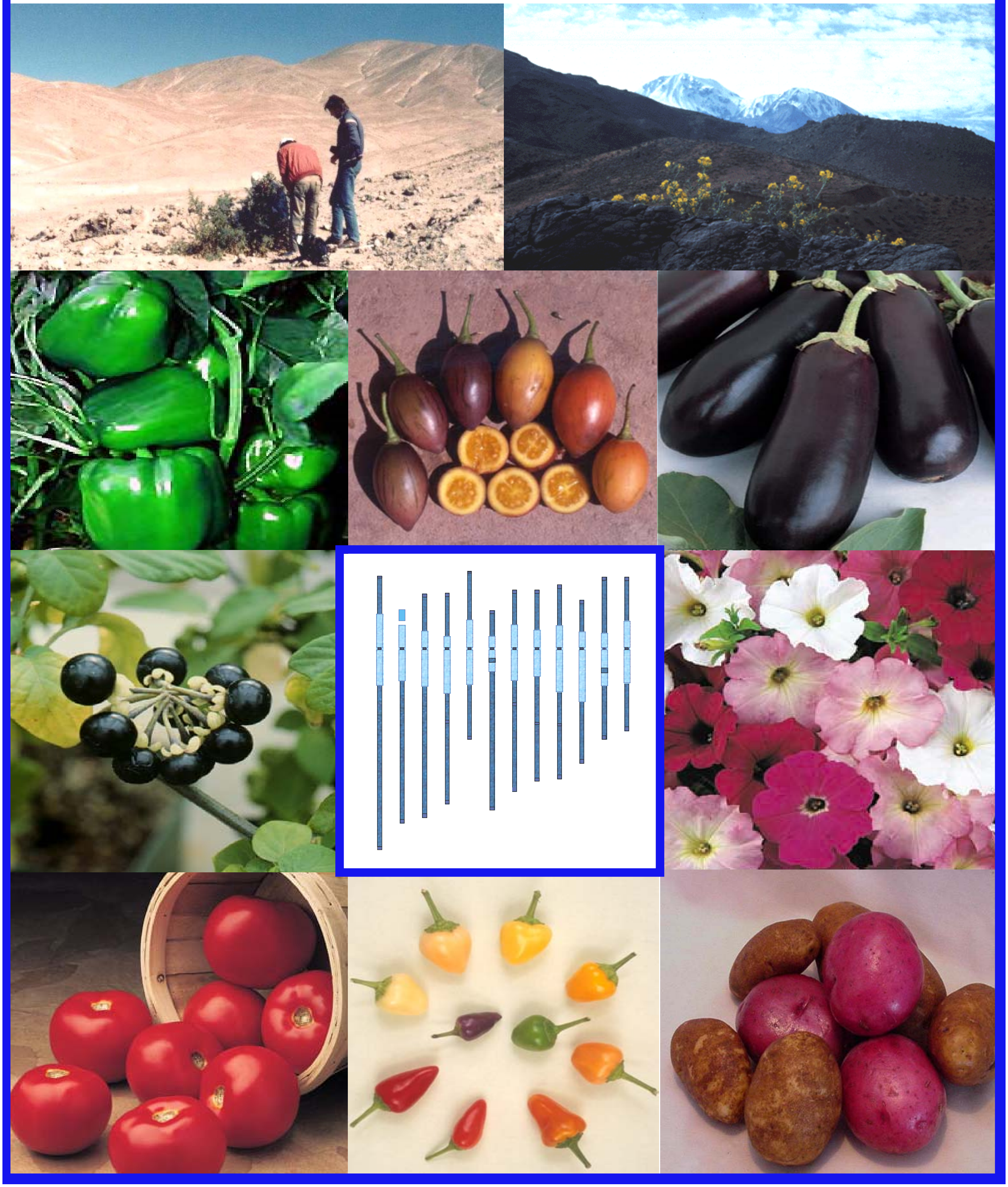
The International Solanaceae Genome Project (SOL): Systems Approach to Diversity and Adaptation



February 15, 2004

PREFACE

Included is the latest version of the SOL document describing the ten-year goals of the International Solanaceae Genome Initiative. This document presents the current international view and will be updated regularly to include new research objectives and additional participating countries.

Happy SOL

Table of Contents **Page**

<u>APPENDIX 1</u> – A TECHNICAL DOCUMENT FOR AN INTERNATIONAL CONSORTIUM TO SEQUENCE THE TOMATO GENOME.

<u>APPENDIX 2</u>- SOLANACEAE PROJECT SEQUENCING AND BIOINFORMATICS STANDARDS AND GUIDELINES.

The International Solanaceae Genome Project (SOL): Systems Approach to Diversity and Adaptation

Summary

The Questions: Modern biology is expanding our view on life from the reductionist approach - analyzing individual components of biological systems, to the holistic view - integrating entire genetic programs and the complex events they dictate. Information generated by large-scale genomic sequencing has led to a major revolution in biological sciences through the revelation of all the genes required to encode major life forms. One of the biggest surprises has been that organisms that are evolutionary and morphologically distinct, share a very similar gene/protein content and even conserved linkage groups (e.g. human and mouse). Over the coming 10 years the International Solanaceae Genome Project (SOL) will integrate diverse disciplines and research groups from around the world to create a coordinated network of knowledge about the Solanaceae family aimed at answering two of the most important questions about life and agriculture:

- How can a common set of genes/proteins give rise to such a wide range of morphologically and ecologically distinct organisms that occupy our planet?

The corollary question of agricultural importance is:

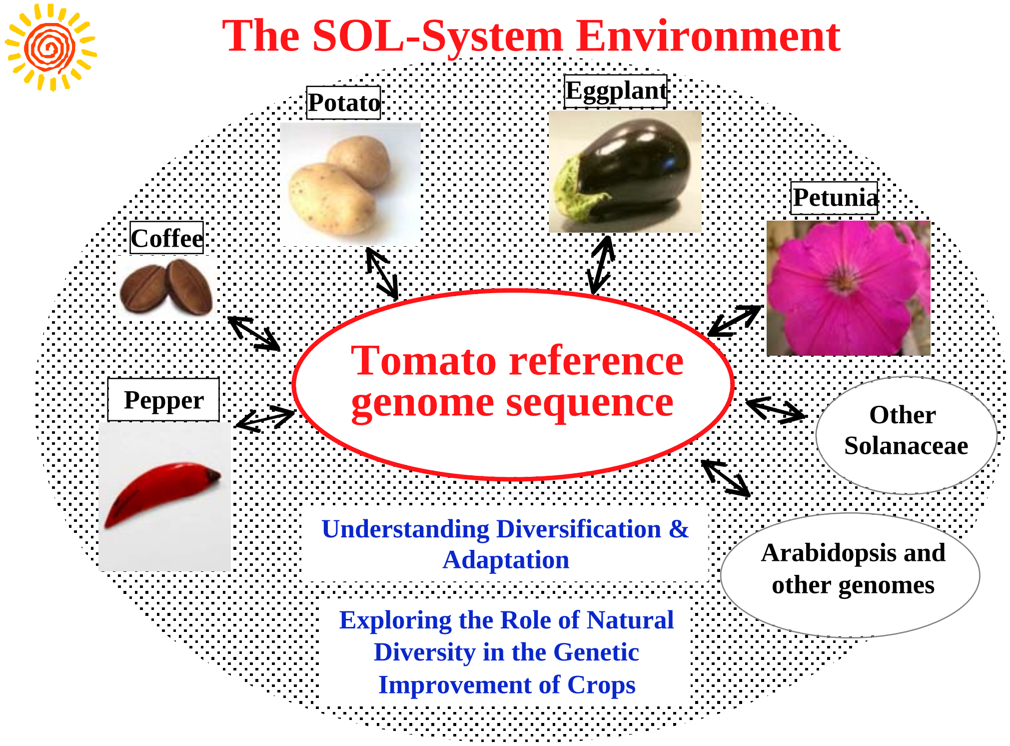
- How can a deeper understanding of the genetic basis of plant diversity be harnessed to better meet the needs of society in an environmentally-friendly and sustainable manner?

The Family: The family Solanaceae is ideally suited to address both of these questions. This taxon includes more than 3000 species many of which evolved in the Andean/Amazonian regions of South America in habitats that vary dramatically and include, rain forests that receive more than 3 meters of rainfall annually, deserts with virtually no rainfall and high mountains with regular snowfall and sub-freezing temperatures. The center of diversity of the Solanaceae is near the equator and thus species were undisturbed by the ice ages and have had time to accumulate adaptive genetic variation for extreme ecological niches.

The Solanaceae is the third most important plant taxon economically and the most valuable in terms of vegetable crops. It encompasses the most variable of crop species in terms of their agricultural utility, as it includes the tuber-bearing potato (a food staple over much of the world), a number of fruit-bearing vegetables (e.g. tomato, eggplant, peppers, husk tomato), ornamental flowers (petunias, *Nicotiana*), edible leaves (*Solanum aethiopicum*, *S. macrocarpon*), and medicinal plants (e.g. *Datura, Capsicum*). Seeds can also be included in this list if we include the closely allied species coffee. Fruit and tubers are major contributors of vitamins, fiber, carbohydrates, and phyto-nutrient compounds in our diet. The nutritional importance of fruit and vegetables is reflected in current USDA recommendations of five or more servings of fruit or vegetables a day for a healthy diet. The World Health Organization and the United Nations Food and Agriculture Organization (FAO) recently launched an effort to enhance fruit and vegetable consumption worldwide as low consumption is considered one of the top ten contributing factors to human mortality. The Solanaceae are unique in that multiple crop species in this family are major contributors to fruit and vegetable consumption and thus to our quality of life.

