, there was an expectation to have lower concentrations of sugars in *L. esculentum* crosses than in inter-taxon. But in the F_2 generation no statistical differences were found between them. Within F_3 intra-taxon crosses the solid soluble content was significantly lower probably due to differences in the environmental conditions during the growing season. The full potential expression of continuous characters could not be determined because of restricted growing conditions at SAC both in space and compost.

In the present study, seed weight was found to be higher for *L. esculentum* in comparison to inter-taxon crosses. Most domesticated plant species produce larger seeds than those from wild relatives (Evans, 1993). During domestication and subsequent plant breeding, plants have been selected for larger seeds to give uniform and high germination and high vigour under field conditions (Doganlar *et al.*, 2000). The incorporation of wild relatives' genes into tomato cultivars resulted in lighter and smaller seeds as phenotypic effect linked to smaller fruits. Seed weight in tomato is quantitatively inherited and determined mainly by additive gene action (Nieuwhof *et al.*, 1989). QTLs for seed weight are often in close proximity to loci for fruit

weight and soluble content (Goldman *et al.*, 1995; Grandillo and Tanksley, 1996), but whether these relationships are due to linkage or pleiotropy has not yet been determined (Doganlar *et al.*, 2000).

Morphological diversity indices

There was no clear tendency within the genetic indices analysed for the three types of data, namely co-dominant markers (morphological and microsatellites) and dominant (RAPD). Genetic variability in small populations is affected by specific phenomena. The effects of genetic drift and selection enhance the risk of losing alleles at selected or unselected genes (de Rochanbeau *et al.*, 2000). The expectation for these genetic indices was a decreasing trend from the F_2 to the F_3 generation. In autogamous species, inbreeding results in homozygosity. The frequency of heterozygous loci over a series of self-pollinated generations will be expected to fall by half in each succeeding generation (H_S) (Srb *et al.*, 1965). Similarly it was also expected to observe an increase in A_e because of its relation with homozygosity.

However, in the morphological character analysis, the intra-taxon crosses for all indices showed similar or identical values from F_2 to F_3 generations. This can be explained by the similarity present among the *L. esculentum* cultivars and the type of characters selected, which are very stable after years of breeding. The little genetic variability found could be due to a few fruit characters, such as shape or size. Conversely, inter-taxon crosses displayed a decreasing tendency for all indices, but H_S not accomplishing the expected half value. Similar situation was observed when comparing A and A_e among groups, where there was no statistical difference between *L. esculentum* and *L. esculentum* var. *cerasiforme* groups, either for H_S and I in F_2 generation, but they were statistically different to *L. parviflorum* and *L. pimpinellifolium* groups.

In the case of bulked populations, from 20 loci examined, inter-taxon crosses tended to be almost three fold higher for both F_2 and F_3 generations than intra-taxon crosses in respect to proportion of polymorphic loci (P). However, there were no differences observed from F_2 to F_3 generations in both types of bulked populations. P is simply the proportion of loci examined that show evidence of more than one allele, but it

suffers from two defects: arbitrariness and imprecision (Ayala and Kiger, 1984). The number of variable loci observed will depend on how many individuals are examined, but still it is a useful measure of variation. The results suggest that crosses involving tomato wild relatives as parents possess higher number of polymorphic loci as effect of heterozygous alleles incorporation into mainly homozygous loci of *L. esculentum* cultivars and they are conserved at high level until F_3. However, it is necessary to explore during more generations to conclude properly whether this index decreases. The effective number of alleles (A_e) from inter- to intra-taxon crosses was higher in 9% for F_2 and 4% for F_3 generations. There were less homozygous loci in inter- than in intra-taxon bulked populations, effect also observed in H_S and I. From F_2 to F_3 generations, these three indices showed decreasing values for inter- and a very slight variation for intra- specific bulked populations. The F_{ST} for inter-taxon crosses showed that 37% of the total allelic variation is apportioned within populations, and nearly 40% in the case of intra-taxon crosses, within the range estimated for predominantly inbreeding species, approximately 43% (Bretting and Goodman, 1989).

These findings suggest that the number of loci analysed was too small and a number of them could be homozygous for the species or accessions utilised as parents, especially between the closely related species *L. esculentum* and *L. esculentum* var. *cerasiforme*. This low level of genetic variation found within self-compatible species may be because of the role of autogamy that drives the decrease of genetic variation and fixation of alleles (Rick, 1984; Peralta and Spooner, 2001).

Molecular diversity indices

Microsatellite and RAPD markers produced very variable results, from decreasing to increasing values in F_2 and F_3 generations. Proportion of polymorphic loci (P) showed in microsatellites an expected tendency between groups for F_2 generation, where *L. parviflorum* group presented the highest value and the lowest *L. esculentum* group. However, RAPD markers, at the same generation, showed completely distorted values in relation to the microsatellites, being the highest *L. esculentum* var. *cerasiforme* and the lowest *L. pimpinellifolium* group. In F_3 generation,

microsatellites displayed 3 identical values (0.57); the exception was *L. pimpinellifolium* group with 0.86. RAPD conserved similar distribution of values. In bulked populations, microsatellites and RAPD markers, in F_2 generation the inter-taxon crosses were almost three fold higher than intra-taxon crosses, but in both markers this difference decreased in F_3 generation. These results are not the best comparison between populations, because they are biased by the selection of the most polymorphic primers in both molecular marker systems utilised in this experiment. However, they give a robust indication of the differences between the created populations, especially between inter- and intra-taxon crosses.

In respect to A and A_e, both markers showed that *L. esculentum* and *L. parviflorum* groups in F_2 generation had differences statistically significant from *L. esculentum* var. *cerasiforme* and *L. pimpinellifolium* groups. In bulked populations, A_e presented differences between inter- and intra-taxon crosses of 27% to 11% for F_2 and F_3 generations in both molecular markers. Considering that the closer the difference between A and A_e, the higher the similarity of allele frequencies between populations; therefore it is likely that less variability exist among the accessions analysed, microsatellites showed the least difference in F_2 generation for *L. esculentum* and *L. esculentum* var. *cerasiforme*. However these differences increased in F_3 generation. This relation was reflected in bulked populations, where intra-taxon crosses showed little difference in comparison to inter-taxon crosses. In the case of RAPD, the trend of the values was similar to microsatellites but the differences were higher. These results can be expected in predominantly self-pollinated species because of their tendency to homozygosity, especially remarkable is the case of intra-taxon *esculentum* crosses where the parents used in the crosses were genetically close. Both molecular markers showed a clear difference between intra- and inter-taxon crosses; this is expected since the difference reflects the lower genetic diversity present in edible tomato accessions and which increases when hybridised to accessions with more variation in their genetic background, such as wild relatives.

Slight differences between F_2 and F_3, in both markers, were found for H_S and I. The H_S and I in *L. esculentum* group was statistically different from all other groups in the F_2 generation for both microsatellites and RAPD. H_S in bulked populations

showed almost two fold higher values in inter- than intra-taxon crosses in F_2 and F_3 generations for both markers. In respect to H_t, inter-taxon bulked populations showed higher values than intra-taxon for microsatellites and RAPD in both generations. However, from F_2 to F_3 generations for microsatellite markers the value decreased in inter-taxon crosses, but the index decreased in intra-taxon and both crosses analysed by RAPD. These measures are the most commonly used to estimate genetic variability. In theory these values should range from 0 to 1 (homozygosity to full heterozygosity), although for dominant markers, like RAPD, the maximum level is 0.5. Co-dominant markers never reach the maximum level of 1 for self-pollinating species; populations in equilibrium can reach 0.5 as maximum. For autogamous species, H_t and I are more useful indices because H_S does not reflect well the amount of genetic variation in such organisms. There will be more homozygotes in a population in which crosses between relatives is common, even though different individuals can carry different alleles if the locus is variable in the population. There will also be more homozygotes in a population in which mating between relative is common than in a population where it does not occur, even when the allelic frequencies are identical in both populations (Ayala, 1982). The lower H_t index of intra-taxon *esculentum* crosses demonstrate the low levels of diversity present in *L. esculentum* accessions but indicate that there is still variability present within landraces and old cultivars. This may be useful for breeding purposes when incorporated into appropriate populations (Saavedra and Spoor, 2002).

Fixation index (F_{ST}) is usually utilised to analyse the differences in genetic variability among populations. The F_2 bulked populations analysed with microsatellite markers in inter-taxon crosses showed that about 20% and 4% in intra-taxon crosses genetic variation can be explained as differences between populations; but 80% and 96% of the genetic variation lie in the differences within populations, respectively. In F_3 generation inter-taxon crosses decreased to 13% the variation due to differences between populations and intra-taxon crosses increased to 8%. For RAPD markers, the values obtained reflect the high differentiation of genetic variability among created populations, 37% for the inter-taxon and 40% for the intra-taxon or *esculentum* group. However, these results fell to 18% and 21% respectively in F_3

generation. The variability of these results can be explained by the number of segregating individuals present in the samples taken in each population. However, these results indicate that a sizeable portion of the genetic variability in created populations lies within populations, but also the variability between is very important; this is one of the objectives in a base broadening approach.

These results were not expected for markers assumed not affected by environmental conditions. However, there are several possible explanations for these observations. The most obvious is sample size, due to which the total diversity potentially present within the created populations may not be represented. However due to time and financial constraints it was not possible to increase the number of samples per created population. An alternative explanation is accidental out-crossing occurred within the greenhouse, in spite of controlled conditions. The out-crossing may have occurring as a result of: the individual plants being grown too close together; contact between flowers; pollen blown away by movement during watering; and/or by insects. The relativity of the indices obtained with each genetic marker, for instance H_S (average gene diversity) will depend on the number of polymorphic loci utilised in the calculations, each monomorphic locus included will decrease the index level. Also the number of polymorphic loci included will change the effective number of alleles (A_e) considering it as the inverse of homozygosity, and the proportion of polymorphic loci (P) that with low number of loci samples the information is locked up in allele frequencies.

Low values for H_S and I in *L. esculentum* crosses were found in comparison to inter-taxon crosses. This fact demonstrates again the low molecular diversity present in tomato cultivars as a result of low genetic variability of ancestral forms (Rick and Chetelat, 1995). Breeding methods utilising pedigree selection, backcrosses and single-seed descent promote homozygosity. The inter-taxon crosses showed statistical differences in comparison to *L. esculentum* group for these indices. Thus, when considering the tomato cultivars as a genetic starting point, incorporation produced an increase in heterogeneity, the desired effect in a genetic base broadening project. This is reflected in the total gene diversity (H_t) where values for inter-taxon crosses are twice as high of the intra-taxon with microsatellite markers. By grouping

data, it may be possible to formulate preliminary recommendations about relative approaches. This is very important because base broadening should act to create pools of diversity, which will be handled as populations.

Comparing results between the 3 types of markers, it is necessary to consider that there may be several reasons for the observed differences in the RAPD assay and the other marker systems. Scoring of RAPD polymorphisms appears to be more subject to error than scoring the other, co-dominant polymorphisms, such as microsatellites or morphological. The presence of a RAPD band of apparently identical molecular weight in different individuals is not evidence that the two individuals share the same homologous fragment, and single bands can sometimes be comprised of several co-migrating amplification products. These limitations do not prevent the estimation of allele frequencies necessary for population genetic analysis, but they do reduce the accuracy of such estimation relative to co-dominant markers such as microsatellites. To increase the degree of statistical power using RAPD 2 – 10 times more individuals need to be sampled per locus (Lynch and Milligan, 1994).

Field trial

A field trial was carried out to analyse the effect of incorporation of genes from wild relatives in tomato cultivars. The experiment was carried out studying the SATP attack of fruits under field conditions. Cultivar resistance/tolerance to the SATP in tomato may provide an efficient alternative method for pest control. Resistance to SATP has been found in several wild *Lycopersicon* species (França *et al.*, 1984; Lourenção *et al.*, 1984). These preliminary results show a significant difference between inter- and intra-taxon crosses, especially those including *L. parviflorum* as parent. Unfortunately, these accessions presented the smaller fruits within all inter-taxon groups, characteristic that have to be improved. However, the effect of natural resistance could be due to secondary compounds produced by the plants, mainly a-tomatine that acts as repellent to moths laying eggs in leaves and green fruits, usually found in greater concentrations in tomato wild relatives (Rick and Chetelat, 1995).

This is a valuable source of information for future genetic base broadening projects because the data show some effect in early segregating generations, which later can

become in integral part of breeding programmes for control of this pest in South America.

Although the observed results of this chapter were not as expected, the information on segregating populations from inter- and intra-taxon crosses has provided a platform to develop further the idea of incorporation. In addition, the information obtained in this project will help answer questions relating to strategies for the conservation of created genetic variability in autogamous plants (Chapter 6).

Chapter 6

Theoretical strategies for conservation of genetic variability in autogamous crops subjected to genetic base broadening

6.1 Introduction

A genetic base broadening programme does not only involve the selection of parents, hybridisation and management of consecutive generations, but also includes a range of strategies determined by constraints such as size of experiment, type of accessions involved, methodology suitable to the reproductive biology of the species, locations, amount of time, and funding.

The utilisation of available genetic resources, such as wild relatives and germplasm temporarily not utilised by breeders or recycled old cultivars are key resources. Wild relatives of crops that have survived under natural selection pressures can be particularly useful as source of genes for specific adaptive traits (Hawtin *et al.*, 1996). The incorporation of this genetic material into domesticated and adapted germplasm through hybridisation can be the starting point for broadening the genetic base, but in the case of autogamous crops it is also necessary to design a strategy for the conservation of the created variability. In highly autogamous crops, such as tomato, the created populations will be at a homozygosity level similar to parents after few generations of self-pollination. Even without human intervention, genetic variability will decrease, however such intervention can exacerbate the decline.

The aims of this chapter are to discuss different strategies to conserve, as best as possible, the high genetic variability created with broad scale hybridisations at the beginning of a genetic base broadening programme. In addition, some of the general questions raised at the start of this project will be examined in the light of information and experience acquired during its course.

6.2 General constraints

Genetic base broadening is often considered to be an activity at the interface between germplasm conservation and utilisation (Cooper *et al.*, 1998). As such there is a lack of clarity as to interest and who will be the key players, the public or private sector. However, the main problem lies in funding the activity, because this point focuses on several questions regarding the scale of the operation and the duration of the project. The scarcity of long-term funding for research, in general, has had negative effects both for maintaining and increasing the utilisation of germplasm stored in genebanks,

and this is reflected in the limited activity in genetic base broadening projects. Depending on the reproductive biology of the crop under study, several generations are required as a minimum in order to achieve a degree of local adaptation.

Large-scale base broadening approaches have been successful in improving crop productivity, for example in maize (Goodman, 1985) and sorghum (Mengesha and Rao, 1982). Simmonds (1993) proposed that genetic base broadening should be on a large scale because there are heavy losses and discards within the genetic material created. For instance, in this study 90 hybridisations including reciprocals were carried out at different development stages of the flowers, but only 49 crosses successfully produced F_1 seeds, from these 49 crosses only 33 produced seeds in F_2, and 32 in F_3. Most of the losses were because of genetic incompatibilities, such as the case of crosses between green- and red-fruited species, but there were also populations presenting susceptibility to greenhouse diseases or weak plant development due to unsuitable genetic combinations.

6.3 Strategies for autogamous crop species

Base broadening is about creating large populations that have good local adaptation, but have not been selected for the other requirements of crops namely pest/disease resistance and quality aspects. So are we creating variability to 'fix the variability' or are we creating populations where the possibilities, exist in future, for further recombination and assortments. So we must remember that initially with autogamous species, we need to maintain diversity or new combinations at early stages (in order to allow local adaptation or natural selection to work). This will be followed by an inevitable collapse as selection kicks in, and what is required then are mechanisms ('natural' or with human intervention) which facilitate further recombination.

There are several ways to conserve the genetic variability in predominantly autogamous plants after hybridisation, from the simple method of self-pollination to more complex system involving facultative out-crossing or male sterility. It is further possible to utilise combinations of methods depending on the reproductive biology of the species.

6.3.1 Self-pollination

Selfing is a simple method for multiplying hybridised genetic material. However, there are some constraints for this system (see Chapter 5). But, variation in the system adopted such as backcrossing to both original parents in a population, or double crosses from F_1 populations could mitigate the loss of genetic variability. Alternatively, forced hybridisation may be used to regularly regenerate variability, but how often and what percentage of population would be involved? It is difficult to estimate. By itself self-pollination cannot be considered as base broadening, since one would need extremely substantial resources in order to sample all possible combinations within a population. Inevitably, if the programme is carried out at a single site, there will be very heavy losses of the variability due to specific natural selection pressure. This might be appropriate, if the base broadening project has reasonably defined aims (e.g. introduction of a range of pest/diseases resistances while maintaining adaptability) and if new base broadening populations are started for other defined projects. Difficulties arise where base broadening is more generic, and where aims are not defined or indeed where they are not known at all. It may be then that other approaches use a large number of sites to maintain a 'large scale' diversity, whilst sacrificing variability at each site. This can be with or without cycles of random deliberate hybridisation within each sub-population, and may also include deliberate hybridisation between sub-populations.

Double crosses of F_1 accessions of different parentage is another possibility of producing populations with wider combinations of alleles, but at the end selfing of these populations will lead to homozygosity in further generations. However, the genetic variability that could be created from these crosses may be higher than in simple crosses because of the wider possibilities of recombination of genes and the subsequent production of novel genotypes.

6.3.2 Facultative out-crossing populations

In autogamous crops there are, within natural occurring populations or cultivars stored in gene-banks, accessions that possess characteristics for out-crossing, such as

exserted styles in *Lycopersicon*. This character is found within wild relatives and can be introgressed into populations presenting inserted styles by simple hybridisation. Exserted styles allow the reception of pollen from different plants and flowers through insects.

By utilising wild type accessions carrying this trait as parent in simple crosses and later, if necessary, in double crosses, populations segregating toward both phenotypes can be created. Therefore out-crossing could occur giving the appropriate environmental conditions for growing and the presence of pollination vectors, such as insects and/or wind. However, it is important to consider possible genetic incompatibilities between selected parents, otherwise it might be necessary to use "bridge" crosses in order to utilise these lineages in genetic base broadening programmes.

A simple analysis of the segregation for style length was carried out in F_2 and F_3 inter- and intra-taxon crosses, and in double crosses. Results of F_2 and F_3 segregation are shown in Table 6.1, and those for double crosses in Table 6.2.

In both generations F_2 and F_3, the *L. esculentum* intra-taxon and inter-taxon crosses containing *L. esculentum* var. *cerasiforme* as parent did not show individuals with exserted styles in the flowers. It is likely that these two species share the same alleles controlling this trait. In the case of those crosses containing *L. parviflorum* and *L. pimpinellifolium*, segregation for this character showed a higher percentage of individuals presenting inserted than exserted styles in F_2 and F_3. There was also a tendency of increased number of individuals with inserted styles from F_2 to F_3 generations, perhaps due to a dominant character controlled by the family genes *Ex*. However, this may also be a reflection of the plants sampled in the F_2 generation.

Double crosses are usually used in hybrid cultivar production to exploit hybrid vigour from four lines. In the present study, two out of seven populations had only individuals with inserted styles. The other populations showed a tendency toward inserted styles, but one cross (5219x3915) displayed higher percentage of individuals possessing exserted styles. However, some populations segregated strangely and these results presented unexplained abnormalities.

Table 6.1 Percentage of individuals presenting inserted or exserted styles in inter- and intra-taxon crosses of F_2 and F_3 generations.

		F_2		F_3	
L. esculentum cross with	N	Inserted style	Exserted style	Inserted style	Exserted style
L. esculentum	20	100%	0%	100%	0%
L. esculentum var. *cerasiforme*	40	100%	0%	100%	0%
L. parviflorum	16	60.0%	40.0%	69.0%	31.0%
L. pimpinellifolium	16	75.8%	24.2%	92.5%	7.5%

N = number of plants observed

In a genetic base broadening project this could be another strategy to create populations with mixed style types exhibiting in- and out-breeding characteristics, and conserving heterozygosity within populations.

Table 6.2 Percentage of individuals presenting inserted or exserted styles in double crosses involving tomato and its wild relatives.

Double cross	N	Inserted styles	Exserted styles
1939x5211	18	57.9%	42.1%
6021x2239	15	100%	0%
3911x1560	12	72.7%	27.3%
5222x1960	19	78.9%	21.1%
5219x3915	20	45.0%	55.0%
1160x5219	16	100%	0%
1539x1560	20	55.0%	45.0%

The first two digits represent the female parent and the second two the male parent. 11=*L. esculentum* cv. Limachino; 15=*L. esculentum* cv. Ace; 19=*L. esculentum* cv. Lukullus; 21=*L. esculentum* cv. Marglobe; 22=*L. esculentum* cv. San Marzano; 39=*L. esculentum* var. *cerasiforme*, LA-1673; 52=*L. parviflorum*, T1264/94; 60=*L. pimpinellifolium*; PI-390739. N = number of plants observed.

Selection in favour of facultative out-crossing, may create problems of utilisation of the material, since: a) the individual maintains a higher levels of heterozygosity and therefore useful superior traits in breeding programmes are masked, and b) the

character of exserted style itself will have to be selected against in order to eventually obtain autogamous cultivars.

6.3.3 Exploitation of male sterility

Exploitation of male sterility is another alternative to maintain genetic variation at higher levels in self-pollinated crops. Male sterility is widely used commercially in F_1 hybrid production. The use of male sterile parents in a genetic base broadening programme could be one solution to the problem of conserving genetic variability within populations. The presence of male sterility alleles allows the identification of male sterile plants in the population, then seeds can be harvested mainly from male sterile plants ensuring higher levels of out-crossing and recombination.

Male sterility may be genetically controlled by nuclear genes; it is usually recessive, thus *MsMs* and *Msms* are male fertile and *msms* male sterile plants. However, it can also be cytoplasmically controlled and in this case is maternally inherited (Kaul, 1991), and then female (S = male sterile) x male (F = fertile) produces female S individuals. Male sterility may also involve both genetic and cytoplasmic control, with both the *msms* genotype and the S cytoplasm needed for male sterility, and the *Ms* genes are epistatic to the S cytoplasmic genes. This allows simple restoration of male fertility (Mayo, 1987).

When selecting male sterile parents the choice of individuals presenting recognisable characteristics within the populations is very important. This methodology has already been utilised by Kannenberg and Falk (1995) in their hierarchical recurrent introgressive population enrichment (RIPE) method for enhancing the genetic base in barley.

In plant breeding an advantage associated with the use of monogenic sterility systems is their inability to generate 100% sterile populations, critical are the early generations (Jensen, 1988). In equilibrium, F_2s yield 25% sterile progeny. However, in plant breeding practice the gradual loss of male sterile individuals in a population is of little consequence, except in long-term research projects (Jensen, 1988). A base broadening programme is not interested in 100% effectiveness, the system is merely a mean of allowing recombination to occur, but the scale of the operation is

important. An ideal genetic male sterility system for a genetic base broadening project should be facultative in the extent that it can also be autogamous under certain conditions or manipulations.

In the case of tomato, there are a number of genetic male sterile mutants in a wide variety of types and genetic backgrounds listed by Stevens and Rick (1986). The degree of androecium reduction in the *ms* series varies from extremely modified stamens to those that can be distinguished from normal only by the absence of viable pollen (Stevens and Rick, 1986). However, it seems that not all male sterile variants are potentially useful, because some of them do not accomplish the requirements of total recessivity, or normal female fertility. In the present study male sterility was not used as germplasm for hybridisation.

Questions on the exploitation of male sterility are very much minor to the comments above. A weak system of male sterility might be as effective as facultative out-crossing as a means of generating and maintaining genetic variability.

In any base broadening programme some selection will inevitably take place, particularly for generally favourable agronomic characters, and for characteristics influencing local adaptability. The use of techniques such as male sterility systems may have value in producing populations where there is little initial focus, but have intrinsic problems in a clearly directed programme with very defined aims.

6.4 Management of created populations

6.4.1 Single-site exploitation of natural selection

Single-site exploitation is an easy way for managing the created populations in a genetic base broadening programme. Regardless of less costs and better control over field trials and data acquisition, this method has a limitation from the point of conserving genetic variability. Single-site exploitation would lead to the adaptation of populations to defined environments, biasing the selection of individuals and narrowing the genetic background of the populations. Conversely, knowledge of the nature and relative magnitude of the various types of genotype-environment interactions only become obvious when the populations are subjected to many environments in different sites.

6.4.2 Multi-site exploitation of natural selection

Genetic base broadening may be desirable for a number of reasons (discussed elsewhere in this thesis), either to create new genotypes to be exploited for existing crop production areas, or indeed to extend the cropping areas and off-season production (saline soils, cold/heat resistance, etc.). How such populations are treated may indeed determine success or failure. For example, development and exploitation of a base broadening programme at one site may produce material adapted to those specific conditions, this is acceptable whether the site is representative of a major cropping area, but of limited or no value if the site does not represent such area.

Multi-site evaluation of populations in a base broadening programme can have several advantages and can operate in a range of environments, depending on the overall scale of the operation.

Environments can vary greatly, so that testing sites cannot cover the whole range of production areas for a crop. The adaptation of a crop, i.e. the ability to survive in particular environment, and the exploitation of its productivity are under an extremely complex genetic control (Hawtin *et al.*, 1996). In genetic base broadening, multi-sites studies intend to exploit the genotype-environment interactions that allow local adaptation in artificially created populations. Simmonds (1993) proposed 'let nature do the work', in populations spread across very diverse environmental conditions, such as countries, regions and sites within a region. The populations will be exposed to diverse environmental stresses and disease pressures and the result being that different genotypes will survive at each site.

Although the scale of a genetic base broadening programme depends absolutely on economic factors, it can be considered at a number of levels: 1) for local needs in a specific environment; 2) for undefined environments lying within a broad eco-geographic region/zone; 3) for undefined (or unknown) quality aspects across a range of environments (specific broad eco-geographic zone); and 4) for undefined environments lying across a wide range of eco-geographical zones.

This net is not exhaustive, there may be other combinations. All, ultimately, depend on the resources (financial) available, but the different levels and scales may require collaboration at national, regional, or greater levels. This inevitably will bring

problems of logistics and collaboration. Management of such complex programmes may involve centralised reassortment of the gene-pool, followed by selection for adaptability at single sites, and followed by reassessment at multiple sites to determine the nature of adaptation. The frequency of such reassortment activities, along with other questions relating to the incorporation of new genetic material will depend on the nature of the crop and the environmental effects on the plants.

Examples for multi-sites experiments in genetic base broadening are found in Latin America and the USA, where locally adapted maize germplasm was distributed and evaluated under different environmental conditions (Sevilla and Holle, 1995). Other examples are potatoes (Simmonds, 1993) and GEM (germplasm enhancement of maize) systems in maize (Goodman, 1985).

The disadvantages of this kind of approach are the high cost, logistic difficulties to find partners around the world and the control over every experimental site, specially the personnel involved at each site.

However, this method, linked with any other quoted in this chapter, could help to solve the problem of utilising genetic variation and accelerate the analysis of progeny performances under different selection pressures in agronomically relevant environments.

6.5 General conclusions

Genetic base broadening is an activity needed in autogamous crops, but the lack of long-term research funds has contributed to neglecting long-term pre-breeding activities (Cooper *et al.*, 1998). This type of activity should last for long periods aiming at the creation of new diversity through continued recombination and selection. There are several sources of germplasm with pest resistance and/or tolerance to environmental stresses that could be incorporated into adapted cultivars through a comprehensive germplasm enhancement programme. The relative success of any effort or programme will depend on the availability and utilisation of the genetic variation. However, the conservation of the created genetic variability at higher levels in autogamous species is the key issue, and there is a need to address the question of useful strategies.

Modelling a strategy for autogamous crops can help to decide when and how often to hybridise again the population to regenerate genetic variability. Autogamous crops, assuming that they are in Hardy-Weinberg equilibrium (distribution of alleles frequencies $p^2 + 2pq + q^2$), decrease heterozygous individuals for a determined single mendelian segregating locus by a half in each self-pollination, therefore increasing the proportion of homozygous individuals for that locus (Table 6.3), then within the population there is a decrease of the genetic variability for that locus. The existence of genetic variation is a necessary condition for evolution (Ayala and Kiger, 1984). If at a certain gene locus all individuals of a population are homozygous for exactly the same allele, selection cannot take place at that locus, because the allelic frequencies cannot change from generation to generation. Critical point in this trend occurs between F_4 and F_5, where the number of heterozygous loci fall below 10%, after this point the decreasing heterozygosity for this locus reaches the lowest levels, almost zero. This model suggest that backcross hybridisation using both parents should be done at this stage, repeated every 4 cycles (Table 6.4) and it must involve a large part of the population. In this way, the equilibrium of the proportion of homozygous (decrease to 50%) and heterozygous (increase to 50%) individuals can be recovered and the genetic variability conserved at higher levels.

The individual merits of the different approaches can be determined, but there is need to stimulate support and funding for these initiatives, which is essential for carrying out and developing appropriate techniques and systems.

Table 6.3 Proportion of heterozygous and homozygous individuals from crosses of heterozygous and homozygous populations in one locus with simple mendelian segregation.

Generations	Homozygous x Heterozygous populations		Heterozygous x Heterozygous populations		Homozygous x Homozygous populations	
	Heterozygous individuals	Homozygous individuals	Heterozygous individuals	Homozygous individuals	Heterozygous individuals	Homozygous individuals
F_1	50.00%	50.00%	50.00%	50.00%	100.00%	
F_2	25.00%	75.00%	25.00%	75.00%	50.00%	50.00%
F_3	12.50%	87.50%	12.50%	87.50%	25.00%	75.00%
F_4	**6.25%**	**93.75%**	**6.25%**	**93.75%**	12.50%	87.50%
F_5	3.13%	96.87%	3.13%	96.87%	**6.25%**	**93.75%**
F_6	1.56%	98.44%	1.56%	98.44%	3.13%	96.87%
F_7	0.78%	99.22%	0.78%	99.22%	1.56%	98.44%
F_8	0.39%	99.61%	0.39%	99.61%	0.78%	99.22%

Table 6.4 Model for a single locus segregating in mendelian ratios for a cross between one heterozygous and one homozygous population, followed by back crosses to both parents .

Generations	AA	AB	BB	
F_1	50.00%	50.00%		
F_2	62.50%	25.00%	12.50%	
F_3	68.75%	12.50%	18.75%	
F_4	71.88%	6.25%	21.87%	
				← cross AA & BB
F_1	37.50%	50.00%	12.50%	
F_2	50.00%	25.00%	25.00%	
F_3	56.25%	12.50%	31.25%	
F_4	59.38%	6.25%	34.37%	
				← cross AA & BB
F_1	31.25%	50.00%	18.75%	
F_2	43.75%	25.00%	31.25%	
F_3	50.00%	12.50%	37.50%	
F_4	53.12%	6.25%	40.63%	
				← cross AA & BB
F_1	28.12%	50.00%	21.88%	
F_2	40.63%	25.00%	34.37%	
F_3	46.88%	12.50%	40.62%	
F_4	50.00%	6.25%	43.75%	
				← cross AA & BB
F_1	26.56%	50.00%	23.44%	
F_2	39.06%	25.00%	35.94%	
F_3	45.31%	12.50%	42.19%	
F_4	48.44%	6.25%	45.31%	
				← cross AA & BB
F_1	25.78%	50.00%	24.22%	

Chapter 7

General Discussion

This project has aimed to investigate some of the problems associated with the initial management of plant populations specifically created to exploit the potential benefits of a wide genetic base using *Lycopersicon* as a model. The narrowness of the genetic base in *Lycopersicon esculentum* was examined using both morphological and molecular markers. Selection of potential parents for such base broadening activity was determined first by morphological characters such as flowering and fruit setting in greenhouse, morphological diversity between potential parents, and then molecular markers to confirm the diversity in the first choice of parents. The behaviour of these specially created populations was examined through subsequent generations using a similar range of investigatory tools. The main results and achievements are as follows: analysis of the genetic diversity in *Lycopersicon* spp germplasm; parent selection, characterisation and hybridisation; behaviour in F_1, F_2 and F_3 generations of created *Lycopersicon* populations; and theoretical considerations for conservation of created genetic variability extrapolated to autogamous crops subjected to genetic base broadening.

The approach of this project was ambitious, attempting to respond to broad questions that cannot be answered in a three year period with limited space, labour and resources. The several questions outlined at the beginning of the project, should be refined and narrowed after the experience obtained during the development of this research. In addition there also are several questions about the research, particularly the methodology, accessions selection, and many more facets that should be analysed with the aim to improve any future investigation in this large topic of genetic base broadening. Research is intended to be perfectly planned and executed, but when working with live organisms it is difficult to achieve the perfect plan because of the behaviour and responses of these organism to the environment. There are several factors that can be controlled such as temperature, light, soil, nutrition, etc., but they also can react and trigger other reactions, which cannot be predicted because of the interaction genotype – environment.

In respect to the species choice for the project, there are other species more suitable for genetic studies; for example the widely utilised *Arabidopsis thaliana*. This is a

small weed of the mustard family (*Brassicaceae*), which in 4 to 6 weeks produces mature plants containing more than 10,000 seeds. However, *Lycopersicon esculentum* was selected from among other autogamous crops such as *Phaseolus* spp because *Lycopersicon* spp satisfied all the basic needs for this type of project; easy to cross; a narrow genetic base; a large number of wild relatives and landraces stored in genebanks ready to use for this kind of research. Wild and unadapted germplasm represents a rich source of variation. Though exotic germplasm can present problems of adaptation and characters not desirable in a breeding programme, it can help to increase response to selection as a result of the improvement of genetic variability. Furthermore, there is a complete list of morphological markers, recognised and described genes, and molecular markers already tested for this species. Considering all these facts, tomato was the right choice to answer the questions stated in this research.

The parents used in the project were selected both for their morphological and molecular characteristics. One accession per wild species was chosen and five accessions of *L. esculentum*. Close relatives such as *L. esculentum* var. *cerasiforme* and *L. esculentum* cultivars did not show huge morphological differences and genetic distances, but there were still more genetic differences present than among tomato accessions. The other non-domesticated accessions represented most of the species related to tomato, however *L. cheesmanii* was not included as parent in the hybridisations because flowering did not coincide with the other accessions. Unfortunately, hybridisations between green-fruited accessions and *L. esculentum* were not successful producing more generations, but this was expected and a more ample vision on the behaviour of wider crosses was not possible to achieve. However, the natural acceptance/rejection of the crosses, without human intervention after hybridisation, was decided as part of a base broadening project to observe the effects and problems that could arise using wide crosses in other projects.

In the case of wide crosses, the use of hybridisation techniques, such as embryo rescue or bud pollination may be desirable; the latter method was utilised in this research but was not successful when crossing the incompatible species selected in this project. The created populations did not represent the huge genetic diversity

present between and within *Lycopersicon* spp, reflecting that only a small part of each species and hence possible genetic combinations were used. It also likely that some of the accessions selected have similar alleles sharing the same locus, when analysed through molecular markers; therefore the genetic indices may not reflect the real diversity between and within populations, nevertheless such techniques were a useful approach to the tendencies existing in the created populations.

Much effort was concentrated on collecting large amounts of morphological and molecular data on individual crosses; this information tended to treat the individuals as a series of crosses instead of populations. Other approaches not explored might include narrowing the scope and examining only two large contrasting populations such as inter- and intra-taxon groups. These two approaches would have given more global information about the development of autogamous populations subjected to genetic base broadening, but the project looked for more specific information about determined crosses and then integrating the data in groups. The experimental design adopted, on reflection, was not necessarily the most appropriate but these views have arisen following the experience obtained during the development of the project, and will be useful when designing future projects related to genetic base broadening.

