STATUS OF SOLANACEAE CROPS RESEARCH	2
TOMATO SOLANUM LYCOPERSICON	2
Genetic resources:	
Genomic resources:	3
POTATO SOLANUM TUBEROSUM	3
Genetic resources:	
Genomic resources:	
PEPPER CAPSICUM SPP	
Genetic resources:	
Genomic resources:	
EGGPLANT (SOLANUM MELONGENA)	
Genetic Resources:	
Genomic Resources:	
PETUNIA PETUNIA HYBRIDA	
TREE TOMATO SOLANUM BETACEUM	
PEPINO SOLANUM MURICATUM	
NARANJILLA SOLANUM QUITOENSE	
COFFEE COFFEA ARABICA	
Genetic resources:	
Genomic resources:	

This section summarizes the status of the research in Solanaceae crops and the available genetic and genomic resources. Although an effort was made to list all the major relevant information, this record is far from complete. We will amend the catalog with additional information as it becomes available. One of the first tasks of the Steering Committee will be to provide the community an updated 'on-line' report of the status of the research for each of the Solanaceae crops.

Most Solanaceae species have the same basic chromosome number (x=12) suggesting that large scale genomic changes have involved chromosomal inversions and/or interchanges - a prediction has been largely born out by comparative mapping studies. The two most closely related crop species, tomato and potato, differ by only five paracentric inversions. The pepper and eggplant genomes have more inversions/exchanges relative to tomato and potato, nonetheless, detailed comparative maps now allow cross species syntenic comparisons among all four genomes (Tanksley *et al.* 1988, Livingstone et al. 1999, Doganlar et al. 2002). These comparative maps have not only shed light on the nature of genome evolution, but have also facilitated comparative mapping studies of qualitative and quantitative traits - especially those involved in domestication. These comparative maps (as well as those now under construction for petunia and coffee) will also provide the infrastructure by which the tomato genome sequence can be directly utilized by researchers working on virtually all Solanaceae species.

TOMATO (Solanum lycopersicon)

Tomato originated in the New World and all its wild relatives are native of the Andean region is South America. Recent data indicates that accessions of *L. esculentum* var. *cerasiforme*, that were imported to Europe in the 15th century, had already reached an advanced stage of cultivation in the center of domestication in Mexico. Initially, in Europe, tomato was thought to be poisonous plant however, selection and breeding, mainly in Italy, resulted in increased popularity of the tomato due to its appearance, taste and its nutritional value. The related wild species of *L. esculentum* are: *L. pimpinellifolium*, *L. cheesmanii*, *L. peruvianum*, *L. chilense*, *L. hirsutum*, *L. parviflorum*, *L. chmielewskii* and *L. pennellii*. The following *Solanum* species can be hybridized to the culativate tomato: *S. lycopersicoides*, *S. ochrantum* and *S. sitiens*. Tomato is the genetic centerpiece of the Solanaceae family due to its modest-sized diploid genome and tolerance to inbreeding. As a result, a vast knowledge has accumulated on tomato genetics, cytogenetics and physiology.

Genetic resources:

A consensus mapping population: The primary mapping resource for the entire SOL project is an F2 population derived from a cross of *L. esculentum* and *L. pennellii* (F2.2000; http://www.sgn.cornell.edu/). This F2 population, which was developed in the Tanksley's lab, includes 84 plants and the population is continually propagated from sterile cuttings. Large amounts of DNA and cuttings of F2.2000 are available for distribution along with data for approximately 2000 segregating markers as well as plant phenotypes. This population is currently being used by Dutch Scientists (KeyGene) to construct a joint physical map for the sequencing project (see technical paper).

Introgression lines: Recent research in tomato has demonstrated that despite their inferior phenotypes, wild species contain genes that can improve performance of cultivated varieties. A major effort in the tomato community has been to construct interspecific introgression lines - a set of nearly isogenic lines each containing marker defined segments from wild species accessions in a uniform genetic background of a cultivated variety. The following genetic

resources were developed and are being distributed through the Tomato Genetics Resource Center at Davis:

- 1) L. pennellii (LA 716; Eshed and Zamir 1995).
- 2) L. hirsutum (LA1777; Monforte and Tanksley 2000).
- 3) Solanum lycopersicoides (LA2951; Ji and Chetelat 2003).