Solanaceous crops have been subjected to intensive human selection, allowing their use as models to study the evolutionary interface between plants and people. The ancient mode of Solanaceae evolution, coupled with an exceptionally high level of conservation of genome organization at the macro and micro levels, makes the family a model to explore the basis of phenotypic diversity and adaptation to natural and agricultural environments.

The Concept: The long-term goal of the SOL program is to create a network of map based resources and information to address key questions in plant adaptation and diversification. This will be done using the tools and philosophy of systems biology which is a multidisciplinary approach to tackle the complex interactions that occur at all levels of biological organization and their functional relationship to the organism as a whole. Moreover, from these studies we wish to provide a new outlook to how we value and utilize natural variation to impact the health and well being of humans in a more environmentally friendly and sustainable manner. Our international effort will not only impact Solanaceae biology but will also set the road map for implementing rational strategies for improvement of other crop species that are important to human nutrition.



Background on the Inception of SOL: On November 3, 2003 researchers, from more than 10 countries, representing academic and government research labs, industry and extension/outreach specialists with interest in the Solanaceae met for a full day in Washington DC to kick off the 10 year initiative entitled "The International Solanaceae Genome Project (SOL)". The forum united around a common set of tools, populations and concepts with a firm commitment to work together to elevate our level of understanding of the network of interactions that lead to population diversity and adaptation. The agreed upon course of action for the first stage of this initiative was: 1) to obtain high quality sequence of the tomato genome as a reference for Solanaceae plants as well as plants from other related taxa, 2) to display all data generated from around the world via a single virtual entry point for Solanaceae genomics, and 3) to establish a Steering Committee that will facilitate and coordinate research and funding for projects under the virtual umbrella of SOL.

This document is designed to assist scientists in preparing their SOL grant proposal. It includes contributions from many members of our community – comments and questions should be addressed to Dani Zamir <Zamir@agri.huji.ac.il>.

DIMENSIONS IN DIVERSITY AND ADAPTATION

How can one genome code for diverse adaptive outcomes?

A basic descriptor in comparative genomics is the degree of homology between genome sequences and phenotypes of evolutionary divergent species. For example, comparisons between sequenced genomes of worms, yeast and flies revealed that many of the proteins encoded by these eukaryotic genomes are similar. These conserved features, which are shared by many living organisms, reflect the unity of life. At the same time, the diversity of characteristics among the earth's millions of species is staggering. Understanding the genetic basis of traits that distinguish closely related taxonomic groups is one of the great challenges in biology. Some of the differences between taxa reflect neutral ticks of the molecular evolutionary clock while others are associated with features of adaptation that enhance survival and reproduction in unique ecosystems. Natural selection is the ultimate determinant of adaptation where over evolutionary time, mutations, recombination and the element of drift resulted in descendent species that are dramatically different from their last common ancestor.

With the sequencing of the human, mouse and chimpanzee genomes, the quest to discover the genetic basis of phenotypes that distinguish us is gaining momentum. Virtually all the protein-coding genes in human align with homologues in mice and the two genomes show a very high level of gene order conservation (synteny). Despite nearly 100 Myr of divergence, mouse and humans share a nearly gene-for-gene match between their genomes, raising the question of how a common set of proteins could lead to such dramatically different outcomes as a mouse and a human. Comparisons of the two genomes have begun to identify regulatory genomic regions, which substantially outnumber the protein repertoire, with potential functional roles in controlling gene expression and chromatin organization. This question of "which sequences make us human" must wait until more mammalian genomes, particularly apes, are fully sequenced to reveal new leads to explore the complexity of phenotypes[1].

It is now becoming clear that most of the traits involved in adaptation and diversification are polygenic and affect continuous or quantitative adaptive phenotypes. Moreover, most of these genetic changes are not "loss of function" mutations of the type induced in the laboratory. Rather, they are genetic variants that change the function of the proteins for which they code, or perhaps more often, the temporal and spatial expression of those genes. Understanding the nature of the genetic changes underlying adaptation and diversification, is a prerequisite to understand the basis of life and evolution. This will also enable us to fully appreciate and utilize the natural variation around us to better adapt plants to the need of humans in an environmentally sound and, sustainable manner. Due to a number of inherent attributes, plants are an ideal model for resolving the genetic basis of quantitative traits related to adaptation and diversification (the cloning of quantitative trait loci – QTL - was achieved in plants first). Some of these features include short generation time, large families and tolerance to inbreeding. Furthermore, the ability to generate segregating populations from divergent plant species that are adapted to different growth conditions facilitates the mapping of numerous QTL that affect fitness[2]. As genome sequences and phenotypes become available from a range of related plant species, we will be able to better understand how a common set of genetic building-blocks regulate the diverse outcomes that affect adaptation.

The Solanaceae family is an ideal model to explore the basis of diversity and adaptation

From the rich diversity of ~300,000 higher plant species on our planet, the genomes of a single dicot (Arabidopsis) and a single monocot (rice) have been sequenced. Solanaceae represent a unique portion of the family tree of plants and sequencing and exploration of tomato will enable comparative analysis for the discovery of distinct and common aspects of plant evolution. The family Solanaceae is anchored in a section of plants' evolutionary tree that is distant from both Arabidopsis and rice (Figure 1). Contained within these clades (Asterids I, II) are not only the Solanaceae crops, but also a number of other major crop plants such as coffee, lettuce, safflower and sunflower. The Solanaceae family includes more than 3000 species. *Solanum* is the largest genus in the family with approximately 1500 species. Extensive current and ongoing knowledge exists about systematics of the family, including a recent generic conspectus and up-to-date family-wide molecular phylogenies.

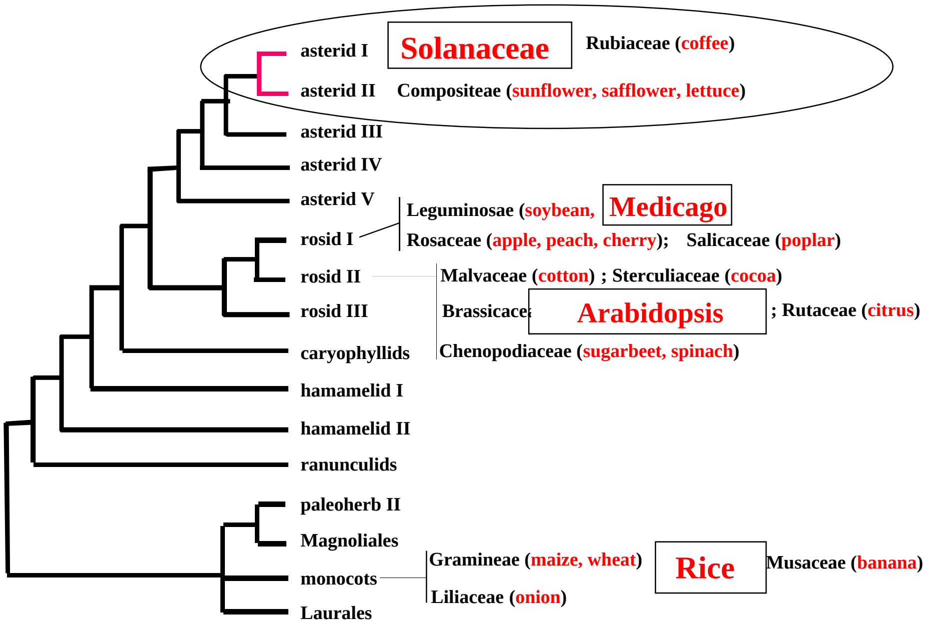


Figure 1: Solanaceae represent a unique portion of the family tree of flowering plants[3] and its sequencing will enable comparative analysis for the discovery of distinct and common aspects of plant evolution.

The Solanaceae is the third most valuable crop family exceeded only by the grasses (e.g. rice, maize, wheat) and legumes (e.g. soybean), and the most valuable in terms of vegetable crops. The family includes the fourth most important crop, potato, as well as a number of fruit-bearing vegetables such as tomato, pepper and eggplant. Moreover, the Solanaceae include species that are grown for their edible leaves (*Solanum aethiopicum*, *S. macrocarpon*), ornamentals like Petunia and many medicinal plants such as *Datura* and *Capsicum*. A crop grown for its seeds can also be included in this list if we adopt the closely allied species coffee. Multiple important species in the family are major contributors to fruit and vegetable consumption and thus human health.

Solanaceae species thrive in some of the most diverse natural habitats that include rain forests, deserts in which plants survive entirely on moisture from fog, to the high elevation Andean mountains where UV radiation is high and temperature plummet to sub freezing on a regular basis (Figure 2). Solanaceae species range in habit from tall forest trees and woody lianas to tiny annual herbs. Being near the equator and thus undisturbed by the ice ages, the Solanaceae have had time to adapt to diverse niches. Yet, despite this high level of phenotypic variation and ecological adaptation, the Solanaceae share very similar genomes and gene repertoire.

Figure 2: Solanaceae species evolved and are adapted to some of the most diverse and extreme habitats on earth.