In this project, two types of markers were used, morphological and molecular, to identify differences between and within species; accessions selected as parents for hybridisation; and created populations. Morphological markers have been, until recently, most used for research studying species and populations, and also in breeding projects. The benefits of such markers are the number of individuals that can be assessed in one generation; the diversity of characters that can be studied in each species; and the ease of scoring. The expression of these, usually qualitative, characters depend on genotype-environment interaction, but the effect of environment can be reduced by carrying out the experiments in standardised conditions. When analysing large segregating populations, morphological markers give accurate information about segregation rates and genetic diversity present.

Most of the morphological markers selected for this research were related to fruit characteristics, which are important in tomato descriptors because of the limited differences among accessions in vegetative traits. There are more traits that could be

utilised as markers, but the ones selected were the most commonly used and easy to observe. The number of morphological markers could be increased but this would not have produced any more accurate results, but an increase in number of individuals sampled per segregating generation, would have produced potentially more valuable results without losing any of the advantages of the approach.

In order to make recommendations on strategies for base broadening, it is essential that appropriate approaches to describe the variability be developed, and how such approaches might change with different management. With molecular markers, both co-dominant and dominant markers were used, requiring two different types of statistical analysis and resulting in enhanced value of the data. The co-dominant microsatellite markers are the product of highly mutable loci, which may be present at many sites in a genome. They fall into the category of site-targeted PCR, where the primers are designed to amplify specific regions of the genome. Conversely, dominant RAPD markers are the products of arbitrary primers in a PCR reaction, which is usually the amplification of many discrete DNA products. Each product will be derived from a region of the genome that contains two short segments which share sequence similarity to the primer and which are on opposite strands and sufficiently close together for the amplification to work. Polymorphisms are detected as presence or absence of bands and mainly result from sequence differences in the primer binding site. Both markers, microsatellites and RAPD, can be visualised by agarose gel electrophoresis, although microsatellites are recommended to be visualised utilising polyacrylamide gel (PAGE). However, tomato genome is large enough to amplify substantial microsatellites bands in agarose gels. The PCR techniques utilised in both markers are relatively simple and time consuming, but microsatellites visualisation is more expensive because of the use of special Metaphor agarose, but still cheaper than PAGE since it is less complicated to prepare and assemble.

Microsatellites are sometimes not representative of the whole genome; this is a limitation because the loci can be located in a specific region and it is only possible to sample the diversity present in that region. However, RAPD has a random spread around the genome, and the loci sampled are more representative of the genome. RAPD is a quick technique, simple and efficient, but band profiles can be difficult to

reproduce, even within laboratories, and more so if personnel, equipment or conditions are changed. However, an important limitation is data quality, since for dominant markers, heterozygosity is not detectable, bands sometimes consist of co-migrating products and band identities are difficult to assign. In spite of these limitations, both techniques generate data that can be analysed and applied to genetic diversity and variability studies through the analysis of genetic relationship between samples, or calculation of population genetics parameters, in particular diversity and its partitioning at different levels. The results demonstrate the robust nature of the information from microsatellites and DNA analysis, uninfluenced by environmental factors.

Other molecular marker techniques that could have been used in this project include Amplified Fragment Length Polymorphism (AFLP), which is a dominant highly reproducible method that combines restriction digestion and PCR. However, such an approach is more demanding technically and expensive than RAPD. Other approaches such as co-dominant Restriction Fragment Length Polymorphism (RFLP) have similar limitations. Newer techniques such as proteomics are in development and offer choices for the future.

Statistical analysis of the populations was carried out utilising the most used population genetic indices, perhaps other software packages could analyse the data in other ways, but at the end the indices are the same. All population genetic indices used in this project were utilised as an approach to test the usefulness of each one. However, there were some better than others and more useful. Mean proportion of polymorphic loci (P), for instance, is a very imprecise and arbitrary measure of genetic variation, but useful for certain purposes, such as quick observations. Average gene diversity (H_S) is a better and more precise measure of genetic variation, because it estimates the probability that two alleles taken at random from a population are different. its quality is limited in self-pollinating populations because most individuals will present more homozygous loci witrhin a population. However, it gives a good approach that it is improved when utilised together with information from total gene diversity (H_t), which is calculated from allelic frequencies as if the individuals in the population were mating with each other at random. Shannon's

information index (I) is similar to H_S but not bounded by 1.0; this index is useful when comparing large populations through co-dominant and dominant markers, but in this case its information was not relevant in comparison to the other indices. Number of polymorphic alleles (A) was not very informative *per se*, but in combination with effective number of polymorphic alleles (A_e) gave valuable information about genetic variability. Partitioning the genetic variability (F_{st}) gave a good approach how the diversity was distributed, within or between populations.

In the light of the results obtained during this research most aims proposed at the beginning were achieved, but whether these results answer the questions stated are now part of the evaluation. The analysis of genetic diversity within and between *Lycopersicon* taxa showed great levels of variability between the tomato wild relatives, but also there was genetic variability between tomato cultivars, especially old cultivars and landraces. After selection of parents and hybridisation, all created populations showed phenotypic uniformity in the F_1 generation for the characters studied, also there was genetic uniformity at molecular level. In selfing F_2 and F_3 generations, these populations segregated as expected for morphological continuous and discontinuous characters, as long as for molecular markers. In general, phenotypic traits tended toward the wild type characters. The genetic indices analysed did not show the expected decline of variability for autogamous crops after consecutive selfing. However, the observed results gave information of segregating populations from inter- and intra-taxon crosses to develop further the idea of incorporation and speculate about strategies for conservation of created genetic variability in autogamous crops. Several strategies, based in the information and experience gathered in this research, were analysed intending to answer the question about "How to maintain this genetic variability in later generations?". From the simple and naturally occurring system of self-pollination to more complex systems involving facultative out-crossing and/or male sterility exploitation were discussed, but any method depend on the reproductive biology of the species and to give a general recommendation of which method rely is difficult to answer.

Questions about the scale of operation and number of parents to be utilised in the projects are more related to economical concerns. However, this research gave an

idea that the genetic base broadening projects should be of large scale, so the number of parents involved. It is important to consider a manageable population size, number of populations and hybridisations. In respect to parental material range, this germplasm should be wide enough to allow the incorporation of novel genetic combinations and characters, but also facilitate the hybridisations without the use of special techniques to produce crosses, especially when working with large populations with rich genetic diversity.

The created populations should be large enough to generate variability for many years, and in this way can overcome the consequences of genetic drift, where the population reaches a "fixation" state and only new mutations or incorporations into the population can reintroduce variation. Most autogamous species still possess some out-crossing rate, but in some cases, such as tomato, it is possible to encourage the use of facultative out-crossing when selecting parental lines that have exserted stigma as characteristic; another alternative discussed in this research is the exploitation of male sterility. About the continuous hybridisation to conserve the created genetic variability, a model was designed for one single locus and showed that there will be a need for hybridisation again in F_4 to F_5 generations in the case of completely autogamous populations.

Genetic base broadening programmes must minimise selection, as part of its philosophy, particularly during initial stages, but natural selection always occur in the populations. There are biotic and abiotic stresses acting over the population producing a natural selection pressure, therefore human intervention should be minimised to allow that these forces produce the effect in the population. With the aim to utilise these forces is recommended a multi-site exploitation of natural selection. The populations exposed to different environmental conditions and biotic pressures will produce many different genotypes surviving at each site maintaining the genetic variability as an overall within the population. However, the high costs and logistic difficulties are great disadvantages for this kind of approach.

The results obtained in this project are interesting, but require validation/comparison with other alternative populations. Future research should include comparison of

populations with differing levels of initial variability against focussed and non-focussed base broadening objectives. Additional questions for autogamous species will focus on the desirability of establishing populations that can continue to provide variability on which selection can take place over an extended period thereby eliminating the need to re-establish populations from original parental sources.

Most domesticated crops need genetic base broadening, some earlier than others to reduce the barriers for future crop improvement. Chilean agriculture is similar to many; the rapid development of F_1 hybrids and genetically narrow modern cultivars over the last 20-25 years, has led to a rapid erosion of the variability once widespread within the species utilised by the agricultural community.. At present, there is still diverse germplasm available in isolated communities and farms, but this is transitory as farmers continue to accept the norms of modern agriculture with high yield potential and uniformity within cultivars. Therefore in the short term there is real risk that this genetic material will be lost forever. Chile has considerable diversity with regard to climate and soil types. There will be a real need to find cultivars, which will be able to exploit this environmental variability. Base broadening activity in a country such as Chile may hold additional advantages, namely the conservation of genetic material in a manner which allows further evolution. Such research in a range of specific crops could be extrapolated/exported to neighbouring countries because of the geographical position and climatic advantages in comparison to other countries. For instance, it is possible to obtain two or three harvests a year in crops such as tomato, beans, broccoli, maize, etc.

For the future, it is hoped to use the experience gained with *Lycopersicon* spp utilising morphological and molecular markers in order to answer some of the questions posed in the introduction. These operational questions need to be examined in order to remove the empiricism that has, by and large, dominated previous base broadening efforts. Such methodologies on their own will not answer all the questions. Some questions will be very much species specific, others are more a matter of resources and finance. Nevertheless, the lessons learnt from pursuing some of the questions, in such an amenable species, will have messages for other autogamous crops in other environments.

Finally, broadening the genetic base of a crop species can take many forms: by creating diverse populations utilising a wide range of parental material (landraces through to progenitor species); by encouraging exploitation of genotypes in space and time (diversification schemes); by utilising deliberate genotype mixtures or designing improved landraces (exploiting agronomic combining ability); by developing farmer participatory selection programmes (allowing farm-based adaptation), to name but a few. All of these approaches have merits for different agricultural systems and all need to be considered in order to avoid an unsustainable dependency on a few genotypes.

References

ADAMS, T. L. and QUIROS, C. F. (1985). Somatic hybridization between *Lycopersicon peruvianum* and *L. pennellii*: regenerating ability and antibiotic resistance as selection systems. *Plant Science* **40**, 209-219.

AGWANDA, C. O. and OWUOR, J. B. O. (1989). Clonal comparative trials in Arabica Coffee (*Coffea arabica* L.). I: The effect of broadening the genetic base on the stability of yield in Kenya. *Kenya Coffee* **54**, 639-643.

AHMAD, M., MCNEIL, D. L., FAUTRIER, A. G., ARMSTRONG, K. F. and PATERSON, A. M. (1995). Genetic relationships in *Lens* species and parentage determination of their inter-specific hybrids using RAPD markers. *Theor Appl Genet* **92**, 1091-1098.

ARENS, P., BREDEMEIJER, G., SMULDERS, M. J. M. and VOSMAN, B. (1995). Identification of tomato cultivars using microsatellites. *Acta Horticulturae* **412**, 49-57.

AYALA, F. J. and KIGER, J. A. Jr. (1984). Modern genetics. Menlo Park, California: The Benjamin/Cummings Publishing Company, Inc.

AYUSA, M. C., NUEZ, F. and CUARTERO, J. (1986). Relationship between *Lycopersicon esculentum* and a compatible form of *Lycopersicon peruvianum*. *Rep Tomato Genet Coop* **36**, 2.

BAHAMONDES, L. A. and MALLEA, A. R. (1969). Biologia en Mendoza de *Scrobipalpula absoluta* (Meyrick) Povolny (Lepidoptera -Gelechiidae), espécie nueva para la República Argentina. *Revista de Facultad de Ciencias Agrárias* **15**, 96-104.

BECKER, H. C., ENGQVIST, G. M. and KARLSSON, B. (1995). Comparison of rapeseed cultivars and resynthesized lines based on allozymes and RFLP markers. *Theor Appl Genet* **91**, 62-67.

BERG, T. (1997). Devolution of plant breeding. In *Using diversity. Enhancing and maintaining genetic resources on-farm*, eds. SPERLING, L. and LOEVINSOHN, M. Ottawa, Canada: IDRC.

BOHM, G. W., SANDERSON, G. and BOYNTON, J. E. (1966). A search for biochemical mutants in the tomato. *Rep Tomato Genet Coop* **16**, 5-7.

BOSWELL, V. R. (1937). Improvements and genetics of tomato, peppers and eggplant. *Yearbook of agriculture (US Dept. Agr.)* 176-206.

BRETO, M. P., ASINS, M. J. and CARBONELL, E. A. (1993). Genetic variability in *Lycopersicon* species and their genetic relationship. *Theor Appl Genet* **86**, 113-120.

BRETTING, P. K. and GOODMAN, M. M. (1989). Genetic variation in crop plants and management of germplasm collections. In *Scientific management of germplasm: characterization, evaluation and enhancement*, eds. STALKER, H. T. and CHAPMAN, C. pp.41-54. Rome : IBPGR.

BROUN, P. and TANKSLEY, S. D. (1996). Characterisation and genetic mapping of simple repeat sequences in the tomato genome. *Mol. Gen. Genet.* **250**, 39-49.

BUTLER, L. (1952). The linkage map of the tomato. *J. Hered.* **43**, 25-35.

BUTLER, L. (1955). Dihybrid F_2 segregations for genes of linkages groups II, III, and IV. *Rep Tomato Genet Coop* **5**, 7-13.

CAETANO-ANOLLÉS, G., BASSAUR, B. J. and GRESSHOFF, P. M. (1991). DNA amplification fingerprinting using very short arbitrary oligonucleotides primers. *Bio/Technology* **9**, 553-557.

CARVALLO, R., BASSO, C., SCATONI, I. and COMOTTO, F. (1981). Ensayo para el control de *Scrobipalpula absoluta* (Meyrick) temporada 1980-81. *Revista Técnica, Montevideo* **50**, 41-46.

CARVER, B. F. and TALIAFERRO, C. M. (1992). Selection theory and procedure in progenies of distant hybrids. In *Distant hybridisation of crop plants*, eds. KALLOO, G. and CHOWDHURY, J. B. pp. 122-148. Heidelberg, Germany: Springer-Verlag.

CHAVE, J. W. (1991). Sugar cane breeding at the W.I. Central Sugar Cane Breeding Station with special reference to the base broadening programme. 63-70.(Abstract)

CHETELAT, R. T. and DEVERNA, J. W. (1991). Expression of unilateral incompatibility in pollen of *Lycopersicon pennellii* is determined by major loci in chromosomes 1, 6 and 10. *Theor Appl Genet* **82**, 704-712.

CHETELAT, R. T., RICK, C. M. and DEVERNA, J. W. (1989). Isozyme analysis, chromosome pairing, and fertility of *Lycopersicon esculentum* x *Solanum lycopersicoides* backcross hybrids. *Genome* **32**, 783-790.

CLAYBERG, C. D., BUTLER, L., KERR, E. A., RICK, C. M. and ROBINSON, R. W. (1967). Supplementary list of tomato genes as of January, 1967. *Rep Tomato Genet Coop* **17**, 2-11.

CLAYBERG, C. D., BUTLER, L., KERR, E. A., RICK, C. M. and ROBINSON, R. W. (1970). Additions to the list of genes, January 1967 to January 1970. *Rep Tomato Genet Coop* **20**, 6-11.

CLAYBERG, C. D., BUTLER, L., KERR, E. A., RICK, C. M. and ROBINSON, R. W. (1973). Additions to the list of genes. *Rep Tomato Genet Coop* **23**, 3-7.

CLAYBERG, C. D., BUTLER, L., RICK, C. M. and YOUNG, P. A. (1960). Second list of known genes in the tomato. *J. Hered.* **51**, 167-174.

COOPER, H. D., SPILLANE, C., KERMALI, I. and ANISHETTY, N. M. (1998). Harnessing plant genetic resources for sustainable agriculture. *Plant Genetic Resources Newsletter* **114**, 1-8.

DE ROCHAMBEAU, H., FOURNET-HANOCQ, F., and VU TIEN KHANG, J. (2000). Measuring and managing genetic variability in small populations. *Ann. Zootech.* **49**, 77-93.

DEVERNA, J. W., CHETELAT, R. T., RICK, C. M. and STEVENS, M. A. (1987). Introgression of *Solanum lycopersicoides* germplasm. In *Tomato Biotechnology*, eds. JONES, R. A. and NEVINS, D. J. pp. 27-37. New York: Alan R. Liss, Inc.

DEVERNA, J. W. and PATERSON, A. H. (1991). Genetics of *Lycopersicon.*. In *Genetic improvement of tomato.* ed. KALLOO, G. pp. 22-38. Berlin: Springer-Verlag.

DOGANLAR, S., FRARY, A. and TANKSLEY, S. D. (2000). The genetic basis of seed-weight variation: tomato as a model system. *Theor Appl Genet* **100**, 1267-1273.

DOOLITTLE, S. P. (1954). The use of wild *Lycopersicon* species for tomato disease control. *Phytopathology* **44**, 404-414.

DUBREUIL, P. and CHARCOSSET, A. (1998). Genetic diversity within and among maize populations: a comparison between isozyme and nuclear RFLP loci. *Theor Appl Genet* **96**, 577-587.

ESPINOZA, W. (1991). *Manual de produçao de tomate industrial no Vale de Sâo Francisco*. Brasilia: CODEVASF/IICA.

ESTAY, P. (1998). Polilla del tomate, su manejo integrado. *Tierra Adentro* **21**, 20-23.

EVANS, D. A. (1987). Somaclonal variation. In *Tomato Biotechnology*, eds. JONES, R. A. and NEVINS, D. J. pp. 59-71. New York: Alan R. Liss, Inc.

EVANS, L. T. (1993). Crop evolution, adaptation and yield. Cambridge: Cambridge University Press.

EXCOFFIER, L., SMOUSE, P. E. and QUATTRO, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes. *Genetics* **131**, 479-491.

FAO (1996a). Global plan of action for the conservation and sustainable utilisation of plant genetic resources for food and agriculture, and the Leipzig Declaration.

FAO (1996b). Report on the state of the world's plant genetic resources.

FAO (1999). Production Yearbook. *Production Yearbook 1999* **52**, 227.

FOOLAD, M. R. (1996). Unilateral incompatibility as a major cause of skewed segregation in the cross between *Lycopersicon esculentum* and *L. pennellii*. *Plant Cell Report* **15**, 627-633.

FRANÇA, F. H., MALUF, W. R., ROSSI, P. E. F., MIRANDA, J. E. C. and COELHO, M. C. F. (1984). Avaliaçâo e seleçâo em tomate visando resistência à traça-do-tomateiro. p143.

FREY, K. J. (1994). Remarking a crop gene pool: the case history of *Avena*. *Proceedings of the National Science Council, Republic of China, part B, Life Sciences* **18**, 85-93.(Abstract)

FRYXELL, P. A. (1954). Genetics of locule number. *Rep Tomato Genet Coop* **4**, 10-11.

GARANKO, I. B. (1991). Germplasm resources in *Lycopersicon*. In *Genetic Improvements of Tomato*. ed. KALLOO, G. pp. 51-57. Berlin: Springer-Verlag.

GEPTS, P. (1993). The use of molecular and biochemical markers in crop evolution studies. In *Evolutionary biology*, ed. HECHT, M. K. E. pp. 51-94. New York: Plenum Press.

GLENDINNING, D. R. (1979). Enriching the potato gene-pool using primitive cultivars. In *Proceedings of the conference: Broadening the genetic base of crops*, ed. ZEVEN, A. C. and VAN HARTEN, A. M. pp. 39-45. Wageningen: Centre for Agricultural Publishing and Documentation.

GOLDMAN, I. L., PARAN, I. and ZAMIR, D. (1995). Quantitative trait locus analysis of a recombinant inbred line population derived from a *Lycopersicon esculentum* x *L. cheesmanii* cross. *Theor Appl Genet* **90**, 925-932.

GOODMAN, M. M. (1985). Exotic maize germplasm: status, prospects and remedies. *Iowa State J Res* **59**, 497-527.

GOWER, J. C. (1985). Measures of similarity, dissimilarity, and distance. In *Encyclopedia of statistical sciences*, eds. KOTZ, S. and JOHNSON, N. L. pp. 397-405. New York: Wiley.

GRADZIEL, T. and ROBINSON, R. W. (1985). Overcoming stylar incompatibility in *Solanum* and *Lycopersicon* by using bud pollination under high humidity. *Plant Cell Incomp Newslett.* **17**, 13.

GRANDILLO, S., KU, H. M. and TANKSLEY, S. D. (1999). Identifying the loci responsible for natural variation in fruit size and shape in tomato. *Theor Appl Genet* **99**, 978-987.

GRANDILLO, S. and TANKSLEY, S. D. (1996). QTL analysis of horticultural traits differentiating cultivated tomato from the closely related species *Lycopersicon pimpinellifolium*. *Theor Appl Genet* **92**, 935-942.

GRIFFITHS, A. J. F., MILLER, J. H., SUZUKI, D. T., LEWONTIN, R. C. and GELBART, W. M. (1996). *An introduction to genetic analysis*. New York: W.H. Freeman and Co.

HACHIZUME, T., SHIMAMOTO, L., HARUSHIME, Y., YUI, M., SATO, T., IMAI, T. and HIRAI, M. (1996). Construction of a linkage map for watermelon (*Citrullus lanatus* (Thumb.) Matsui and Nakai) using random amplified polymorphic DNA (RAPD). *Euphytica* **90**, 265-273.

HALSTED, B. D., OWEN, E. J. and SHAW, J. K. (1905). Experiments with tomatoes. NJ Agr Exp Stn Annu Rep **26**, 447-477.

HAMADA, H., PETRINO, M. G. and KAKUNAGA, T. (1982). A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proc Natl Acad Sci USA* **79**, 6465-6469.

HAMRICK, J. L. and GODT, M. J. W. (1997). Allozyme diversity in cultivated crops. *Crop Sci* **37**, 26-30.

HANDLEY, L. W., NICKELS, R. L., CAMERON, M. W., MOORE, P. P. and SINK, K. C. (1986). Somatic hybrids plants between *Lycopersicon esculentum* and *Solanum lycopersicoides*. *Theor Appl Genet* **71**, 691-697.

HANSON, M. R. and MCCLEAN, P. M. (1987). Tomato plastid and mitochondrial genomes: changes and conservation. In *Tomato Biotechnology*, eds. JONES, R. A. and NEVINS, D. J. pp. 215-228. New York: Alan R. Liss, Inc.

HARDON, J. J. (1967). Unilateral incompatibility between *Solanum pennellii* and *Lycopersicon esculentum*. *Genetics* **57**, 795-808.

HARTL, D.L. and CLARK, A.G. (1997). *Principles of population genetics*. Sunderland, MA, USA: Sinauer Assoc.

HAWKES, J. G. (1979). Genetic poverty of the potato in Europe. In *Proceedings of the conference: Broadening the genetic base of crops,* ed. ZEVEN, A. C. and VAN HARTEN, A. M. pp. 19-36. Wageningen: Centre for Agricultural Publishing and Documentation.

HAWKES, J. G. (1991). Centros de diversidad genética vegetal en Latinoamérica. *Diversity* **7**, 7-9.

HAWTIN, G., IWANAGA, M. and HODGKIN, T. (1996). Genetic resources in breeding for adaptation. *Euphytica* **92**, 255-266.

HECTOR, J. M. (1936). Introduction to the botany of field crops, vol II. Johannesburg, South Africa: Central News Agency Ltd.

HEDGE, S. G., VALKOUN, J. and WAINES, J. G. (2000). Genetic diversity in wild wheats and goat grass. *Theor Appl Genet* **101**, 309-316.

HEWITT, J. D. and GARVEY, T. C. (1987). Wild sources of high soluble solids in tomato. In *Tomato Biotechnology*, eds. JONES, R. A. and NEVINS, D. J. pp. 45-55. New York: Alan R. Liss, Inc.

HOELZEL, A. R. and GREEN, A. (1998). PCR protocols and population analysis by direct DNA sequencing and PCR-based DNA fingerprinting. In *Molecular genetic analysis of populations. A practical approach*, ed. HOELZEL, A. R. pp. 201-236. Oxford: IRL Press.

HOGENBOOM, N. G. (1979). Incompatibility and incongruity. In *The Biology and taxonomy of the Solanaceae*, eds. HAWKES, J. G., LESTER, R. N. and SKELDING, A. D. pp. 435-444. London: Academic Press.

HOSTIKA, L. P. and HANSON, M. R. (1984). Induction of plastid mutations in tomatoes by nitromethylurea. *J Hered.* **75**, 242-246.

HU, J. and QUIROS, C. F. (1991). Identification of broccoli and cauliflower cultivars with RAPD markers. *Plant Cell Reports* **10**, 505-511.

IPGRI (1996). Descriptors for tomato (*Lycopersicon spp*). International Plant Genetic Resources Institute, Rome Italy.

JENKINS, J. A. (1948). The origin of the cultivated tomato. *Econ. Bot.* **2**, 379-392.

JENSEN, N. F. (1988). *Plant breeding methodology*. New York: John Wiley and Sons.

JONES, D. F. (1917). Linkage in *Lycopersicum*. *Am. Nat.* **51**, 608-621.

KANNENBERG, L. W. and FALK, D. E. (1995). Models for activation of plant genetic resources for crop breeding programs. *Canadian J Plant Sci* **75**, 45-53.

KAUL, M. L. H. (1991). Reproductive biology in tomato. In *Genetic improvement of tomato*, ed. KALLOO, G. pp. 39-50. Berlin: Springer-Verlag.

KAWCHUK, L. M., LYNCH, D. R., HACHEY, J., BAINS, P. S. and KULCSAR, F. (1994). Identification of a codominant amplified polymorphic DNA marker linked to the *Verticillium* wilt resistance gene in tomato. *Theor Appl Genet* **89**, 661-664.

KNAPP, S. and SPOONER, D. M. (1999). A new name for a common Ecuadorean and Peruvian wild tomato species. *NOVON* **9**, 375-376.

KRESOVICH, S., WILIAMS, J. G. K., MCFERSON, J. R., ROUTMAN, E. J. and SCHAAL, B. A. (1992). Characterization of genetic identities and relationships of *Brassica oleracea* L. via a random amplified polymorphic DNA assay. *Theor Appl Genet* **85**, 190-196.

KRUSKAL, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* **29**, 1-27.

KRUSKAL, W. H. and WALLIS, W. A. (1952). Use of ranks in one criterion analysis of variance. *J. Amer. Statist. Assoc.* **47**, 583-621.

KU, H. M., DOGANLAR, S., CHEN, K.-Y. and TANKSLEY, S. D. (1999). The genetic basis of pear-shaped tomato fruit. *Theor Appl Genet* **9**, 844-850.

LARRAIN, P. (1986). Eficacia de insecticidas y frecuencia de aplicación basada en niveles poblacionales críticos de *Scrobipalpula absoluta* (Meyrick) en tomates. *Agricultura Técnica* **46**, 329-333.

LEHMANN, L. C. (1991). The use of genetic resources for isolating disease resistance for barley cultivar development. *Barley Genetics* **VI**, 650-652.

LI, C. P., CHEN, Y. H., CHERN, C. G., TSENG, T. H., CHEN, C. C., LAI, M. S. and KUO, Y. C. (1994). Broadening genetic diversity of cultivated rice through interspecific hybridisation within genus *Oryza*.I. *Journal of Agricultural Reaserch of China* **43**, 261-274.(Abstract)

LINDHOUT, P., VERKERK, R., MULLAART, E., UITTERLINDEN, A. and VIJG, J. (1991). Cultivar identification of tomato by DNA analysis with repetitive probes. *Rep Tomato Genet Coop* **41**, 31.

LOBO, M., ALVARAN, I. and CARDONA, M. P. (1990). Fruit storage life in *Lycopersicon pimpinellifolium*. *Rep Tomato Genet Coop* **40**, 20.

LOURENÇAO, A. L., NAGAI, H. and ZULLO, M. A. T. (1984). Fontes de resistência a *Scrobipalpula absoluta* (Meyrick, 1917) em tomateiro. *Bragantia* **43**, 569-577.

LYNCH, M. and MILLIGAN, B. G. (1994). Analysis of population genetic structure with RAPD markers. *Molecular Ecology* **3**, 91-99.

MACARTHUR, J. W. and BUTLER, L. (1938). Size inheritance and geometric growth processes in the tomato fruit. *Genetics* **23**, 253-268.

MACGILLIVRAY, J. H. and CLEMENTE, L. J. (1956). Effect of tomato size on solid content. *Proc. Am. Soc. Hort. Sci.* **68**, 466-469.

MANTEL, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**, 209-220.

MARTIN, F. W. (1962). Distribution and interrelationship of incompatibity barriers with *Lycopersicon hirsutum* Humb. and Bonpl. complex. *Evolution* **17**, 519-528.

MAYO, O. (1987). *The theory of plant breeding*. Oxford: Clarendon Press.

MCCLEAN, P. E. and HANSON, M. R. (1986). Mitochondrial DNA sequence divergence among *Lycopersicon* and related *Solanum* species. *Genetics* **112**, 649-667.

MCGRATH, D. J., GILLESPIE, D. and VAWDREW, L. (1987). Inheritance of resistance to *Fusarium oxysporum* f.sp. *lycopersici* races 2 and 3 in *Lycopersicon pennellii*. *Aust J Agric Res* **38**, 729-733.

MENDOZA, H. A. (1989). Population breeding as a tool for germplasm enhancement. *American Potato Journal* **66**, 639-653.

MENGESHA, M. H. and PRASADA-RAO, K. E. (1982). Current situation and future of sorghum germplasm. In *Sorghum in the Eighties*, ed. ICRISAT pp 323-335. India: ICRISAT.

MENGISTU, L. W., MUELLER-WARRANT, G. W. and BARKER, R. E. (2000). Genetic diversity of *Poa annua* in western Oregon grass seed crops. *Theor Appl Genet* **101**, 70-79.

MILLER, J. C. and TANKSLEY, S. D. (1990). RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theor Appl Genet* **80**, 437-448.

MILLIGAN, B. G. (1992). Plant DNA isolation. In *Molecular genetic analysis of populations. A practical approach.* ed. HOELZEL, A. R. pp. 59-88. Oxford, UK: IRL Press.

MITRICK, A. J., SKROCH, P. W., NIENHUIS, J., HINRICHSEN, P., BASCUR, G. and MUÑOZ-SCHICK, C. (1997). Gene pool classification of common bean landraces from Chile based in RAPD and morphological data. *Crop Sci* **37**, 605-613.

MOREIRA, J. O. T., LARA, F. M. and CHURATA-MASCA, M. G. C. (1981). Ocorrência de *Scrobipalpula absoluta* (Meyrick) (Lepidoptera - Gelechiidae) danificando tomate rastreiro em Jaboticabal-SP. 58.

MORGANTE, M. and OLIVIERI, A. M. (1993). PCR-amplified microsatellites as markers in plant genetics. *Plant J* **3**, 175-182.

MÜLLER, C. H. (1940). A revision of the genus *Lycopersicon*. *USDA Misc. Publ.* **328**, 29.

NASSAR, N. M. A. (1989). Broadening the genetic base of cassava, *Manihot esculenta* Krantz, by interspecific hybridisation. *Can J Plant Sci* **69**, 1071-1073.

NCBI (2001). Taxonomy browser. In: *National Center for Biotechnology Information.* Bethesda, MD, USA. http://www.ncbi.nlm.nih.gov/htbin-post/Taxonomy.

NEI, M. (1987). *Molecular evolutionary genetics*. New York: Columbia University Press.

NIENHUIS, J. and BOSCO, J. D. (1994). Genetic distances among toamto cultivars as measured by RAPD molecular markers. *Acta Horticulturae* **376**, 59-66.

NIEUWHOF, M., GARRETSEN, F. and VAN OEVEREN, J. C. (1989). Maternal and genetic effects on seed weight of tomato, and effects of seed weight on growth of genotypes of tomato (*Lycopersicon esculentum* Mill.). *Plant Breed* **102**, 248-254.

O'CONNELL, M. A. and HANSON, M. R. (1985). Somatic hybridization between *Lycopersicon esculentum* and *Lycopersicon pennellii*. *Theor Appl Genet* **70**, 1-12.

O'CONNELL, M. A. and HANSON, M. R. (1987). Construction and characterisation of somatic hybrids in *Lycopersicon*. In *Tomato Biotechnology*, eds. JONES, R. A. and NEVINS, D. J. pp. 71-80. New York: Alan R. Liss, Inc.

OLMSTEAD, R. G. and PALMER, J. D. (1997). Implications for the phylogeny, classification, and biogeography of *Solanum* from cpDNA restriction site variation. *Systematic Botany* **22**, 19-29.

ORTIZ, R. (1998). Genetic enhancement and base broadening efforts. In *Implementation of GPA in Europe - Braunschweig Proceedings*, pp. 191-203.

ORTIZ, R., MADSEN, S., RUIZ-TAPIA, E. N., JACOBSEN, S. E., MUJICA-SANCHEZ, A., CHRISTIANSEN, J. and STOLEN, O. (1999). Validating a core collection of Peruvian quinoa germplasm. *Genetic resources and crop evolution* **46**, 285-290.

PALMER, J. D. and ZAMIR, D. (1982). Chloroplast DNA evolution and phylogenetic relationships in *Lycopersicon*. *Proc. Natl. Acad. Sci.* **79**, 5006-5010.

PASQUET, R. S. (2000). Allozyme diversity of cultivated cowpea *Vigna unguiculata* (L.) Walp. *Theor Appl Genet* **101**, 211-219.

PEJIC, I., AJNONE-MARSON, P., MORGANTE, M., KOZUMPLIK, V., CASTIGLIONI, P., TARAMINO, G. and MOTTO, M. (1998). Comparative analysis of genetic similarities among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. *Theor Appl Genet* **97**, 1248-1255.

PERALTA, I. E. and SPOONER, D. M. (2001). Granule-Bound Starch Synthase (GBSSI) gene phylogeny of wild tomatoes (*Solanum L.*section *Lycopersicon* (Mill.) Wettst. subsection *Lycopersicon*). In press: *Am J Bot.*.

PLAISTED, R. L. (1982). Progress and future plans for the use of neotuberosum populations. In *Utilisation of the genetic resources of the potato III*, pp. 65-79. Lima, Peru: International Potato Center.

POVOLNY, D. (1975). On three Neotropical species of Gnorimoschemini (Lepidoptera - Gelechiidae) mining Solanaceae. *Acta Universitatis Agriculturae Facultas, Brno* **2**, 379-393.

PROKOPY, R. J. and OWENS, E. D. (1983). Visual detection of plants by herbivorous insects. *Ann. Rev. Entomol.* **28**, 337-364.

PROVAN, J., POWELL, W. and WAUGH, R. (1996). Microsatellite analysis of relationships within cultivated potato (*Solanum tuberosum*). *Theor Appl Genet* **92**, 1078-1084.

QUIROS, C. F. (1974). Phylogeny of the tomato species. *Rep Tomato Genet Coop* **24**, 19-20.

RICK, C. M. (1958). The role of natural hybridisation in the derivation of cultivated tomatoes of western South America. *Econ. Bot.* **12**, 346-367.

RICK, C. M. (1969). Controlled introgression of chromosomes of *Solanum pennellii* into *Lycopersicon esculentum*: segregation and recombination. *Genetics* **62**, 753-768.

RICK, C. M. (1973). Potential genetic resources in tomato species: clues from observations in native habitats. In *Genes, enzymes, and populations*, ed. SRB, A. M. New York: Plenum Press.