Mutants and other stocks: More than 2000 mutants have been characterized in tomato and are publicly available (http://tgrc.ucdavis.edu/) along with numerous cytogenetic stocks and wild species accessions. A saturated mutagenesis resource was developed recently in the genetic background of the inbred variety M82. A total of 13,000 M2 families, derived from EMS and fast-neutron mutagenesis, were phenotyped and 3,417 mutations can be searched according to phenotypic categories and seed ordered on line (http://zamir.sgn.cornell.edu/mutants/). Further insertional mutants and others are available if the dwarf variety Microtom. Germplasm are also available through the USDA GRIN resource at http://www.ars-grin.gov/cgi-bin/npgs/html/splist.pl?7056. The USDA GRIN system comprises the 4830 L. esculentum accessions, and 267 v. cerasiforme accessions curated and distributed at the Plant Genetic Resources Unit in Geneva, NY.

Genomic resources:

Bioinformatics: SGN (Solanaceae Genome Network; http://www.sgn.cornell.edu/) is a website that provides a virtual workbench for scientists doing research on species in the family Solanaceae. This relational database (MySQL) hosts all published gene/EST sequences from Solanaceae species in a comparative genomics format. SGN is designed to provide a common entry point to various sources of data and analysis tools, integrating both local and remote resources.

BAC libraries: A BAC (17X *Hind* III) library from the variety Heinz 1706 was constructed by the lab of Rod Wing who also constructed, based on this resource, a Finger Print Contig (FPC) physical map (www.genome.arizona.edu/fpc/tomato).

ESTs: A total of 160,000 ESTs were analyzed from 15 different libraries of tomato, which are annotated in SGN and can be blasted on this site or in a tomato site in TIGR. In addition, microarray hybridization slides of 10,000 unigenes were made available to the community and the first transcript profiles of tomato fruits are made available in SGN.

Sequencing of chromosome 6 of tomato was already initiated in Holland and assignments of chromosomes to other countries is shown in Figure 10 (Page 28).

POTATO Solanum tuberosum

The cultivated potato Solanum tuberosum belong to a very large genus that includes 160 tuber bearing species eight of which are cultivated. Potato (x=12) has a number of ploidy levels ranging from diploid (2n=24) to triploids, tetraploids, pentaploids and hexaploids; most cultivated varieties are autotetraploid (4n=48). The tetrasomic inheritance of potato is responsible for the fact that the genetic basis of the majority of the agronomic traits is poorly understood. The primary evolution of the potato, which is a highly heterozygous species, has been at the diploid level. The high level of heterozygosity is attributed to the self-incompatibility mechanism, which has been enhanced by the asexual propagation by tubers.

Domestication of the potato took place in the Bolivian-Peruvian Andes as early as 8,000 years ago. When the Spanish conquerors reached the Andes in the early 1500s, they found the

Incas growing potatoes; by the end of the 18th century potatoes had become the food staple in most of Europe. In 1845 and 1846, late blight disease virtually destroyed the Irish potato crops causing a famine that claimed 1 million lives.

Genetic resources:

The Potato Genebank aims to facilitate improvements in the potato of the future by promoting the use of valuable exotic genes found in wild potato germplasm. Small tubers are typical of wild potato species but they represent an invaluable treasure of diversity for potentially useful traits that may someday be bred into new varieties that will be able to overcome the challenges of pests and stresses with less dependence on chemical fertilizers, insecticides and fungicides. NRSP-6 is using a 5-fold approach: introduction, classification, preservation, evaluation and distribution of potato germplasm http://www.ars-grin.gov/ars/MidWest/NR6/. Valuable potato germplasm is being collected, maintained and investigated in CIP http://www.cipotato.org/index2.htm. Potato germplams available through USDA GRIN can be obtained at http://www.ars-grin.gov/cgi-bin/npgs/html/splist.pl?11264.