The Solanaceae genome is highly conserved

In plants, the use of comparative genetic molecular mapping has revealed a high level of conservation of gene content and order within the grasses, crucifers, legumes and Solanaceae species[4]. Sequencing of Arabidopsis and rice has shown that more than 80% of the genes that have been annotated in Arabidopsis were also found in rice; however, nearly 50% of the predicted rice genes do not have a match in Arabidopsis thus providing a basis for the specificities of the two organisms[5]. Numerous episodes of polyploidy within both the grasses and Brassicaceae have led to segmental duplications, selective gene losses and significant genome reshuffling. As a result, species in the grasses and crucifers are characterized by different chromosome numbers coupled with extensive loss of microsynteny between the paralogous segments of Brassica chromosomes, and between those and their Arabidopsis homoeologs. The Solanaceae family is unique in that there have been no large-scale duplication events (e.g. polyploidy) early in the radiation of this family. The polyploidy events (e.g. tetraploid potatoes and tetraploid tobacco) are all recent events and the diploid forms of both of these species are still in existence. As a result, microsynteny conservation amongst the genomes of tomato, potato, pepper and eggplant is very high (Fig 6 Technical Sequencing Document- Appendix 1). This allows us to predict regions between genomes that are identical by descent and to study the evolution of sequence and function of orthologous genes – a key to understanding diversification and adaptation. The highly conserved genome organization, both at the macro- and microsyntenic levels, allows extension the information basis beyond the individual species thus creating a common map-based framework of knowledge. Hence, the first goal of SOL is to determine, with great precision, the nucleotide sequence of the tomato genome and link it to the Solanaceae map.

The tomato map-based genome will provide a reference to interpret the sequence organization of other Solanaceae crops and their wild relatives as the basis of understanding how plants diversify and adapt to new and adverse environments. Tomato was selected as a reference since it provides the smallest diploid genome (950 Mb) for which homozygous inbreds are available, as well as an advanced BAC based physical map to start the sequencing. Tomato is also the most intensively researched Solanaceae genome with simple diploid genetics, short generation time, routine transformation technology, and the availability of rich genetic and genomic resources. The tomato genome encodes approx. 35,000 genes[6], which are largely sequestered in contiguous euchromatic regions corresponding to less than 25% of the total DNA in the tomato nucleus (220~250 Mb of gene rich regions). Presently the *Solanaceae Genome Network* (SGN; http://www.sgn.cornell.edu/) hosts multiple information from diverse sources around the world in a (MySQL) relational database. SGN currently contains approximately 200,000 gene/EST sequences from tomato, potato, eggplant, pepper and petunia. As part of the SOL project, SGN will integrate with other related plant-genome-databases to provide a virtual workbench to explore phenotypic diversity in the highly conserved genomes of the Solanaceae.

Networks in physiology and biochemistry

The underlying genetic diversity in the Solanaceae is arrayed on a broad canvas of phenotypic variation, where the richer the genetic pool, the more diverse the resulting cellular processes and organismal complexity. Accordingly, genome-analysis projects that are undertaken within a framework of genetically diverse germplasm will, by definition, result in a far more profound and comprehensive understanding of biochemistry and physiology. Members of the Solanaceae collectively comprise precisely the germplasm diversity, and crucially the phenotypic diversity, that will prove essential for pushing forward the frontiers of plant biochemistry and physiology.

The plant kingdom is estimated to produce approximately 200,000 metabolites, many of which play specific roles in allowing adaptation to specific ecological niches. The spectrum of Solanaceous species collectively occupy almost every conceivable niche, presenting a remarkable opportunity to access the breadth of phenotypic variation and underlying biochemical diversity. In addition, many Solanaceae crops have been selected for specific characteristics related to biochemical composition and physiological traits. These include composition of sugars, organic acids, volatiles, a structurally diverse array of secondary metabolites and tolerance of environmental stresses. As a result of this biological specialization, a formidable knowledge base of plant biochemistry and physiology has been accumulated, in which members of the Solanaceae have been used as pioneering species in discovery-based research across the metabolic map[7]. Examples include cell wall and storage polysaccharide synthesis and metabolism, volatile production, vitamin biosynthesis, biosynthesis and action of the hormones ethylene and brassinosteroids and biosynthesis of flavonoids and carotenoids. Fruits of tomato and peppers accumulate high amounts of carotenoid pigments. Carotenoids with beta-ring, known as provitamin A, are indispensable in the human diet because they are the only source of vitamin A. Epidemiological studies indicate carotenoids to be preventive agents against specific diseases such as prostate cancer (lycopene) and age-related macular degeneration (lutein/zeaxanthin). Carotenoids show protective activity *in vitro* and *in vivo* against a variety of degenerative diseases, possibly through their activity as antioxidants[8].

Solanaceous species have been an excellent source of all major classes of secondary metabolites, including alkaloids, terpenoids, flavonoids, amino acid and fatty acid derivatives, and sugars. For example, atropine, found in Datura (devil's apple) was used for many years as a shamanistic tool and now has an important role in modern medicine. Another example is the capsaicinoid compounds, which are synthesized in the placental septum of the pod or fruit of only pungent *Capsicum* species. These compounds, which contain moieties derived from both the phenylpropanoid and amino acid/fatty acid pathways, underlie the familiar burning sensation of hot peppers, and their human uses range from traditional painkilling to chemo-preventative agents in cancer treatment. While some progress has been made in elucidating a few of the secondary metabolites biochemical pathways, there are many other types of metabolites in the rich biodiversity of the Solanaceae for which little information is available regarding their synthesis and physiological function.

Figure 3: The genomic tools resulting from SOL will facilitate comparative biology of common phenotypes such as fruit development and carotenoids content.

Another attribute of Solanaceae species that makes this taxon an attractive target for study of secondary metabolites is the presence of glandular trichomes, specialized structures that are easily isolated and are a major site of production, storage and secretion of certain metabolites. These trichomes are found on the surface of most aerial organs and consist of long stems with a two- or four-celled gland at the tip (Figure 4). In tobacco, these glands produce and secrete the diterpene cembratriene-diol (CBT-diol), which plays a major role in the interaction between the plant and colonizing aphids. Similar anatomical structures in *Lycopersicon* are the site of biosynthesis of monoterpenes and sesquiterpenes, acylsugars, and methylketones. Procedures for separating these glands from the leaves and then isolating chemicals, proteins or mRNAs from them have been developed, and so large chemical and molecular databases (e.g., EST) have been obtained for glands from several Solanaceae species. Combined with genomic sequence information, these tools to study the biology of a single type of cell are extremely powerful and will guarantee quick progress in elucidating the biochemical pathways leading to the synthesis of a myriad of secondary compounds[9].

The scientific literature extends far beyond this shortlist, which could be readily extended to include the many physiological processes for which members of the Solanaceae represent model species, including source-sink relations, temperature, drought and salinity stress tolerances. In addition Solanaceae species are in many cases experimentally far more tractable and attractive as models for biochemical and physiological research than smaller organisms such as Arabidopsis. It is important to recognize that members of the Solanaceae represent model species for large sectors of plant biochemistry and physiology and will continue to occupy this central position in plant science.

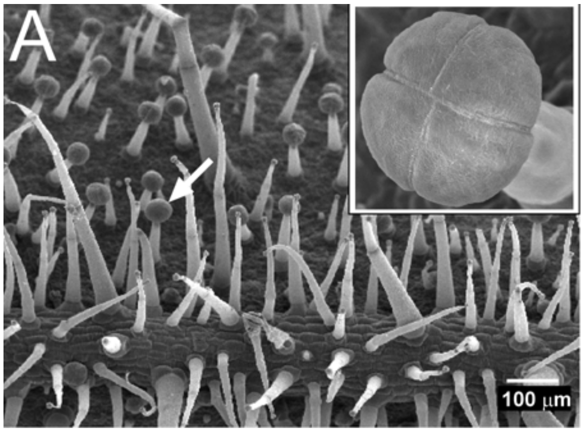


Figure 4: Scanning electron micrograph of the lower surface of the leaf of *L. hirsutum*. Type VI glands (four-celled; identified by white arrow, and in inset) predominate.

It is clear that biological hypotheses will continue to be developed and tested in the Solanaceae. The genetic, phenotypic, biochemical and physiological diversity, coupled with the wealth of literature and expert knowledge base, when supported by a genome sequence, make this taxonomic group a clear front-runner when electing a model group of organisms in which to develop the paradigm of plant systems biology.

Fruit and tuber biology provide a key to understand agricultural yield

In higher plants, sugars are produced photosynthetically, primarily in the leaves where carbon is fixed. These sugars are transported to other plant organs that are involved in active growth and development (such as roots, tubers, flowers, seed and fruits) where they are metabolized or stored. Because of the basic importance of the source-sink balance to human food and nutrition, the partitioning of assimilates is of key concern to basic and applied plant biology. Understanding the genetic mechanisms regulating source-sink relationships is a prerequisite for optimal modulation of yield through crop design and breeding. The Solanaceae are a unique model to explore the basis of yield: in tomato, peppers and eggplant fruits represent the major sink tissue while in potato, tubers are the underground stems that are uniquely modified as starch-storing organs. The physiological and developmental processes that regulate sink strength are expected to be largely independent of the anatomical identity of the organ. Comparison of the phylogenetically close potato with the fruit bearing Solanaceae is an active area of research aimed at revealing molecular modulators and communication networks that connect source with sink.

By anatomical definition the fruit is a mature ovary and therefore typically includes carpel tissues in part or in whole. Evolutionary pressures have resulted in a variety of developmental manifestations of fruit tissues, resulting in structures that range in design and function from hardened fruit capsules or pods that forcefully expel seed at maturation, to forms optimized for seed movement by wind, water, animal fur or gravity, and finally those implementing developmental programs yielding succulent and flavorful tissues for organisms that consume and disperse the associated seed[10]. In recent years, particular emphasis has been placed on tomato as an especially tractable system for molecular genetic analysis of fleshy fruit development and ripening. Fleshy fruits such as tomato undergo a ripening process in which the biochemistry, physiology and structure of the organ are developmentally altered to influence appearance, texture, flavor and aroma in ways designed to attract seed dispersing organisms. The importance of tomato as an agricultural commodity has resulted in decades of public and private breeding efforts, which have yielded numerous genomic resources including mapping populations, mapped DNA markers, extensive EST collections and publicly available microarrays. The sequence conservation of Solanaceae genes facilitated the use of the tomato microarrays to explore gene expression in fruits of other species (Figure 5) thus revealing a high commonality in fruit development among the family members.