RICK, C. M. (1976). Tomato (family *Solanaceae*). In *Evolution of crop plants*, ed. SIMMONDS, N. W. pp. 262-273. London: Longman.

RICK, C. M. (1977). A key for the tomato species. *Rep Tomato Genet Coop* **27**, 21.

RICK, C. M. (1979a). Biosystematic studies in *Lycopersicon* and closely related species of *Solanum*. In *The biology and taxonomy of the Solanaceae*, eds. HAWKES, J. G., LESTER, R. N. and SKELDING, A. D. pp. 667-677. London: Academic Press.

RICK, C. M. (1979b). Potential improvement of tomatoes by controlled introgression of genes from wild species. 167-173.

RICK, C. M. (1982). *Lycopersicon esculentum* (tomato). In *Genetic maps*, ed. O'BRIEN, S. J. pp. 360-367. Frederick, MD: National Cancer Institute.

RICK, C. M. (1983). Tomato (*Lycopersicon*). In *Tomato biotechnology*, eds. TANKSLEY, S. D. and ORTON, T. J. pp. 17-26. New York: Alan R. Liss.

RICK, C. M. (1984). Evolution of mating systems: evidence from allozyme variation. In *Genetics: new frontiers, Proceedings of the XV International Congress of Genetics*, eds. CHOPRA, V. L., JOSHI, B. C., SHARMA, R. P., and BANSAL, H. C. **4**, 215-221. Essex; Epping.

RICK, C. M. (1984). Stock list. *Rep Tomato Genet Coop* **34**, 22.

RICK, C. M. (1987). Genetic resources in *Lycopersicon*. In *Tomato Biotechnology*, eds. NEVINS, D. J. and JONES, R. A. pp. 17-27. New York: Alan R. Liss, Inc.

RICK, C. M. (1991). Tomato resources of South America reveal many genetic treasures. *Diversity* **7**, 54-56.

RICK, C. M. and BUTLER, L. (1956). Cytogenetics of the tomato. *Adv. Genet.* **8**, 267-382.

RICK, C. M. and CHETELAT, R. T. (1995). Utilisation of related wild species for tomato improvement. *Acta Horticulturae* **412**, 21-38.

RICK, C. M., CHETELAT, R. T. and DEVERNA, J. W. (1988). Recombination in sesquidiploid hybrids of *Lycopersicon esculentum* x *Solanum lycopersicoides* and derivatives . *Theor Appl Genet* **76**, 647-655.

RICK, C. M., DEVERNA, J. W., CHETELAT, R. T. and STEVENS, M. A. (1986). Meiosis in sesquidiploid hybrids of *Lycopersicon esculentum* and *Solanum lycopersicoides*. *Proc Natl Acad Sci USA* **83**, 3580-3583.

RICK, C. M. and FOBES, J. F. (1975). Allozyme variation in the cultivated tomato and closely related species. *Bull. Torrey Bot. Club* **102**, 376-386.

RICK, C. M. and LAMM, R. (1955). Biosystematic studies on the status of *Lycopersicon chilense*. *Am J Bot.* **42**, 663-675.

RICK, C. M., TANKSLEY, S. D. and FOBES, J. F. (1979). A pseudoduplication in *Lycopersicon pimpinellifolium*. *Proc Natl Acad Sci* **76**, 3435-3439.

RUS-KORTEKAAS, W., SMULDERS, M. J. M., ARENS, P. and VOSMAN, B. (1994). Direct comparison of levels of genetic variation in tomato detected by a GACA-containing microsatellite probe and by random amplified polymorphic DNA. *Genome* **37**, 375-381.

SAAVEDRA, G. and SPOOR, W. (2002). Genetic base broadening in autogamous crops: *Lycopersicon esculentum* Mill. As a model. In Managing plant genetic resources. ed. ENGELS, J. M. M., BROWN, A. H. D., and JACKSON, M. T. pp. 291-299. Wallingford, UK: CABI Publishing.

SALHUANA, W., POLLAK, L. and TIFFANY, D. (1993). Public/private collaboration proposed to strengthen quality and production of U.S. corn through maize germplasm enhancement. *Diversity* **9-10**, 77-79.

SCARDINI, D. M. B., FERREIRA, L. R. and GALVEAS, P. A. O. (1983). Ocorrênciada traça do tomateiro *Scrobipalpula absoluta* (Meyr.) no Estado do Espírito Santo. In *Congresso Brasileiro de Entomologia, 8 Resumos* p72. Brasilia, Brasil.

SCHIFFMAN, S. S., REYNOLDS, M. L. and YOUNG, F. W. (1981). *Introduction to multidimensional scaling.* New York: Academic Press.

SCHLÖTTERER, C. (1998). Microsatellites. In *Molecular genetic analysis of population. A practical approach.* ed. HOELZEL, A. R. pp. 237-261. Oxford, UK: IRL Press.

SCHLÖTTERER, C. and TAUTZ, D. (1992). Slippage synthesis of simple sequence DNA. *Nucleic Acids Res* **20**, 211-215.

SCHOONHOVEN, L. M., JERMY, T. and VAN LOON, J. J. A. (1998). Insects and flowers: the beauty of mutualism. In *Insect-Plant biology from physiology to evolution*, Anonymous pp. 315-342. London: Chapman and Hall.

SCOTT, J. W. and JONES, J. P. (1991). Genetic control of resistance to races 1,2, and 3 of *Fusarium* wilt. *Rep Tomato Genet Coop* **41**, 47.

SEVILLA, R. and HOLLE, M. (1995). Recursos genéticos vegetales. La Molina, Lima, Peru.

SIMMONDS, N. W. (1979). *Principles of crop improvements*. London: Longman.

SIMMONDS, N. W. (1993). Introgression and incorporation. Strategies for the use of crop genetic resources. *Biol. Rev.* **68**, 539-562.

SMARTT, J. and SIMMONDS, N. W. (1994). *Evolution in crop plants*. London: Longman.

SMULDERS, M. J. M., BREDEMEIJER, G., RUS-KORTEKAAS, W., ARENS, P. and VOSMAN, B. (1997). Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. *Theor Appl Genet* **94**, 264-272.

SNEATH, P. H. A. and SOKAL, R. O. (1973). *Numerical taxonomy*. San Francisco, USA: Freeman.

SNEDECOR, G. W. (1934). *Calculation and interpretation of analysis of variance and covariance*. Ames, Iowa, USA: Collegiate Press.

SNEEP, J. (1979). The breeder's point of view on broadening the genetic base. In *Proceedings of the conference: Broadening the genetic base of crops,* ed. ZEVEN,A.C. and VAN HARTEN,A.M. pp. 13-18. Wageningen: Centre for Agricultural Publishing and Documentation.

SPOONER, D. M., ANDERSON, G. J. and JANSEN, R. K. (1993). Chloroplast DNA evidence for the interrelationships of tomatoes, potatoes, and pepino (*Solanaceae*). *Am J Bot* **80**, 676-688.

SRB, A. M., OWEN, R. D. and EDGAR, R. S. (1965). *General Genetics*. San Francisco: W.H. Freeman and Company.

STEVENS, M. A. (1980). Utilization of natural varieties to develop resistence to biotic and environmental stress in processing tomato cultivars. *Acta Horticulturae* **100**, 405-410.

STEVENS, M. A. and RICK, C. M. (1986). 2. Genetics and breeding. In *The Tomato Crop. A scientific basis for improvement*, eds. ATHERTON, J. G. and RUDICH, J. pp. 35-110. Cambridge: Chapman and Hall Ltd.

STEVENS, M. A. and RUDICH, J. (1978). Genetic potential for overcoming physiological limitations on adaptability, yield, and quality in the tomato. *HortScience* **13**, 673-678.

SUNESON, C. A. (1956). An evolutionary plant breeding method. *Agron J* **56**, 188-191.

TANKSLEY, S. D., GANAL, M. W., PRINCE, J. P., DE VICENTE, M. C., BONIERBALE, M. W., BROUN, P., FULTON, T. M., GIOVANNONI, J. J., GRANDILLO, S., MARTIN, G. B., MESSEGUER, R., MILLER, J. C., PATERSON, A. H., PINEDA, O., RODER, M., WING, R. A., WU, W. and YOUNG, N. D. (1992). High density molecular linkage maps of the tomato and potato genomes. *Genetics* **132**, 1141-1160.

TANKSLEY, S. D. and NELSON, J. C. (1996). Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into breeding lines. *Theor Appl Genet* **92**, 191-203.

TAUTZ, D. and RENZ, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acid Res.* **12**, 4127-4138.

TAYLOR, B. H., MANHART, J. R. and AMASINO, R. M. (1993). Isolation and characterisation of plant DNA. In *Methods in plant molecular biology and biotechnology*, eds. GLICK, B. R. and THOMPSON, J. E. pp. 37-47. Boca Raton, Florida: CRC Press Inc.

TAYLOR, I. B. (1986). 1. Biosystematics of the tomato. In *The Tomato Crop. A scientific basis for improvement*, eds. ATHERTON, J. G. and RUDICH, J. pp. 1-34. Cambridge: Chapman and Hall Ltd.

TIKOO, S. K. and ANAND, N. (1982). Stable phenotypic expression of stigma position in tomato. *Rep Tomato Genet Coop* **32**, 48.

TUKEY, J. W. (1953). *The problem of multiple comparisons*. Princeton, USA: Princeton University.

VAN BEUNINGEN, L. T. and BUSCH, R. H. (1997). Genetic diversity among North American spring wheat cultivars: II. Ancestor contributions to gene pools of different eras and regions. *Crop Sci.* **37**, 580-585.

VAN DER BEEK, H., VERKERK, R., ZABEL, P. and LINDHOUT, P. (1992). Mapping strategy for resistance genes in tomato based on RFLPs between cultivars:

CF9 (resistance to *Cladosporium fulvum*) on chromosome 1. *Theor Appl Genet* **84**, 106-112.

VAVILOV, N. I. (1926). Studies on the origin of cultivated plants. *Bull. Appl. Bot. Pl. Breed.* **16**, 1-248.

VETELAINEN, M., NISSILA, E., TIGERSTEDT, P. M. A. and VON BOTHMER, R. (1996). Utilisation of exotic germplasm in Nordic barley breeding and its consequence for adaptation. *Euphytica* **92**, 267-273.

VILLAND, J. (1995). Comparison of molecular marker and morphological data to determine genetic distances among tomato cultivars. Dissertation, University of Wisconsin, Madison.

VILLAND, J., SKROCH, P. W., LAI, I., HANSON, P., KUO, C. G. and NIENHUIS, J. (1998). Genetic variation among tomato accessions from primary and secondary centers of diversity. *Crop Sci* **38**, 1339-1347.(Abstract)

VOSMAN, B. and ARENS, P. (1997). Molecular characterisation of GATA/GACA microsatellite repeats in tomato. *Genome* **40**, 25-33.

WANG, J. L. (1994). Broadening soybean genetic basis in the northeast of China. *Journal of Northeast Agricultural University* **1**, 1-8.(Abstract)

WARNOCK, S. J. (1988). A review of taxonomy and phylogeny of the genus *Lycopersicon*. *HortScience* **23**, 669-673.

WAUGH, R. and POWELL, W. (1992). Using RAPD markers for crop improvement. *Tibtech* **10**, 186-191.

WELSH, J. and MCCLELLAND, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acid Res* **18**, 7213-7218.

WELSH, W., BUSHUK, W., ROCA, W. and SINGH, S. P. (1995). Characterisation of agronomic traits and markers of recombinant inbred lines from intra-and interracial populations of *Phaseolus vulgaris* L. *Theor Appl Genet* **91**, 169-177.

WILKIE, S., ISAAC, P. G. and SLATER, R. J. (1993). Random amplified polymorphic DNA (RAPD) markers for genetic analysis in *Allium*. *Theor Appl Genet* **86**, 497-504.

WILLIAMS, C. E. and ST.CLAIR, D. A. (1993). Phenetic relationship and levels of variability detected by restriction fragment length polymorphism and random amplified polymorphic DNA analysis of cultivated and wild accessions of *Lycopersicon esculentum*. *Genome* **36**, 619-630.

WILLIAMS, J. G. K., KUBLELIK, A. R., LIVAK, K. J., RAFALSKI, J. A. and TINGEY, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleis Acid Res* **18**, 6531-6535.

WINKLER, H. (1909). Über die experimentelle Erzeugung von Pflanzen mit abweichenden Chromosomenzahlen. *Z Bot.* **8**, 417-531.

WRICKE, G. and WEBER, W. E. (1986). *Quantitative genetics and selection in plant breeding*. Berlin: Walter de Gruyter.

WRIGHT, S. (1965). The interpretation of population structure by F-statistics with special regard to system mating. Evolution **19,** 395-420.

YEH, F., YANG, R. C., BOYLE, T., YE, Z. H. and MAO, J. (1997). Popgene, the user friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada.

Appendix 1

APPENDIX 1 : MORPHOLOGICAL CHARACTERIZATION OF TOMATO (IPGRI, 1996)

Plant descriptors

1) Vegetative

1.1 Seedling

Records should be taken when the seedling primary leaves are fully opened and the terminal bud is around 5 mm in size.

1.1.1 Hypocotyl colour

1 Green
2 1/4 purple from the base
3 1/2 purple from the base
4 Purple

1.1.2 Hypocotyl pubescence

0 Absent
1 Present

1.2 Plant characteristics

Records should be taken when the fruits of the 2nd and 3rd truss are ripened.

1.2.1 Plant growth type

1 Dwarf
2 Determinate
3 Semi-determinate
4 Indeterminate

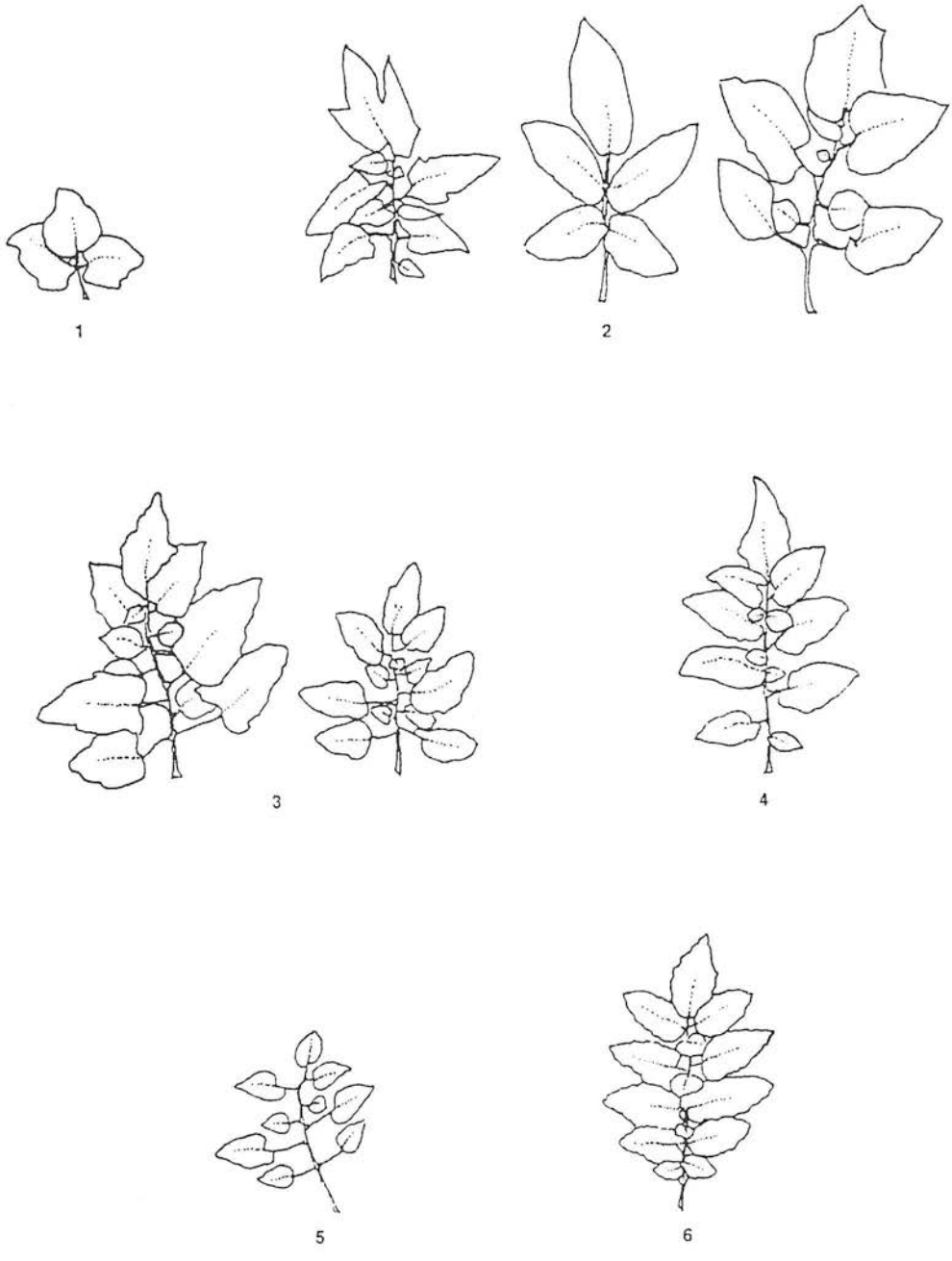
1.2.2 Leaf attitude

3 Semi-erect
5 Horizontal
7 Drooping

1.2.3 Leaf type (see figure 1)

1 Dwarf
2 Potato leaf type
3 Standard
4 Peruvianum
5 Pimpinellifolium
6 Hirsutum
7 Other

Figure 1. Leaf type.



2) Inflorescence and fruit

Unless otherwise indicated, all observations on the fruit should be taken, when possible, on the 3rd fruit of the 2nd and/or 3rd truss at the full matutity stage, provide normal fertilization has occurred.

2.1 Inflorescence descriptor

2.1.1 Inflorescence type

Observe the 2nd and 3rd truss of at least 10 plants

1 Generally uniparous
2 Both (partly uniparous, partly multiparous)
3 Generally multiparous

2.1.2 Corolla colour

1 White
2 Yellow
3 Orange
4 Other

2.1.3 Style position

The relative position of the style compared with the stamens.
Average of 10 styles from different flowers of different plants.

1 Inserted
2 Same level as stamen
3 Slightly exserted
4 Highly exserted

2.2 Fruit descriptors

Unless otherwise indicated, all observations on the fruit should be taken, when possible, on the 3rd fruit of the 2nd and/or 3rd truss at the full maturity stage, provided normal fertilization has occurred. Record the average of 10 fruits from different plants.

2.2.1 Exterior colour of immature fruit

1 Greenish-white
3 Light green
5 Green
7 Dark green
9 Very dark green

2.2.2 Fruit pubescence

3 Sparse (*L. esculentum)*
5 Intermediate (*L. pennelii*)
7 Dense (*L. hirsutum)*

2.2.3 Predominant fruit shape

Recorded after fruits turn colour (see figure 2)

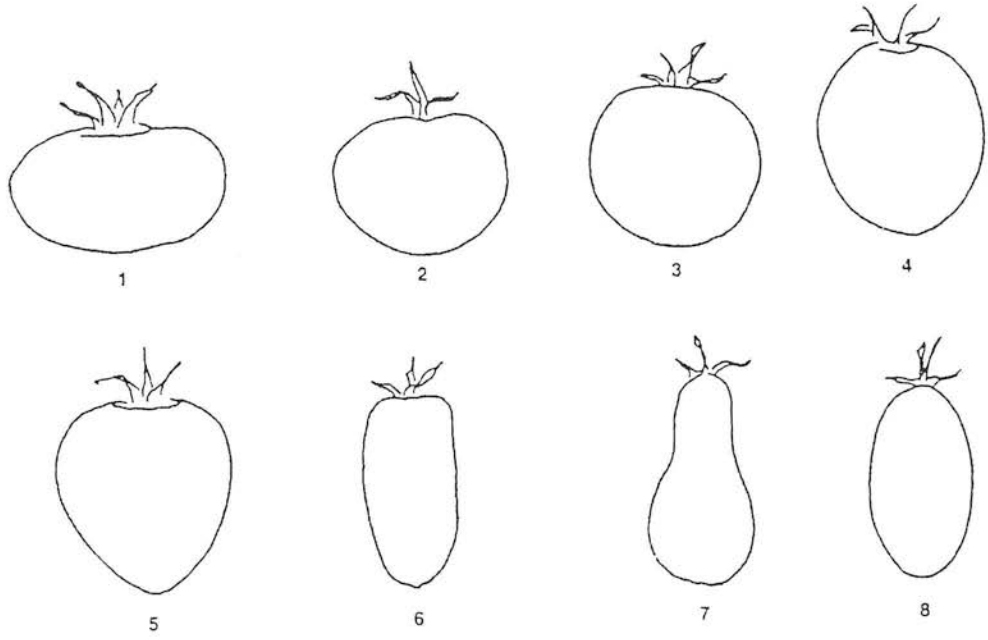
1 Flattened (oblate)
2 Slightly flattened
3 Rounded
4 High rounded
5 Heart shaped
6 Cylindrical (long oblong)
7 Pyriform
8 Ellipsoid (plum shaped)
9 Other

2.2.4 Fruit size

At maturity

1 Very small (<3 cm)
2 Small (3 - 5 cm)
3 Intermediate (5.1 - 8 cm)
4 Large (8.1 - 10 cm)
5 Very large (>10 cm)

Figure 2. Predominant fruit shape.



2.2.5 Exterior colour of mature fruit

Recorded at maturity

1 Green
2 Yellow
3 Orange
4 Pink
5 Red
6 Other

2.2.6 Flesh colour of pericarp (interior)

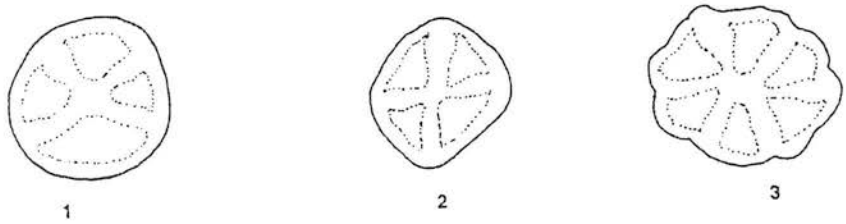
1 Green
2 Yellow
3 Orange
4 Pink
5 Red
6 Other

2.2.7 Fruit cross-sectional shape

(See figure 3)

1 Round
2 Angular
3 Irregular

Figure 3. Fruit cross-sectional shape.



2.2.8 Number of locules

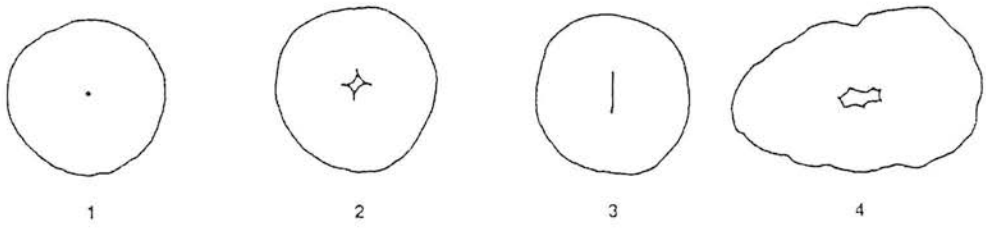
Counted on at least 10 fruits

2.2.9 Shape of pistil scar

(See figure 4)

1 Dot
2 Stellate
3 Linear
4 Irregular

Figure 4. Shape of pistil scar.

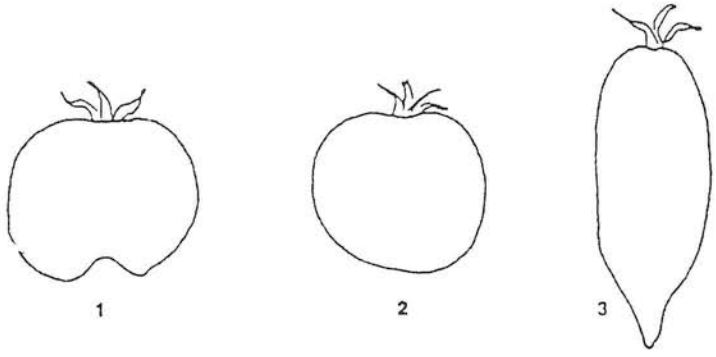


2.2.10 Fruit blossom end shape

(See figure 5)

1 Indented
2 Flat
3 Pointed

Figure 5. Fruit blossom end shape.



Appendix 2

Appendix 2 Part 1 Genetic similarity matrix from 16 morphological characters of 35 accessions of *Lycopersicon* accessions (Accessions codes in Part 4).

Lchees1	1.00																																				
Lchees2	0.54	1.00																																			
Lchees3	0.69	0.77	1.00																																		
Lcheesm1	0.54	0.77	0.69	1.00																																	
Lcheesm2	0.39	0.85	0.54	0.85	1.00																																
Lescul1	0.31	0.39	0.15	0.31	0.31	1.00																															
Lescul2	0.46	0.39	0.46	0.46	0.39	0.31	1.00																														
Lescul3	0.39	0.54	0.31	0.46	0.39	**0.54**	0.46	1.00																													
Lescul5	0.15	0.23	0.00	0.31	0.31	0.31	0.31	0.31	1.00																												
Lescul6	0.39	0.39	0.39	0.31	0.23	0.46	0.54	0.39	0.54	1.00																											
Lescul7	0.39	0.23	0.31	0.31	0.23	**0.23**	0.62	**0.39**	0.46	0.46	1.00																										
Lescul8	0.54	0.46	0.46	0.39	0.31	0.77	0.62	0.69	0.31	0.54	0.62	1.00																									
Lescul9	0.46	0.46	0.39	0.46	0.39	**0.54**	0.62	**0.62**	0.23	0.46	**0.46**	0.77	1.00																								
Lescul10	0.23	0.46	0.15	0.39	0.31	**0.46**	0.46	**0.62**	0.46	0.54	**0.54**	0.62	**0.46**	1.00																							
Lescul11	0.39	0.15	0.23	0.23	0.15	0.31	0.54	0.31	0.69	0.62	0.62	0.39	0.31	0.39	1.00																						
Lescul12	0.39	0.23	0.23	0.23	0.15	0.46	0.46	0.39	0.62	0.62	0.54	0.46	0.39	0.46	0.77	1.00																					
Lescul13	0.46	0.31	0.31	0.39	0.31	0.31	0.69	0.54	0.39	0.39	0.69	0.69	0.54	0.54	0.62	0.62	1.00																				
Lescul14	0.46	0.31	0.23	0.31	0.31	0.39	0.54	0.54	0.46	0.46	0.54	0.62	0.54	0.62	0.46	0.54	0.77	1.00																			
Lescul15	0.15	0.39	0.23	0.39	0.31	0.54	0.39	0.54	0.39	0.46	0.39	0.39	0.39	0.69	0.31	0.39	0.46	0.54	1.00																		
Lescul16	0.54	0.54	0.46	0.54	0.46	0.54	0.62	0.62	0.23	0.46	0.54	0.69	0.85	0.46	0.31	0.39	0.54	0.54	0.46	1.00																	
Lescul17	0.46	0.31	0.23	0.31	0.23	0.62	0.54	0.62	0.39	0.54	0.54	0.69	0.69	0.54	0.54	0.62	0.62	0.62	0.39	0.77	1.00																
Lescul18	0.46	0.31	0.23	0.31	0.23	0.39	0.31	0.54	0.54	0.31	0.62	0.62	0.54	0.46	0.62	0.54	0.54	0.54	0.46	0.69	0.62	1.00															
Lescer	0.54	0.54	0.46	0.46	0.39	**0.23**	0.62	**0.69**	0.46	0.46	**0.39**	0.69	**0.46**	0.46	0.54	0.54	0.62	0.54	0.39	0.54	0.54	0.54	1.00														
Lhirs1	0.54	0.54	0.69	0.62	0.39	0.08	0.39	0.31	0.00	0.23	0.31	0.31	0.23	0.23	0.15	0.23	0.39	0.31	0.23	0.31	0.23	0.15	**0.31**	1.00													
Lhirs2	0.54	0.54	0.69	0.62	0.46	0.08	0.39	0.23	0.00	0.23	0.31	0.31	0.23	0.23	0.15	0.23	0.39	0.31	0.23	0.31	0.23	0.15	0.31	1.00	1.00												
Lhirsg1	0.54	0.54	0.69	0.62	0.39	0.08	0.39	0.23	0.00	0.23	0.31	0.31	0.23	0.23	0.15	0.23	0.39	0.31	0.23	0.31	0.23	0.15	0.31	1.00	1.00	1.00											
Lhirsg2	0.54	0.54	0.69	0.62	0.39	0.08	0.39	0.23	0.00	0.23	0.31	0.31	0.23	0.23	0.15	0.23	0.39	0.31	0.23	0.31	0.23	0.15	0.31	1.00	1.00	1.00	1.00										
Lhirsg3	0.31	0.31	0.39	0.31	0.15	**0.08**	0.31	**0.23**	0.15	0.31	**0.31**	0.23	**0.15**	0.31	0.31	0.39	0.39	0.31	0.46	0.15	0.15	0.23	**0.39**	0.69	0.69	0.69	0.69	1.00								
Lparvi1	0.54	0.54	0.54	0.62	0.54	0.15	0.31	0.23	0.15	0.23	0.46	0.39	0.31	0.23	0.31	0.31	0.39	0.23	0.23	0.39	0.15	0.39	0.23	0.62	0.62	0.62	0.62	0.54	1.00							
Lparvi3	0.46	0.62	0.54	0.46	0.31	**0.08**	0.23	**0.46**	0.39	0.46	**0.23**	0.39	**0.31**	0.46	0.31	0.31	0.31	0.31	0.23	0.23	0.23	0.39	**0.62**	0.31	0.31	0.31	0.31	**0.39**	0.46	1.00						
Lpenn1	0.39	0.39	0.31	0.46	0.39	0.23	0.31	0.31	0.23	0.39	0.23	0.31	0.31	0.23	0.23	0.31	0.23	0.31	0.23	0.31	0.23	0.31	0.31	0.46	0.46	0.54	0.54	0.54	0.62	0.46	1.00					
Lpennp	0.39	0.39	0.39	0.46	0.15	**0.15**	0.23	**0.23**	0.23	0.31	**0.15**	0.23	**0.23**	0.15	0.23	0.31	0.23	0.39	0.23	0.23	0.15	0.31	**0.31**	0.39	0.39	0.31	0.39	**0.62**	0.54	**0.54**	0.77	1.00				
Lpimpi1	0.39	0.62	0.39	0.46	0.46	0.15	0.46	0.31	0.15	0.15	0.23	0.31	0.39	0.31	0.15	0.23	0.39	0.31	0.23	0.39	0.39	0.23	0.46	0.46	0.46	0.46	0.46	0.39	0.39	0.39	0.46	0.39	1.00			
Lpimpi2	0.39	0.69	0.54	0.46	0.39	0.15	0.23	0.39	0.08	0.23	0.00	0.31	0.23	0.15	0.00	0.08	0.15	0.23	0.15	0.23	0.15	0.08	0.39	0.39	0.39	0.39	0.39	0.31	0.46	0.62	0.39	0.46	0.54	1.00		
Lpimpi3	0.39	0.62	0.39	0.46	0.46	**0.15**	0.31	**0.39**	0.15	0.23	**0.08**	0.31	**0.39**	0.15	0.08	0.15	0.23	0.31	0.15	0.39	0.31	0.23	**0.39**	0.23	0.23	0.23	0.23	**0.15**	0.39	**0.46**	0.54	**0.46**	0.69	0.92	1.00	

Appendix 2 Part 2 Genetic similarity matrix from 55 microsatellite markers of 38 *Lycopersicon* accessions (Accessions codes in Part 4).