Genomic resources:

BAC libraries:

- Canada (Qin Chen): Two BAC libraries (6X coverage) were constructed from the wild Mexican diploid potato, *Solanum pinnatisectum*, as a step in the characterization of disease and insect resistance genes important for potato (Chen et al. 2003).
- The Netherlands (Visser): Potato BAC libraries are being ordered and sequenced by Biosystems Genomics (http://www.biosystemsgenomics.nl/)

ESTs:

- US (Baker et al.) Approximately 100,000 ESTs from diverse libraries yielding 25,000 unigenes were generated for potato http://www.potatogenome.org/nsf3/ and hosted by The TIGR Potato Gene Index (StGI) StGI Release 7.0 April 24, 2003 and by SGN.
- Canada (Flinn) 100,000 ESTs are being generated both from 3' and 5' from 16 diverse libraries including normalized libraries.
- SNP and InDel data from diverse di- and tetraploid potato strains is accessible via GabiPD from the GABI-CONQUEST project. The data can be searched via GreenCards. https://gabi.rzpd.de/projects/pomamo.shtml.

PEPPER Capsicum spp

The basic chromosome number of the genus *Capsicum* is 12 and all reported species are diploids. Pepper originated in the New World, where at least 20 species have been described. Columbus introduced chiles to Europe and this unique vegetable was almost immediate incorporation to local recipes. From there, chiles quickly expanded to Africa and Asia where it has become an important component of the local diet. Domestication of the *Capsicum* species has taken place in Mexico and Guatemala through to Colombia, and independently in the Andean regions of Peru and Bolivia. There are five domesticated *Capsicum* species: *C. annuum*, *C. frutescens*, *C. chinense*, *C. baccatum* and *C. pubescens*.

Genetic resources:

- Brazil (Leonardo Silva Boiteux)

C. annuum 'AG 672' x C. chinense 'CNPH 679'

C. annuum 'Majorca' x C. annuum 'CNPH 148'

C. annuum 'Azeth' x C. annuum 'CNPH 143'

C. baccatum x C. baccatum

- France (Alain Palloix)

C. annuum 'H3' x 'Vania'

'Perennial' x 'Yolo Wonder'

'Yolo Wonder' x 'Criollo de Morelos 334'

'Yolo Wonder' x 'Criollo de Morelos 334'

Vania x C. baccatum 'Pen 79'

- Israel (Ilan Paran)

C. frutescens 'BG 2816' x C.annuum 'Maor'

C.annuum 'Maor' x C.annuum 'Perrenial'

C.annuum '5226' x C.chinense 'PI 159234'

C.chinense 'PI 152225' in background of C.annuum

- Korea (Byung-Dong Kim)

C.chinense 'Habanero' x C. annuum 'TF68'

C.annuum 'CM334' x C. annuum 'Chilsung'

C.annuum 'ECW123R' x C. annuum 'CM334'

- USA (Molly Jahn)

C. annuum 'NuMex RNaky' x C. chinense 'PI 159234' F2; RILs in progress.

C. frutescens BG2914-6 x C. annuum RNaky F2; RILs in progress

C. chinense 159234 x C. chinense Habanero F2

C. frutescens BG2814-6 x C. chinense 149234

- USA (USDA) refer to http://www.ars-grin.gov/cgi-bin/npgs/html/splist.pl?2056
- The Chile Pepper Institute at New Mexico State University provides information and Chiles germplasm http://www.chilepepperinstitute.org/.

Genomic resources:

ESTs:

- Korea (Doil Choi) 30,354 ESTs (http://plant.pdrc.re.kr/Gene/index.html).
- USA (Molly Jahn) 234 contigs from fruit-specific SSH library.

BAC libraries:

- Korea (Byung-Dong Kim, http://plaza.snu.ac.kr/~cpmgbr/)

Capsicum annuum 'CM334', 313,336 clones, 15 x haploid genome equivalents

- France (Abdelhafid Bendahmane)

Doubled haploid line 'HD208', 239,232 clones, 10x haploid genome equivalents

- USA (Jim Giovannoni and Molly Jahn). Inbred line from Mexican accession *C. frutescens* BG2816 220,000 9x haploid genome equivalents

EGGPLANT (Solanum melongena)

Unlike the other domesticated Solanum species which are endemic to the Americas, eggplant is of Old World origin. The exact origin of *Solanum melongena* is uncertain but it may have

been indirectly derived from the African wild species *Solanum incanum* and was domesticated in India and South-East China. Cultivation of the crop spread to the Mediterrenean during Arab conquests of the area starting in the 7th century. Evolutionary relationships between *Solanum melongena* and more than 300 other Solanum species that are considered as eggplant remain mostly unclear.