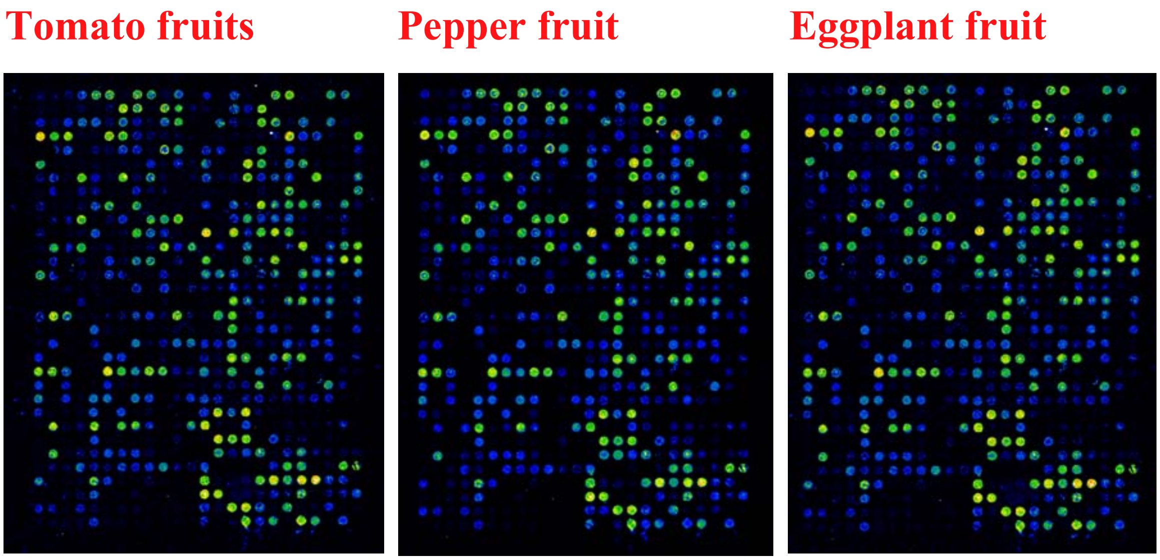


Figure 5: The high degree of sequence similarity among Solanaceae facilitates cross utilization of microarrays resources and tools - thus demonstrating that a single genome sequence will have a broad impact across the family.

Tuber formation is the most critical physiological function involved in potato production[11]. It involves a number of biological processes at the stolon tip, such as carbon partitioning, starch biosynthesis, signal transduction, and meristem determination. It is hypothesized that each physiological stage in tuber development is controlled by specific gene interactions. A number of genes has been found to be associated with early tuber formation including tubulins, S-adenosylmethionine decarboxylase, MADS box genes, acyl carrier protein thioesterase, and lipoxygenases. Due to the complexity of the tuber development process, it is difficult to fully understand the molecular mechanism of tuberization through the examination of individual genes. A firm understanding of the molecular mechanisms that regulate cell growth during tuber development are critical to address the improvement of numerous tuber traits including starch and protein content, internal quality, tuber shape, fast bulking, and early maturity. Given the high degree of similarity of genomes among the Solanaceae it will be especially interesting to assess how these genes have changed through evolution and selection to bring about the developmental and anatomical diversity displayed between members of this family. Through comparative genomics approaches, it should be possible to determine the evolutionary changes that enable plants like potato to produce tubers.

Diversity in Solanaceae defense responses

The Solanaceae species provide many unique advantages for studies of plant defense responses[12,13]. First, they are hosts to many well-characterized and economically important pathogens and insects (viruses, bacteria, fungi, nematodes, chewing/sucking insects). Second, many of the Solanaceae (e.g., tobacco) have large leaves that are especially amenable to pathogen/insect assays. Large leaves are also important for the isolation of proteins that are present in low abundance in plant tissues. Third, Solanaceae species, unlike Arabidopsis, possess multicellular trichomes that play an important role in insect defense (Figure 4 pg 10). Fourth, several of the Solanaceae species permit high efficiency Agrobacterium-mediated transient expression that expedites characterization of disease resistance genes and defense-signaling genes, an approach that has been especially useful in tobacco and tomato. Fifth, the family encompasses extensive natural genetic diversity that has been instrumental in the identification of host resistance genes and other defense factors. Sixth, over 100 disease resistance (*R*) genes are known in various Solanaceous species and many have been cloned (see below). Finally, virus-induced gene silencing (VIGS) has, to date, been found to be most efficient in the Solanaceae (*Nicotiana benthamiana*, tomato, tobacco)[14].

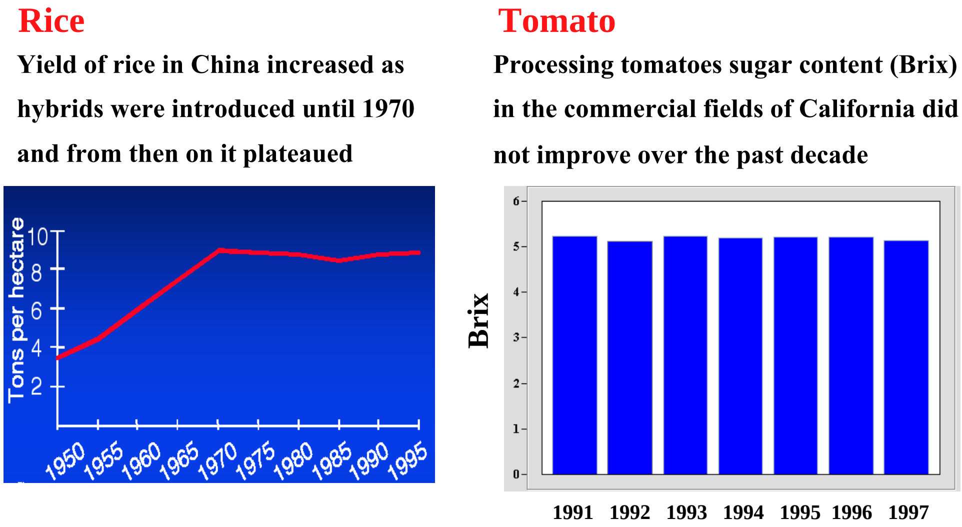
Tomato was the first plant from which a "gene-for-gene" class of *R* gene was cloned and more than 12 R genes now have been isolated from tomato. These include genes conferring resistance to fungi, nematodes, aphids, bacteria and viruses. Additional R genes, some of which have been shown to function when transferred into tomato, have been isolated from related species including pepper, potato, and tobacco. Remarkably, four *R* genes from tomato are unique (i.e. *Pto, Ve, Asc, Cf* genes); no similar *R* genes have been identified yet from Arabidopsis, rice or any other plant species.

In addition to yielding many *R* genes, tomato and related species have been used as a model system for other important advances in host defenses. Some of the major accomplishments include: 1) Discovery of systemin (present only in Solanaceous species). Systemin, which plays a key role in defense against herbivorous insects, was the first peptide hormone discovered in plants; 2) Isolation of the first mitogen-activated protein kinases (WIPK and SIPK); 3) Isolation of the first nitric oxide synthase gene from plants (iNOS); 4) Isolation of the first salicylic acid-binding proteins (SABPs) from plants; and 5) Some of the first insights into the mechanism of viral cross-protection and the development of virus-induced gene silencing (VIGS).

The many cloned *R* genes from tomato and other Solanaceae species and the abundance of information and resources related to diverse plant defense responses provides an unparalleled foundation for using the tomato genome sequence to increase our understanding of plant disease resistance and susceptibility. Finally, it is important to note that potato, pepper, and tobacco are also susceptible to many of the same pathogens as tomato (e.g. *Phytophthora* spp., *Fusarium* spp.). The availability of genome sequences for a range of Solanaceae plants will broadly benefit the study of plant defense responses and their role in adaptation to natural and agricultural environments.

What is the role of natural diversity in the genetic improvement of plants?

On the agricultural side, Solanaceae provide an opportunity to explore natural diversity as a sustainable resource to enrich the genetic basis of cultivated plants with novel genes that increase productivity. This approach is both a complement and an alternative to the GMO strategy for improving the quality and quantity of food.



Rice

Yield of rice in China increased as hybrids were introduced until 1970 and from then on it plateaued

Tomato

Processing tomatoes sugar content (Brix) in the commercial fields of California did not improve over the past decade

Susan McCouch

The California Tomato Grower (1998)

Figure 6: Depletion of the genetic variation for yield-associated traits in elite germplasm is the major factor for the slow progress in improving yield potentials.

Plant evolution under domestication has lead to increased productivity, but at the same time has narrowed the genetic basis of crop species[15] (Figure 6). The challenges facing modern plant breeders are to develop higher yielding, nutritious and environmentally friendly varieties that improve our quality of life by not harnessing additional natural habitats for agricultural production[16]. Genetic variation is the engine that propels breeding to meet future challenges. The observation that wild genetic resources can contribute to crop improvement, combined with the alarming rate at which locally adapted landraces are being lost and natural habitats are being damaged, has led to the establishment of large germplasm collections. These seed banks initiate collection missions, maintain and characterize accessions, and make them available to the breeding community. The task facing us is to devise the tools and concepts that will allow us to rapidly utilize the genetic potential that exists in the rich genetic diversity of wild species.