Lchees1	1.00																																					
Lchees2	0.59	1.00																																				
Lchees3	0.50	0.71	1.00																																			
Lcheem1	0.50	0.78	0.83	1.00																																		
Lcheem2	0.50	0.53	0.33	0.36	1.00																																	
Lescul1	0.48	0.45	0.43	0.48	0.28	1.00																																
Lescul2	0.36	0.39	0.33	0.29	0.27	0.67	1.00																															
Lescul3	0.46	0.39	0.42	0.33	0.27	**0.73**	0.84	1.00																														
Lescul4	0.46	0.39	0.42	0.33	0.27	0.73	0.84	1.00	1.00																													
Lescul5	0.35	0.33	0.37	0.28	0.15	0.50	0.65	0.65	0.65	1.00																												
Lescul6	0.32	0.31	0.34	0.26	0.14	0.52	0.73	0.73	0.73	0.83	1.00																											
Lescul7	0.28	0.27	0.30	0.23	0.14	**0.47**	0.67	**0.67**	0.67	0.83	0.83	1.00																										
Lescul8	0.30	0.29	0.32	0.25	0.17	0.44	0.62	0.62	0.62	0.69	0.77	0.84	1.00																									
Lescul9	0.28	0.27	0.30	0.23	0.14	**0.47**	0.67	**0.67**	0.67	0.75	0.83	0.91	0.77	1.00																								
Lescul10	0.28	0.27	0.34	0.26	0.14	**0.47**	0.55	**0.55**	0.55	0.62	0.69	**0.76**	0.77	0.83	1.00																							
Lescul11	0.31	0.34	0.38	0.33	0.17	0.55	0.53	0.53	0.53	0.59	0.67	0.67	0.62	0.73	0.73	1.00																						
Lescul12	0.33	0.32	0.40	0.31	0.19	0.54	0.63	0.63	0.63	0.64	0.72	0.79	0.73	0.79	0.79	0.76	1.00																					
Lescul13	0.23	0.31	0.30	0.26	0.18	0.47	0.61	0.50	0.50	0.56	0.63	0.69	0.70	0.76	0.91	0.67	0.72	1.00																				
Lescul14	0.22	0.31	0.30	0.26	0.16	0.43	0.52	0.46	0.46	0.41	0.43	0.48	0.50	0.54	0.67	0.58	0.56	0.74	1.00																			
Lescul15	0.39	0.47	0.45	0.41	0.26	0.70	0.74	0.68	0.68	0.69	0.70	0.64	0.60	0.64	0.64	0.74	0.73	0.64	0.50	1.00																		
Lescul16	0.34	0.52	0.45	0.41	0.31	0.59	0.74	0.68	0.68	0.69	0.70	0.64	0.60	0.64	0.53	0.62	0.61	0.53	0.40	0.85	1.00																	
Lescul17	0.39	0.57	0.55	0.50	0.36	0.70	0.62	0.68	0.68	0.57	0.59	0.53	0.50	0.53	0.53	0.62	0.61	0.53	0.45	0.85	0.85	1.00																
Lescul18	0.41	0.45	0.39	0.35	0.35	0.62	0.66	0.71	0.71	0.50	0.52	0.47	0.44	0.47	0.38	0.45	0.44	0.38	0.34	0.63	0.75	0.75	1.00															
Lescer	0.41	0.43	0.29	0.33	0.36	**0.67**	0.70	**0.70**	0.70	0.43	0.50	**0.55**	0.52	0.55	**0.50**	0.53	0.57	0.55	0.52	0.62	0.57	0.62	0.60	1.00														
Lhirs1	0.43	0.34	0.29	0.29	0.19	0.25	0.24	0.28	0.28	0.23	0.22	0.18	0.24	0.18	0.22	0.24	0.22	0.18	0.21	0.31	0.27	0.27	0.30	0.24	1.00													
Lhirs2	0.33	0.30	0.29	0.29	0.25	0.29	0.23	0.28	0.28	0.13	0.13	0.16	0.19	0.16	0.20	0.23	0.25	0.20	0.23	0.23	0.19	0.27	0.30	0.37	0.37	1.00												
Lhirsg1	0.36	0.39	0.33	0.38	0.32	0.41	0.35	0.35	0.35	0.26	0.24	0.28	0.26	0.28	0.28	0.40	0.38	0.28	0.32	0.39	0.34	0.39	0.33	0.50	0.40	0.65	1.00											
Lhirsg2	0.25	0.27	0.21	0.22	0.20	0.18	0.14	0.17	0.17	0.12	0.11	0.14	0.14	0.18	0.18	0.24	0.15	0.18	0.20	0.17	0.17	0.20	0.23	0.28	0.52	0.55	0.42	1.00										
Lhirsg3	0.23	0.36	0.40	0.33	0.15	**0.38**	0.36	**0.36**	0.36	0.33	0.33	**0.43**	0.40	0.46	**0.58**	0.54	0.58	0.58	0.58	0.46	0.36	0.46	0.43	**0.50**	0.31	0.45	0.50	0.38	1.00									
Lparvi1	0.21	0.30	0.29	0.29	0.25	0.24	0.31	0.31	0.31	0.22	0.21	0.24	0.23	0.28	0.24	0.27	0.21	0.28	0.32	0.23	0.26	0.26	0.29	0.40	0.20	0.42	0.36	0.37	0.46	1.00								
Lparvi2	0.21	0.30	0.29	0.29	0.25	0.24	0.31	0.31	0.31	0.22	0.21	0.24	0.23	0.28	0.24	0.27	0.21	0.28	0.32	0.23	0.26	0.26	0.29	0.40	0.20	0.42	0.36	0.37	0.46	1.00	1.00							
Lparvi3	0.26	0.27	0.26	0.23	0.21	**0.42**	0.61	**0.55**	0.55	0.43	0.50	**0.55**	0.57	0.61	**0.55**	0.45	0.52	0.55	0.43	0.44	0.53	0.44	0.47	**0.55**	0.15	0.28	0.32	0.18	**0.43**	0.43	0.43	1.00						
Lpenn1	0.19	0.21	0.17	0.17	0.08	0.16	0.23	0.23	0.23	0.21	0.19	0.19	0.18	0.23	0.20	0.23	0.17	0.20	0.28	0.19	0.22	0.19	0.25	0.23	0.31	0.26	0.22	0.32	0.50	0.42	0.42	0.32	1.00					
Lpenn2	0.19	0.25	0.29	0.24	0.19	0.24	0.24	0.24	0.24	0.18	0.17	0.17	0.16	0.21	0.24	0.28	0.21	0.24	0.28	0.23	0.23	0.27	0.28	0.23	0.24	0.38	0.31	0.29	0.40	0.45	0.45	0.42	0.67	1.00				
Lpennp	0.10	0.13	0.12	0.08	0.24	**0.00**	0.04	**0.04**	0.04	0.04	0.04	**0.04**	0.04	0.04	**0.04**	0.04	0.04	0.04	0.05	0.04	0.08	0.08	0.08	**0.08**	0.17	0.12	0.09	0.19	**0.20**	0.17	0.17	**0.18**	0.42	0.42	1.00			
Lpimp1	0.36	0.27	0.31	0.23	0.26	0.36	0.43	0.48	0.48	0.42	0.44	0.46	0.43	0.52	0.54	0.52	0.50	0.48	0.50	0.43	0.37	0.38	0.40	0.39	0.29	0.29	0.25	0.26	0.27	0.25	0.25	0.37	0.28	0.36	0.16	1.00		
Lpimp2	0.17	0.24	0.27	0.19	0.14	0.19	0.27	0.27	0.27	0.36	0.33	0.35	0.32	0.41	0.43	0.41	0.32	0.36	0.24	0.38	0.42	0.38	0.31	0.23	0.22	0.20	0.17	0.26	0.13	0.22	0.22	0.39	0.14	0.25	0.17	0.47	1.00	
Lpimp3	0.12	0.21	0.28	0.21	0.17	**0.30**	0.38	**0.33**	0.33	0.38	0.40	**0.42**	0.38	0.48	**0.55**	0.46	0.43	0.55	0.43	0.42	0.37	0.37	0.31	**0.35**	0.13	0.27	0.23	0.24	**0.33**	0.19	0.19	**0.43**	0.24	0.33	**0.18**	0.67	0.56	1.00

Appendix 2 Part 3 Genetic similarity matrix from 268 RAPD markers of 38 *Lycopersicon* accessions (Accessions codes in Part 4).

Lchees1	1.00																																												
Lchees2	0.60	1.00																																											
Lchees3	0.50	0.59	1.00																																										
Lcheem1	0.53	0.64	0.57	1.00																																									
Lcheem2	0.55	0.61	0.57	0.69	1.00																																							
Lescul1	0.49	0.48	0.54	0.48	0.56	1.00																																						
Lescul2	0.49	0.46	0.57	0.61	0.62	0.63	1.00																																					
Lescul3	0.42	0.47	0.49	0.47	0.54	**0.70**	0.76	1.00																																				
Lescul4	0.36	0.43	0.42	0.49	0.56	0.57	0.64	0.66	1.00																																			
Lescul5	0.42	0.39	0.44	0.42	0.51	0.58	0.59	0.66	0.66	1.00																																		
Lescul6	0.42	0.46	0.45	0.46	0.56	0.56	0.59	0.62	0.63	0.71	1.00																																	
Lescul7	0.32	0.34	0.38	0.42	0.49	**0.51**	0.59	**0.62**	0.54	0.61	0.71	1.00																																
Lescul8	0.33	0.37	0.40	0.38	0.47	0.53	0.54	0.60	0.59	0.63	0.67	0.76	1.00																															
Lescul9	0.37	0.38	0.42	0.42	0.52	**0.55**	0.62	**0.62**	0.60	0.64	0.73	**0.79**	0.72	1.00																														
Lescul10	0.33	0.35	0.39	0.45	0.49	**0.43**	0.52	**0.53**	0.59	0.56	0.61	**0.67**	0.70	**0.69**	1.00																													
Lescul11	0.38	0.36	0.41	0.40	0.49	0.55	0.57	0.56	0.51	0.63	0.71	0.76	0.80	0.74	0.54	1.00																												
Lescul12	0.32	0.37	0.37	0.40	0.42	0.50	0.52	0.54	0.55	0.59	0.63	0.73	0.79	0.72	0.62	0.88	1.00																											
Lescul13	0.29	0.33	0.45	0.40	0.41	0.44	0.55	0.52	0.42	0.51	0.57	0.69	0.65	0.65	0.58	0.70	0.69	1.00																										
Lescul14	0.25	0.32	0.39	0.36	0.41	0.42	0.48	0.52	0.46	0.48	0.49	0.65	0.69	0.62	0.62	0.67	0.68	0.68	1.00																									
Lescul15	0.35	0.38	0.45	0.38	0.44	0.46	0.49	0.51	0.43	0.52	0.68	0.68	0.66	0.68	0.52	0.65	0.65	0.70	0.63	1.00																							
Lescul16	0.32	0.36	0.44	0.37	0.42	0.44	0.47	0.52	0.40	0.48	0.54	0.59	0.56	0.64	0.55	0.59	0.60	0.67	0.62	0.63	1.00																					
Lescul17	0.36	0.36	0.43	0.39	0.44	0.37	0.46	0.46	0.39	0.45	0.50	0.54	0.57	0.60	0.48	0.63	0.60	0.66	0.63	0.63	0.76	1.00																				
Lescul18	0.27	0.33	0.39	0.36	0.41	0.33	0.41	0.43	0.38	0.42	0.48	0.60	0.58	0.55	0.56	0.62	0.63	0.63	0.60	0.65	0.67	0.73	1.00																			
Lescer	0.28	0.35	0.35	0.33	0.40	**0.34**	0.37	**0.44**	0.34	0.43	0.50	**0.59**	0.63	**0.57**	**0.53**	0.61	0.62	0.63	0.58	0.69	0.63	0.72	0.73	1.00																		
Lhirs1	0.26	0.30	0.36	0.35	0.31	0.24	0.24	0.24	0.21	0.24	0.27	0.27	0.32	0.28	0.30	0.32	0.32	0.36	0.35	0.39	0.35	0.36	0.39	0.38	1.00																	
Lhirs2	0.23	0.27	0.31	0.30	0.30	0.24	0.24	0.25	0.25	0.25	0.27	0.31	0.30	0.26	0.27	0.28	0.30	0.34	0.26	0.31	0.30	0.35	0.39	0.34	0.42	1.00																
Lhirsg1	0.23	0.28	0.27	0.25	0.27	0.21	0.22	0.24	0.20	0.22	0.26	0.28	0.29	0.28	0.26	0.28	0.29	0.29	0.28	0.31	0.26	0.31	0.34	0.31	0.42	0.61	1.00														
Lhirsg2	0.20	0.26	0.28	0.29	0.32	0.18	0.24	0.22	0.19	0.20	0.25	0.29	0.27	0.26	0.27	0.27	0.27	0.29	0.29	0.30	0.27	0.31	0.34	0.30	0.46	0.53	0.64	1.00													
Lhirsg3	0.21	0.26	0.34	0.28	0.29	**0.24**	0.31	**0.27**	0.23	0.24	0.28	**0.33**	0.32	**0.31**	**0.38**	0.31	0.32	0.37	0.34	0.35	0.32	0.34	0.37	**0.35**	0.35	0.35	0.29	0.33	1.00												
Lparvi1	0.34	0.40	0.44	0.36	0.38	0.33	0.33	0.32	0.30	0.29	0.30	0.33	0.35	0.31	0.30	0.31	0.29	0.35	0.34	0.35	0.37	0.36	0.37	0.38	0.38	0.29	0.30	0.27	0.40	1.00										
Lparvi2	0.31	0.36	0.45	0.35	0.40	0.31	0.35	0.30	0.30	0.31	0.31	0.36	0.37	0.31	0.28	0.34	0.30	0.41	0.37	0.39	0.36	0.35	0.40	0.40	0.41	0.34	0.32	0.33	0.41	0.82	1.00									
Lparvi3	0.30	0.33	0.31	0.30	0.37	**0.31**	0.31	**0.32**	0.31	0.32	0.41	**0.44**	0.46	**0.47**	**0.37**	0.52	0.44	0.48	0.45	0.48	0.44	0.47	0.50	**0.51**	0.30	0.32	0.28	0.28	**0.41**	0.43	0.48	1.00								
Lpenn1	0.20	0.24	0.24	0.23	0.26	0.24	0.27	0.26	0.24	0.26	0.26	0.26	0.29	0.29	0.23	0.28	0.28	0.27	0.26	0.29	0.28	0.29	0.33	0.28	0.34	0.26	0.34	0.33	0.29	0.38	0.36	0.35	1.00							
Lpenn2	0.23	0.25	0.33	0.25	0.28	0.32	0.28	0.28	0.27	0.28	0.27	0.28	0.30	0.29	0.21	0.28	0.28	0.29	0.26	0.29	0.28	0.28	0.29	0.27	0.29	0.27	0.32	0.31	0.23	0.29	0.33	0.29	0.61	1.00						
Lpennp	0.26	0.28	0.31	0.32	0.31	**0.24**	0.30	**0.24**	0.28	0.18	0.23	**0.25**	0.29	**0.25**	**0.26**	0.29	0.28	0.29	0.31	0.28	0.27	0.28	0.33	**0.25**	0.37	0.31	0.28	0.31	**0.31**	0.41	0.34	**0.35**	0.51	0.49	1.00					
Lpimp1	0.37	0.35	0.37	0.39	0.42	0.36	0.44	0.42	0.40	0.41	0.45	0.48	0.49	0.48	0.43	0.47	0.46	0.51	0.49	0.48	0.50	0.50	0.47	0.49	0.36	0.26	0.25	0.34	0.36	0.36	0.39	0.47	0.32	0.29	0.38	1.00				
Lpimp2	0.35	0.36	0.41	0.42	0.49	0.46	0.57	0.53	0.55	0.52	0.51	0.59	0.59	0.58	0.49	0.56	0.55	0.58	0.65	0.55	0.56	0.54	0.53	0.53	0.34	0.27	0.26	0.28	0.30	0.35	0.38	0.46	0.30	0.32	0.31	0.66	1.00			
Lpimp3	0.39	0.36	0.36	0.37	0.40	**0.34**	0.41	**0.38**	0.35	0.39	0.37	**0.44**	0.43	**0.46**	**0.43**	0.40	0.40	0.48	0.43	0.41	0.45	0.41	0.38	**0.39**	0.33	0.30	0.29	0.28	**0.31**	0.35	0.34	**0.39**	0.31	0.30	**0.34**	0.57	0.59	1.00		

Appendix 2 Part 4 *Lycoperiscon* accessions codes of identification.

Lchees1	=	*L. cheesmanii*				CGN-15820
Lchees2	=	*L. cheesmanii*				LA-0166
Lchees3	=	*L. cheesmanii*				PI-379035
Lcheem1	=	*L. cheesmanii*	var.	*minor*		LA-0317
Lcheem2	=	*L. cheesmanii*	var.	*minor*		PI-379040
Lescul1	**=**	***L. esculentum***			**cv.**	**Limachino**
Lescul2	=	*L. esculentum*			cv.	Ailsa Craig
Lescul3	**=**	***L. esculentum***			**cv.**	**Ace**
Lescul4	=	*L. esculentum*			cv.	Cal Ace
Lescul5	=	*L. esculentum*			cv.	Earliana
Lescul6	=	*L. esculentum*			cv.	Edkawi
Lescul7	**=**	***L. esculentum***			**cv.**	**Lukullus**
Lescul8	=	*L. esculentum*			cv.	Moneymaker
Lescul9	**=**	***L. esculentum***			**cv.**	**Marglobe**
Lescul10	**=**	***L. esculentum***			**cv.**	**San Marzano**
Lescul11	=	*L. esculentum*			cv.	Pearson
Lescul12	=	*L. esculentum*			cv.	Stone
Lescul13	=	*L. esculentum*			cv.	Red Top
Lescul14	=	*L. esculentum*			cv.	Roma
Lescul15	=	*L. esculentum*			cv.	Super Roma
Lescul16	=	*L. esculentum*			cv.	1702 F_1
Lescul17	=	*L. esculentum*			cv.	Boa F_1
Lescul18	=	*L. esculentum*			cv.	Cobra F_1
Lescer	**=**	***L. esculentum***	**var.**	***cerasiforme***		**LA-1673**
Lhirs1	=	*L. hirsutum*				LA-1353
Lhirs2	=	*L. hirsutum*				LYC 4/88
Lhirsg1	=	*L. hirsutum*	var.	*glabratum*		G-29255
Lhirsg2	=	*L. hirsutum*	var.	*glabratum*		LA-1223
Lhirsg3	**=**	***L. hirsutum***	**var.**	***glabratum***		**PI-199381**
Lparvi1	=	*L. parviflorum*				LA-1322
Lparvi2	=	*L. parviflorum*				LA-1326
Lparvi3	**=**	***L. parviflorum***				**T-1264/94**
Lpenn1	=	*L. pennellii*				LA-0716
Lpenn2	=	*L. pennellii*				PI-473422
Lpennp	**=**	***L. pennellii***	**var.**	***puberulum***		**LA-1926**
Lpimp1	=	*L. pimpinellifolium*				PI-230327
Lpimp2	=	*L. pimpinellifolium*				PI-270449
Lpimp3	**=**	***L. pimpinellifolium***				**PI-390739**

Appendix 3

Appendix 3 Statistical analyses of morphological and molecular markers data of *Lycopersicon* spp accessions.

Part 1 Statistical analyses of genetic indices from microsatelite markers data given in Tables 3.3.1 and 3.3.2. Non-parametric test Kruskal-Wallis was carried out for non-normally distributed indices *P*, *A* and A_e ; ANOVA was carried out for predominantly normally distributed genetic indices H_s and *I*.

Kruskal-Wallis Test for proportion of polymorphic loci (*P*) in microsatellite markers data of *Lycopersicon* spp.

```
Kruskal-Wallis Test

Species      N     Median      Ave Rank     Z
L.escul      1     29.00          8.0      1.53
L.chees      1     14.00          3.5     -0.44
L.cheesm     1     14.00          3.5     -0.44
L.hirs       1     17.00          6.0      0.65
L.hirsg      1     18.00          7.0      1.09
L.parv       1     16.00          5.0      0.22
L.penn       1     10.00          1.0     -1.53
L.pimp       1     11.00          2.0     -1.09
Overall      8                    4.5

H = 6.92  DF = 7  P = 0.437
H = 7.00  DF = 7  P = 0.429 (adjusted for ties)
```

Kruskal-Wallis Test for number of polymorphic alleles per locus (*A*) in microsatellite data of *Lycopersicon* spp.

```
Kruskal-Wallis Test

Species     N      Median     Ave Rank      Z
L. chees    55     1.000       212.0      -0.53
L. cheesm   55     1.000       212.0      -0.53
L. escul    55     2.000       272.0       3.21
L. hirs     55     1.000       224.0       0.22
L. hirsg    55     1.000       228.0       0.47
L. parvi    55     1.000       220.0      -0.03
L. penn     55     1.000       196.0      -1.53
L. pimp     55     1.000       200.0      -1.28
Overall    440                 220.5

H = 13.22  DF = 7  P = 0.067
H = 21.26  DF = 7  P = 0.003 (adjusted for ties)
```

Kruskal-Wallis Test for effective number of alleles (A_e) in microsatellite markers data of *Lycopersicon* spp.

```
Kruskal-Wallis Test

Species     N      Median     Ave Rank      Z
L. chees    55     1.000       211.7      -0.55
L. cheesm   55     1.000       213.3      -0.45
L. escul    55     1.058       266.8       2.89
L. hirs     55     1.000       225.5       0.31
L. hirsg    55     1.000       230.4       0.62
L. parvi    55     1.000       218.4      -0.13
L. penn     55     1.000       196.1      -1.52
L. pimp     55     1.000       201.8      -1.17
Overall    440                 220.5
```

```
H = 11.37  DF = 7  P = 0.123
H = 17.62  DF = 7  P = 0.014 (adjusted for ties)
```

One-Way Analysis of Variance of average gene diversity (H_s) in microsatellite markers data of *Lycopersicon* spp.

```
Analysis of Variance
Source     DF        SS        MS        F        P
Factor      7    0.4366    0.0624     1.90    0.068
Error     432   14.1877    0.0328
Total     439   14.6243

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  -------+---------+---------+-----
L. chees   55    0.0968    0.1738       (-------*-------)
L. cheesm  55    0.1054    0.1821         (-------*-------)
L. escul   55    0.1769    0.1988                   (------*------)
L. hirs    55    0.1280    0.1932            (-------*-------)
L. hirsg   55    0.1332    0.1973             (-------*-------)
L. parvi   55    0.1077    0.1768          (-------*-------)
L. penn    55    0.0703    0.1553   (-------*-------)
L. pimp    55    0.0805    0.1679    (-------*-------)
                                    -------+---------+---------+-----
Pooled StDev =   0.1812              0.060     0.120     0.180
```

One-Way Analysis of Variance of Shannon's information index (*I*) in microsatellite markers data of *Lycopersicon* spp.

```
Analysis of Variance
Source     DF        SS        MS        F        P
Factor      7    0.9952    0.1422     2.05    0.048
Error     432   29.9248    0.0693
Total     439   30.9200

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  -------+---------+---------+-----
L. chees   55    0.1436    0.2534      (------*------)
L. cheesm  55    0.1539    0.2658       (------*------)
L. escul   55    0.2658    0.2854                  (------*------)
L. hirs    55    0.1869    0.2820           (------*------)
L. hirsg   55    0.1946    0.2853           (------*------)
L. parvi   55    0.1610    0.2593        (------*------)
L. penn    55    0.1039    0.2262  (------*------)
L. pimp    55    0.1176    0.2418    (------*------)
                                    -------+---------+---------+-----
Pooled StDev =   0.2632              0.10      0.20      0.30
```

Part 2 Equation regression and analysis of variance of regression analysis done in section 3.3.1.1.4 Figure 3.3.3 between sample size and population genetic indices based in microsatellite markers data.

Sample size versus observed number of alleles(*A*).

- *L. cheesmanii*

$A = 1.10 + 0.0544$ Sample size

S = 0.05829 R-Sq = 50.6% R-Sq = 47.1%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	0.048733	0.048733	14.34**	0.002
Error	14	0.047576	0.003398		
Total	15	0.096309			

- *L. esculentum*

$A = 1.03 + 0.0600$ Sample size

S = 0.06490 R-Sq = 64.7% R-Sq = 63.4%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	0.21596	0.21596	51.27**	0.000
Error	28	0.11795	0.00421		
Total	29	0.33391			

- *L. hirsutum*

$A = 1.12 + 0.0934$ Sample size

S = 0.05187 R-Sq = 72.8% R-Sq = 70.3%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	0.079124	0.079124	29.41**	0.000
Error	11	0.029591	0.002690		
Total	12	0.108715			

Sample size versus effective number of alleles (A_e).

- *L. cheesmanii*

A_e = 1.08 + 0.0357 Sample size

S = 0.04405 R-Sq = 43.5% R-Sq = 39.5%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	0.020908	0.020908	10.78**	0.005
Error	14	0.027161	0.001940		
Total	15	0.048068			

- *L. esculentum*

A_e = 1.04 + 0.0357 Sample size

S = 0.04698 R-Sq = 55.2% R-Sq = 53.6%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	0.076270	0.076270	34.56**	0.000
Error	28	0.061787	0.002207		
Total	29	0.138057			

- *L. hirsutum*

A_e = 1.10 + 0.0537 Sample size

S = 0.03515 R-Sq = 65.8% R-Sq = 62.7%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	0.026184	0.026184	21.19**	0.000
Error	11	0.013594	0.001236		
Total	12	0.039778			

Sample size versus average gene diversity (H_s)

- *L. cheesmanii*

$H_S = 0.0459 + 0.0194$ Sample size

S = 0.02454 R-Sq = 42.3% R-Sq = 38.2%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	0.0061867	0.0061867	10.28**	0.006
Error	14	0.0084285	0.0006020		
Total	15	0.0146151			

- *L. esculentum*

$H_S = 0.0225 + 0.0202$ Sample size

S = 0.02617 R-Sq = 56.1% R-Sq = 54.5%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	0.024519	0.024519	35.80**	0.000
Error	28	0.019178	0.000685		
Total	29	0.043697			

- *L. hirsutum*

$H_S = 0.0601 + 0.0304$ Sample size

S = 0.02012 R-Sq = 65.4% R-Sq = 62.3%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	0.0084133	0.0084133	20.79**	0.000
Error	11	0.0044516	0.0004047		
Total	12	0.0128649			

Sample size versus Shannon's information index (*I*)

- *L. cheesmanii*

$I = 0.0660 + 0.0288$ Sample size

S = 0.03549 R-Sq = 43.6% R-Sq = 39.6%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	0.013632	0.013632	10.82**	0.005
Error	14	0.017636	0.001260		
Total	15	0.031268			

- *L. esculentum*

$I = 0.0309 + 0.0304$ Sample size

S = 0.03807 R-Sq = 57.7% R-Sq = 56.2%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	0.055304	0.055304	38.16**	0.000
Error	28	0.040583	0.001449		
Total	29	0.095887			

- *L. hirsutum*

$I = 0.0837 + 0.0467$ Sample size

S = 0.02996 R-Sq = 66.7% R-Sq = 63.7%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	0.019826	0.019826	22.08**	0.000
Error	11	0.009876	0.000898		
Total	12	0.029702			

Sample size versus number of polymorphic loci.

- *L. cheesmanii*

$P = 5.20 + 3.02$ Sample size

S = 3.174 R-Sq = 51.5% R-Sq = 48.0%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	149.44	149.44	14.84**	0.002
Error	14	141.00	10.07		
Total	15	290.44			

- *L. esculentum*

$P = 1.63 + 3.30$ Sample size

S = 3.570 R-Sq = 64.7% R-Sq = 63.4%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	653.40	653.40	51.28**	0.000
Error	28	356.77	12.74		
Total	29	1010.17			

- *L. hirsutum*

$P = 6.27 + 5.16$ Sample size

S = 2.869 R-Sq = 72.8% R-Sq = 70.3%

Analysis of Variance

Source	DF	Sum of Squares	Medium Square	F	P
Regression	1	241.77	241.77	29.38**	0.000
Error	11	90.53	8.23		
Total	12	332.31			

Part 3 Statistical analyses of genetic indices from microsatellite markers data given in Tables 3.3.5. Non-parametric test Kruskal-Wallis was carried out for non-normally distributed indices A and A_e; ANOVA was carried out for predominantly normally distributed genetic indices H_s and I. Tested levels were groups of species in red-fruited and green-fruited.

Non-parametric Kruskal-Wallis Test for A in *Lycopersicon* spp fruit colour groups.

```
Kruskal-Wallis Test

Groups       N     Median    Ave Rank      Z
Red         55      2.000        54.5  -0.33
Green       55      2.000        56.5   0.33
Overall    110                   55.5

H = 0.11  DF = 1  P = 0.742
H = 0.26  DF = 1  P = 0.608 (adjusted for ties)
```

Non-parametric Kruskal-Wallis Test for A_e in *Lycopersicon* spp fruit colour groups.

```
Kruskal-Wallis Test

Groups       N     Median    Ave Ran       Z
Red         55      1.360       54.2   -0.42
Green       55      1.284       56.8    0.42
Overall    110                   55.5

H = 0.17  DF = 1  P = 0.678
H = 0.17  DF = 1  P = 0.677 (adjusted for ties)
```

One-Way Analysis of Variance for H_s in *Lycopersicon* spp fruit colour groups.

```
Analysis of Variance
Source     DF        SS         MS        F        P
Groups      1   0.00880    0.00880     1.11    0.294
Error     108   0.85432    0.00791
Total     109   0.86312

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+------
Red        55   0.09192   0.08825  (-----------*-----------)
Green      55   0.10981   0.08962            (----------*----------)
                                    ------+---------+---------+------
Pooled StDev =  0.08894              0.080     0.100     0.120
```

One-Way Analysis of Variance for I in *Lycopersicon* spp fruit colour groups.

```
Analysis of Variance
Source     DF        SS         MS        F        P
Groups      1    0.0095     0.0095     0.16    0.690
Error     108    6.3886     0.0592
Total     109    6.3981

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  ---------+---------+---------+---
Red        55    0.3705    0.2480  (------------*------------)
Green      55    0.3890    0.2383      (------------*------------)
                                    ---------+---------+---------+---
Pooled StDev =  0.2432                    0.350     0.400     0.450
```

Part 4 Statistical analyses of genetic indices given in Tables 3.3.6 and 3.3.7. Non-parametric test Kruskal-Wallis was carried out for non-normally distributed indices *P*, *A* and A_e ; ANOVA was carried out for predominantly normally distributed genetic indices H_s and *I*.

Kruskal-Wallis Test for proportion of polymorphic loci (*P*) in RAPD markers data of *Lycopersicon* spp.

Kruskal-Wallis Test

Species	N	Median	Ave Rank	Z
L.chees	1	0.3358	6.0	0.65
L.cheesm	1	0.1493	1.0	-1.53
L.escul	1	0.6231	8.0	1.53
L.hirs	1	0.2089	2.0	-1.09
L.hirg	1	0.3433	7.0	1.09
L.parv	1	0.2948	5.0	0.22
L.penne	1	0.2388	3.0	-0.65
L.pimpi	1	0.2799	4.0	-0.22
Overall	8		4.5	

H = 7.00 DF = 7 P = 0.429

* NOTE * One or more small samples

Kruskal-Wallis Test for number of polymorphic alleles per locus (*A*) in RAPD markers data of *Lycopersicon* spp.

Kruskal-Wallis Test on A

Cl	N	Median	Ave Rank	Z
L.chees	268	1.000	281.0	0.48
L.cheesm	268	1.000	245.0	-1.54
L.escul	268	2.000	365.0	5.19
L.hirs	268	1.000	241.0	-1.77
L.hirsug	268	1.000	281.0	0.48
L.parv	268	1.000	273.0	0.03
L.penne	268	1.000	265.0	-0.42
L.pimpi	268	1.000	229.0	-2.44
Overall	2144		272.5	

H = 34.12 DF = 7 P = 0.000
H = 49.92 DF = 7 P = 0.000 (adjusted for ties)

Kruskal-Wallis Test for effective number of alleles (A_e) in RAPD markers data of *Lycopersicon* spp.

Kruskal-Wallis Test on Ae

Cl	N	Median	Ave Rank	Z
L.chees	268	1.000	281.9	0.53
L.cheesm	268	1.000	247.5	-1.40
L.escul	268	1.269	351.2	4.42
L.hirs	268	1.000	243.4	-1.63
L.hirsug	268	1.000	283.6	0.62
L.parv	268	1.000	276.2	0.21
L.penne	268	1.000	267.1	-0.30
L.pimpi	268	1.000	229.0	-2.44
Overall	2144		272.5	

H = 27.02 DF = 7 P = 0.000
H = 37.30 DF = 7 P = 0.000 (adjusted for ties)

One-Way Analysis of Variance of average gene diversity (H_s) in RAPD markers data of *Lycopersicon* spp.

```
Analysis of Variance
Source      DF        SS        MS        F        P
Hs           7    2.7190    0.3884    12.13    0.000
Error     2136   68.4206    0.0320
Total     2143   71.1396

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
L.chees   268    0.1321    0.1923                (---*----)
L.cheesm  268    0.0618    0.1479  (---*----)
L.escul   268    0.1888    0.1937                          (----*---)
L.hirs    268    0.0865    0.1687       (---*----)
L.hirsug  268    0.1300    0.1866                (---*---)
L.parvi   268    0.1153    0.1848             (---*---)
L.penne   268    0.0922    0.1696        (---*----)
L.pimpi   268    0.1102    0.1836             (---*---)
                                        --+---------+---------+---------+----
Pooled StDev =    0.1790              0.050     0.100     0.150     0.200
```

One-Way Analysis of Variance of Shannon's information index (*I*) in RAPD markers data of *Lycopersicon* spp.

```
Analysis of Variance
Source      DF        SS        MS        F        P
I            7    2.7190    0.3884    12.13    0.000
Error     2136   68.4206    0.0320
Total     2143   71.1396

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
L.chees   268    0.1321    0.1923                (---*----)
L.cheesm  268    0.0618    0.1479  (---*----)
L.escul   268    0.1888    0.1937                          (----*---)
L.hirs    268    0.0865    0.1687       (---*----)
L.hirsug  268    0.1300    0.1866                (---*---)
L.parvi   268    0.1153    0.1848             (---*---)
L.penne   268    0.0922    0.1696        (---*----)
L.pimpi   268    0.1102    0.1836             (---*---)
                                        --+---------+---------+---------+----
Pooled StDev =    0.1790              0.050     0.100     0.150     0.200
```

Part 5 Equation regression and analysis of variance of regression analysis done in section 3.3.2.1.4 Figure 3.3.10 between sample size and population genetic indices based in RAPD markers data.

Sample size versus observed number of alleles (*A*).

- L. cheesmanii

```
The regression equation is
A = 1.06 + 0.108 Sample size

Predictor       Coef       StDev          T        P
Constant     1.06052     0.02860      37.08    0.000
Sample s    0.108149    0.008971      12.06    0.000

S = 0.03468     R-Sq = 90.1%     R-Sq(adj) = 89.5%

Analysis of Variance

Source       DF          SS          MS          F        P
Regression    1     0.17479     0.17479     145.34    0.000
Error        16     0.01924     0.00120
Total        17     0.19403

Unusual Observations
Obs  Sample s          A        Fit  StDev Fit   Residual    St Resid
  5      4.00    1.56060    1.49312    0.01177    0.06748       2.07R

R denotes an observation with a large standardized residual
```

- L. hirsutum

```
The regression equation is
A = 0.456 + 0.276 Sample size

Predictor       Coef       StDev          T        P
Constant      0.4563      0.2614       1.75    0.106
Sample s     0.27561     0.08237       3.35    0.006

S = 0.2461      R-Sq = 48.3%     R-Sq(adj) = 44.0%

Analysis of Variance

Source       DF          SS          MS          F        P
Regression    1     0.67821     0.67821     11.20    0.006
Error        12     0.72692     0.06058
Total        13     1.40513

Unusual Observations
Obs  Sample s          A        Fit  StDev Fit   Residual    St Resid
 14      2.00     0.2623     1.0075     0.1101    -0.7452      -3.39R

R denotes an observation with a large standardized residual
```

- L. esculentum

The regression equation is
A = 1.09 + 0.0781 Sample size

Predictor	Coef	StDev	T	P
Constant	1.09319	0.07344	14.88	0.000
Sample s	0.07810	0.02289	3.41	0.004

S = 0.07153 R-Sq = 43.7% R-Sq(adj) = 39.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.059563	0.059563	11.64	0.004
Error	15	0.076747	0.005116		
Total	16	0.136310			

Sample size versus effective number of alleles (A_e).