Genetic Resources:

Gene banks for eggplant and related species have been collected at the National Burea for Plant Genetic Resources of India, the National Gene Bank of China, the National Institute and Agrobiological Resources in Japan, the Asian Vegetable Research and Development Center in Taiwan, the Vavilov Institute in Russia, INRA in France, the University of Birmingham in the UK, the Nijmegen Botanical Garden in the Netherlands and the USDA Beltsville Research Station in the USA. Sexual compatibility between *Solanum melongena* and many other Solanum species has been extensively studied. Although many interspecific crosses produce viable, fertile hybrids, the use of wild germplasm for eggplant breeding has been very limited.

A molecular genetic linkage map based on tomato cDNA, genomic DNA, and EST markers was constructed for eggplant in a *Solanum linnaeanum* x *S. melongena* F2. The map consists of 12 linkage groups, spans 1480 cM, and contains 233 markers. Comparison of the eggplant and tomato maps revealed conservation of large tracts of colinear markers, a common feature of genome evolution in the Solanaceae and other plant families. Overall, eggplant and tomato were differentiated by 28 rearrangements, which could be explained by 23 paracentric inversions and five translocations during evolution from the species' last common ancestor.

Genomic Resources:

Genomic resources for eggplant are limited compared to other Solanaceae crops. A comparative RFLP based molecular genetic linkage map has recently been constructed, 5000 ESTs sequences are available (SGN) and BAC library has been constructed.

PETUNIA Petunia hybrida

Petunia is a native to South America and the MesoAmerica and was derived from a cross between *P. axillaire* and *P. violancea*. This important ornamental plant (n=7) has a genome size that is lightly larger than tomato (1.2.Mbp). Petunias make very popular hanging basket petunia, also called surfinia, as developed by Suntory (Japan). Yearly turn over worldwide can be estimated at around \$500 million. Petunia plants can be propagated by seed or by vegetative production.

In research, Petunia has played an important role in a range of topics, although never really as a frontrunner. Research on flower colour has started in the late fifties in The Netherlands and France enabled the definition of over thirty genes, involved in flavonoid synthesis. Petunia was used by Monsanto's Rob Horsch to develop the leaf disk genetic transformation system. The first mapping of transgenes, the discovery of antisense and co-suppression effects in plants and, more recently the development of refined transposon-tagging and transposon-display approaches have been major developments and achievements. The petunia system contributes significantly to research in areas as the comparative biology of floral and plant development and -architecture, flavonoid synthesis, meiosis and gamete formation and -function and pollination syndrome analysis.

Most of the groups that use petunia as a major research subject have joined an as yet loosely organized structure, the major activity of which is the organisation of the so-called World Petunia Days: a gathering of petunia researchers to present their latest research, to discuss technological developments and exchanges of material and people. The WPDs will be organised for the sixth time in march 2004 (see http://www.botany.unibe.ch/petunia/). Germplasm is available through the USDA GRIN system at http://www.ars-grin.gov/cgibin/npgs/html/splist.pl?9183.

TREE TOMATO Solanum betaceum

The tree tomato was originally named *Solanum betaceum*, then was transferred to the genus *Cyphomandra*, and recently has been returned to *Solanum* (Bohs 1995; Bohs and Olmstead 1997; Olmstead and Palmer 1997). The tree tomato has long been popular in the Andes, where it is native, and in many other parts of the tropics. It has been grown successfully in New Zealand and has been exported to the north temperature zone for around two decades. The fruit is eaten raw, as juice, preserves, jellies, and as vegetable, either cooked or raw in salads. Several wild species of the tree tomato in South America also have edible fruits. Various parts of the plants, particularly the leaves, of the tree tomato and wild species have been used in folk medicine in Latin America (Heiser and Anderson 1999). A wild ancestor of the tree tomato has not yet been identified unequivocally but recent studies indicate that it most likely had its origin in the southern Andes, perhaps Bolivia, and that the cultivated tree tomato was derived from a wild form of the same species (Bohs 1998).