To enhance the rate of progress of such introgression breeding, the Solanaceae community pioneered 15 years ago the development of permanent genetic resources, which are comprised of marker-defined genomic regions taken from wild species and introgressed onto the background of elite crop lines. These introgression populations (lines with individual introgressed chromosomes or inbred backcross populations with a few small segments of wild species DNA in an otherwise isogenic background), which are now available for a number of tomato and pepper species, serve as unique powerful reagents for the discovery and characterization of genes that underlie traits of agricultural value[17] (Figure 7). The populations are being phenotyped by different labs for a range of yield-associated traits, including biotic and abiotic stresses, as well as for metabolic profiles of hundreds of distinct compounds. A user-friendly bioinformatics management system has been established such that this QTL data is available to the community over the Internet <http://zamir.sgn.cornell.edu/Qtl/Html/home.htm>. The challenge facing SOL for the coming years is to develop methodologies that will enable genomic information to be associated with phenotypes of interest for crop improvement. The framework for this data organization is the highly conserved genetic map of the Solanaceae that will allow us to extend the information basis beyond individual species. These exotic introgression line (IL) populations make a wide array of previously unexplored genetic variation rapidly available to plant breeders and geneticists. Either in combination with GM technology or without it, exotic genetic libraries represent a dynamic new resource base that can substantially enrich traditional crop improvement programs for many years to come. The vision uniting the SOL is to take advantage of the distinctive exotic breeding populations as a springboard to unite systems biology concepts with the practical discipline of plant breeding. This effort will not only impact Solanaceae biology but will also set the road map for implementing rational strategies for improvement of other crop plants that are important to human nutrition.

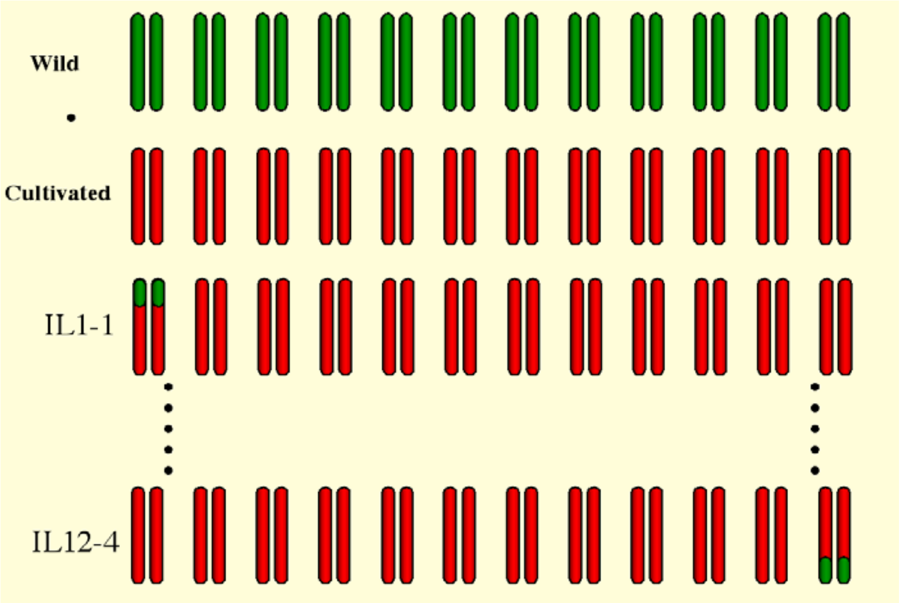


Figure 7: The construction of a set of interspecific introgression resources in elite genetic backgrounds will make a range of biodiversity available for breeders. This will allow scientists to unravel the basis of 'hidden' wild genes for productivity, adaptation and other phenotypes that affect human well-being.

***How can a system-level-approach in the Solanaceae help in resolving some of life's complexity*?**

The genomics era has largely matured around sequencing and studying the function of genes within a single organism (e.g. Drosophila, yeast, Arabidopsis). The tools for study most often involve under- (e.g. gene knockouts) or over-expression of genes to deduce their function. Increasingly, the approach that is taken is to study the response of the entire set of genes/proteins to perturbations in the expression of single genes. In reality, evolution of form and adaptation occurs not through the radical loss of gene function, but through the modification of gene function through changes in protein structure/activity and quantitative, temporal and special changes in gene expression[18,19]. Moreover, the genetic variation that fuels these adaptive changes has passed through thousands if not millions of years of selection and cannot be replicated in the laboratory. Through segregation and recombination that reshuffles the genomes of diverse donor parents in genetic crosses, the Solanaceae community has generated populations of offspring with multiple combinations of allelic variants, many of which date back thousands of years and have been honed through selective pressures. Such populations facilitate the study of traits that are determined by many genes, which almost always interact with each other and with the environment. Systems biology requires quantitative data that are of high quality, all-inclusive and taken simultaneously at different stages of development from defined cell lineages of appropriate germplasm resources. Looking ahead, genomics, quantitative genetics and computing sciences will be integrated in a comprehensive strategy of designing, modeling and analyzing complex biological data.

Systems biology is an approach to tackle the complex interactions that occur at all levels of biological organization and their functional relationship to the organism as a whole[20,21,22]. This new branch of biology, though yet somewhat fuzzy, is gaining in popularity in recent years, as modern genomic technologies are capable of generating comprehensive data sets on a wide spectrum of attributes of the organism. The high throughput technologies describe components of living systems that include whole genome DNA sequence, RNA transcription and processing, protein synthesis, post translational modifications, the formation of protein complexes, metabolic networks as well as simple and complex morphological phenotypes. Progress in future biological research will depend on our ability to find ways to start tying together the independent components into higher order complexity with multiple dimensions. It is also becoming apparent that multidisciplinary research efforts, involving the increased input of chemistry, physics, statistics, mathematics and computing sciences, is crucial for the success of a multifactorial approach. The key question is how to obtain the greatest knowledge about complex biological systems through clever experimental designs, models and methods of analysis.

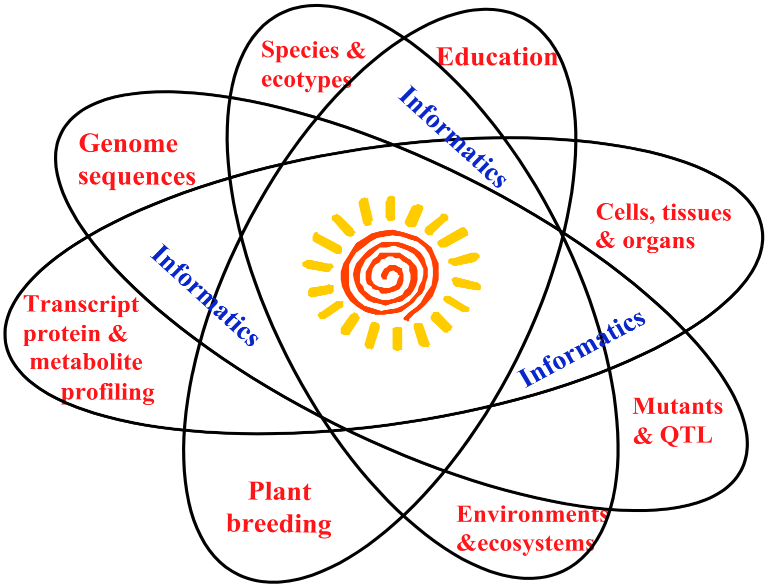


Figure 8: The systems approach aims to tackle the complex interactions that occur at all levels of biological organization and their functional relationship to the organism as a whole.

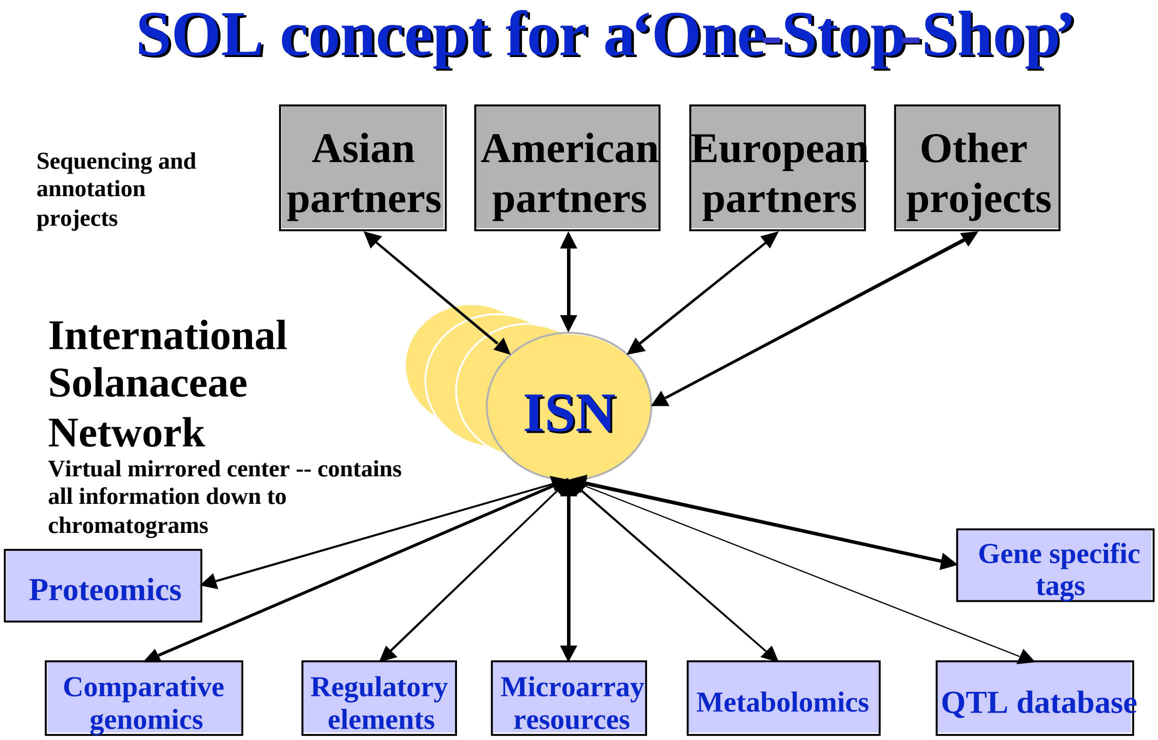
How can bioinformatics evolve to accommodate systems biology on the scale of the SOL project?