- L. cheesmanii

The regression equation is
A_e = 1.07 + 0.0648 Sample size

Predictor	Coef	StDev	T	P
Constant	1.06776	0.02002	53.32	0.000
Sample s	0.064775	0.006280	10.31	0.000

S = 0.02428 R-Sq = 86.9% R-Sq(adj) = 86.1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.062703	0.062703	106.38	0.000
Error	16	0.009430	0.000589		
Total	17	0.072134			

- L. hirsutum

The regression equation is
A_e = 0.506 + 0.221 Sample size

Predictor	Coef	StDev	T	P
Constant	0.5056	0.2835	1.78	0.100
Sample s	0.22134	0.08934	2.48	0.029

S = 0.2670 R-Sq = 33.8% R-Sq(adj) = 28.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.43743	0.43743	6.14	0.029
Error	12	0.85516	0.07126		
Total	13	1.29258			

Unusual Observations

Obs	Sample s	Ae	Fit	StDev Fit	Residual	St Resid
12	2.00	0.1286	0.9483	0.1194	-0.8197	-3.43R

R denotes an observation with a large standardized residual

- *L. esculentum*

The regression equation is
A_e = 0.984 + 0.0579 Sample size

Predictor	Coef	StDev	T	P
Constant	0.9841	0.2675	3.68	0.002
Sample s	0.05790	0.08337	0.69	0.498

S = 0.2605 R-Sq = 3.1% R-Sq(adj) = 0.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.03273	0.03273	0.48	0.498
Error	15	1.01794	0.06786		
Total	16	1.05067			

Unusual Observations

Obs	Sample s	Ae	Fit	StDev Fit	Residual	St Resid
7	3.00	0.1955	1.1578	0.0639	-0.9623	-3.81R

R denotes an observation with a large standardized residual

Sample size versus average gene diversity (H_s)

- *L. cheesmanii*

The regression equation is
H_s = 0.110 + 0.0323 Sample size

Predictor	Coef	StDev	T	P
Constant	0.1098	0.2026	0.54	0.595
Sample s	0.03231	0.06354	0.51	0.618

S = 0.2457 R-Sq = 1.6% R-Sq(adj) = 0.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.01560	0.01560	0.26	0.618
Error	16	0.96552	0.06035		
Total	17	0.98112			

Unusual Observations

Obs	Sample s	Hs	Fit	StDev Fit	Residual	St Resid
7	3.00	1.1601	0.2067	0.0580	0.9534	3.99R

R denotes an observation with a large standardized residual

- *L. hirsutum*

The regression equation is
H_s = - 0.0555 + 0.0644 Sample size

Predictor	Coef	StDev	T	P
Constant	-0.05550	0.03166	-1.75	0.105
Sample s	0.064397	0.009976	6.46	0.000

S = 0.02981 R-Sq = 77.6% R-Sq(adj) = 75.8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.037027	0.037027	41.67	0.000
Error	12	0.010663	0.000889		
Total	13	0.047690			

Unusual Observations

Obs	Sample s	Hs	Fit	StDev Fit	Residual	St Resid
12	2.00	0.00753	0.07330	0.01333	-0.06577	-2.47R

R denotes an observation with a large standardized residual

- *L. esculentum*

The regression equation is
H_s = 0.0500 + 0.0256 Sample size

Predictor	Coef	StDev	T	P
Constant	0.05001	0.02789	1.79	0.093
Sample s	0.025643	0.008691	2.95	0.010

S = 0.02716 R-Sq = 36.7% R-Sq(adj) = 32.5%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.0064208	0.0064208	8.70	0.010
Error	15	0.0110642	0.0007376		
Total	16	0.0174850			

Sample size versus Shannon's information index (*I*)

- *L. cheesmanii*

The regression equation is
I = 0.0597 + 0.0541 Sample size

Predictor	Coef	StDev	T	P
Constant	0.05968	0.01650	3.62	0.002
Sample s	0.054073	0.005175	10.45	0.000

S = 0.02000 R-Sq = 87.2% R-Sq(adj) = 86.4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.043696	0.043696	109.20	0.000
Error	16	0.006402	0.000400		
Total	17	0.050099			

- *L. hirsutum*

The regression equation is
I = - 0.0405 + 0.0836 Sample size

Predictor	Coef	StDev	T	P
Constant	-0.04053	0.03335	-1.22	0.248
Sample s	0.08359	0.01051	7.96	0.000

S = 0.03140 R-Sq = 84.1% R-Sq(adj) = 82.7%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.062393	0.062393	63.28	0.000
Error	12	0.011831	0.000986		
Total	13	0.074224			

Unusual Observations

Obs	Sample s	I	Fit	StDev Fit	Residual	St Resid
7	3.00	0.27350	0.21025	0.00843	0.06325	2.09R

R denotes an observation with a large standardized residual

- *L. esculentum*

The regression equation is
I = 0.0707 + 0.0389 Sample size

Predictor	Coef	StDev	T	P
Constant	0.07074	0.04116	1.72	0.106
Sample s	0.03891	0.01283	3.03	0.008

S = 0.04009 R-Sq = 38.0% R-Sq(adj) = 33.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.014782	0.014782	9.20	0.008
Error	15	0.024108	0.001607		
Total	16	0.038890			

Sample size versus number of polymorphic loci (*P*).

- *L. cheesmanii*

The regression equation is
P = 0.0457 + 0.109 Sample size

Predictor	Coef	StDev	T	P
Constant	0.04565	0.02805	1.63	0.123
Sample s	0.109391	0.008798	12.43	0.000

S = 0.03401 R-Sq = 90.6% R-Sq(adj) = 90.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.17883	0.17883	154.60	0.000
Error	16	0.01851	0.00116		
Total	17	0.19734			

- *L. hirsutum*

```
The regression equation is
P = - 0.123 + 0.160 Sample size

Predictor       Coef       StDev          T        P
Constant    -0.12301     0.05670      -2.17    0.051
Sample s     0.15975     0.01787       8.94    0.000

S = 0.05339     R-Sq = 86.9%     R-Sq(adj) = 85.9%

Analysis of Variance

Source       DF          SS          MS          F        P
Regression    1     0.22786     0.22786      79.94    0.000
Error        12     0.03420     0.00285
Total        13     0.26207
```

- *L. esculentum*

```
The regression equation is
P = 0.0851 + 0.0776 Sample size

Predictor       Coef       StDev          T        P
Constant     0.08506     0.06721       1.27    0.225
Sample s     0.07760     0.02095       3.70    0.002

S = 0.06546     R-Sq = 47.8%     R-Sq(adj) = 44.3%

Analysis of Variance

Source       DF          SS          MS          F        P
Regression    1    0.058798    0.058798      13.72    0.002
Error        15    0.064275    0.004285
Total        16    0.123073
```

Part 6 Statistical analyses of genetic indices from RAPD markers data given in Tables 3.3.10. Non-parametric test Kruskal-Wallis was carried out for non-normally distributed indices A and A_e; ANOVA was carried out for predominantly normally distributed genetic indices H_s and I. Tested levels were groups of species in red-fruited and green-fruited.

Non-parametric Kruskal-Wallis Test for A in *Lycopersicon* spp fruit colour groups.

```
Kruskal-Wallis Test

A               N     Median     Ave Rank      Z
Red fruited   168      2.000        170.0   0.28
Green fruited 168      2.000        167.0  -0.28
Overall       336                   168.5

H = 0.08  DF = 1  P = 0.777
H = 0.22  DF = 1  P = 0.638 (adjusted for ties)
```

Non-parametric Kruskal-Wallis Test for A_e in *Lycopersicon* spp fruit colour groups.

```
Kruskal-Wallis Test

Ae              N      Median     Ave Ran        Z
Red fruited     167     1.385       162.5    -1.13
Green fruited   168     1.473       173.9     1.02
Overall         336                 168.5

H = 2.07  DF = 2  P = 0.355
H = 2.08  DF = 2  P = 0.354 (adjusted for ties)

* NOTE * One or more small samples
```

One-Way Analysis of Variance for H_s in *Lycopersicon* spp fruit colour groups.

```
Analysis of Variance
Source      DF        SS         MS         F        P
Hs           1    0.0340     0.0340      1.07    0.302
Error      334   10.6160     0.0318
Total      335   10.6500
                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level           N      Mean     StDev  ---+---------+---------+---------+---
Red fruited     168    0.2716   0.1777    (----------*---------)
Green fruited   168    0.2918   0.1789              (----------*----------)
                                       ---+---------+---------+---------+---
Pooled StDev =        0.1783              0.250     0.275     0.300     0.325
```

One-Way Analysis of Variance for I in *Lycopersicon* spp fruit colour groups.

```
Analysis of Variance
Source      DF        SS         MS         F        P
I            1    0.0512     0.0512      0.89    0.347
Error      334   19.2777     0.0577
Total      335   19.3289
                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level           N      Mean     StDev  -----+---------+---------+---------+-
Red fruited     168    0.4131   0.2396    (-----------*-----------)
Green fruited   168    0.4378   0.2409              (-----------*-----------)
                                       -----+---------+---------+---------+-
Pooled StDev =        0.24                0.390     0.420     0.450     0.480
```

Part 7 Statistical analyses of genetic indices from microsatellite and RAPD markers data given in Table 3.4.1. Non-parametric test Kruskal-Wallis was carried out for non-normally distributed indices *A* and A_e; ANOVA was carried out for predominantly normally distributed genetic indices H_s and *I*. Tested levels were *L. esculentum* accessions grouped in landraces, old varieties, modern varieties OP, and F_1 hybrids.

Non-parametric Kruskal-Wallis Test for *A* in *L. esculentum* groups for RAPD.

```
Kruskal-Wallis Test

A                     N      Median   Ave Rank      Z
Landrace            168      1.000      291.5  -3.47
Old Varieties.      168      1.000      383.5   3.62
Modern Varieties OP 168      1.000      361.5   1.93
F1 hybrids          168      1.000      309.5  -2.08
Overall             672                 336.5

H = 24.91  DF = 3  P = 0.000
H = 38.26  DF = 3  P = 0.000 (adjusted for ties)
```

Non-parametric Kruskal-Wallis Test for A_e in *L. esculentum* groups for RAPD.

```
Kruskal-Wallis Test

Ae                    N      Median   Ave Rank      Z
Landrace            168      1.000      292.3  -3.41
Old Varieties       168      1.000      376.1   3.05
Modern varieties OP 168      1.000      368.5   2.47
F1 hybrids          168      1.000      309.2  -2.11
Overall             672                 336.5

H = 23.60  DF = 3  P = 0.000
H = 34.35  DF = 3  P = 0.000 (adjusted for ties)
```

One-Way Analysis of Variance for H_s in *L. esculentum* groups for RAPD.

```
Analysis of Variance
Source     DF        SS        MS        F        P
Hs          3    0.7632    0.2544     7.82    0.000
Error     668   21.7434    0.0326
Total     671   22.5066

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev  --------+---------+---------+--------
Landrace   168    0.0764    0.1612   (------*------)
Old Var.   168    0.1517    0.1919                    (------*------)
Modern OP  168    0.1484    0.1975                   (------*------)
F1 hybrids 168    0.0904    0.1685      (------*-----)
                                    --------+---------+---------+--------
Pooled StDev =    0.1804                 0.080     0.120     0.160
```

One-Way Analysis of Variance for *I* in *L. esculentum* groups for RAPD.

```
Analysis of Variance
Source     DF        SS        MS        F        P
I           3    1.7863    0.5954     8.73    0.000
Error     668   45.5818    0.0682
Total     671   47.3681

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  --------+---------+---------+--------
Landrace   168    0.1116    0.2353   (------*-----)
Old Var.   168    0.2294    0.2756                      (-----*------)
Modern OP  168    0.2197    0.2850                    (------*-----)
F1 hybrids 168    0.1342    0.2458        (-----*------)
                                      --------+---------+---------+-------
Pooled StDev =    0.2612                 0.120     0.180     0.240
```

Part 8 ANOVA of means from continuous characters in fruits of *Lycopersicon* spp accessions selected as parents showed in Table 4.3.1 and 4.3.2.

Fruit diameter (cm)

One-Way Analysis of Variance

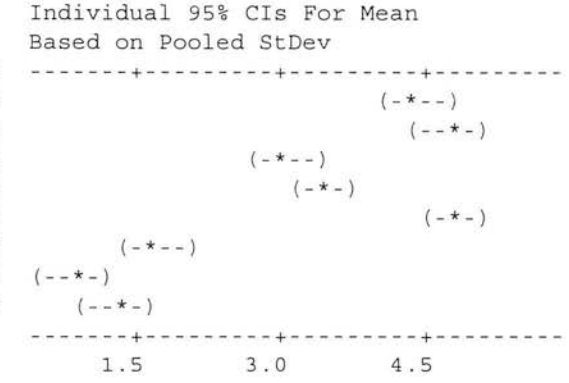
```
Analysis of Variance
Source     DF        SS        MS        F        P
Parents     7   299.476    42.782   117.09    0.000
Error      72    26.307     0.365
Total      79   325.783

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  -------+---------+---------+---------
Limachino   10    5.8630    0.6689                            (-*-)
Ace         10    6.7580    1.4093                                 (-*-)
Lukullus    10    3.6470    0.3576                 (-*-)
Marglobe    10    3.8130    0.4489                  (-*-)
San Marzano10     2.8230    0.3206              (-*-)
L.esc.ceras10     1.7970    0.1571        (-*-)
L.parviflo  10    1.0290    0.0472   (-*-)
L.pimpinel  10    1.4380    0.1740     (-*-)
                                      -------+---------+---------+---------
Pooled StDev =    0.6045                 2.0       4.0       6.0
```

Fruit length (cm)

One-Way Analysis of Variance

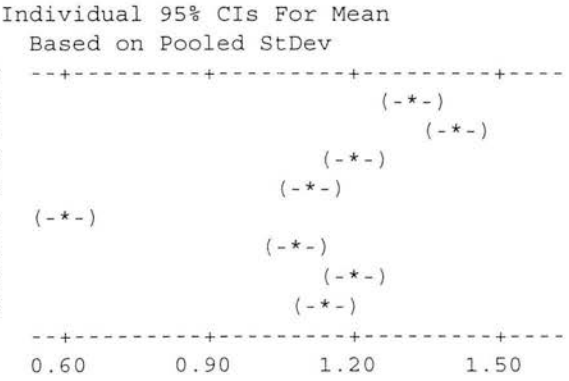
```
Analysis of Variance
Source      DF        SS        MS        F        P
Parents      7   177.690    25.384    77.95    0.000
Error       72    23.446     0.326
Total       79   201.136
```



Fruit ratio (d/l)

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Parents      7   3.98852   0.56979    60.22    0.000
Error       72   0.68120   0.00946
Total       79   4.66972
```



Fruit weight (g)

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Parents      7    145967     20852    26.21    0.000
Error       72     57283       796
Total       79    203249

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N    Mean StDev  ----+---------+---------+---------+--
Limachino   10   86.21 30.77                    (--*---)
Ace         10  127.07 72.09                             (--*---)
Lukullus    10   24.89  6.26         (---*---)
Marglobe    10   30.42 10.94           (--*---)
San Marzano10    22.79  7.75         (---*--)
L.esc.ceras10     3.46  1.05    (---*--)
L.parviflo 10     0.68  0.09    (--*---)
L.pimpinel 10     1.98  0.67    (--*---)
                              ----+---------+---------+---------+--
Pooled StDev =    28.21           0        50       100       150
```

Solid soluble content (°brix)

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Parents      7   256.300    36.614    50.83    0.000
Error       72    51.860     0.720
Total       79   308.160

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  --------+---------+---------+--------
Limachino   10     5.100     0.408   (-*--)
Ace         10     5.770     1.442      (--*--)
Lukullus    10     6.680     0.755           (-*--)
Marglobe    10     5.230     0.849   (--*--)
San Marzano10      6.200     0.510        (--*--)
L.esc.ceras10      7.090     0.761             (-*--)
L.parviflo 10     10.760     0.986                                (--*-)
L.pimpinel 10      8.590     0.644                     (--*--)
                                    --------+---------+---------+--------
Pooled StDev =     0.849             6.0       8.0       10.0
```

Weight of 1,000 seeds (g)

One-Way Analysis of Variance

```
Analysis of Variance
Source     DF        SS        MS        F        P
Parents     7  25.12230   3.58890   490.86    0.000
Error      24   0.17548   0.00731
Total      31  25.29777

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev  ----------+---------+---------+------
Limachino    4    2.9700    0.0476                                   (*-)
Ace          4    3.0000    0.0924                                    (*)
Lukullus     4    3.1750    0.1676                                      (*-)
Marglobe     4    2.4500    0.0622                          (*)
San Marzano  4    1.5900    0.0383              (-*)
L.esc.ceras  4    1.2625    0.0960         (*)
L.parviflo   4    1.1300    0.0258       (*)
L.pimpinel   4    0.8550    0.0661   (*)
                                    ----------+---------+---------+------
Pooled StDev =    0.0855                  1.40      2.10      2.80
```

Part 9.1 ANOVA of means in crosses and parent accessions of continuous characters in F_1 generation of inter- and intra-taxon crosses presented in Tables 4.7.1 and 4.7.2.

Fruit diameter (cm)

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3   17.1319    5.7106    74.65    0.000
Error       36    2.7540    0.0765
Total       39   19.8858

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev  ----+---------+---------+---------+--
I 3919      10    2.6690    0.2020                (-*--)
I 1939      10    2.6950    0.3356                (--*--)
L.esc.cer.  10    1.7970    0.1571   (--*--)
Lukullus    10    3.6470    0.3576                                  (--*--)
                                    ----+---------+---------+---------+--
Pooled StDev =    0.2766              1.80      2.40      3.00      3.60
```

One-Way Analysis of Variance

```
Analysis of Variance for C4
Source      DF        SS        MS        F        P
Accessions   3   146.843    48.948   372.14    0.000
Error       36     4.735     0.132
Total       39   151.578

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev -----+---------+---------+---------+-
I 5211      10    1.5530    0.2234     (*-)
I 1152      10    1.9120    0.1628       (-*)
L.parvif    10    1.0290    0.0472 (-*)
Limachino   10    5.8630    0.6689                                  (*-)
                                    -----+---------+---------+---------+-
Pooled StDev =    0.3627              1.5       3.0       4.5       6.0
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   2    29.126    14.563   131.31    0.000
Error       27     2.994     0.111
Total       29    32.120

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev -----+---------+---------+---------+-
I 6021      10    2.2430    0.3194          (--*--)
L.pimpin    10    1.4380    0.1740 (--*--)
Marglobe    10    3.8110    0.4477                                (--*-)
                                    -----+---------+---------+---------+-
Pooled StDev =    0.3330              1.60      2.40      3.20      4.00
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   2   152.865    76.433   112.81    0.000
Error       27    18.293     0.678
Total       29   171.159

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev ------+---------+---------+---------+
I 1560      10    2.7930    0.1270       (--*--)
L.pimpin    10    1.4380    0.1740 (-*--)
Ace         10    6.7580    1.4093                             (--*-)
                                    ------+---------+---------+---------+
Pooled StDev =    0.8231              2.0       4.0       6.0       8.0
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3     7.210     2.403    16.45    0.000
Error       36     5.258     0.146
Total       39    12.469

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev  ---------+---------+---------+-------
E 1922      10    3.9660    0.2855                          (----*----)
E 2219      10    3.2950    0.5216             (----*----)
Lukullus    10    3.6470    0.3576                    (----*----)
San Marzano10    2.8230    0.3206    (---*----)
                                    ---------+---------+---------+-------
Pooled StDev =    0.3822                3.00      3.50      4.00
```

Fruit length (cm)

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3     9.9438    3.3146    44.28    0.000
Error       36     2.6951    0.0749
Total       39    12.6389

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev  -+---------+---------+---------+-----
I 3919      10    2.4770    0.1808                    (---*--)
I 1939      10    2.4840    0.3317                    (---*--)
L.esc.cer.  10    1.6720    0.1536   (--*---)
Lukullus    10    3.0740    0.3649                                (--*---)
                                    -+---------+---------+---------+-----
Pooled StDev =    0.2736          1.50      2.00      2.50      3.00
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3    77.1775   25.7258   363.94    0.000
Error       36     2.5447    0.0707
Total       39    79.7222

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev  -----+---------+---------+---------+-
I 5211      10    1.3000    0.2320     (-*)
I 1152      10    1.6790    0.1738        (*)
L.parvif    10    0.8680    0.0494   (*-)
Limachino   10    4.4210    0.4430                                 (-*)
                                    -----+---------+---------+---------+-
Pooled StDev =    0.2659             1.2       2.4       3.6       4.8
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   2    23.798    11.899    87.51    0.000
Error       27     3.671     0.136
Total       29    27.470

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev -------+---------+---------+---------
I 6021      10    2.1280    0.3266           (--*--)
L.pimpin    10    1.2860    0.2147 (--*--)
Marglobe    10    3.4500    0.5052                          (--*--)
                                    -------+---------+---------+---------
Pooled StDev =    0.3688          1.60      2.40      3.20
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   2    62.133    31.067   126.17    0.000
Error       27     6.648     0.246
Total       29    68.781

                                Individual 95% CIs For Mean
                                Based on Pooled StDev
Level        N      Mean     StDev --+---------+---------+---------+----
I 1560      10    2.3840    0.1645          (--*--)
L.pimpin    10    1.2860    0.2147 (--*-)
Ace         10    4.7360    0.8158                            (-*--)
                                --+---------+---------+---------+----
Pooled StDev =    0.4962          1.2       2.4       3.6       4.8
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3    17.269     5.756    14.49    0.000
Error       36    14.302     0.397
Total       39    31.571

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev --+---------+---------+---------+----
E 1922      10    3.9020    0.1621             (-----*-----)
E 2219      10    3.4010    0.3832       (-----*-----)
Lukullus    10    3.0740    0.3649 (-----*-----)
San Marzano10     4.8150    1.1326                          (-----*-----)
                                    --+---------+---------+---------+----
Pooled StDev =    0.6303          2.80      3.50      4.20      4.90
```

Fruit Ratio D/L

One-Way Analysis of Variance

```
Analysis of Variance

Source      DF        SS        MS        F        P
Accessions   3   0.10302   0.03434     4.04    0.014
Error       36   0.30609   0.00850
Total       39   0.40911

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  -----+---------+---------+---------+-
I 3919      10    1.0778    0.0345   (-------*-------)
I 1939      10    1.0880    0.0773     (-------*--------)
L.esc.cer.  10    1.0756    0.0353  (--------*-------)
Lukullus    10    1.1972    0.1600                    (-------*-------)
                                     -----+---------+---------+---------+-
Pooled StDev =    0.0922               1.050     1.120     1.190     1.260
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3   0.18569   0.06190    10.20    0.000
Error       36   0.21847   0.00607
Total       39   0.40417

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  ----+---------+---------+---------+--
I 5211      10    1.2016    0.0727          (-----*-----)
I 1152      10    1.1441    0.0971   (-----*-----)
L.parvif    10    1.1869    0.0456        (-----*------)
Limachino   10    1.3271    0.0865                          (-----*-----)
                                     ----+---------+---------+---------+--
Pooled StDev =    0.0779               1.120     1.200     1.280     1.360
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   2   0.02762   0.01381     3.87    0.033
Error       27   0.09642   0.00357
Total       29   0.12404

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  -------+---------+---------+---------
I 6021      10    1.0563    0.0365  (-------*-------)
L.pimpin    10    1.1274    0.0718                (-------*-------)
Marglobe    10    1.1104    0.0650             (-------*-------)
                                     -------+---------+---------+---------
Pooled StDev =    0.0598               1.050     1.100     1.150
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   2   0.47891   0.23946    40.72    0.000
Error       27   0.15877   0.00588
Total       29   0.63769

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev  -+---------+---------+---------+-----
I 1560      10    1.1743    0.0587       (---*---)
L.pimpin    10    1.1274    0.0718    (---*---)
Ace         10    1.4158    0.0951                          (---*---)
                                    -+---------+---------+---------+-----
Pooled StDev =    0.0767          1.08      1.20      1.32      1.44
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3    1.7945    0.5982    40.87    0.000
Error       36    0.5269    0.0146
Total       39    2.3214

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev  ----+---------+---------+---------+--
E 1922      10    1.0174    0.0757                      (---*---)
E 2219      10    0.9672    0.0797                   (---*---)
Lukullus    10    1.1972    0.1600                            (---*---)
San Marzano10     0.6131    0.1444    (---*---)
                                    ----+---------+---------+---------+--
Pooled StDev =    0.1210              0.60      0.80      1.00      1.20
```

Fruits Weight (g)

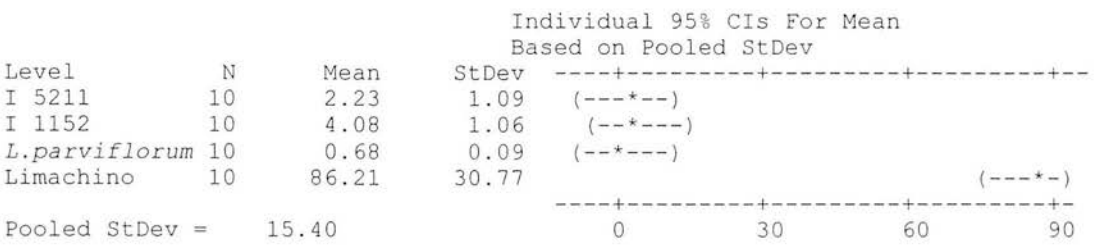
One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3    2369.5     789.8    50.78    0.000
Error       36     559.9      15.6
Total       39    2929.4

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level           N      Mean     StDev  ---------+---------+---------+-------
I 3919         10    11.225     2.189            (--*--)
I 1939         10    11.819     4.135             (--*--)
L.escul.cer.   10     3.459     1.051   (--*--)
Lukullus       10    24.894     6.262                              (--*--)
                                    ---------+---------+---------+-------
Pooled StDev =     3.944                    8.0      16.0      24.0
```

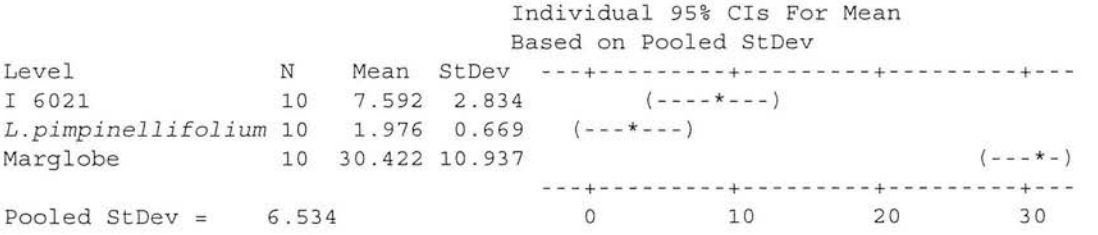
One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3     52828     17609    74.22    0.000
Error       36      8541       237
Total       39     61369
```



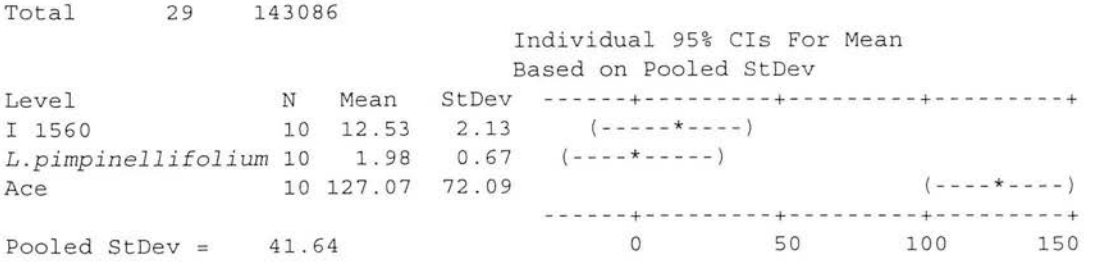
One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   2    4539.7    2269.9    53.16    0.000
Error       27    1152.8      42.7
Total       29    5692.5
```



One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   2     96263     48131    27.75    0.000
Error       27     46823      1734
Total       29    143086
```



One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3     774.2     258.1     4.87    0.006
Error       36    1906.9      53.0
Total       39    2681.1

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level           N    Mean    StDev  ----------+---------+---------+------
E 1922         10  33.830    4.008                      (------*-------)
E 2219         10  23.769    9.822     (-------*------)
Lukullus       10  24.894    6.262       (------*-------)
San Marzano    10  22.788    7.754  (-------*-------)
                                     ----------+---------+---------+------
Pooled StDev =    7.278              24.0      30.0      36.0
```

Solid Soluble Content (°Brix)

One-Way Analysis of Variance

```
Analysis of Variance

Source      DF        SS        MS        F        P
Accessions   3     27.16      9.05     4.57    0.008
Error       36     71.29      1.98
Total       39     98.46

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level        N      Mean     StDev  --+---------+---------+---------+----
I 3919      10     6.860     1.391    (------*-------)
I 1939      10     8.750     2.200                (-------*------)
L.esc.cer.  10     7.090     0.761      (------*-------)
Lukullus    10     6.680     0.755  (-------*------)
                                     --+---------+---------+---------+----
Pooled StDev =    1.407             6.0       7.2       8.4       9.6
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3   175.789    58.596   130.05    0.000
Error       36    16.221     0.451
Total       39   192.010

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level        N      Mean     StDev  -------+---------+---------+---------
I 5211      10     7.680     0.476              (-*--)
I 1152      10     9.270     0.662                      (-*--)
L.parvif    10    10.760     0.986                              (-*-)
Limachino   10     5.100     0.408  (-*--)
                                     -------+---------+---------+---------
Pooled StDev =    0.671            6.0       8.0      10.0
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   2     66.86     33.43    15.87    0.000
Error       27     56.87      2.11
Total       29    123.74

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev --+---------+---------+---------+----
I 6021      10     8.160     2.277                    (-----*------)
L.pimpin    10     8.590     0.644                       (-----*------)
Marglobe    10     5.230     0.849  (-----*-----)
                                    --+---------+---------+---------+----
Pooled StDev =     1.451           4.5       6.0       7.5       9.0
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   2     64.59     32.30    31.15    0.000
Error       27     28.00      1.04
Total       29     92.59

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev ----------+---------+---------+------
I 1560      10     5.250     0.785  (---*---)
L.pimpin    10     8.590     0.644                       (---*----)
Ace         10     5.770     1.442     (---*----)
                                    ----------+---------+---------+------
Pooled StDev =     1.018             6.0       7.5       9.0
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3      4.70      1.57     1.50    0.232
Error       36     37.69      1.05
Total       39     42.40

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev --------+---------+---------+--------
E 1922      10     6.190     1.706        (--------*---------)
E 2219      10     5.710     0.669  (---------*--------)
Lukullus    10     6.680     0.755                (--------*---------)
San Marzano10      6.200     0.510        (---------*--------)
                                    --------+---------+---------+--------
Pooled StDev =     1.023            5.60      6.30      7.00
```

Weight of 1000 seeds (g)

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3    7.6527    2.5509   200.79    0.000
Error       12    0.1524    0.0127
Total       15    7.8052
```

```
                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev --+---------+---------+---------+----
I 3919       4    2.4975    0.0974                      (-*-)
I 1939       4    2.5200    0.0632                      (-*-)
L.esc.cer.   4    1.2625    0.0960  (-*-)
Lukullus     4    3.1750    0.1676                                  (-*-)
                                    --+---------+---------+---------+----
Pooled StDev =    0.1127            1.20      1.80      2.40      3.00
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3   7.94187   2.64729   319.27    0.000
Error       12   0.09950   0.00829
Total       15   8.04137
```

```
                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev ---+---------+---------+---------+---
I 5211       4    1.6600    0.1657          (-*)
I 1152       4    2.4250    0.0526                      (*-)
L.parvif     4    1.1300    0.0258 (-*)
Limachino    4    2.9700    0.0476                                (*-)
                                    ---+---------+---------+---------+---
Pooled StDev =    0.0911             1.20      1.80      2.40      3.00
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   2   5.45807   2.72903   312.48    0.000
Error        9   0.07860   0.00873
Total       11   5.53667
```

```
                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev ------+---------+---------+---------+
I 6021       4    2.0250    0.1340                        (-*--)
L.pimpin     4    0.8550    0.0661  (-*-)
Marglobe     4    2.4500    0.0622                                  (-*-)
                                    ------+---------+---------+---------+
Pooled StDev =    0.0935              1.00      1.50      2.00      2.50
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
C7           2    9.7361    4.8680   461.18    0.000
Error        9    0.0950    0.0106
Total       11    9.8311

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  ----------+---------+---------+------
I 1560       4    2.3750    0.1370                      (-*-)
L.pimpin     4    0.8550    0.0661   (*-)
Ace          4    3.0000    0.0924                                (-*-)
                                      ----------+---------+---------+------
Pooled StDev =    0.1027              1.40      2.10      2.80
```

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
Accessions   3    7.7141    2.5714   150.52    0.000
Error       12    0.2050    0.0171
Total       15    7.9191

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  ------+---------+---------+---------+
E 1922       4    3.2050    0.1578                          (-*--)
E 2219       4    3.2000    0.1178                          (-*--)
Lukullus     4    3.1750    0.1676                          (-*-)
San Marzano  4    1.5900    0.0383  (--*-)
                                      ------+---------+---------+---------+
Pooled StDev =    0.1307            1.80      2.40      3.00      3.60
```

Part 9.2 ANOVA of means in crosses and parent accessions of continuous characters in F_1 generation of inter- and intra-taxon crosses presented in Appendix 5.