PEPINO Solanum muricatum

The pepino (cachum in Quechua), Solanum muricatum, shows some similarities to the tree tomato in that it is also native to the Andes where it is an important fruit and it has also recently been grown in New Zealand (Hewett 1993) where it is exported to the United States, Europe and Japan. Like the tree tomato, its future as a crop outside of the Andes is uncertain. There are many differences—the pepino is eaten raw and it was probably domesticated much earlier than the tree tomato; in addition, its origin is more complex and the two plants are very different in appearance. Pepino is the Spanish word for cucumber, and the plant was so named by the Spanish because of a slight resemblance to some cucumbers. Contrary to the situation in the tree tomato, and similar to the naranjilla (see following) there is no clearly defined wild ancestral form of the pepino. Several studies have identified the wild species S. caripense, S. tabanoense, and S. basendopogon that are most closely related to the pepino, and that have been treated as possible progenitors (Anderson 1979; Anderson et al. 1996; Anderson and Jansen 1998; Jansen et al. 1998). The wild species are found in Costa Rica, Colombia, Ecuador, and Peru (Heiser and Anderson 1999).

NARANJILLA Solanum quitoense

The naranjilla or lulo (S. quitoense) is a tall plant (ca. 2 m), bears globose berries of about 5 cm in diameter, orange and covered with short stiff hairs which have usually rubbed off by the time the fruits have reached the markets. The pulp is green and gives a green juice, the form in which it is usually used. The taste is unique, but has been described by some as like

that of a mixture of pineapples and strawberries. The plant was found by the Spanish in Ecuador and Colombia (Patiño 1963) where most of it is still grown. It was introduced to Panama and Costa Rica in the middle of this century where it is also grown today. No wild ancestor has been reported and it is most unlikely that *S. quitoense* descends directly from any other species in section *Lasiocarpa*. In Colombia and Central America the plants usually have prickles on the stems and leaves whereas in Ecuador plants are unarmed.

Root knot nematodes and various insect pests and fungal diseases limit its production in all of the countries where it is presently grown (National Research Council 1989). There are two promising recent developments toward its improvement: the use of interspecific hybrids with the cocona (S. sessiliflorum) and the introduction of a nematode resistant variety. Although the naranjilla is a perennial and ought to yield for several years, because of these parasites it is often grown as an annual in many places. Resistance to nematodes was discovered in plants of S. hirtum grown in the greenhouse at Indiana University (Heiser 1971). Seeds of hybrids were sent to plant breeders in several countries and a resistant cultivar has now been produced in Colombia from F_2 seeds (Bernal et al. 1998a, b). The plant was backcrosssed to S. quitoense twice, and material from F_2 plants of the BC_2 was multiplied by tissue culture. Tests in various parts of the country showed good performance. In addition the new plant, called "lulo la selva" grows better in full sun, has fruit of better quality, outyields the traditional lulo and is also spineless. The resistant cultivar is propagated vegetatively and it was formally released in June of 1998 (Heiser and Anderson 1999).

COFFEE Coffea arabica

Today, coffee is the most valuable agricultural commodity in world trade, commanding a turnover of US\$ 10 billion annually and its ranks second just behind petrol on international trade exchanges. Coffee is the world's most important beverage and upon which the economies of many producing countries from the south depend. Much of the world coffee production is made from 20-30 traditional cultivars of *Coffea arabica* (66%) and *Coffea canephora* (34%), which have been released some years ago from simple selection and breeding programmes and generally propagated by seedlings. Unfortunately, coffee appears as an "orphan crop" due to the lack of resources and the very few public or private institutes invested in coffee research and especially in coffee genomics. The International Solanaceae Genome Project (SOL) is certainly the best opportunity for coffee research network to develop collaborative work in this particular area. Coffee belongs to Rubiaceae and is closely related to the Solanaceae family, as these two families are included in the Asterid I class.

Genetic resources:

The IAC in Brazil possess a significant germplasm collection of *Coffea* species, collected at their centers of origin. Also, the Institute is in charge of a significant Breeding Program, which developed most of all coffee cultivars planted in Brazil. Therefore, a large collection of selected hybrids and evaluated progenies is available for future analysis such as comparative mapping, identification of genes of agronomic importance, analysis of significant traits as overall cup quality.