Presently, bioinformatics has most of its efforts focused on storage and retrieval of information in a format usable by as many scientists as possible. In the future, bioinformatics must evolve into a network of information that will become the driving force for creative ideas. In this new paradigm, the bioinformatics network (and its associated tools), will lead investigators into new hypotheses that can be tested *in silico* or through predictive laboratory experimentation[23]. The proposed 10 year international SOL initiative will turn this vision into reality by making sure that the independent research conducted in the participating laboratories will focus on a defined set of similarly grown genotypes, tissues and cells in a manner that will facilitate multifactorial analysis of combined datasets.

There is strong and broad agreement in SOL to develop a single virtual entry point (One-stop-shop) for Solanaceae genome sequence and related genomics and systems information. The systems biology framework that SOL has adopted poses new challenges to bioinformatics and in particular to data management. Bioinformatics is the discipline that analyses large-scale genome data and makes predictions on entities such as gene structure, gene function, protein structure, expression levels, and phenotypes and it also deals with data management and visualization issues. In the past few years, bioinformatics has been very successful in solving very specific problems such as gene finding and protein domain identification. Systems biology builds on these successes as the discipline that analyzes the networks that can be ‘overlaid’ onto these and other types of data, such as metabolic networks, gene expression networks, regulatory pathways, developmental pathways and even networks of interactions among organisms. Systems biology makes inferences about how these different networks interact and predictions about outcomes when these networks are disturbed. Obviously, the description of these extensive networks is an information-intensive task[24]. A prerequisite to work intelligently with such data is a universal way to describe these networks. Second, we need these data accessible in one location (not necessarily physical) to be able to integrate and analyze them. Therefore, the ultimate aim of the bioinformatics strategy is to provide scientists with a knowledge environment that will enable them to generate new hypotheses in an exploratory fashion. The Solanaceae bioinformatics effort will work with the major model organism databases to make a more unified and advanced querying of information possible.

The importance of bioinformatics to the overall success of the project cannot be overestimated. Bioinformatics is really the glue that holds a project of this magnitude together. Our view is therefore that such an effort should be closely coordinated, more so than in other genomics projects in the past. It is particularly important to establish, early on, standards and coordinates to ensure efficient use of resources, the generation of consistent, high quality data, and to prevent major duplications of efforts. The collaborative spirit in the Solanaceae community, that was evident at the recent Solanaceae meeting in Washington, will also be reflected in the bioinformatics strategy. At the meeting, the representatives of a number of bioinformatics centers have agreed that the most desirable deliverable of such a project is a unified data-center on the web that holds all data and connects to other model organism databases. Although many projects have advocated an open source approach in the past, in practice, few projects really achieve a true collaborative environment. The prerequisites for a collaborative environment are common technical standards and a basic infrastructure to share codes and data[25]. The centerpiece of the collaboration will be a repository (The International Solanaceae Network; Figure 9) where all centers share their pipelines, programs and data to the extent possible in common standard formats. This common repository will be the basis for database implementations at the participating bioinformatics centers that mirror the complete information available. The standards will be established by a bioinformatics committee that is already working on standards for the first phase of the Solanaceae project - the sequencing of the tomato genome.

Figure 9: The concept for organization of the international SOL bioinformatics network



SOL GOALS AND OBJECTIVES

Over the coming decade the International Solanaceae Genome Project (SOL) will create a coordinated network of knowledge about the Solanaceae family aimed at answering two of the most important questions about life and agriculture:

How can a common set of genes/proteins give rise to such a wide range of morphologically and ecologically distinct organisms that occupy our planet?

The corollary question of agricultural importance is:

How can a deeper understanding of the genetic basis of plant diversity be harnessed to better meet the needs of society in an environmentally-friendly and sustainable manner?

Key components and milestones for the SOL project:

1) Sequence the reference tomato genome on a BAC by BAC basis (see the technical sequencing paper; Appendix 1) tying this information together in a common framework with Arabidopsis, rice and Medicago.

2) Develop deep EST databases from various Solanaceae tissues and shotgun genomic sequencing of other Solanaceae with data integration. Align EST/shotgun sequence from other Solanaceae species against the contiguous tomato genome sequence so as to provide the basis for identifying sets of orthologs across as many species as possible. In addition, this process will allow the identification of both conserved genes (and sequence motifs – e.g. promoters) as well as genes/regions of the genome that are evolving rapidly under positive selection and may hence be related to species diversification. The sequence data will also facilitate the development of sequence repertoire resources for large-scale functional genomics experimentation across all Solanaceae species. The planned resources will include complete sets of Gene Specific Tags (GSTs) and ORFs for all tomato genes cloned in versatile vectors, and which will enable the systematic gene silencing and over expression of genes.

3) Complete a set of high-resolution comparative genetic maps for Solanaceae species and related taxa (e.g. coffee), based on the analysis of Conserved Orthologous Set (COS) markers.

4) Construct a set of interspecific introgression resources (e.g. introgression lines, backcross inbred lines etc.) for all Solanaceae crop species in order to provide the genetic material from which the genes/QTL underlying species divergence, biochemistry, environmental adaptation, evolution of mating systems etc. can be identified and the development of associated phenotypic databases.

5) Establish saturation mutagenesis genetic resources, methods for high throughput identification of mutants associated with a specific sequence and comparative biology of developmental pathways between taxonomic groups.

6) Construct a comprehensive phylogenetic and geographical distribution information network that can be used by both researchers and educators – including online information about natural distribution, habitats, species images, germplasm banks, botanical gardens and herbarium collections.

7) Application, on a basis of a broad phylogenetic sampling, of transcription (using whole genome chips and high-throughput real time PCR), proteomic and metabolic profiling to begin understanding the range of evolution of plant chemicals and to provide a base line for follow-up studies in chemical ecology. This will include the application of non-destructive 'Real Time' physiology and phenotyping platforms including imaging techniques to dynamically monitor changes in cells and organs throughout development.

8) Improve the efficiency of plant breeding based on the use of wild species variation, marker assisted selection, and mutagenesis. Develop deep genetic variation databases from representative accessions of cultivated lines and wild relatives for different Solanaceae species. The genomic survey of natural variation will encompass several thousand genes with either proven or suspected agronomic potential. The database will serve the dual purpose of generating useful markers for marker assisted selection, as well as give an insight in the common haplotypes present in cultivated and wild material, and hence provide a rational framework for exploiting genetic variation. This component of the SOL project will be conducted in collaboration with industries that rely on Solanaceae species in a manner that will allow dissemination of the results to the entire community.

9) Develop an education package, based on the Solanaceae, with which the public and students (of all ages) can become more aware of the natural diversity of plants, the process of domestication, and the role of these in sustainable agriculture.

10) Develop an international bioinformatics platform, which will not only allow storage and retrieval of all information from the SOL project, but also the development and application of new algorithmic tools which will promote and facilitate a systems approach to the study of Solanaceae diversity and adaptation.

Organization and Coordination of the SOL project on an International Level:

The Steering Committee of SOL will function as a virtual 'umbrella organization' for the project by coordinating and facilitating research along the vision proposed in this document. The members of the committee (24 – one for each tomato chromosome) will include representatives of the participating countries as well as scientists that are expert in specific fields (such as: systematics, bioinformatics, profiling technologies, education and theory and practice of systems biology). Committee membership will rotate as the project moves forward in time and knowledge - **from sequence to systems** - and the Committee will meet once a year in the International SOL Genomic Symposium (**The First Solanaceae Genomics Symposium will take place in Wageningen, The Netherlands, on September 19-21, 2004. It is being organized by Willem Stiekema**). SOL will promote and actively seek additional scientists, countries and funding agencies to participate in this expedition into the higher order organization of biological information. Scientist are encouraged to freely use this whole document, or parts of it, for their own proposals. In addition a power point presentation describing the project can be downloaded from SGN < http://sgn.cornell.edu/solanaceae-project/>. Members of the scientific community, from around the world, who would like to submit local or international collaborative research grants under the umbrella of SOL would receive a warm letter of support from SOL Steering Committee. Grants that will be endorsed by SOL will have to include a section detailing how the research is complementary with SOL goals and how the generated data will be integrated with SOL bioinformatics system.

STATUS OF SOLANACEAE CROPS RESEARCH

This section summarizes the status of the research in Solanaceae crops and the available genetic and genomic resources[23]. 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[27,28,29]. 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 provide the infrastructure by which the tomato genome sequence can be directly utilized by researchers working on virtually all Solanaceae species.