Fruit Diameter (cm)

One-Way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS        F        P
C1           3    96.781    32.260   158.45    0.000
Error       36     7.330     0.204
Total       39   104.111

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  ----------+---------+---------+------
I 3911      10    2.7340    0.3093        (-*-)
I 1139      10    2.5620    0.4965       (-*-)
Ceras       10    1.7970    0.1571  (-*-)
Limachi     10    5.8630    0.6689                          (-*-)
                                      ----------+---------+---------+------
Pooled StDev =    0.4512              3.0       4.5       6.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C4
Source     DF        SS        MS        F        P
C3          3   153.450    51.150    85.26    0.000
Error      36    21.598     0.600
Total      39   175.048

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
I 3915     10    3.2260    0.4612          (-*--)
I 1539     10    2.1750    0.4198    (--*-)
Ceras      10    1.7970    0.1571   (-*-)
Ace        10    6.7580    1.4093                          (--*-)
                                    ----+---------+---------+---------+--
Pooled StDev =   0.7746              2.0       4.0       6.0       8.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C6
Source     DF        SS        MS        F        P
C5          3   35.9603   11.9868   283.28    0.000
Error      36    1.5233    0.0423
Total      39   37.4836

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---------+---------+---------+-------
I 5219     10    1.9780    0.1596            (-*)
I 1952     10    1.8860    0.1169           (-*)
Parvifl    10    1.0290    0.0472  (-*-)
Lukullus   10    3.6470    0.3576                                   (-*)
                                    ---------+---------+---------+-------
Pooled StDev =   0.2057            1.60      2.40      3.20
```

One-Way Analysis of Variance

```
Analysis of Variance for C8
Source     DF        SS        MS        F        P
C7          3   16.6762    5.5587   118.86    0.000
Error      36    1.6836    0.0468
Total      39   18.3598

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
I 5222     10    1.6780    0.1749            (-*-)
I 2252     10    1.6910    0.2269            (-*-)
Parvifl    10    1.0290    0.0472  (-*-)
SanMarz    10    2.8230    0.3206                                (-*-)
                                    ------+---------+---------+---------+
Pooled StDev =   0.2163            1.20      1.80      2.40      3.00
```

One-Way Analysis of Variance

```
Analysis of Variance for C10
Source     DF        SS        MS        F        P
C9          2   105.962    52.981   300.52    0.000
Error      27     4.760     0.176
Total      29   110.723

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---+---------+---------+---------+---
I 1160     10    2.5510    0.2261        (-*-)
Limachi    10    5.8630    0.6689                                (-*-)
Pimpinel   10    1.4380    0.1740  (-*)
                                    ---+---------+---------+---------+---
Pooled StDev =   0.4199            1.5       3.0       4.5       6.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C12
Source     DF        SS        MS        F        P
C11         2   25.6339   12.8170   158.20    0.000
Error      27    2.1875    0.0810
Total      29   27.8215

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---+---------+---------+---------+---
I 1960     10    2.1120    0.2914            (-*--)
Lukullus   10    3.6470    0.3576                                  (--*--)
Pimpinel   10    1.4380    0.1740   (--*-)
                                    ---+---------+---------+---------+---
Pooled StDev =   0.2846             1.40      2.10      2.80      3.50
```

One-Way Analysis of Variance

```
Analysis of Variance for C14
Source     DF        SS        MS        F        P
C13         2     7.952     3.976    37.55    0.000
Error      27     2.859     0.106
Total      29    10.810

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---------+---------+---------+-------
I 2239     10    2.9450    0.4361                          (---*---)
SanMarz    10    2.8230    0.3206                      (---*----)
Ceras      10    1.7970    0.1571   (---*---)
                                    ---------+---------+---------+-------
Pooled StDev =   0.3254             2.00      2.50      3.00
```

One-Way Analysis of Variance

```
Analysis of Variance for C16
Source     DF        SS        MS        F        P
C15         2     9.6189    4.8094    54.75    0.000
Error      27     2.3716    0.0878
Total      29    11.9905

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
I 2260     10    2.0660    0.3612              (---*---)
SanMarz    10    2.8230    0.3206                                (--*---)
Pimpinel   10    1.4380    0.1740   (---*---)
                                    ------+---------+---------+---------+
Pooled StDev =   0.2964            1.50      2.00      2.50      3.00
```

One-Way Analysis of Variance

```
Analysis of Variance for C18
Source     DF        SS        MS        F        P
C17         3    27.962     9.321     9.50    0.000
Error      31    30.407     0.981
Total      34    58.368

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
E 1115     10    4.4410    0.4616  (------*------)
E 1511      5    5.2960    1.2832        (--------*--------)
Limachi    10    5.8630    0.6689                 (------*------)
Ace        10    6.7580    1.4093                          (------*------)
                                    --+---------+---------+---------+----
Pooled StDev =   0.9904            4.0       5.0       6.0       7.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C20
Source     DF        SS        MS        F        P
C19         3    34.797    11.599    33.83    0.000
Error      36    12.343     0.343
Total      39    47.140

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --------+---------+---------+--------
E 1119     10    3.9520    0.4986      (---*--)
E 1911     10    3.5970    0.7399  (---*---)
Limachi    10    5.8630    0.6689                          (---*--)
Lukullus   10    3.6470    0.3576   (--*---)
                                    --------+---------+---------+--------
Pooled StDev =    0.5855            4.0       5.0       6.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C22
Source     DF        SS        MS        F        P
C21         3    40.532    13.511    47.87    0.000
Error      36    10.160     0.282
Total      39    50.691

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
E 1121     10    3.7690    0.5478        (---*--)
E 2111     10    3.2220    0.4240   (--*---)
Limachi    10    5.8630    0.6689                              (---*--)
Marglobe   10    3.8130    0.4489         (--*---)
                                    --+---------+---------+---------+----
Pooled StDev =    0.5312            3.0       4.0       5.0       6.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C24
Source     DF        SS        MS        F        P
C23         3    49.518    16.506    23.91    0.000
Error      36    24.852     0.690
Total      39    74.369

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  -----+---------+---------+---------+-
E 1519     10    5.0100    0.7497             (----*---)
E 1915     10    4.7960    0.2918           (---*---)
Ace        10    6.7580    1.4093                              (---*----)
Lukullus   10    3.6470    0.3576   (---*----)
                                    -----+---------+---------+---------+-
Pooled StDev =    0.8309              3.6       4.8       6.0       7.2
```

One-Way Analysis of Variance

```
Analysis of Variance for C26
Source     DF        SS        MS        F        P
C25         3    82.678    27.559    33.22    0.000
Error      36    29.866     0.830
Total      39   112.544

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
E 1522     10    4.3440    0.5704           (---*---)
E 2215     10    3.9180    0.9508         (---*---)
Ace        10    6.7580    1.4093                              (---*---)
SanMarz    10    2.8230    0.3206   (---*---)
                                    ------+---------+---------+---------+
Pooled StDev =    0.9108              3.0       4.5       6.0       7.5
```

One-Way Analysis of Variance

```
Analysis of Variance for C28
Source     DF        SS        MS        F        P
C27         3     7.210     2.403    16.45    0.000
Error      36     5.258     0.146
Total      39    12.469

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---------+---------+---------+-------
E 1922     10    3.9660    0.2855                          (----*----)
E 2219     10    3.2950    0.5216             (----*----)
Lukullus   10    3.6470    0.3576                    (----*----)
SanMarz    10    2.8230    0.3206   (---*----)
                                    ---------+---------+---------+-------
Pooled StDev =    0.3822                  3.00      3.50      4.00
```

One-Way Analysis of Variance

```
Analysis of Variance for C30
Source     DF        SS        MS        F        P
C29         2    50.14     25.07    24.64    0.000
Error      26    26.45      1.02
Total      28    76.59

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
E 1521      9    4.241     0.919      (----*-----)
Ace        10    6.758     1.409                            (----*-----)
Marglobe   10    3.813     0.449  (-----*----)
                                    ----+---------+---------+---------+--
Pooled StDev =    1.009             3.6       4.8       6.0       7.2
```

One-Way Analysis of Variance

```
Analysis of Variance for C32
Source     DF        SS        MS        F        P
C31         2    49.761    24.880    85.15    0.000
Error      27     7.890     0.292
Total      29    57.650

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----------+---------+---------+------
E 2211     10    3.6130    0.5713        (--*--)
SanMarz    10    2.8230    0.3206   (--*-)
Limachi    10    5.8630    0.6689                              (--*--)
                                    ----------+---------+---------+------
Pooled StDev =    0.5406            3.6       4.8       6.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C34
Source     DF        SS        MS        F        P
C33         2     5.339     2.669    14.39    0.000
Error      25     4.639     0.186
Total      27     9.978

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----------+---------+---------+------
E 2221      8    3.5950    0.5211              (-----*-----)
SanMarz    10    2.8230    0.3206   (----*-----)
Marglobe   10    3.8130    0.4489                       (----*-----)
                                    ----------+---------+---------+------
Pooled StDev =    0.4307            3.00      3.50      4.00
```

Fruit Length (cm)

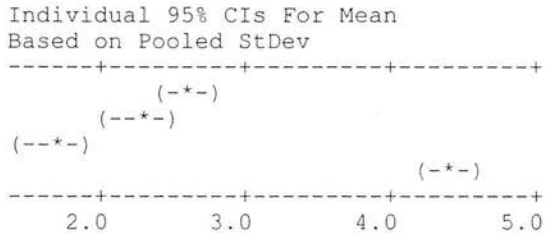
One-Way Analysis of Variance

Analysis of Variance for C2

Source	DF	SS	MS	F	P
C1	3	42.260	14.087	114.36	0.000
Error	36	4.434	0.123		
Total	39	46.695			

Level	N	Mean	StDev
I 3911	10	2.5830	0.3581
I 1139	10	2.2550	0.3802
Ceras	10	1.6720	0.1536
Limachi	10	4.4210	0.4430

Pooled StDev = 0.3510



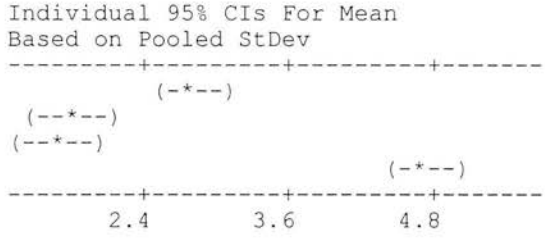
One-Way Analysis of Variance

Analysis of Variance for C4

Source	DF	SS	MS	F	P
C3	3	60.753	20.251	71.91	0.000
Error	36	10.139	0.282		
Total	39	70.891			

Level	N	Mean	StDev
I 3915	10	2.8030	0.3557
I 1539	10	1.7670	0.5576
Ceras	10	1.6720	0.1536
Ace	10	4.7360	0.8158

Pooled StDev = 0.5307



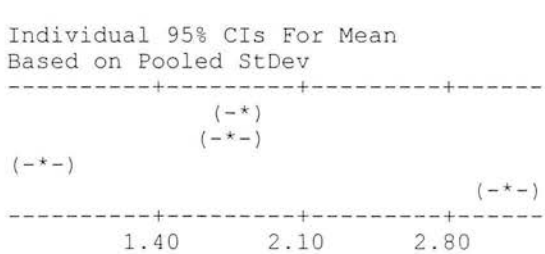
One-Way Analysis of Variance

Analysis of Variance for C6

Source	DF	SS	MS	F	P
C5	3	24.8031	8.2677	183.72	0.000
Error	36	1.6200	0.0450		
Total	39	26.4232			

Level	N	Mean	StDev
I 5219	10	1.7850	0.1802
I 1952	10	1.7300	0.1093
Parvifl	10	0.8680	0.0494
Lukullus	10	3.0740	0.3649

Pooled StDev = 0.2121



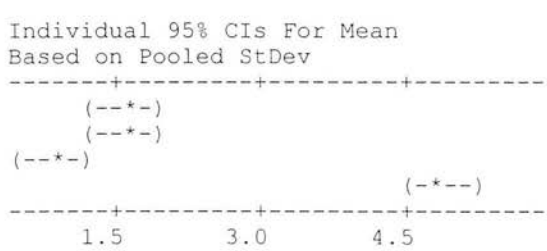
One-Way Analysis of Variance

Analysis of Variance for C8

Source	DF	SS	MS	F	P
C7	3	93.190	31.063	91.62	0.000
Error	36	12.206	0.339		
Total	39	105.396			

Level	N	Mean	StDev
I 5222	10	1.5870	0.1704
I 2252	10	1.6230	0.2047
Parvifl	10	0.8680	0.0494
SanMarz	10	4.8150	1.1326

Pooled StDev = 0.5823



One-Way Analysis of Variance

```
Analysis of Variance for C10
Source     DF        SS        MS        F        P
C9          2   51.2355   25.6178   289.15    0.000
Error      27    2.3921    0.0886
Total      29   53.6277

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----------+---------+---------+------
I 1160     10    2.2930    0.1530            (-*-)
Limachi    10    4.4210    0.4430                                  (-*-)
Pimpinel   10    1.2860    0.2147   (-*-)
                                    ----------+---------+---------+------
Pooled StDev =   0.2977                      2.0       3.0       4.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C12
Source     DF        SS        MS        F        P
C11         2   16.194     8.097    80.71    0.000
Error      27    2.708     0.100
Total      29   18.902

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
I 1960     10    2.0030    0.3488            (--*---)
Lukullus   10    3.0740    0.3649                                (--*---)
Pimpinel   10    1.2860    0.2147  (--*---)
                                    --+---------+---------+---------+----
Pooled StDev =   0.3167             1.20      1.80      2.40      3.00
```

One-Way Analysis of Variance

```
Analysis of Variance for C14
Source     DF        SS        MS        F        P
C13         2   49.817    24.909    54.54    0.000
Error      27   12.331     0.457
Total      29   62.149

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----------+---------+---------+------
I 2239     10    2.9910    0.2523            (---*---)
SanMarz    10    4.8150    1.1326                              (---*---)
Ceras      10    1.6720    0.1536  (---*---)
                                    ----------+---------+---------+------
Pooled StDev =   0.6758                      2.4       3.6       4.8
```

One-Way Analysis of Variance

```
Analysis of Variance for C16
Source     DF        SS        MS        F        P
C15         2   66.310    33.155    65.50    0.000
Error      27   13.667     0.506
Total      29   79.977

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  -----+---------+---------+---------+-
I 2260     10    2.2720    0.4354         (--*--)
SanMarz    10    4.8150    1.1326                          (--*--)
Pimpinel   10    1.2860    0.2147  (---*--)
                                    -----+---------+---------+---------+-
Pooled StDev =   0.7115               1.5       3.0       4.5       6.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C18
Source     DF        SS        MS        F        P
C17         3     7.798     2.599     5.71    0.003
Error      31    14.103     0.455
Total      34    21.901

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---------+---------+---------+-------
E 1115     10    3.5520    0.6556   (------*------)
E 1511      5    3.9820    0.7871       (---------*----------)
Limachi    10    4.4210    0.4430                 (-------*------)
Ace        10    4.7360    0.8158                        (------*------)
                                    ---------+---------+---------+-------
Pooled StDev =    0.6745                  3.60      4.20      4.80
```

One-Way Analysis of Variance

```
Analysis of Variance for C20
Source     DF        SS        MS        F        P
C19         3    10.560     3.520    14.92    0.000
Error      36     8.492     0.236
Total      39    19.052

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
E 1119     10    3.5220    0.3387          (----*----)
E 1911     10    3.2810    0.7067     (-----*----)
Limachi    10    4.4210    0.4430                          (-----*----)
Lukullus   10    3.0740    0.3649  (----*----)
                                    ----+---------+---------+---------+--
Pooled StDev =    0.4857             3.00      3.60      4.20      4.80
```

One-Way Analysis of Variance

```
Analysis of Variance for C22
Source     DF        SS        MS        F        P
C21         3    10.687     3.562    15.55    0.000
Error      36     8.247     0.229
Total      39    18.933

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
E 1121     10    3.6170    0.5397            (----*----)
E 2111     10    2.9890    0.4166   (----*----)
Limachi    10    4.4210    0.4430                           (----*----)
Marglobe   10    3.4500    0.5052          (-----*----)
                                    ------+---------+---------+---------+
Pooled StDev =    0.4786               3.00      3.60      4.20      4.80
```

One-Way Analysis of Variance

```
Analysis of Variance for C24
Source     DF        SS        MS        F        P
C23         3    13.848     4.616    16.74    0.000
Error      36     9.926     0.276
Total      39    23.774

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  -+---------+---------+---------+-----
E 1519     10    3.9290    0.4056              (----*----)
E 1915     10    3.9790    0.3739               (----*----)
Ace        10    4.7360    0.8158                         (----*---)
Lukullus   10    3.0740    0.3649  (----*----)
                                    -+---------+---------+---------+-----
Pooled StDev =    0.5251          2.80      3.50      4.20      4.90
```

One-Way Analysis of Variance

```
Analysis of Variance for C26
Source     DF        SS        MS        F        P
C25         3     5.834     1.945     2.36    0.088
Error      36    29.687     0.825
Total      39    35.521

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
E 1522     10    4.2690    0.4177         (---------*---------)
E 2215     10    3.8680    1.0843   (--------*---------)
Ace        10    4.7360    0.8158                 (---------*---------)
SanMarz    10    4.8150    1.1326                   (--------*---------)
                                      ------+---------+---------+---------+
Pooled StDev =    0.9081                   3.60      4.20      4.80      5.40
```

One-Way Analysis of Variance

```
Analysis of Variance for C28
Source     DF        SS        MS        F        P
C27         3    17.269     5.756    14.49    0.000
Error      36    14.302     0.397
Total      39    31.571

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
E 1922     10    3.9020    0.1621            (-----*-----)
E 2219     10    3.4010    0.3832      (-----*----)
Lukullus   10    3.0740    0.3649 (-----*-----)
SanMarz    10    4.8150    1.1326                          (-----*-----)
                                      --+---------+---------+---------+----
Pooled StDev =    0.6303             2.80      3.50      4.20      4.90
```

One-Way Analysis of Variance

```
Analysis of Variance for C30
Source     DF        SS        MS        F        P
C29         2     8.273     4.137     7.29    0.003
Error      26    14.757     0.568
Total      28    23.030

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level       N      Mean     StDev  --------+---------+---------+--------
E 1521      9    4.0667    0.8993           (------*------)
Ace        10    4.7360    0.8158                    (------*------)
Marglobe   10    3.4500    0.5052 (------*------)
                                      --------+---------+---------+--------
Pooled StDev =    0.7534             3.50      4.20      4.90
```

One-Way Analysis of Variance

```
Analysis of Variance for C32
Source     DF        SS        MS        F        P
C31         2     6.654     3.327     5.67    0.009
Error      27    15.843     0.587
Total      29    22.497

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level       N      Mean     StDev  -------+---------+---------+---------
E 2211     10    3.6790    0.5302 (-------*--------)
SanMarz    10    4.8150    1.1326                   (-------*--------)
Limachi    10    4.4210    0.4430           (--------*-------)
                                      -------+---------+---------+---------
Pooled StDev =    0.7660             3.60      4.20      4.80
```

One-Way Analysis of Variance

```
Analysis of Variance for C34
Source     DF        SS        MS        F        P
C33         2     9.388     4.694     7.21    0.003
Error      25    16.270     0.651
Total      27    25.659

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---------+---------+---------+-------
E 2221      8    4.0200    0.5889          (-------*--------)
SanMarz    10    4.8150    1.1326                    (-------*------)
Marglobe   10    3.4500    0.5052   (------*-------)
                                     ---------+---------+---------+-------
Pooled StDev =   0.8067                    3.50      4.20      4.90
```

Fruit Ratio (d/l)

One-Way Analysis of Variance

```
Analysis of Variance for C2
Source     DF        SS        MS        F        P
C1          3   0.44923   0.14974    30.04    0.000
Error      36   0.17947   0.00499
Total      39   0.62870

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---------+---------+---------+-------
I 3911     10    1.0620    0.0656   (---*----)
I 1139     10    1.1350    0.0828         (---*----)
Ceras      10    1.0750    0.0360    (---*----)
Limachi    10    1.3270    0.0865                              (----*---)
                                     ---------+---------+---------+-------
Pooled StDev =   0.0706                    1.10      1.20      1.30
```

One-Way Analysis of Variance

```
Analysis of Variance for C4
Source     DF        SS        MS        F        P
C3          3   0.68345   0.22782    23.73    0.000
Error      36   0.34563   0.00960
Total      39   1.02908

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---+---------+---------+---------+---
I 3915     10    1.1510    0.0769       (---*---)
I 1539     10    1.2700    0.1493              (----*---)
Ceras      10    1.0750    0.0360  (----*---)
Ace        10    1.4210    0.0943                        (---*---)
                                     ---+---------+---------+---------+---
Pooled StDev =   0.0980              1.05      1.20      1.35      1.50
```

One-Way Analysis of Variance

```
Analysis of Variance for C6
Source     DF        SS        MS        F        P
C5          3   0.08930   0.02977     3.88    0.017
Error      36   0.27594   0.00766
Total      39   0.36524

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --------+---------+---------+--------
I 5219     10    1.1110    0.0436     (--------*---------)
I 1952     10    1.0900    0.0394  (---------*--------)
Parvifl    10    1.1880    0.0452                   (--------*--------)
Lukullus   10    1.1990    0.1586                    (---------*--------)
                                     --------+---------+---------+--------
Pooled StDev =   0.0875              1.080     1.140     1.200
```

One-Way Analysis of Variance

```
Analysis of Variance for C8
Source     DF        SS        MS        F        P
C7          3   1.88651   0.62884    94.53    0.000
Error      36   0.23949   0.00665
Total      39   2.12600

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---+---------+---------+---------+---
I 5222     10    1.0580    0.0501                      (--*--)
I 2252     10    1.0430    0.0440                      (-*--)
Parvifl    10    1.1880    0.0452                              (-*--)
SanMarz    10    0.6120    0.1419   (--*-)
                                    ---+---------+---------+---------+---
Pooled StDev =   0.0816             0.60      0.80      1.00      1.20
```

One-Way Analysis of Variance

```
Analysis of Variance for C10
Source     DF        SS        MS        F        P
C9          2   0.28941   0.14470    28.85    0.000
Error      27   0.13541   0.00502
Total      29   0.42482

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
I 1160     10    1.1120    0.0496   (---*----)
Limachi    10    1.3270    0.0865                      (----*---)
Pimpinel   10    1.1260    0.0714    (----*---)
                                    ----+---------+---------+---------+--
Pooled StDev =   0.0708              1.10      1.20      1.30      1.40
```

One-Way Analysis of Variance

```
Analysis of Variance for C12
Source     DF        SS        MS        F        P
C11         2    0.0967    0.0483     4.45    0.021
Error      27    0.2933    0.0109
Total      29    0.3900

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
I 1960     10    1.0600    0.0483  (--------*-------)
Lukullus   10    1.1990    0.1586                    (--------*-------)
Pimpinel   10    1.1260    0.0714          (--------*-------)
                                    ------+---------+---------+---------+
Pooled StDev =   0.1042             1.040     1.120     1.200     1.280
```

One-Way Analysis of Variance

```
Analysis of Variance for C14
Source     DF        SS        MS        F        P
C13         2    1.2130    0.6065    40.34    0.000
Error      27    0.4059    0.0150
Total      29    1.6189

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
I 2239     10    0.9890    0.1539                    (---*---)
SanMarz    10    0.6120    0.1419   (---*---)
Ceras      10    1.0750    0.0360                             (---*---)
                                    ----+---------+---------+---------+--
Pooled StDev =   0.1226              0.60      0.80      1.00      1.20
```

One-Way Analysis of Variance

```
Analysis of Variance for C16
Source     DF        SS        MS        F        P
C15         2   1.33331   0.66665    73.30    0.000
Error      27   0.24556   0.00909
Total      29   1.57887

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  ---+---------+---------+---------+---
I 2260     10    0.9120    0.0454                 (--*--)
SanMarz    10    0.6120    0.1419   (--*--)
Pimpinel   10    1.1260    0.0714                          (--*--)
                                    ---+---------+---------+---------+---
Pooled StDev =   0.0954             0.60      0.80      1.00      1.20
```

One-Way Analysis of Variance

```
Analysis of Variance for C18
Source     DF        SS        MS        F        P
C17         3    0.1090    0.0363     2.22    0.106
Error      31    0.5074    0.0164
Total      34    0.6164

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  -+---------+---------+---------+-----
E 1115     10    1.2760    0.1837  (--------*-------)
E 1511      5    1.3280    0.1186    (-----------*----------)
Limachi    10    1.3270    0.0865       (--------*-------)
Ace        10    1.4210    0.0943                (-------*-------)
                                    -+---------+---------+---------+-----
Pooled StDev =   0.1279           1.20      1.30      1.40      1.50
```

One-Way Analysis of Variance

```
Analysis of Variance for C20
Source     DF        SS        MS        F        P
C19         3    0.3081    0.1027     8.05    0.000
Error      36    0.4595    0.0128
Total      39    0.7676

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  --------+---------+---------+--------
E 1119     10    1.1250    0.1120     (------*-------)
E 1911     10    1.1020    0.0766   (------*------)
Limachi    10    1.3270    0.0865                          (-------*------)
Lukullus   10    1.1990    0.1586              (------*------)
                                    --------+---------+---------+--------
Pooled StDev =   0.1130                 1.10      1.20      1.30
```

One-Way Analysis of Variance

```
Analysis of Variance for C22
Source     DF        SS        MS        F        P
C21         3   0.48193   0.16064    26.37    0.000
Error      36   0.21931   0.00609
Total      39   0.70124

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  --------+---------+---------+--------
E 1121     10    1.0450    0.0624   (---*---)
E 2111     10    1.0820    0.0932     (---*---)
Limachi    10    1.3270    0.0865                          (----*---)
Marglobe   10    1.1110    0.0656       (----*---)
                                    --------+---------+---------+--------
Pooled StDev =   0.0781                 1.08      1.20      1.32
```

One-Way Analysis of Variance

```
Analysis of Variance for C24
Source     DF        SS        MS        F        P
C23         3    0.3079    0.1026     5.73    0.003
Error      36    0.6453    0.0179
Total      39    0.9532

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --------+---------+---------+--------
E 1519     10    1.2750    0.1405         (------*------)
E 1915     10    1.2140    0.1338    (------*------)
Ace        10    1.4210    0.0943                      (------*-------)
Lukullus   10    1.1990    0.1586   (------*------)
                                     --------+---------+---------+--------
Pooled StDev =   0.1339                1.20      1.32      1.44
```

One-Way Analysis of Variance

```
Analysis of Variance for C26
Source     DF        SS        MS        F        P
C25         3    3.2790    1.0930    35.98    0.000
Error      36    1.0938    0.0304
Total      39    4.3728

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
E 1522     10    1.0190    0.1073               (---*---)
E 2215     10    1.0470    0.2846                (---*---)
Ace        10    1.4210    0.0943                                (--*---)
SanMarz    10    0.6120    0.1419   (--*---)
                                     ----+---------+---------+---------+--
Pooled StDev =   0.1743               0.60      0.90      1.20      1.50
```

One-Way Analysis of Variance

```
Analysis of Variance for C28
Source     DF        SS        MS        F        P
C27         3    1.8119    0.6040    42.21    0.000
Error      36    0.5152    0.0143
Total      39    2.3271

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---------+---------+---------+-------
E 1922     10    1.0180    0.0742                   (--*--)
E 2219     10    0.9680    0.0802                 (--*--)
Lukullus   10    1.1990    0.1586                          (--*--)
SanMarz    10    0.6120    0.1419  (--*---)
                                     ---------+---------+---------+-------
Pooled StDev =   0.1196              0.75      1.00      1.25
```

One-Way Analysis of Variance

```
Analysis of Variance for C30
Source     DF        SS        MS        F        P
C29         2   0.77009   0.38504    43.23    0.000
Error      26   0.23158   0.00891
Total      28   1.00167

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  -----+---------+---------+---------+-
E 1521      9    1.0500    0.1187   (---*---)
Ace        10    1.4210    0.0943                          (---*---)
Marglobe   10    1.1110    0.0656       (---*---)
                                     -----+---------+---------+---------+-
Pooled StDev =   0.0944               1.05      1.20      1.35      1.50
```

One-Way Analysis of Variance

```
Analysis of Variance for C32
Source     DF        SS        MS        F        P
C31         2    2.5601    1.2801    62.91    0.000
Error      27    0.5494    0.0203
Total      29    3.1095

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level       N      Mean     StDev  ----------+---------+---------+------
E 2211     10    0.9940    0.1828                (---*--)
SanMarz    10    0.6120    0.1419   (--*---)
Limachi    10    1.3270    0.0865                            (---*---)
                                      ----------+---------+---------+------
Pooled StDev =   0.1426             0.75      1.00      1.25
```

One-Way Analysis of Variance

```
Analysis of Variance for C34
Source     DF        SS        MS        F        P
C33         2    1.2535    0.6267    51.46    0.000
Error      25    0.3044    0.0122
Total      27    1.5579

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level       N      Mean     StDev  ---+---------+---------+---------+---
E 2221      8    0.9000    0.1099             (---*---)
SanMarz    10    0.6120    0.1419  (---*--)
Marglobe   10    1.1110    0.0656                       (---*--)
                                      ---+---------+---------+---------+---
Pooled StDev =   0.1104              0.60      0.80      1.00      1.20
```

Fruit Weight (g)

One-Way Analysis of Variance

```
Analysis of Variance for C2
Source     DF        SS        MS        F        P
C1          3     45792     15264    61.83    0.000
Error      36      8887       247
Total      39     54679

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level       N      Mean     StDev  ---+---------+---------+---------+---
I 3911     10     12.42      4.39      (--*---)
I 1139     10      9.41      4.51     (--*--)
Ceras      10      3.46      1.05   (--*---)
Limachi    10     86.21     30.77                            (---*--)
                                      ---+---------+---------+---------+---
Pooled StDev =   15.71               0         30        60        90
```

One-Way Analysis of Variance

```
Analysis of Variance for C4
Source     DF        SS        MS        F        P
C3          3    105271     35090    26.71    0.000
Error      36     47302      1314
Total      39    152573

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
I 3915     10     18.35      6.24     (----*---)
I 1539     10      6.11      4.26   (---*----)
Ceras      10      3.46      1.05  (----*---)
Ace        10    127.07     72.09                          (---*----)
                                      ----+---------+---------+---------+--
Pooled StDev =   36.25               0         50        100       150
```

One-Way Analysis of Variance

```
Analysis of Variance for C6
Source     DF        SS        MS        F        P
C5          3    3638.5    1212.8   118.34    0.000
Error      36     369.0      10.2
Total      39    4007.5

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
I 5219     10     4.597     1.214       (--*-)
I 1952     10     4.174     0.546       (-*--)
Parvifl    10     0.679     0.090  (--*-)
Lukullus   10    24.894     6.262                                  (-*--)
                                    --+---------+---------+---------+----
Pooled StDev =    3.201             0.0       8.0      16.0      24.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C8
Source     DF        SS        MS        F        P
C7          3    3158.6    1052.9    68.05    0.000
Error      36     556.9      15.5
Total      39    3715.5

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  ---+---------+---------+---------+---
I 5222     10     2.988     0.786     (--*--)
I 2252     10     3.595     1.062     (--*---)
Parvifl    10     0.679     0.090  (--*--)
SanMarz    10    22.788     7.754                              (--*---)
                                    ---+---------+---------+---------+---
Pooled StDev =    3.933              0.0       8.0      16.0      24.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C10
Source     DF        SS        MS        F        P
C9          2     43388     21694    68.44    0.000
Error      27      8558       317
Total      29     51946

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
I 1160     10      9.65      1.95     (---*---)
Limachi    10     86.21     30.77                                (---*---)
Pimpinel   10      1.98      0.67   (---*---)
                                    ----+---------+---------+---------+--
Pooled StDev =    17.80              0        30        60        90
```

One-Way Analysis of Variance

```
Analysis of Variance for C12
Source     DF        SS        MS        F        P
C11         2    2977.7    1488.8    92.10    0.000
Error      27     436.5      16.2
Total      29    3414.1

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  -+---------+---------+---------+-----
I 1960     10     6.174     2.972       (---*--)
Lukullus   10    24.894     6.262                                (--*--)
Pimpinel   10     1.976     0.669  (--*---)
                                    -+---------+---------+---------+-----
Pooled StDev =    4.021           0.0       8.0      16.0      24.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C14
Source     DF        SS        MS        F        P
C13         2    1887.4     943.7    37.78    0.000
Error      27     674.4      25.0
Total      29    2561.8

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----------+---------+---------+------
I 2239     10    14.825     3.702                  (---*----)
SanMarz    10    22.788     7.754                            (----*---)
Ceras      10     3.459     1.051  (----*----)
                                    ----------+---------+---------+------
Pooled StDev =    4.998             7.0       14.0      21.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C16
Source     DF        SS        MS        F        P
C15         2    2458.9    1229.5    54.75    0.000
Error      27     606.3      22.5
Total      29    3065.3

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
I 2260     10     5.750     2.607      (---*---)
SanMarz    10    22.788     7.754                          (--*---)
Pimpinel   10     1.976     0.669   (--*---)
                                    --+---------+---------+---------+----
Pooled StDev =    4.739             0.0       8.0       16.0      24.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C18
Source     DF        SS        MS        F        P
C17         3     38774     12925     6.28    0.002
Error      31     63783      2058
Total      34    102557

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --------+---------+---------+--------
E 1115     10    40.38     13.28   (------*------)
E 1511      5    68.10     41.53      (---------*---------)
Limachi    10    86.21     30.77              (-------*------)
Ace        10   127.07     72.09                       (-------*------)
                                    --------+---------+---------+--------
Pooled StDev =   45.36              40        80        120
```

One-Way Analysis of Variance

```
Analysis of Variance for C20
Source     DF        SS        MS        F        P
C19         3     25371      8457    25.29    0.000
Error      36     12039       334
Total      39     37410

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  -----+---------+---------+---------+-
E 1119     10    32.86      9.71     (----*----)
E 1911     10    27.54     16.05   (----*----)
Limachi    10    86.21     30.77                        (---*----)
Lukullus   10    24.89      6.26  (----*----)
                                    -----+---------+---------+---------+-
Pooled StDev =   18.29              25        50        75        100
```

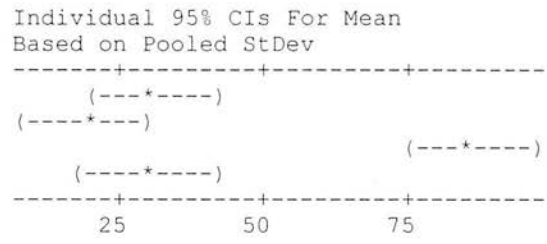
One-Way Analysis of Variance

Analysis of Variance for C22

Source	DF	SS	MS	F	P
C21	3	27371	9124	27.24	0.000
Error	36	12058	335		
Total	39	39429			

Level	N	Mean	StDev
E 1121	10	30.51	14.41
E 2111	10	19.26	8.11
Limachi	10	86.21	30.77
Marglobe	10	30.42	10.94

Pooled StDev = 18.30



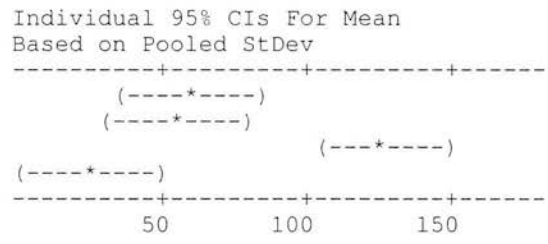
One-Way Analysis of Variance

Analysis of Variance for C24

Source	DF	SS	MS	F	P
C23	3	55797	18599	13.00	0.000
Error	36	51502	1431		
Total	39	107299			

Level	N	Mean	StDev
E 1519	10	58.61	19.46
E 1915	10	55.65	10.34
Ace	10	127.07	72.09
Lukullus	10	24.89	6.26

Pooled StDev = 37.82



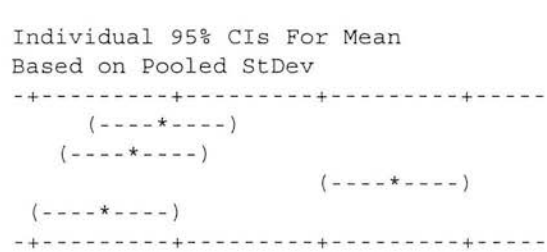
One-Way Analysis of Variance

Analysis of Variance for C2

Source	DF	SS	MS	F	P
C1	3	66416	22139	14.51	0.000
Error	36	54928	1526		
Total	39	121344			

Level	N	Mean	StDev
E 1522	10	44.61	12.92
E 2215	10	36.74	26.05
Ace	10	127.07	72.09
SanMarz	10	22.79	7.75

Pooled StDev = 39.06



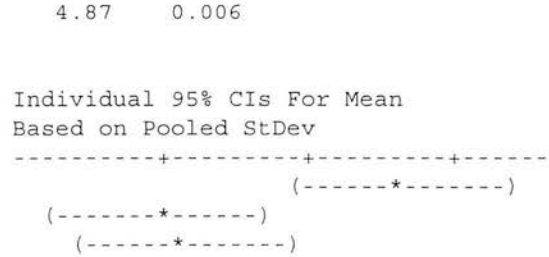
One-Way Analysis of Variance

Analysis of Variance for C4

Source	DF	SS	MS	F	P
C3	3	774.2	258.1	4.87	0.006
Error	36	1906.9	53.0		
Total	39	2681.1			

Level	N	Mean	StDev
E 1922	10	33.830	4.008
E 2219	10	23.769	9.822
Lukullus	10	24.894	6.262
SanMarz	10	22.788	7.754

Pooled StDev = 7.278



One-Way Analysis of Variance

```
Analysis of Variance for C26
Source      DF        SS        MS        F        P
C25          2     52497     26248    12.42    0.000
Error       26     54928      2113
Total       28    107424

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level        N      Mean     StDev  ----------+---------+---------+------
E 1521       9     48.19     29.74     (------*------)
Ace         10    127.07     72.09                    (-----*-----)
Marglobe    10     30.42     10.94  (-----*-----)
                                     ----------+---------+---------+------
Pooled StDev =     45.96                      50        100       150
```

One-Way Analysis of Variance

```
Analysis of Variance for C28
Source      DF        SS        MS        F        P
C27          2     25097     12548    33.46    0.000
Error       27     10126       375
Total       29     35223

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level        N      Mean     StDev  ------+---------+---------+---------+
E 2211      10     27.16     10.88    (----*----)
SanMarz     10     22.79      7.75  (----*----)
Limachi     10     86.21     30.77                        (----*-----)
                                     ------+---------+---------+---------+
Pooled StDev =     19.37            25        50        75       100
```

One-Way Analysis of Variance

```
Analysis of Variance for C30
Source      DF        SS        MS        F        P
C29          2     291.5     145.8     1.69    0.206
Error       25    2162.2      86.5
Total       27    2453.7

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level        N      Mean     StDev  ---+---------+---------+---------+---
E 2221       8    26.764     8.820       (-----------*----------)
SanMarz     10    22.788     7.754   (---------*---------)
Marglobe    10    30.422    10.937               (---------*---------)
                                     ---+---------+---------+---------+---
Pooled StDev =     9.300            18.0      24.0      30.0      36.0
```

Solid Soluble Content (°Brix)

One-Way Analysis of Variance

```
Analysis of Variance for C2
Source      DF        SS        MS        F        P
C1           3    87.473    29.158    42.10    0.000
Error       36    24.932     0.693
Total       39   112.406