In coffee, molecular marker technology has already been implemented in germplasm characterization and management, detecting genetically divergent breeding populations (e.g. to predict hybrid vigour), describing coffee phylogeny with related species and in marker-assisted breeding. Successful genetic transformation is still limited to characters controlled by

major genes and transgenic coffee plants have been produced for insect resistance and putative low caffeine content in seeds.

The main point of interest for coffee (which is of great interest Nestlé that invests tremendous resources in coffee genetics and genomics) is the construction of a syntheny network between the coffee and Solanaceae genomes such as tomato, potato, pepper and eggplant. In this regard, coffee will take advantage of the sequencing of the tomato genome due to the generation of shared markers. These markers are COS (Conserved Orthologous Set) come from the identification of unique genes which are highly conserved between plant species and their functions have been conserved during evolution. These genes are mainly involved in the primary processes of the plant cell as energy metabolism, transcription or cellular communication and these COS markers will be use as anchor markers to link the tomato and coffee genetic maps. The location of the syntheny region between the coffee and tomato genomes will determine, in the end, a "virtual sequencing" of these parts of coffee genome and consequently greatly improve the knowledge of the coffee genome. This work will lead to the discovery or genetic elucidation of the pathways and precursors determining, for example, quality traits in coffee.

Germplasms in USDA GRIN can be ordered online, the coffee germplasms can be viewed at http://www.ars-grin.gov/cgi-bin/npgs/html/splist.pl?2779 .

Genomic resources:

The Brazilian Coffee Genome Project is funded by three Institutions and programs; Fapesp, CBP&D do Café (Brazilian Consortium for Coffee Research and Development) and Embrapa/ Cenargen (Brazilian Agriculture Research Company / Center of Genetic Resources). A total of US\$ 700,000 was invested in the project. Most of the infrastructure utilized for sequencing was already set up, by the AEG group (Agriculture and Environment Genomes), funded by Fapesp. This group alone includes 24 laboratories, at several research institutions, each one equipped with an automated sequencer ABIPrism 377. Five central sequencing labs also possess an ABI3700. The CBP&D is also a national consortium of 44 Universities and Public Research Institutes that funds projects related with all aspects of coffee culture. The central coordination of the project is under responsibility of the IAC (Agronomic Institute of Campinas / SP). The Bioinformatic processing and analysis is under the responsibility of Dr. Gonçalo Guimarães Pereira at Unicamp (State University of Campinas/ SP), and the corresponding web page is www.lge.ibi.unicamp.br.

The project has a goal of 200,000 ESTs to be sequenced, and the deadline for the project is March/04. At present, 85% of all the ESTs were already sequenced. Clustering and annotation of those sequences are also been completed, and Blast searches performed in the database so far can identify genes of all biological categories and functions. Interestingly, the most significant hits are always with species such as tomato, potato and tobacco. A total of 30 distinct cDNA libraries were prepared for sequencing. Most of these are of *C. arabica cv. Mundo Novo*, and all possible organs, tissues and treatments were included, as follows:

- Leaves different developmental stages, from different branches (plagiotropic and orthotropic), treated with BION and arachidonic acid (non-specific defense inducers), inoculated with the fungus *Hemileia vastatrix*, the agent of leaf-rust; *Leucoptera coffeella*, a leaf-miner; and *Xylella*, a bacteria that infects branches.
- Hypocotyls tissues- regular and treated with Bion
- Roots different developmental stages, treated with bion and Al⁺².
- Flowers immature buds, and 4 different stages of development
- Fruits 4 different developmental stages
- \bullet Cells in suspension regular, treated with Bion, brassinosteroids, mannose, NaCl, KCl and Al^{+2} .

• Callus – embriogenic and non-embriogenic, with and without 2,4-D

Also, other libraries were prepared from:

- Seedlings of C. eugenioides, C. canephora, and C. racemosa.
- Callus of *C. canephora*.

Based on this information, we believe that including coffee, as part of the integrated Solanaceae project would represent a major contribution for improving the knowledge of coffee molecular basis, once the background information available at this moment is not much, and no genetic map is yet available for any *Coffea* species.