TOMATO *Lycopersicon esculentum*

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 by sterile cuttings in tissue culture. 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)[30].
2) *L. hirsutum* (LA1777)[31].
3) *Solanum lycopersicoides* (LA2951)[32].

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 induced mutants are available in the dwarf variety Microtom[33,34].

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 (http://www.tigr.org/). 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 are shown in Figure 10 pg 29.

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>.

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[35].
- 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 and Regan) – 100,000 ESTs are being generated both from 3’ and 5’ from 16 diverse libraries including normalized libraries.
- Germany (Gebhardt) - 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.
- The Netherlands (Visser) – Method for transcriptome maps based on AFLPs[36].

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
- 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[29].

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 *Petunia hybrida* 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 baskets, 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[37]. 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 organization 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/).

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* [38]. 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[39]. 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.

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 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. The wild species are found in Costa Rica, Colombia, Ecuador, and Peru[39].

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 where it is still grown[39,40]. 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. 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. 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, 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.

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 (Figure 1 pg 6).

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 synteny 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 synteny 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.

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
- 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

Other libraries were prepared from:

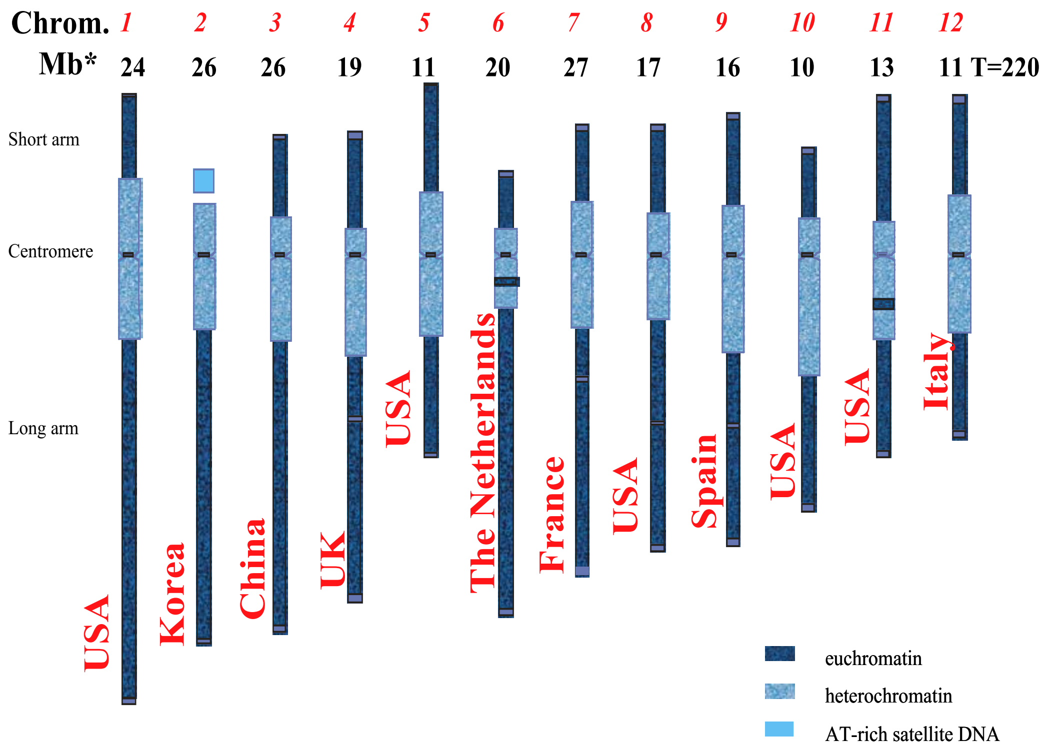
- *C. eugenioides*, *C. canephora*, and *C. racemosa*.

Based on this information, we believe that including coffee, as part of the integrated Solanaceae project would represent a major contribution to the program.

STATUS OF SOLANACEAE RESEARCH FROM AROUND THE WORLD

Scientists from 19 countries provided written summaries of their present Solanaceae research and future grant submission plans. The sequencing of all the euchromatic regions of the tomato genome is being proposed by eight different countries on a chromosome-by-chromosome basis (Figure 10). The vision of SOL, however, goes much beyond the DNA sequence and thus we aim to extend the number of countries that will participate in this 10-year initiative. Some of the projects listed in this document are by nature less expensive than high-throughput genomic undertakings, but are as important, particularly when we aim to explore issues of diversity and adaptation. This will allow many more countries to contribute and participate in the expedition into the higher levels of organization of biological systems.

Figure 10: Assignments of chromosome sequencing tasks to eight countries.



*** Euchromatin portion (based on deJong and Stack, personal comm)**

Solanaceae research in different countries:

1) ARGENTINA

Centers of Solanaceae research in Argentina

Research Institutes

National Institute for Agricultural Technology (INTA):
Instituto de Biotecnología CICVyA INTA Castelar, Province of Buenos Aires
Instituto de Recursos Biológicos INTA Castelar, Province of Buenos Aires
EEA INTA Balcarce, Province of Buenos Aires
EEA INTA La Consulta, Province of Mendoza

INGEBI (Instituto Nacional de Investigaciones en Ingeniería Genética y Biotecnología)
IIBL (Instituto de Investigaciones Bioquímicas Leloir)

Universities

University of Buenos Aires, Faculty of Sciences
University of Mendoza, Faculty of Agronomy
University of Córdoba, Faculty of Agronomy

The Northwestern part of Argentina (”Puna”) belongs both biologically and culturally (southern part of the Inca civilization) to the center of origin and domestication of Solanaceae, particularly in the case of potato.

Sources of Support

Most basic research (including some small scale genomic efforts) is supported in part by ANPCyT (Agencia Nacional de Promoción Científica y Tecnológica) and CONICET (both dependent of the Secretary of Science and Technology) and by international funding like INCO-EU. Breeding and conservation of genetic resources are financed by INTA (dependent of the Secretary of Agriculture).

Topics of Research

The major topics are:

- *Sequencing metabolic genes from wild tomato species (L. pennellii)* (at INTA Castelar, by Hopp E. in collaboration with the group of Fernie A/Willmitzer L, Berlin, Germany)
- *Host pathogen molecular interaction between potato and its pathogenic viruses and tomato and its pathogenic bacteria* (at INTA Castelar, IIBL: headed by Dr. Ceriani F.)

- *Molecular biology of fruit development and stress resistance in tomato* (University of Buenos Aires, headed by Dr. Norberto Iusem)
- *Polen-pistillum molecular interaction in tomato* (INGEBI, headed by Dr. Jorge Muschietti)
- *Characterization of genetic resources and breeding:* This research is focused on potato (INTA Balcarce) tomato and pepper (INTA La Consulta). It includes the use of molecular markers (SSR). There is a large germplasm bank which includes a replica of the one in CIP (Internatinal Potato Center) at INTA Castelar
- *Phylogenetic and genetic diversity studies using molecular markers, sequencing of some markers (ITS, etc.) of Solanum species* (INTA Castelar, University of Córdoba and Mendoza)
- Molecular characterization of protein kinases of *Solanum tuberosum* induced at the onset of tuber development (INGEBI Leaded by Dr. Téllez-Iñón)
- *Plant propagation and genetic transformation:* Research is focused primarily on potato and tomato transformation for different breeding purposes and development of germplasm maintenance protocols (INTA Castelar).
- *Molecular farming.* Research is focused on the development of vaccine products through expression systems in potato (INTA Castelar headed by Dr. M. Borca).

Potential Contribution of Argentina to the Tomato Sequencing Project

At this time, it is unlikely that major funds are available in Argentina to provide support to a tomato sequencing program. However, basic infrastructure and research facilities and human resources are already available for this purpose due to the ongoing sunflower genomic project.

Sequencing metabolically significant region of the wild species *L. pennellii* should allow rapid identification of genetic factors underlying metabolic changes in *L. esculentum* lines harbouring introgressions of this species.

2. Deadlines for Submission of Genomics Submissions

There is no specific genomic-oriented funding, although it is prioritised among general submission programs. October 2004.

2) BRAZIL

Coffee is the most valuable agricultural commodity in world trade, commanding a turnover of US$ 10 billion annually. Coffee is the world's most important beverage and upon which the economies of many producing countries from the south depend. Unfortunately, coffee appears as an "orphan crop" at the taxonomic level. The International Solanaceae Genome Project (SOL) is certainly the best opportunity for coffee research network to develop collaborative work in the area of genomics. Coffee belongs to Rubiaceae and is closely related to the Solanaceae family, as these two families are included in the Asterid I class (Figure 1).

Coffee genomics in Brazil: 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 to date. Most of the infrastructure utilized for sequencing was already set up prior to the initiation of the Coffee Genome Project, 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 has 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. 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 genetics. The inclusion of Coffee as an integral part of SOL will extend the reference yardstick to evaluate issues of diversity and adaptation.