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level        N      Mean     StDev  ----------+---------+---------+------
I 3911      10    9.1100    0.9938                            (---*--)
I 1139      10    8.0710    1.0184                     (---*--)
Ceras       10    7.0900    0.7608               (--*---)
Limachi     10    5.1000    0.4082  (---*---)
                                     ----------+---------+---------+------
Pooled StDev =    0.8322             6.0       7.5       9.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C4
Source     DF        SS        MS        F        P
C3          3     20.70      6.90     4.27    0.011
Error      36     58.15      1.62
Total      39     78.86

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  -+---------+---------+---------+-----
I 3915     10     7.410     1.392                  (-------*-------)
I 1539     10     7.620     1.365                    (-------*-------)
Ceras      10     7.090     0.761                (-------*-------)
Ace        10     5.770     1.442   (-------*-------)
                                    -+---------+---------+---------+-----
Pooled StDev =    1.271           5.0       6.0       7.0       8.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C6
Source     DF        SS        MS        F        P
C5          3    84.075    28.025    51.78    0.000
Error      36    19.485     0.541
Total      39   103.560

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---------+---------+---------+-------
I 5219     10     8.940     0.638               (---*--)
I 1952     10     8.530     0.464             (--*--)
Parvifl    10    10.760     0.986                                 (--*--)
Lukullus   10     6.680     0.755  (---*--)
                                    ---------+---------+---------+-------
Pooled StDev =    0.736             7.5       9.0      10.5
```

One-Way Analysis of Variance

```
Analysis of Variance for C8
Source     DF        SS        MS        F        P
C7          3   109.865    36.622    68.12    0.000
Error      36    19.354     0.538
Total      39   129.219

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
I 5222     10     7.590     0.684          (---*--)
I 2252     10     8.470     0.672                (--*---)
Parvifl    10    10.760     0.986                                   (--*--)
SanMarz    10     6.200     0.510  (--*--)
                                    --+---------+---------+---------+----
Pooled StDev =    0.733             6.0       7.5       9.0      10.
```

One-Way Analysis of Variance

```
Analysis of Variance for C10
Source     DF        SS        MS        F        P
C9          2     66.36     33.18    19.61    0.000
Error      27     45.69      1.69
Total      29    112.05

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
I 1160     10     7.750     2.120                    (-----*----)
Limachi    10     5.100     0.408  (-----*-----)
Pimpinel   10     8.590     0.644                           (----*-----)
                                    --+---------+---------+---------+----
Pooled StDev =    1.301             4.5       6.0       7.5       9.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C12
Source     DF        SS        MS        F        P
C11         2     24.07     12.03    11.91    0.000
Error      27     27.29      1.01
Total      29     51.35

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----------+---------+---------+------
I 1960     10     6.700     1.431  (------*------)
Lukullus   10     6.680     0.755  (------*-----)
Pimpinel   10     8.590     0.644                    (------*-----)
                                    ----------+---------+---------+------
Pooled StDev =    1.005           7.0       8.0       9.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C14
Source     DF        SS        MS        F        P
C13         2     7.274     3.637     5.74    0.008
Error      27    17.094     0.633
Total      29    24.368

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
I 2239     10    7.3500    1.0298                     (--------*-------)
SanMarz    10    6.2000    0.5099  (-------*--------)
Ceras      10    7.0900    0.7608                (-------*--------)
                                    ------+---------+---------+---------+
Pooled StDev =    0.7957          6.00      6.60      7.20      7.80
```

One-Way Analysis of Variance

```
Analysis of Variance for C16
Source     DF        SS        MS        F        P
C15         2    28.789    14.394    17.39    0.000
Error      27    22.345     0.828
Total      29    51.134

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
I 2260     10    7.5800    1.3448               (-----*-----)
SanMarz    10    6.2000    0.5099  (-----*-----)
Pimpinel   10    8.5900    0.6437                         (-----*-----)
                                    ----+---------+---------+---------+--
Pooled StDev =    0.9097            6.0       7.0       8.0       9.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C18
Source     DF        SS        MS        F        P
C17         3     5.080     1.693     1.95    0.142
Error      31    26.910     0.868
Total      34    31.990

                                        Individual 95% CIs For Mean
Based on Pooled StDev
Level       N      Mean     StDev  ---+---------+---------+---------+---
E 1115     10    5.4300    0.6929             (-------*------)
E 1511      5    4.6200    0.7694  (----------*---------)
Limachi    10    5.1000    0.4082          (-------*------)
Ace        10    5.7700    1.4423                   (------*-------)
                                    ---+---------+---------+---------+---
Pooled StDev =    0.9317           4.00      4.80      5.60      6.40
```

One-Way Analysis of Variance

```
Analysis of Variance for C20
Source     DF        SS        MS        F        P
C19         3    19.507     6.502     9.29    0.000
Error      36    25.184     0.700
Total      39    44.691

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  -----+---------+---------+---------+-
E 1119     10    4.9400    0.8631  (------*-----)
E 1911     10    5.9400    1.1472            (-----*------)
Limachi    10    5.1000    0.4082    (------*-----)
Lukullus   10    6.6800    0.7554                      (------*-----)
                                    -----+---------+---------+---------+-
Pooled StDev =   0.8364              4.80      5.60      6.40      7.20
```

One-Way Analysis of Variance

```
Analysis of Variance for C22
Source     DF        SS        MS        F        P
C21         3     18.33      6.11     3.42    0.027
Error      36     64.38      1.79
Total      39     82.71

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --------+---------+---------+--------
E 1121     10    5.322     0.696     (-------*--------)
E 2111     10    6.770     2.405               (--------*-------)
Limachi    10    5.100     0.408  (--------*--------)
Marglobe   10    5.230     0.849    (-------*--------)
                                    --------+---------+---------+--------
Pooled StDev =   1.337              5.0       6.0       7.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C24
Source     DF        SS        MS        F        P
C23         3      7.48      2.49     2.10    0.120
Error      31     36.77      1.19
Total      34     44.25

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  -----+---------+---------+---------+-
E 1519      5    5.420     0.554  (-----------*-----------)
E 1915     10    5.740     1.139        (--------*--------)
Ace        10    5.770     1.442        (--------*--------)
Lukullus   10    6.680     0.755                 (--------*-------)
                                    -----+---------+---------+---------+-
Pooled StDev =   1.089              4.80      5.60      6.40      7.20
```

One-Way Analysis of Variance

```
Analysis of Variance for C26
Source     DF        SS        MS        F        P
C25         3     17.97      5.99     2.20    0.105
Error      36     98.17      2.73
Total      39    116.14

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
E 1522     10    5.410     1.729  (--------*--------)
E 2215     10    7.200     2.362               (--------*--------)
Ace        10    5.770     1.442    (--------*--------)
SanMarz    10    6.200     0.510       (--------*-------)
                                    ----+---------+---------+---------+--
Pooled StDev =   1.651              4.8       6.0       7.2       8.4
```

One-Way Analysis of Variance

```
Analysis of Variance for C28
Source     DF        SS        MS        F        P
C27         3      4.70      1.57     1.50    0.232
Error      36     37.69      1.05
Total      39     42.40

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---------+---------+---------+--------
E 1922     10     6.190     1.706         (--------*---------)
E 2219     10     5.710     0.669  (---------*--------)
Lukullus   10     6.680     0.755                  (--------*---------)
SanMarz    10     6.200     0.510         (---------*--------)
                                    ---------+---------+---------+--------
Pooled StDev =    1.023                   5.60      6.30      7.00
```

One-Way Analysis of Variance

```
Analysis of Variance for C30
Source     DF        SS        MS        F        P
C29         2      2.12      1.06     0.92    0.413
Error      24     27.70      1.15
Total      26     29.82

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---------+---------+---------+-------
E 1521      7     5.143     0.645   (-------------*-------------)
Ace        10     5.770     1.442               (-----------*-----------)
Marglobe   10     5.230     0.849     (-----------*-----------)
                                    ---------+---------+---------+-------
Pooled StDev =    1.074                   4.80      5.40      6.00
```

One-Way Analysis of Variance

```
Analysis of Variance for C32
Source     DF        SS        MS        F        P
C31         2    23.546    11.773    14.42    0.000
Error      27    22.041     0.816
Total      29    45.587

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  -----+---------+---------+---------+-
E 2211     10    7.2700    1.4221                        (-----*-----)
SanMarz    10    6.2000    0.5099              (-----*-----)
Limachi    10    5.1000    0.4082  (-----*-----)
                                    -----+---------+---------+---------+-
Pooled StDev =   0.9035               5.0       6.0       7.0       8.0
```

One-Way Analysis of Variance

```
Analysis of Variance for C34
Source     DF        SS        MS        F        P
C33         2      5.63      2.82     2.32    0.119
Error      25     30.31      1.21
Total      27     35.94

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
E 2221      8     5.312     1.752  (-----------*----------)
SanMarz    10     6.200     0.510               (----------*---------)
Marglobe   10     5.230     0.849  (----------*---------)
                                    ------+---------+---------+---------+
Pooled StDev =    1.101                 4.90      5.60      6.30      7.00
```

Weight 1,000 Seeds (g)

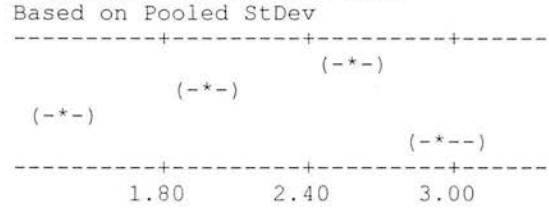
One-Way Analysis of Variance

Analysis of Variance for C2

Source	DF	SS	MS	F	P
C1	3	5.9527	1.9842	160.37	0.000
Error	12	0.1485	0.0124		
Total	15	6.1012			

Level	N	Mean	StDev
I 3911	4	2.5600	0.0678
I 1139	4	1.9800	0.0632
Ceras	4	1.3525	0.1965
Limachi	4	2.9700	0.0476

Pooled StDev = 0.1112



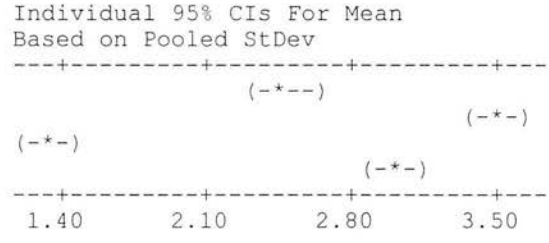
One-Way Analysis of Variance

Analysis of Variance for C4

Source	DF	SS	MS	F	P
C3	3	10.0256	3.3419	190.17	0.000
Error	12	0.2109	0.0176		
Total	15	10.2365			

Level	N	Mean	StDev
I 3915	4	2.4850	0.0900
I 1539	4	3.4800	0.1386
Ceras	4	1.3525	0.1965
Ace	4	3.0050	0.0661

Pooled StDev = 0.1326



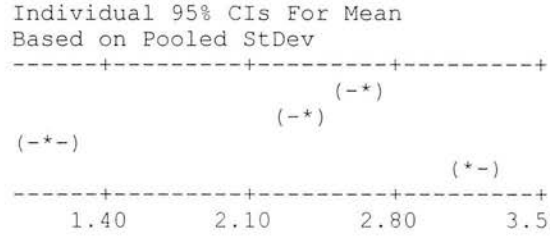
One-Way Analysis of Variance

Analysis of Variance for C6

Source	DF	SS	MS	F	P
C5	3	9.0026	3.0009	240.39	0.000
Error	12	0.1498	0.0125		
Total	15	9.1524			

Level	N	Mean	StDev
I 5219	4	2.6400	0.0632
I 1952	4	2.3550	0.1310
Parvifl	4	1.1300	0.0258
Lukullus	4	3.1750	0.1676

Pooled StDev = 0.1117



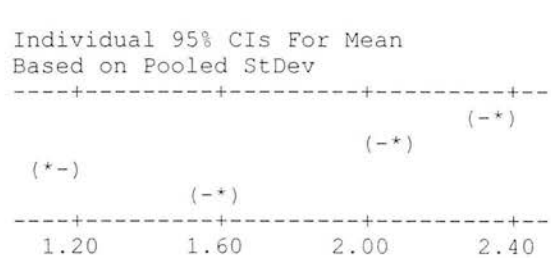
One-Way Analysis of Variance

Analysis of Variance for C8

Source	DF	SS	MS	F	P
C7	3	3.48702	1.16234	343.97	0.000
Error	12	0.04055	0.00338		
Total	15	3.52757			

Level	N	Mean	StDev
I 5222	4	2.3525	0.0780
I 2252	4	2.0725	0.0727
Parvifl	4	1.1300	0.0258
SanMarz	4	1.5900	0.0383

Pooled StDev = 0.0581



One-Way Analysis of Variance

```
Analysis of Variance for C10
Source      DF        SS         MS         F        P
C9           2   8.96247    4.48123   1353.39    0.000
Error        9   0.02980    0.00331
Total       11   8.99227

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev   ---------+---------+---------+-------
I 1160       4    1.8350    0.0574                (*)
Limachi      4    2.9700    0.0476                                    (*)
Pimpinel     4    0.8550    0.0661   (*)
                                      ---------+---------+---------+-------
Pooled StDev =    0.0575             1.40      2.10      2.80
```

One-Way Analysis of Variance

```
Analysis of Variance for C12
Source      DF        SS         MS         F        P
C11          2   10.7889    5.3944    443.78    0.000
Error        9    0.1094    0.0122
Total       11   10.8983

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev   ----------+---------+---------+------
I 1960       4    1.9200    0.0632                (*-)
Lukullus     4    3.1750    0.1676                                      (*-)
Pimpinel     4    0.8550    0.0661   (-*-)
                                      ----------+---------+---------+------
Pooled StDev =    0.1103             1.40      2.10      2.80
```

One-Way Analysis of Variance

```
Analysis of Variance for C14
Source      DF        SS         MS         F        P
C13          2    1.6581    0.8291     60.69    0.000
Error        9    0.1229    0.0137
Total       11    1.7811

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev   ------+---------+---------+---------+
I 2239       4    2.2325    0.0299                            (---*---)
SanMarz      4    1.5900    0.0383           (--*---)
Ceras        4    1.3525    0.1965    (---*--)
                                      ------+---------+---------+---------+
Pooled StDev =    0.1169              1.40      1.75      2.10      2.45
```

One-Way Analysis of Variance

```
Analysis of Variance for C16
Source      DF        SS         MS         F        P
C15          2   2.65127    1.32563    235.32    0.000
Error        9   0.05070    0.00563
Total       11   2.70197

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev   -+---------+---------+---------+-----
I 2260       4    1.9900    0.1052                                    (-*-)
SanMarz      4    1.5900    0.0383                        (-*-)
Pimpinel     4    0.8550    0.0661   (-*-)
                                      -+---------+---------+---------+-----
Pooled StDev =    0.0751            0.80      1.20      1.60      2.00
```

One-Way Analysis of Variance

```
Analysis of Variance for C18
Source     DF        SS        MS        F        P
C17         3   0.79500   0.26500    35.25    0.000
Error      12   0.09020   0.00752
Total      15   0.88520

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ---+---------+---------+---------+---
E 1115      4    3.3900    0.1281                          (---*--)
E 1511      4    2.7750    0.0839  (---*---)
Limachi     4    2.9700    0.0476          (---*---)
Ace         4    3.0050    0.0661           (---*---)
                                   ---+---------+---------+---------+---
Pooled StDev =   0.0867             2.75      3.00      3.25      3.50
```

One-Way Analysis of Variance

```
Analysis of Variance for C20
Source     DF        SS        MS        F        P
C19         3    1.9647    0.6549    53.28    0.000
Error      12    0.1475    0.0123
Total      15    2.1122

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
E 1119      4    3.4100    0.0987                              (--*---)
E 1911      4    2.4600    0.0952   (--*---)
Limachi     4    2.9700    0.0476               (---*--)
Lukullus    4    3.1750    0.1676                    (---*--)
                                   ----+---------+---------+---------+--
Pooled StDev =   0.1109             2.45      2.80      3.15      3.50
```

One-Way Analysis of Variance

```
Analysis of Variance for C22
Source     DF        SS        MS        F        P
C21         3   2.86612   0.95537    99.80    0.000
Error      12   0.11488   0.00957
Total      15   2.98099

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
E 1121      4    3.6175    0.1740                              (-*--)
E 2111      4    3.2200    0.0432                     (--*-)
Limachi     4    2.9700    0.0476                (-*--)
Marglobe    4    2.4500    0.0622   (-*--)
                                   --+---------+---------+---------+----
Pooled StDev =   0.0978            2.40      2.80      3.20      3.60
```

One-Way Analysis of Variance

```
Analysis of Variance for C24
Source     DF        SS        MS        F        P
C23         3    7.1032    2.3677   198.97    0.000
Error      12    0.1428    0.0119
Total      15    7.2460

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
E 1519      4    3.4550    0.1050          (-*-)
E 1915      4    4.7050    0.0640                              (-*-)
Ace         4    3.0050    0.0661  (-*-)
Lukullus    4    3.1750    0.1676     (-*-)
                                   --+---------+---------+---------+----
Pooled StDev =   0.1091            3.00      3.60      4.20      4.80
```

One-Way Analysis of Variance

```
Analysis of Variance for C26
Source     DF        SS        MS        F        P
C25         3    6.5377    2.1792   167.10    0.000
Error      12    0.1565    0.0130
Total      15    6.6942

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
E 1522      4    2.9550    0.0719                      (-*-)
E 2215      4    3.1950    0.2029                          (-*-)
Ace         4    3.0050    0.0661                       (-*-)
SanMarz     4    1.5900    0.0383  (--*-)
                                    ------+---------+---------+---------+
Pooled StDev =    0.1142              1.80      2.40      3.00      3.60
```

One-Way Analysis of Variance

```
Analysis of Variance for C28
Source     DF        SS        MS        F        P
C27         3    7.7141    2.5714   150.52    0.000
Error      12    0.2050    0.0171
Total      15    7.9191

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
E 1922      4    3.2050    0.1578                           (-*--)
E 2219      4    3.2000    0.1178                           (-*--)
Lukullus    4    3.1750    0.1676                           (-*-)
SanMarz     4    1.5900    0.0383  (--*-)
                                    ------+---------+---------+---------+
Pooled StDev =    0.1307              1.80      2.40      3.00      3.60
```

One-Way Analysis of Variance

```
Analysis of Variance for C30
Source     DF        SS        MS        F        P
C29         2    3.1511    1.5755    93.30    0.000
Error       9    0.1520    0.0169
Total      11    3.3030

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
E 1521      4    3.7025    0.2060                          (--*--)
Ace         4    3.0050    0.0661              (--*--)
Marglobe    4    2.4500    0.0622  (--*--)
                                    ----+---------+---------+---------+--
Pooled StDev =    0.1299             2.50      3.00      3.50      4.00
```

One-Way Analysis of Variance

```
Analysis of Variance for C32
Source     DF        SS        MS        F        P
C31         2   7.45920   3.72960   527.77    0.000
Error       9   0.06360   0.00707
Total      11   7.52280

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
E 2211      4    3.4500    0.1322                                   (-*)
SanMarz     4    1.5900    0.0383   (-*)
Limachi     4    2.9700    0.0476                        (*-)
                                    ------+---------+---------+---------+
Pooled StDev =    0.0841              1.80      2.40      3.00      3.60
```

One-Way Analysis of Variance

```
Analysis of Variance for C34
Source     DF        SS        MS         F        P
C33         2    6.4843    3.2421     37.74    0.000
Error       9    0.7732    0.0859
Total      11    7.2575

                                  Individual 95% CIs For Mean
                                  Based on Pooled StDev
Level       N      Mean     StDev  ---+---------+---------+---------+---
E 2221      4    3.3900    0.5024                              (---*----)
SanMarz     4    1.5900    0.0383   (----*---)
Marglobe    4    2.4500    0.0622              (----*----)
                                  ---+---------+---------+---------+---
Pooled StDev =    0.2931           1.40      2.10      2.80      3.50
```

Part 10 ANOVA of continuous characters from inter- and intra-taxon crosses in F_2 and F_3 generations presented in Table 5.2.

One-Way Analysis of Variance for Fruit Weight in F_2

```
Analysis of Variance
Source     DF        SS        MS         F        P
Groups      3    5315.5    1771.8     61.84    0.000
Error      27     773.6      28.7
Total      30    6089.2

                                  Individual 95% CIs For Mean
                                  Based on Pooled StDev
Level       N      Mean     StDev  -+---------+---------+---------+-----
L.esc.cera. 7    12.527     2.576          (--*---)
L.parvifl.  7     3.979     0.540   (--*---)
L.pimpinel. 5     8.234     3.016      (---*---)
L.esculen. 12    34.477     7.953                              (--*-)
                                  -+---------+---------+---------+-----
Pooled StDev =    5.353            0         12        24        36
```

One-Way Analysis of Variance for Fruit Diameter in F_2

```
Analysis of Variance
Source     DF        SS        MS         F        P
Groups      3    26.026     8.675     65.16    0.000
Error      27     3.595     0.133
Total      30    29.621

                                  Individual 95% CIs For Mean
                                  Based on Pooled StDev
Level       N      Mean     StDev  -+---------+---------+---------+-----
L.esc.cera. 7    2.7200    0.2228             (---*---)
L.parvifl.  7    1.8557    0.0443   (--*---)
L.pimpinel. 5    2.3320    0.3849        (---*---)
L.esculen. 12    4.0833    0.4947                               (--*--)
                                  -+---------+---------+---------+-----
Pooled StDev =    0.3649         1.60      2.40      3.20      4.00
```

One-Way Analysis of Variance for Fruit Length in F_2

```
Analysis of Variance
Source       DF        SS        MS        F        P
Groups        3     17.8357    5.9452   162.74    0.000
Error        27      0.9864    0.0365
Total        30     18.8221

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
L.esc.cera. 7    2.5071    0.1868                (--*-)
L.parvifl.  7    1.6129    0.1063  (--*-)
L.pimpinel. 5    2.1200    0.3147          (--*--)
L.esculen. 12    3.5050    0.1687                                    (*-)
                                    ------+---------+---------+---------+
Pooled StDev =   0.1911              1.80      2.40      3.00      3.60
```

One-Way Analysis of Variance for Fruit Ratio in F_2

```
Analysis of Variance
Source       DF        SS        MS        F        P
Groups        3     0.0180    0.0060     0.40    0.756
Error        27     0.4074    0.0151
Total        30     0.4254

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  -----+---------+---------+---------+-
L.esc.cera. 7    1.0971    0.0945  (-----------*-----------)
L.parvifl.  7    1.1557    0.0770          (-----------*-----------)
L.pimpinel. 4    1.1650    0.0592       (---------------*--------------)
L.esculen. 13    1.1485    0.1601            (--------*-------)
                                    -----+---------+---------+---------+-
Pooled StDev =   0.1228             1.040     1.120     1.200     1.280
```

One-Way Analysis of Variance for Soluble Solid Content in F_2

```
Analysis of Variance
Source       DF        SS        MS        F        P
Groups        3      7.19      2.40     1.08    0.373
Error        27     59.74      2.21
Total        30     66.92

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  --------+---------+---------+--------
L.esc.cera. 7     6.243     0.382  (---------*---------)
L.parvifl.  7     7.057     2.947          (---------*---------)
L.pimpinel. 4     7.575     1.164           (------------*------------)
L.esculen. 13     6.315     0.474       (------*------)
                                    --------+---------+---------+--------
Pooled StDev =    1.487               6.0       7.2       8.4
```

One-Way Analysis of Variance for 1,000 Seeds Weight in F_2

```
Analysis of Variance
Source       DF        SS        MS        F        P
Groups        3     5.127     1.709     7.58    0.001
Error        27     6.087     0.225
Total        30    11.213

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  -----+---------+---------+---------+-
L.esc.cera.  7    2.5257    0.2722          (-------*------)
L.parvifl.   7    2.2277    0.1484    (-------*------)
L.pimpinel.  4    2.2840    0.2425   (---------*--------)
L.esculen.  13    3.1527    0.6667                          (----*----)
                                      -----+---------+---------+---------+-
Pooled StDev =    0.4748               2.00      2.50      3.00      3.50
```

One-Way Analysis of Variance for Fruit Weight in F_3

```
Analysis of Variance
Source       DF        SS        MS        F        P
Groups        3     18159      6053    31.04    0.000
Error        28      5459       195
Total        31     23618

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  ---+---------+---------+---------+---
L.esc.cer.   6     17.82      4.58        (-----*-----)
L.parvifl.   7      6.46      1.61   (----*-----)
L.pimpin.    5     13.20      1.20     (------*-----)
L. escul.   14     59.57     20.25                                (---*---)
                                      ---+---------+---------+---------+---
Pooled StDev =    13.96               0        20        40        60
```

One-Way Analysis of Variance for Fruit Diameter in F_3

```
Analysis of Variance
Source       DF        SS        MS        F        P
Groups        3    40.979    13.660    51.56    0.000
Error        28     7.418     0.265
Total        31    48.397

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  -+---------+---------+---------+-----
L.esc.cer.   6    3.1333    0.3351          (---*----)
L.parvifl.   7    2.3443    0.4134   (---*---)
L.pimpin.    5    2.7660    0.1234      (----*---)
L. escul.   14    4.9486    0.6662                                (-*--)
                                      -+---------+---------+---------+-----
Pooled StDev =    0.5147           2.0       3.0       4.0       5.0
```

One-Way Analysis of Variance for Fruit Length in F_3

```
Analysis of Variance
Source      DF        SS        MS        F        P
Groups       3    29.504     9.835    62.40    0.000
Error       28     4.413     0.158
Total       31    33.917

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----------+---------+---------+------
L.esc.cer.  6    2.8600    0.3350              (---*---)
L.parvifl.  7    1.9171    0.1835  (---*---)
L.pimpin.   5    2.6120    0.4455          (----*---)
L. escul.  14    4.2550    0.4687                                (--*--)
                                        ----------+---------+---------+------
Pooled StDev =   0.3970                        2.40      3.20      4.00
```

One-Way Analysis of Variance for Fruit Ratio in F_3

```
Analysis of Variance
Source      DF        SS        MS        F        P
Groups       3    0.0786    0.0262     1.04    0.392
Error       28    0.7079    0.0253
Total       31    0.7865

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
L.esc.cer.  6    1.1150    0.1280      (----------*----------)
L.parvifl.  7    1.2329    0.1779                (----------*---------)
L.pimpin.   5    1.0880    0.1530   (-----------*-----------)
L. escul.  14    1.1764    0.1623               (------*------)
                                        --+---------+---------+---------+----
Pooled StDev =   0.1590                0.96      1.08      1.20      1.32
```

One-Way Analysis of Variance for Solid Soluble Content in F_3

```
Analysis of Variance
Source      DF        SS        MS        F        P
Groups       3    17.203     5.734    30.84    0.000
Error       28     5.206     0.186
Total       31    22.409

                                        Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
L.esc.cer.  6    6.2667    0.2251         (-----*----)
L.parvifl.  7    7.2429    0.4158                          (---*----)
L.pimpin.   5    7.1600    0.3286                        (----*-----)
L. escul.  14    5.5786    0.5177  (---*--)
                                        ----+---------+---------+---------+--
Pooled StDev =   0.4312              5.60      6.30      7.00      7.70
```

One-Way Analysis of Variance for Seed Weight in F_3

```
Analysis of Variance
Source      DF        SS        MS        F        P
Groups       3     8.476     2.825    22.79    0.000
Error       28     3.471     0.124
Total       31    11.948

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev  ---------+---------+---------+-------
L.esc.cer.   6    2.2933    0.4307            (-----*-----)
L.parvifl.   7    1.9229    0.3447     (----*-----)
L.pimpin.    5    1.9040    0.1426    (-----*------)
L. escul.   14    3.0429    0.3669                              (---*---)
                                    ---------+---------+---------+-------
Pooled StDev =    0.3521                   2.00      2.50      3.00
```

Part 11 Statistical analyses of genetic indices for morphological characters in F_2 and F_3 generation presented in Table 5.3. Non-parametric test Kruskal-Wallis was carried out for non-normally distributed indices A and A_e ; ANOVA was carried out for predominantly normally distributed genetic indices H_s and I.

Kruskal-Wallis Test for A in F_2

```
Kruskal-Wallis Test
Groups       N    Median     Ave Rank      Z
L.esc.cera.  7     1.300         9.6   -1.80
L.parvifl.   7     1.500        18.8    1.59
L.pimpin.    5     1.600        22.1    2.28
L.escul.     9     1.300        10.7   -1.67
Overall     28                  14.5

H = 10.51  DF = 3  P = 0.015
H = 10.67  DF = 3  P = 0.014 (adjusted for ties)
```

Kruskal-Wallis Test for A_e in F_2

```
Kruskal-Wallis Test
Groups       N     Median    Ave Rank      Z
L.esc.cera.  7     1.251         9.1   -1.99
L.parvifl.   7     1.393        18.6    1.51
L.pimpin.    5     1.516        23.0    2.55
L.escul.     9     1.277        10.8   -1.65
Overall     28                  14.5

H = 11.87  DF = 3  P = 0.008
H = 11.87  DF = 3  P = 0.008 (adjusted for ties)
```

One-Way Analysis of Variance for H_s in F_2

```
Analysis of Variance
Source     DF        SS        MS        F        P
Group       3   0.08139   0.02713     6.04    0.003
Error      24   0.10785   0.00449
Total      27   0.18924

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level       N      Mean     StDev  ---------+---------+---------+-------
L.esc.cera. 7   0.13400   0.06962   (------*-------)
L.parvifl.  7   0.23814   0.07339                  (------*------)
L.pimpin.   5   0.26880   0.05940                     (-------*--------)
L.escul.    9   0.15322   0.06357      (------*-----)
                                      ---------+---------+---------+-------
Pooled StDev =  0.06703                    0.140     0.210     0.280
```

One-Way Analysis of Variance for I in F_2

```
Analysis of Variance
Source     DF        SS        MS        F        P
Group       3   0.16401   0.05467     5.90    0.004
Error      24   0.22251   0.00927
Total      27   0.38652

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level       N      Mean     StDev  ---------+---------+---------+-------
L.esc.cera. 7   0.19343   0.10236   (------*-------)
L.parvifl.  7   0.34214   0.10248                  (------*-------)
L.pimpin.   5   0.38340   0.08288                    (--------*--------)
L.escul.    9   0.22000   0.09298      (------*------)
                                      ---------+---------+---------+-------
Pooled StDev =  0.09629                    0.20      0.30      0.40
```

Kruskal-Wallis Test for A in F_3

```
Kruskal-Wallis Test
Groups      N    Median     Ave Rank      Z
L.escul.   11     1.300         12.5  -1.01
L.esc.cera. 5     1.200          7.7  -2.04
L.parvifl.  7     1.500         19.2   1.75
L.pimpin.   5     1.500         19.0   1.35
Overall    28                   14.5

H = 7.83  DF = 3  P = 0.050
H = 7.97  DF = 3  P = 0.047 (adjusted for ties)
```

Kruskal-Wallis Test for A_e in F_3

```
Kruskal-Wallis Test
Groups      N    Median     Ave Rank      Z
L.escul.   11     1.266         13.3  -0.64
L.esc.cera. 5     1.129          8.0  -1.95
L.parvifl.  7     1.354         18.6   1.51
L.pimpin.   5     1.327         18.0   1.05
Overall    28                   14.5

H = 5.99  DF = 3  P = 0.112
```

One-way Analysis of Variance for I in F_3

```
Analysis of Variance for I
Source     DF        SS        MS        F        P
Groups      3    0.0975    0.0325     3.17    0.043
Error      24    0.2461    0.0103
Total      27    0.3436

                                  Individual 95% CIs For Mean
                                  Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
L.escul.   11    0.2265    0.0886            (------*-----)
L.esc.cer.  5    0.1576    0.1273  (---------*--------)
L.parvifl.  7    0.3154    0.0993                    (-------*------)
L.pimpin.   5    0.3114    0.1044                  (--------*--------)
                                  ----+---------+---------+---------+--
Pooled StDev =   0.1013             0.10      0.20      0.30      0.40
```

One-way Analysis of Variance for H_s in F_3

```
Analysis of Variance for Hs
Source     DF        SS        MS        F        P
Groups      3   0.04368   0.01456     2.83    0.060
Error      24   0.12334   0.00514
Total      27   0.16702

                                  Individual 95% CIs For Mean
                                  Based on Pooled StDev
Level       N      Mean     StDev  ----+---------+---------+---------+--
L.escul.   11   0.15755   0.06196            (------*-----)
L.esc.cer.  5   0.11020   0.09110  (---------*--------)
L.parvifl.  7   0.21600   0.06898                    (-------*-------)
L.pimpin.   5   0.21400   0.07617                  (---------*--------)
                                  ----+---------+---------+---------+--
Pooled StDev =  0.07169             0.070     0.140     0.210     0.280
```

Part 12 Statistical analyses of genetic indices for microsatellite markers in F_2 and F_3 generation presented in Table 5.5. Non-parametric test Kruskal-Wallis was carried out for non-normally distributed indices A and A_e ; ANOVA was carried out for predominantly normally distributed genetic indices H_s and I.

Kruskal-Wallis Test for A in F_2

```
Kruskal-Wallis Test on Na
Groups           N     Median     Ave Rank      Z
L. escul.       14     1.000         19.0    -2.52
L.esc.cers.     14     2.000         31.0     0.66
L.parvif.       14     2.000         35.0     1.72
L.pimpinel.     14     2.000         29.0     0.13
Overall         56                   28.5

H =  7.32  DF = 3  P = 0.062
H = 10.40  DF = 3  P = 0.015 (adjusted for ties)
```

Kruskal-Wallis Test for A_e in F_2

```
Kruskal-Wallis Test on Ne
Groups           N     Median     Ave Rank      Z
L. escul.       14     1.000         19.0    -2.51
L.esc.cers.     14     1.975         35.2     1.77
L.parvif.       14     1.926         35.3     1.79
L.pimpinel.     14     1.662         24.5    -1.05
Overall         56                   28.5

H = 10.29  DF = 3  P = 0.016
H = 10.88  DF = 3  P = 0.012 (adjusted for ties)
```

One-way ANOVA for H_s in F_2

```
Analysis of Variance
Source     DF        SS         MS         F        P
Groups      3    0.5650     0.1883      4.06    0.011
Error      52    2.4116     0.0464
Total      55    2.9766
                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level            N      Mean     StDev  ---------+---------+---------+------
-
L. escul.       14    0.2379    0.2264  (-------*-------)
L.esc.cers.     14    0.3523    0.2313                  (------*-------)
L.parvif.       14    0.4042    0.1740                     (-------*-------)
L.pimpinel.     14    0.2724    0.2247           (-------*-------)
                                         ---------+---------+---------+------
-
Pooled StDev =   0.2154                        0.15      0.30      0.45
```

One-way ANOVA for I in F_2

```
Analysis of Variance
Source     DF        SS         MS         F        P
Groups      3    1.1081     0.3694      4.07    0.011
Error      52    4.7226     0.0908
Total      55    5.8307
                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level            N      Mean     StDev  ---------+---------+---------+------
-
L. escul.       14    0.3230    0.3168   (-------*-------)
L.esc.cers.     14    0.4902    0.3218                  (--------*-------)
L.parvif.       14    0.5691    0.2432                     (-------*--------
)
L.pimpinel.     14    0.3924    0.3165            (-------*-------)
                                         ---------+---------+---------+------
-
Pooled StDev =   0.3014                        0.20      0.40      0.60
```

Kruskal-Wallis Test for A in F_3

```
Kruskal-Wallis Test

Groups         N     Median     Ave Rank      Z
L.escul.      14      2.000         26.5  -0.53
L.esc.ceras.14      2.000         26.5  -0.53
L.parvif.     14      2.000         26.5  -0.53
L.pimpinel.   14      2.000         34.5   1.59
Overall       56                    28.5

H = 2.53  DF = 3  P = 0.471
H = 3.67  DF = 3  P = 0.300 (adjusted for ties)
```

Kruskal-Wallis Test for A_e in F_3

```
Kruskal-Wallis Test
Groups         N     Median     Ave Rank      Z
L.escul.      14      1.471         21.8  -1.78
L.esc.ceras.14      1.499         27.9  -0.15
L.parvif.     14      1.612         27.7  -0.21
L.pimpinel.   14      1.835         36.6   2.14
Overall       56                    28.5

H = 5.85  DF = 3  P = 0.119
H = 6.15  DF = 3  P = 0.105 (adjusted for ties)
```

One-way ANOVA for H_s in F_3

```
Analysis of Variance for Hs
Source      DF        SS        MS        F        P
Groups       3    0.2364    0.0788     1.71    0.176
Error       52    2.3913    0.0460
Total       55    2.6277

                      Individual 95% CIs For Mean
                                 Based on Pooled StDev
Level         N      Mean     StDev  -+---------+---------+---------+-----
L.escul.     14    0.2233    0.2051  (---------*--------)
L.esc.ceras.14    0.2570    0.2370     (--------*---------)
L.parvif.    14    0.2604    0.2381     (---------*--------)
L.pimpinel.  14    0.3932    0.1703                (---------*--------)
                                 -+---------+---------+---------+-----
Pooled StDev = 0.2144            0.12      0.24      0.36      0.48
```

One-way ANOVA for I in F_3

```
Analysis of Variance
Source      DF        SS        MS        F        P
Groups       3    0.4435    0.1478     1.59    0.202
Error       52    4.8277    0.0928
Total       55    5.2712

                                 Individual 95% CIs For Mean
                                 Based on Pooled StDev
Level         N      Mean     StDev  ---------+---------+---------+-------
L.escul.     14    0.3304    0.3004  (----------*----------)
L.esc.ceras.14    0.3658    0.3337    (----------*----------)
L.parvif.    14    0.3697    0.3355    (----------*----------)
L.pimpinel.  14    0.5578    0.2392                (----------*----------)
                                 ---------+---------+---------+-------
Pooled StDev = 0.3047                  0.30      0.45      0.60
```

Part 13 Statistical analyses of genetic indices for RAPD markers in F_2 and F_3 generation presented in Table 5.6. Non-parametric test Kruskal-Wallis was carried out for non-normally distributed indices A and A_e ; ANOVA was carried out for predominantly normally distributed genetic indices H_s and I.

Kruskal-Wallis Test for A in F_2

```
Kruskal-Wallis Test

Groups              N      Median     Ave Rank      Z
L. escul.         118      1.000        217.0    -1.79
L.esc.ceras.      118      1.000        273.0     3.36
L. parvif.        118      1.000        247.0     0.97
L. pimpinel.      118      1.000        209.0    -2.53
Overall           472                   236.5

H = 16.36  DF = 3  P = 0.001
H = 25.82  DF = 3  P = 0.000 (adjusted for ties)
```

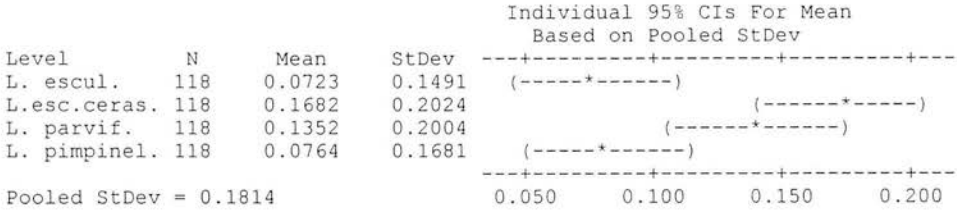
Kruskal-Wallis Test for A_e in F_2

```
Kruskal-Wallis Test
Groups              N      Median     Ave Rank      Z
L. escul.         118      1.000        213.1    -2.15
L.esc.ceras.      118      1.000        271.4     3.21
L. parvif.        118      1.000        249.3     1.18
L. pimpinel.      118      1.000        212.2    -2.24
Overall           472                   236.5

H = 15.97  DF = 3  P = 0.001
H = 24.15  DF = 3  P = 0.000 (adjusted for ties)
```

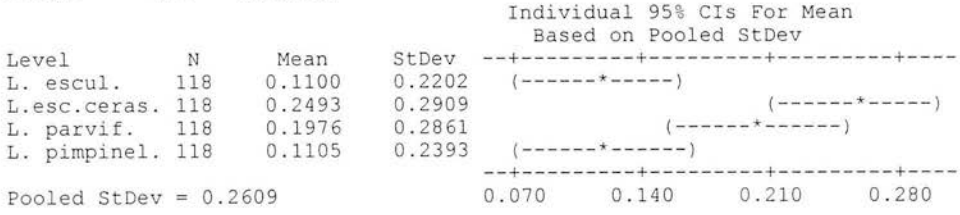
One-way ANOVA for H_s in F_2

```
Analysis of Variance
Source      DF        SS        MS         F        P
Groups       3    0.7710    0.2570      7.81    0.000
Error      468   15.4005    0.0329
Total      471   16.1715
```



One-way ANOVA for I in F_2

```
Analysis of Variance
Source      DF        SS        MS         F        P
Groups       3    1.6695    0.5565      8.18    0.000
Error      468   31.8456    0.0680
Total      471   33.5151
```



Kruskal-Wallis Test for A in F_3

```
Kruskal-Wallis Test
Groups              N      Median     Ave Rank      Z
L.escul.          118      1.000        221.0    -1.43
L.esc.ceras.      118      1.000        265.0     2.62
L.parvif.         118      1.000        241.0     0.41
L.pimpinel.       118      1.000        219.0    -1.61
Overall           472                   236.5

H =  8.75  DF = 3  P = 0.033
H = 13.22  DF = 3  P = 0.004 (adjusted for ties)
```

Kruskal-Wallis Test for A_e in F_3

```
Kruskal-Wallis Test
Groups              N      Median     Ave Rank       Z
L.escul.          118      1.000        221.5    -1.38
L.esc.ceras.      118      1.000        266.3     2.74
L.parvif.         118      1.000        241.4     0.45
L.pimpinel.       118      1.000        216.9    -1.80
Overall           472                   236.5

H =  9.64  DF = 3  P = 0.022
H = 13.83  DF = 3  P = 0.003 (adjusted for ties)
```

One-way ANOVA for H_s in F_3

```
Analysis of Variance
Source      DF        SS        MS         F        P
Groups       3    0.4966    0.1655      4.03    0.008
Error      468   19.2288    0.0411
Total      471   19.7254

                                      Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level         N      Mean     StDev  -------+---------+---------+---------
L.escul.    118    0.1101    0.1925    (------*------)
L.esc.ceras.118    0.1840    0.2235                  (-------*------)
L.parvif.   118    0.1472    0.2085           (------*-------)
L.pimpinel. 118    0.1027    0.1841  (-------*------)
                                     -------+---------+---------+---------
Pooled StDev = 0.2027                   0.100     0.150     0.200
```

One-way ANOVA for I in F_3

```
Analysis of Variance
Source      DF        SS        MS         F        P
Groups       3    1.0034    0.3345      4.03    0.008
Error      468   38.8794    0.0831
Total      471   39.8828

                                      Individual 95% CIs For Mean
                                        Based on Pooled StDev
Level         N      Mean     StDev  ----+---------+---------+---------+--
L.escul.    118    0.1588    0.2741    (--------*--------)
L.esc.ceras.118    0.2645    0.3144                      (--------*--------)
L.parvif.   118    0.2125    0.2975             (--------*--------)
L.pimpinel. 118    0.1495    0.2642  (---------*--------)
                                     ----+---------+---------+---------+--
Pooled StDev = 0.2882                0.120     0.180     0.240     0.300
```

Part 14 ANOVA of means from percentage of damaged fruits by South American tomato pinworm presented in Table 5.8.

One-way ANOVA: Percentage of damaged fruits by tomato pinworm.

```
Analysis of Variance
Source      DF        SS        MS         F        P
Groups       4     11909      2977     18.03    0.000
Error       34      5615       165
Total       38     17524

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level         N      Mean     StDev  --------+---------+---------+--------
L.esculen.   16     54.76     13.27                          (--*---)
L.esc.cera.   7     30.90     14.70            (---*----)
L.parvifl.    7     14.74      8.57   (----*----)
L. pimpin.    5     24.08      7.15       (-----*-----)
Tomato Par.   4     62.98     18.56                            (------*------)
                                     --------+---------+---------+--------
Pooled StDev =     12.85                   20        40        60
```

Appendix 4

Appendix 4 Part 1 Morphological characteristics of F1 generation Lycopersicon spp inter- and intra-taxon crosses.

Crosses	Hypocotyl Colour	Growth Type	Leaf Type	Corolla Colour	Style Position	Immature Fruit Colour	Fruit Pubescence
I – 3911	½ purple	Indeterminate	Standard	Yellow	Inserted	Greenish-white	Sparse
I – 1139	½ purple	Indeterminate	Standard	Yellow	Same level of stamens	Green	Sparse
I – 3915	½ purple	Indeterminate	Standard	Yellow	Inserted	Light green	Sparse
I – 1539	½ purple	Indeterminate	Standard	Yellow	Inserted	Light green	Sparse
I – 5219	½ purple	Indeterminate	Parviflorum	Yellow	Slightly exserted	Light green	Sparse
I – 1952	Purple	Indeterminate	Parviflorum	Yellow	Slightly exserted	Light green	Sparse
I – 5222	½ purple	Indeterminate	Parviflorum	Yellow	Same level of stamens	Light green	Sparse
I – 2252	Purple	Indeterminate	Parviflorum	Yellow	Slightly exserted	Light green	Sparse
I – 1160	½ purple	Indeterminate	Standard	Yellow	Inserted	Light green	Sparse

I = inter-taxon cross; E = intra-taxon cross; first two digits correspond to female parent and seconds two to male parent; details of each cross in Table 4.5; details of morphological character in Appendix 1.

Appendix 4 Part 1 Morphological characteristics of F1 generation Lycopersicon spp inter- and intra-taxon crosses.

Crosses	Fruit Shape	Fruit Size	Mature Fruit Colour	Flesh Colour	Fruit Cross Shape	Shape of Pistil Scar	Number of locules	Blossom End Shape
I – 3911	Rounded	Very small	Red	Orange	Round	Dot	2	Flat
I – 1139	Slightly flattened	Very small	Red	Orange	Round	Dot	3	Flat
I – 3915	Slightly flattened	Small	Red	Orange	Round	Dot	4	Flat
I – 1539	Slightly flattened	Very small	Red	Orange	Angular	Dot	3	Flat
I – 5219	Rounded	Very small	Red	Orange	Round	Dot	2	Flat
I – 1952	Rounded	Very small	Red	Orange	Round	Dot	2	Flat
I – 5222	Rounded	Very small	Red	Red	Round	Dot	2	Flat
I – 2252	Rounded	Very small	Red	Red	Round	Dot	2	Flat
I – 1160	Slightly flattened	Very small	Red	Orange	Round	Dot	2	Flat

I = inter-taxon cross; E = intra-taxon cross; first two digits correspond to female parent and seconds two to male parent; details of each cross in Table 4.5; details of morphological character in Appendix 1.

Appendix 4 Part 2 Morphological characteristics of F1 generation Lycopersicon spp inter- and intra-taxon crosses.

Crosses	Hypocotyl Colour	Growth Type	Leaf Type	Corolla Colour	Style Position	Immature Fruit Colour	Fruit Pubescence
I – 1960	½ purple	Indeterminate	Pimpinellifolium	Yellow	Inserted	Light green	Sparse
I – 2239	½ purple	Indeterminate	Standard	Yellow	Same level of stamens	Greenish-white	Sparse
I – 2260	Purple	Indeterminate	Pimpinellifolium	Yellow	Inserted	Greenish-white	Sparse
E – 1115	½ purple	Indeterminate	Standard	Yellow	Same level of stamens	Green	Sparse
E – 1511	½ purple	Indeterminate	Standard	Yellow	Inserted	Green	Sparse
E – 1119	½ purple	Indeterminate	Standard	Yellow	Inserted	Green	Sparse
E – 1911	½ purple	Indeterminate	Standard	Yellow	Inserted	Light green	Sparse
E – 1121	½ purple	Indeterminate	Standard	Yellow	Same level of stamens	Light green	Sparse
E – 2111	½ purple	Indeterminate	Standard	Yellow	Same level of stamens	Green	Sparse

I = inter-taxon cross; E = intra-taxon cross; first two digits correspond to female parent and seconds two to male parent; details of each cross in Table 4.5; details of morphological character in Appendix 1.

Appendix 4 Part 2 Morphological characteristics of F1 generation Lycopersicon spp inter- and intra-taxon crosses.

Crosses	Fruit Shape	Fruit Size	Mature Fruit Colour	Flesh Colour	Fruit Cross Shape	Shape of Pistil Scar	Number of locules	Blossom End Shape
I – 1960	Rounded	Very small	Red	Red	Round	Dot	2	Flat
I – 2239	Rounded	Very small	Red	Orange	Round	Dot	2	Flat
I – 2260	Cylindrical	Very small	Red	Red	Round	Dot	2	Indented
E – 1115	Slightly flattened	Small	Red	Orange	Angular	Stellate	>3	Flat
E – 1511	Slightly flattened	Intermediate	Red	Orange	Angular	Dot	>4	Indented
E – 1119	Rounded	Small	Red	Red	Round	Dot	3	Flat
E – 1911	Rounded	Small	Red	Red	Angular	Dot	2	Pointed
E – 1121	Rounded	Small	Red	Orange	Angular	Dot	2	Pointed
E – 2111	Rounded	Small	Red	Red	Angular	Dot	2	Pointed

I = inter-taxon cross; E = intra-taxon cross; first two digits correspond to female parent and seconds two to male parent; details of each cross in Table 4.5; details of morphological character in Appendix 1.

Appendix 4 Part 3 Morphological characteristics of F1 generation Lycopersicon spp inter- and intra-taxon crosses.

Crosses	Hypocotyl Colour	Growth Type	Leaf Type	Corolla Colour	Style Position	Immature Fruit Colour	Fruit Pubescence
E – 1519	½ purple	Indeterminate	Standard	Yellow	Inserted	Green	Sparse
E – 1915	½ purple	Indeterminate	Standard	Yellow	Inserted	Light green	Sparse
E – 1522	½ purple	Indeterminate	Standard	Yellow	Same level of stamens	Greenish-white	Sparse
E – 2215	½ purple	Indeterminate	Standard	Yellow	Inserted	Green	Sparse
E – 1922	½ purple	Indeterminate	Standard	Yellow	Inserted	Light green	Sparse
E – 2219	½ purple	Indeterminate	Standard	Yellow	Inserted	Light green	Sparse
E – 1521	½ purple	Indeterminate	Standard	Yellow	Same level of stamens	Green	Sparse
E – 2211	½ purple	Indeterminate	Standard	Yellow	Same level of stamens	Green	Sparse
E – 2221	½ purple	Indeterminate	Standard	Yellow	Inserted	Green	Sparse

I = inter-taxon cross; E = intra-taxon cross; first two digits correspond to female parent and seconds two to male parent; details of each cross in Table 4.5; details of morphological character in Appendix 1.

Appendix 4 Part 3 Morphological characteristics of F1 generation Lycopersicon spp inter- and intra-taxon crosses.

Crosses	Fruit Shape	Fruit Size	Mature Fruit Colour	Flesh Colour	Fruit Cross Shape	Shape of Pistil Scar	Number of locules	Blossom End Shape
E – 1519	Slightly flattened	Intermediate	Red	Orange	Angular	Stellate	4	Flat
E – 1915	Slightly flattened	Small	Red	Orange	Angular	Stellate	4	Flat
E – 1522	Rounded	Small	Red	Pink	Irregular	Stellate	3	Pointed
E – 2215	Rounded	Small	Red	Pink	Angular	Stellate	3	Flat
E – 1922	Rounded	Small	Red	Orange	Angular	Dot	2	Flat
E – 2219	Rounded	Small	Red	Orange	Angular	Dot	2	Flat
E – 1521	Rounded	Small	Red	Red	Angular	Dot	2	Flat
E – 2211	Rounded	Small	Red	Orange	Angular	Dot	2	Pointed
E – 2221	Rounded	Small	Red	Orange	Angular	Dot	2	Flat

I = inter-taxon cross; E = intra-taxon cross; first two digits correspond to female parent and seconds two to male parent; details of each cross in Table 4.5; details of morphological character in Appendix 1.

Appendix 5

Appendix 5. Means of crosses and parental accessions of continuous traits in F_1 generation of inter- and intra-taxon crosses.

Cross/Accession	Fruit diameter (cm)	Fruit length (cm)	Fruit ratio (D/L)
	**	**	**
I – 3911	2.73 bc (± 0.10)	2.58 bc (± 0.11)	1.06 b (± 0.02)
I – 1139	2.56 bc (± 0.16)	2.26 bc (± 0.12)	1.14 b (± 0.03)
Expected value	**3.81**	**3.05**	**1.20**
39 – *L. esculentum* var. *cerasiforme*	1.80 bd (± 0.05)	1.67 bd (± 0.05)	1.08 b (± 0.01)
11 – Limachino	5.86 a (± 0.21)	4.42 a (± 0.14)	1.33 a (± 0.03)
	**	**	**
I – 3915	3.23 bc (± 0.15)	2.80 bc (± 0.11)	1.15 bd (± 0.02)
I – 1539	2.18 bd (± 0.13)	1.78 bd (± 0.18)	1.27 bc (± 0.05)
Expected value	**4.28**	**3.20**	**1.25**
39 – *L. esculentum* var. *cerasiforme*	1.80 bd (± 0.05)	1.67 bd (± 0.05)	1.08 bd (± 0.01)
15 - Ace	6.76 a (± 0.45)	4.74 a (± 0.26)	1.42 a (± 0.02)
	**	**	*
I – 5219	1.98 bc (± 0.05)	1.79 bc (± 0.06)	1.11 ac (± 0.01)
I – 1952	1.89 bc (± 0.04)	1.73 bc (± 0.03)	1.09 bc (± 0.01)
Expected value	**2.34**	**1.97**	**1.19**
52 – *L. parviflorum*	1.03 bd (± 0.02)	0.87 bd (± 0.02)	1.19 ac (± 0.01)
19 – Lukullus	3.65 a (± 0.11)	3.07 a (± 0.12)	1.20 a (± 0.05)
	**	**	**
I – 5222	1.68 bc (± 0.06)	1.59 bc (± 0.05)	1.06 bc (± 0.02)
I - 5222	1.69 bc (± 0.07)	1.62 bc (± 0.06)	1.04 bc (± 0.01)
Expected value	**1.93**	**2.84**	**0.90**
52 – *L. parviflorum*	1.03 bd (± 0.02)	0.87 bd (± 0.02)	1.19 a (± 0.01)
22 – San Marzano	2.82 a (± 0.10)	4.82 a (± 0.36)	0.61 bd (± 0.05)
	**	**	**
I – 1160	2.55 bc (± 0.07)	2.29 bc (± 0.05)	1.11 b (± 0.02)
Expected value	**3.65**	**2.85**	**1.23**
11- Limachino	5.86 a (± 0.21)	4.42 a (± 0.14)	1.33 a (± 0.03)
60 – *L. pimpinellifolium*	1.44 bd (± 0.06)	1.29 bd (± 0.07)	1.13 b (± 0.02)

Appendix 5. Means of crosses and parental accessions of continuous traits in F_1 generation of inter- and intra-taxon crosses.

Cross/Accession	Fruit diameter (cm)	Fruit length (cm)	Fruit ratio (D/L)
	**	**	**
I – 1960	2.11 bc (± 0.09)	2.00 bc (± 0.11)	1.06 bc (± 0.02)
Expected value	**2.54**	**2.18**	**1.16**
19 – Lukullus	3.65 a (± 0.11)	3.07 a (± 0.12)	1.20 a (± 0.05)
60 – *L. pimpinellifolium*	1.44 bd (± 0.06)	1.29 bd (± 0.07)	1.13 ac (± 0.02)
	**	**	**
I – 2239	2.95 a (± 0.14)	2.99 bc (± 0.08)	0.99 ac (± 0.05)
Expected value	**2.31**	**3.24**	**0.84**
22 – San Marzano	2.82 a (± 0.10)	4.82 a (± 0.36)	0.61 bd (± 0.05)
39 – *L. esculentum* var. *cerasiforme*	1.80 b (± 0.05)	1.67 bd (± 0.05)	1.08 a (± 0.01)
	**	**	**
I – 2260	2.07 bc (± 0.11)	2.27 bc (± 0.14)	0.91 bc (± 0.01)
Expected value	**2.13**	**3.05**	**0.87**
22 – San Marzano	2.82 a (± 0.10)	4.82 a (± 0.36)	0.61 bd (± 0.05)
60 – *L. pimpinellifolium*	1.44 bd (± 0.06)	1.29 bd (± 0.07)	1.13 a (± 0.02)
	**	**	ns
E – 1115	4.44 bd (± 0.15)	3.55 bc (± 0.20)	1.28 (± 0.06)
E - 1511	5.30 ac (± 0.57)	3.98 ac (± 0.35)	1.33 (± 0.05)
Expected value	**6.31**	**4.58**	**1.37**
11- Limachino	5.86 a (± 0.21)	4.42 a (± 0.14)	1.33 (± 0.03)
15 - Ace	6.76 a (± 0.45)	4.74 a (± 0.26)	1.42 (± 0.02)
	**	**	**
I – 1119	3.95 b (± 0.16)	3.52 b (± 0.11)	1.13 bc (± 0.04)
I – 1911	3.60 b (± 0.23)	3.28 b (± 0.22)	1.10 bc (± 0.02)
Expected value	**4.76**	**3.75**	**1.26**
11- Limachino	5.86 a (± 0.21)	4.42 a (± 0.14)	1.33 a (± 0.03)
19 – Lukullus	3.65 b (± 0.11)	3.07 b (± 0.12)	1.20 ac (± 0.05)

Appendix 5. Means of crosses and parental accessions of continuous traits in F_1 generation of inter- and intra-taxon crosses.

Cross/Accession	Fruit diameter (cm)	Fruit length (cm)	Fruit ratio (D/L)
	**	**	**
E –1121	3.77 b (± 0.05)	3.62 bc (± 0.17)	1.05 b (± 0.02)
E – 2111	3.22 b (± 0.13)	2.99 bde (± 0.13)	1.08 b (± 0.03)
Expected value	**4.84**	**3.94**	**1.22**
11 – Limachino	5.86 a (± 0.21)	4.42 a (± 0.14)	1.33 a (± 0.03)
21 – Marglobe	3.81 b (± 0.14)	3.45 bce (± 0.16)	1.11 b (± 0.02)
	**	**	**
E – 1519	5.01 bc (± 0.23)	3.93 bc (± 0.13)	1.28 ac (± 0.04)
E – 1915	4.80 bc (± 0.09)	3.98 bc (± 0.12)	1.21 bc (± 0.04)
Expected value	**5.20**	**3.91**	**1.31**
15 - Ace	6.76 a (± 0.45)	4.74 a (± 0.26)	1.42 a (± 0.02)
19 – Lukullus	3.65 bd (± 0.11)	3.07 bd (± 0.12)	1.20 bc (± 0.05)
	**	ns	**
E – 1522	4.34 bc (± 0.18)	4.27 (± 0.13)	1.02 bc (± 0.03)
E – 2215	3.92 bce (± 0.30)	3.87 (± 0.34)	1.05 bc (± 0.09)
Expected value	**4.79**	**4.78**	**1.02**
15 - Ace	6.76 a (± 0.45)	4.74 (± 0.26)	1.42 a (± 0.02)
22 – San Marzano	2.82 bde (± 0.10)	4.82 (± 0.36)	0.61 bd (± 0.05)
	**	**	**
E – 1922	3.97 a (± 0.09)	3.90 bc (± 0.05)	1.02 bc (± 0.02)
E - 2219	3.30 bc (± 0.16)	3.40 bce (± 0.12)	0.97 bc (± 0.03)
Expected value	**3.24**	**3.94**	
19 – Lukullus	3.65 ac (± 0.11)	3.07 bde (± 0.12)	1.20 a (± 0.05)
22 – San Marzano	2.82 bd (± 0.10)	4.82 a (± 0.36)	0.61 bd (± 0.05)
	**	**	**
E – 1521	4.24 b (± 0.29)	4.07 ac (± 0.30)	1.05 b (± 0.04)
Expected value	**5.29**	**4.09**	**1.27**
15 - Ace	6.76 a (± 0.45)	4.74 a (± 0.26)	1.42 a (± 0.02)
21 – Marglobe	3.81 b (± 0.14)	3.45 bc (± 0.16)	1.11 b (± 0.02)

Appendix 5. Means of crosses and parental accessions of continuous traits in F_1 generation of inter- and intra-taxon crosses.

Cross/Accession	Fruit diameter (cm)	Fruit length (cm)	Fruit ratio (D/L)
	**	**	**
E – 2211	3.61 bc (± 0.18)	3.68 bc (± 0.17)	0.99 bc (± 0.06)
Expected value	**4.34**	**4.62**	**0.97**
22 – San Marzano	2.82 bd (± 0.10)	4.82 a (± 0.36)	0.61 bd (± 0.05)
11 – Limachino	5.86 a (± 0.21)	4.42 ac (± 0.14)	1.33 a (± 0.03)
	**	**	**
E – 2221	3.60 a (± 0.16)	4.02 ac (± 0.17)	0.90 bc (± 0.04)
Expected value	**3.32**	**4.13**	**0.86**
22 – San Marzano	2.82 b (± 0.10)	4.82 a (± 0.36)	0.61 bd (± 0.05)
21 – Marglobe	3.81 a (± 0.14)	3.45 bc (± 0.16)	1.11 a (± 0.02)

Appendix 5. Means of crosses and parental accessions of continuous traits in F_1 generation of inter- and intra-taxon crosses.

Cross/Accession	Fruit weight (g)	Solid soluble content (brix°)	Weight 1,000 seeds (g)
	**	**	**
I – 3911	12.42 b (± 1.38)	9.11 a (± 0.31)	2.56 bc (± 0.03)
I – 1139	9.41 b (± 1.42)	8.07 bc (± 0.32)	1.98 bde (± 0.03)
Expected value	**44.93**	**6.10**	**2.16**
39 – *L. esculentum* var. *cerasiforme*	3.46 b (± 0.33)	7.09 bc (± 0.24)	1.26 bdf (± 0.05)
11 – Limachino	86.21 a (± 9.70)	5.10 bd (± 0.13)	2.97 a (± 0.02)
	**	*	**
I – 3915	18.35 b (± 1.97)	7.41 a (± 0.44)	2.49 bde (± 0.04)
I – 1539	6.11 b (± 1.34)	7.62 a (± 0.43)	3.48 a (± 0.07)
Expected value	**65.27**	**6.43**	**2.18**
39 – *L. esculentum* var. *cerasiforme*	3.46 b (± 0.33)	7.09 a (± 0.24)	1.26 bdf (± 0.05)
15 - Ace	127.07 a (± 22.79)	5.77 b (± 0.46)	3.00 bde (± 0.01)
	**	**	**
I – 5219	4.60 bc (± 0.38)	8.94 bc (± 0.20)	2.64 bc (± 0.03)
I – 1952	4.17 bce (± 0.17)	8.53 bce (± 0.15)	2.36 bde (± 0.07)
Expected value	**12.79**	**8.72**	**2.15**
52 – *L. parviflorum*	0.68 bde (± 0.03)	10.76 a (± 0.31)	1.13 bdf (± 0.02)
19 – Lukullus	24.90 a (± 1.98)	6.68 bde (± 0.24)	3.18 a (± 0.05)
	**	**	**
I – 5222	2.99 b (± 0.25)	7.59 bc (± 0.22)	2.35 a (± 0.04)
I - 2252	3.60 b (± 0.34)	8.47 bc (± 0.21)	2.07 bc (± 0.04)
Expected value	**11.73**	**8.48**	**1.36**
52 – *L. parviflorum*	0.68 b (± 0.03)	10.76 a (± 0.31)	1.13 bdf (± 0.02)
22 – San Marzano	22.79 a (± 2.45)	6.20 bd (± 0.16)	1.59 bde (± 0.02)
	**	**	**
I – 1160	9.65 b (± 0.62)	7.75 ac (± 0.67)	1.84 bc (± 0.03)
Expected value	**44.10**	**6.85**	**1.91**
11- Limachino	86.21 a (± 9.70)	5.10 bd (± 0.13)	2.97 a (± 0.02)
60 – *L. pimpinellifolium*	1.44 b (± 0.06)	1.29 a (± 0.07)	1.13 bd (± 0.02)

Appendix 5. Means of crosses and parental accessions of continuous traits in F_1 generation of inter- and intra-taxon crosses.

Cross/Accession	Fruit weight (g)	Solid soluble content (brix°)	Weight 1,000 seeds (g)
	**	**	**
I – 1960	6.17 b (± 0.94)	6.70 b (± 0.45)	1.92 b (± 0.03)
Expected value	**13.44**	**7.64**	**2.02**
19 – Lukullus	24.90 a (± 1.98)	6.68 b (± 0.24)	3.18 a (± 0.05)
60 – *L. pimpinellifolium*	1.98 b (± 0.21)	8.59 a (± 0.20)	0.86 c (± 0.03)
	**	**	**
I – 2239	14.83 b (± 1.17)	7.35 a (± 0.32)	2.23 a (± 0.01)
Expected value	**13.12**	**6.65**	**1.47**
22 – San Marzano	22.79 a (± 2.45)	6.20 bd (± 0.16)	1.59 b (± 0.02)
39 – *L. esculentum* var. *cerasiforme*	3.46 c (± 0.33)	7.09 ac (± 0.24)	1.26 c (± 0.05)
	**	**	**
I – 2260	5.75 b (± 0.82)	7.58 b (± 0.42)	1.99 a (± 0.05)
Expected value	**12.38**	**7.40**	**1.23**
22 – San Marzano	22.79 a (± 2.45)	6.20 c (± 0.16)	1.59 b (± 0.02)
60 – *L. pimpinellifolium*	1.98 b (± 0.21)	8.59 a (± 0.20)	0.86 c (± 0.03)
	**	**ns**	**
E – 1115	40.38 bc (± 4.20)	5.43 (± 0.22)	3.39 a (± 0.06)
E - 1511	68.10 bc (± 18.57)	4.62 (± 0.34)	2.78 bd (± 0.04)
Expected value	**106.64**	**5.44**	**2.99**
11- Limachino	86.21 ac (± 9.70)	5.10 (± 0.13)	2.97 bc (± 0.02)
15 - Ace	127.07 a (± 22.79)	5.77 (± 0.46)	3.00 bc (± 0.01)
	**	**	**
I – 1119	32.86 b (± 3.07)	4.94 bc (± 0.27)	3.41 a (± 0.03)
I – 1911	27.54 b (± 5.08)	5.94 ac (± 0.36)	2.46 bd (± 0.03)
Expected value	**55.55**	**5.89**	**3.07**
11- Limachino	86.21 a (± 9.70)	5.10 bc (± 0.13)	2.97 bc (± 0.02)
19 – Lukullus	24.90 b (± 1.98)	6.68 a (± 0.24)	3.18 bc (± 0.05)

Appendix 5. Means of crosses and parental accessions of continuous traits in F_1 generation of inter- and intra-taxon crosses.

Cross/Accession	Fruit weight (g)	Solid soluble content (brix°)	Weight 1,000 seeds (g)
	**	*	**
E –1121	30.51 b (± 4.56)	5.32 ac (± 0.22)	3.62 a (± 0.09)
E – 2111	19.26 b (± 2.56)	6.77 a (± 0.76)	3.22 b (± 0.02)
Expected value	**58.32**	**5.17**	**2.71**
11- Limachino	86.21 a (± 9.70)	5.10 bc (± 0.13)	2.97 c (± 0.02)
21 – Marglobe	30.42 b (± 3.46)	5.23 ac (± 0.27)	2.45 d (± 0.03)
	**	ns	**
E – 1519	58.61 b (± 6.15)	5.42 (± 0.18)	3.46 bc (± 0.05)
E – 1915	55.65 b (± 3.27)	5.74 (± 0.36)	4.71 a (± 0.03)
Expected value	**75.98**	**6.23**	**3.09**
15 - Ace	127.07 a (± 22.79)	5.77 (± 0.46)	3.00 bd (± 0.01)
19 – Lukullus	24.90 b (± 1.98)	6.68 (± 0.24)	3.18 bd (± 0.05)
	**	ns	**
E – 1522	44.61 b (± 4.08)	5.41 (± 0.55)	2.96 bc (± 0.04)
E – 2215	36.74 b (± 8.24)	7.20 (± 0.75)	3.20 a (± 0.10)
Expected value	**74.93**	**5.99**	**2.30**
15 - Ace	127.07 a (± 22.79)	5.77 (± 0.46)	3.00 ac (± 0.01)
22 – San Marzano	22.79 b (± 2.45)	6.20 (± 0.16)	1.59 bd (± 0.02)
	**	ns	**
E – 1922	33.83 a (± 1.27)	6.19 (± 0.54)	3.21 a (± 0.08)
E - 2219	23.77 b (± 3.11)	5.71 (± 0.21)	3.20 a (± 0.06)
Expected value	**23.84**	**6.44**	**2.38**
19 – Lukullus	24.90 b (± 1.98)	6.68 (± 0.24)	3.18 a (± 0.05)
22 – San Marzano	22.79 b (± 2.45)	6.20 (± 0.16)	1.59 b (± 0.02)
	**	ns	**
E – 1521	48.19 b (± 9.40)	5.14 (± 0.20)	3.70 a (± 0.10)
Expected value	**78.75**	**5.50**	**2.73**
15 - Ace	127.07 a (± 22.79)	5.77 (± 0.46)	3.00 b (± 0.01)
21 – Marglobe	30.42 b (± 3.46)	5.23 (± 0.27)	2.45 c (± 0.03)

Appendix 5. Means of crosses and parental accessions of continuous traits in F_1 generation of inter- and intra-taxon crosses.

Cross/Accession	Fruit weight (g)	Solid soluble content (brix°)	Weight 1,000 seeds (g)
	**	**	**
E – 2211	27.16 b (± 3.44)	7.27 a (± 0.45)	3.45 a (± 0.07)
Expected value	**54.50**	**5.65**	**2.28**
22 – San Marzano	22.79 b (± 2.45)	6.20 b (± 0.16)	1.59 c (± 0.02)
11- Limachino	86.21 a (± 9.70)	5.10 c (± 0.13)	2.97 b (± 0.02)
	ns	**ns**	**
E – 2221	26.76 (± 2.79)	5.31 (± 0.62)	3.39 a (± 0.25)
Expected value	**26.60**	**5.72**	**2.02**
22 – San Marzano	22.79 (± 2.45)	6.20 (± 0.16)	1.59 c (± 0.02)
21 – Marglobe	30.42 (± 3.46)	5.23 (± 0.27)	2.45 b (± 0.03)