3) CANADA

Centers of *Solanaceae* research:

Federal Research Institutions

AAFC Crops and Livestock Research Centre (Charlottetown-Nappan)
AAFC Atlantic Food and Horticulture Research Centre (Kentville-Bouctouche)
AAFC Potato Research Centre (Fredericton)
AAFC Horticulture Research and Development Centre (Saint-Jean-sur-Richelieu)
AAFC Food Research Program Centre (Guelph)
AAFC Southern Crop Protection and Food Research Centre (London-Vineland-Delhi)
AAFC Greenhouse and Processing Crops Research Centre (Harrow)
AAFC Lethbridge Research Centre (Lethbridge)
AAFC Pacific Agri-Food Research Centre (Summerland-Agassiz)

Provincial Research Institutions

- Alberta Department of Agriculture, Food and Rural Development Crop Diversification Centre South (Brooks)
- Saskatchewan Department of Agriculture, Food and Rural Revitalization,
- New Brunswick Department of Agriculture, Fisheries and Aquaculture (Bon Accord-Fredericton-Wicklow)

Universities

Carleton University (Ottawa)
University of Alberta (Edmonton)
University of Manitoba (Winnipeg)

University of Guelph (Guelph)
University of Waterloo (Waterloo)
University of Western Ontario (London)
University of Montreal (Montreal)
University of Quebec in Montreal (Montreal)
McGill University (Montreal)
Laval University (Quebec City)
Nova Scotia Agricultural College (Truro)
St. Mary's University (Halifax)

Genomics companies
Solanum Genomics International Inc. (Fredericton)

Sources of Support

Most university research is supported in part by NSERC (the Natural Sciences and Engineering Research Council of Canada). Large-scale genomics efforts are funded to a maximum of 50% through Genome Canada, with additional support provided through provincial funding initiatives or other sources.

The major topics of research are:

- *Production of genetic resources and mapping:* This research is focused on potato (AAFC Fredericton, AAFC Lethbridge) and tomato, petunia (University of Guelph).

- *Development and implementation of genomics platforms:* This occurs through the Canadian Potato Genome Project, a Genome Canada-funded project encompassing five institutions (based primarily in Fredericton), and involves EST generation, microarray expression analysis, activation tagged mutant production and bioinformatics.

- *Pre-and post-harvest physiology:* This research is focused on effects associated with processing quality traits (primarily with potato), as well as senescence, aging and the effects of a variety of greenhouse growth conditions on tomato/pepper production. Major efforts for potato exist at AAFC Fredericton, AAFC Charlottetown-Nappan, NSAC Truro, NBDAFA Wicklow and University of Alberta. Major tomato and pepper programs exist at University of Guelph, University of Waterloo, AAFC Harrow and AAFC Summerland-Agassiz.

- *Metabolism:* This research is focused on carbon metabolism and source-sink relationships. Efforts are geared towards potato starch synthesis (University of Guelph) and starch quality (AAFC Guelph), and carbon metabolism during tomato fruit development (University of Guelph and Laval University).

- *Biotic and abiotic stress:* Biotic stress research is focused on plant pest/pathogen interactions, metabolite profiling of these interactions, resistance genes and mechanisms, pest/pathogen screening and control strategies. Major research programs are in progress at several AAFC Centres (Charlottetown-Nappan, Fredericton, Sain-Jean-sur-Richelieu, London-Vineland-Delhi, Lethbridge), as well as University of Guelph, University of Manitoba, McGill University, University of Montreal, Laval University, Alberta Crop Diversification Centre South, NBDAFA Wicklow and SGII. Abiotic stress research is focussed on drought, salinity and nutrient stresses, with research programs at AAFC Fredericton, Alberta Crop Diversification Centre South, University of Montreal and University of Quebec in Montreal.

- *Plant propagation:* Research is focused primarily on potato organogenesis, somatic embryogenesis, development of germplasm maintenance protocols and nuclear stock production. Programs are in effect at NBDAFA Wicklow-Fredericton, AAFC Fredericton and University of Montreal.

- *Molecular farming.* Research is focused on the development of vaccine products through expression systems in tobacco and potato. Programs are ongoing at AAFC London-Vineland-Delhi, Laval University and SGII.

- *Reproductive development:* Research is focused on self-incompatability in potato and tomato fruit development. Research programs are in effect at University of Montreal and University of Guelph.

Potential Contribution of Canada to the Tomato Sequencing Project

At this time, it is unlikely that funds are available in Canada to provide support to a tomato sequencing program. Given the importance of potato, and the major potato research efforts in Canada, a contribution towards a potato genome sequencing effort is a more distinct possibility.

Deadlines For Submission Of Genomics Submissions

Funding for the next round of large-scale genomics research via Genome Canada has not yet been confirmed. However, discussions with the upper levels of Genome Canada have indicated an interest in a potato genome sequencing effort for any future rounds of funding. Preliminary discussions have taken place between Canadian researchers and European/American researchers towards this end.

4) CHINA

Highlights of Solanaceae Research in China

The Solanaceae is one of the most economically important crop families in China. There is a large scientific community including breeders, geneticists and molecular biologists with research focused on tomato and potato, the two major Solanaceae species in China.

Centers of Solanaceae research in China

Institute of Genetics and Developmental Biology in China (IGDB), Chinese Academy of Sciences, Beijing, Beijing Genomics Institute, Beijing

Sources of Support

Chinese Academy of Sciences, Ministry of Science and technology in China, National Natural Science foundation of China

Topics of research

1) Tomato as a model to study plant-insect interactions: This research is aimed at understanding the biochemical and molecular basis of plant resistance to insects, with a long-term goal of developing novel crop protection strategies that rely on the plant's "built-in" resistance capabilities, thereby reducing the use of environmentally damaging pesticides. Wound-inducible proteinase inhibitors in tomato provide an attractive model system to study the signaling pathways that regulate this form of defense. An ever-increasing body of evidence has indicated that the fatty acid derived signal jasmonic acid is an intercellular signal for host defense response, and that the biosynthesis of jasmonic acid is regulated by a peptide signal called systemin (Li et al., 2003; Li et al., 2002). Sequence of tomato genome will greatly facilitate gene identification in the systemin/JA signaling pathway and broadly increase our understanding of induced resistance of plants to insects attack.

***2)** Using tomato as model to study molecular biology of plant nutrition:*

The main consideration to use tomato as a model for this aspect of research is lethal mutants of the root-specific house-keeping genes can be easily rescued and maintained by grafting onto a wild type compared with *Arabidopsis*. Additionally, some genes involved in mineral uptake and metabolism in tomato genome show less genetic and biochemical redundancy compare to those in *Arabidopsis* genome, for example, *CHLN* gene. Dr. Hongqing Ling's group is studying molecular mechanisms of Fe and P uptake and metabolism. In last several years, they have isolated *CHLN* (a gene encoding nicotinanamine synthase catalyzing nicotianamine synthesis from methionine in plants. Nicotanamine is a polyamine involved in iron uptake and metabolism), *FER* (a gene encoding a protein containing bHLH DNA-binding domain, it is as a putative transcription factor controlling iron deficiency responses and iron uptake in tomato roots) and *LeFRO1* (a gene encoding a Fe^{3+}-chelate reductase reducing ferric to ferrous iron on root surface) from tomato genome using map-based cloning and PCR-based homology cloning approaches (Ling et al., 2002, Ling et al., 1999). Additionally, they are trying to isolate new genes involved in iron and phosphate uptake and metabolism using a forward genetic approach.

Potential Contribution of China to the tomato and potato sequencing project

1) As proposed before, the Institute of Genetics and Developmental Biology, Chinese Academy of Sciences could do the sequence of the euchromatin region of tomato chromosome 3. We are now actively seeking sources of funds from relevant agencies in China.

2) The other major potential contribution of China to the genomics study of Solanaceae is that the Bejing Genomics Institute potentially could sequence the potato genome using a whole-genome shotgun sequencing strategy. If an ordered sequence map can be created for tomato, then other Solanaceous species such as potato could be shotgun sequenced and ordered using the tomato genome as a guide (through synteny). This would also allow fantastic comparative genomics studies. The Beijing Genomics Institute has track record for its sequencing and genomics capability (they completed 1% of the human genome sequencing, the draft sequence of rice genome, and a draft sequence of silkworm genome, etc.)

5) COLOMBIA

Centers of Solanaceae research

Research Institutions

Corporacion Colombiana de Investigacion Agropecuaria, CORPOICA
Corporación para Investigaciones Biológicas, CIB

Universities
Universidad Nacional de Colombia
Universidad de Antioquia
Universidad Catolica de Oriente

Others
Centro virtual de investigacion de la cadena agroalimentaria de la papa, CEVIPAPA

Centers of coffee research
Centro Nacional de Investigaciones de Cafe, CENICAFE

Topics of Solanaceae research

Germplasm collection and characterization. Mainly conducted at CORPOICA, which is the institution responsible for maintaining the National Germplasm Bank for food and agriculture. At present, there are more than 22000 accessions of plant genetic resources representing more than 75 species. In the case of the Solanaceae family, the germplasm bank comprises 44 species of the genera Capsicum, Lycopersicum, Cyphomandra, Physalis, Nicotiana and Solanum. The Solanum collection is the second largest of the world after the CIP collection in Peru. CORPOICA has developed a solid ex-situ conservation platform including field, refrigeration, and in vitro conservation systems supported by a complete database. Research efforts are oriented to improve the knowledge of the genetic variability existing in the germplasm banks and to advance in the molecular characterization of these materials (through RFLP, AFLP, SSRs) to adding value to these collections and allowing a sustainable use of these genetic resources. Important information regarding morphological, physiological and biochemical characterization of different accessions is available in CORPOICA's databases.

Other sources of Solanaceae germplasm are in the National Colombian Herbarium (COL, Universidad Nacional de Colombia), which comprises 520.000 collections of the Colombian flora. COL exchanges material with national and international herbaria and it is one of the most important sources of information in South America.

Systematics and taxonomy. Study of phylogenetic relationships among the most widely distributed groups in Solanum, Cestrum, Lycianthes, and some genera of the subtribe Withanineae through morphological, chemical and molecular systematics (Universidad Nacional de Colombia, Universidad de Antioquia in collaboration with the Natural History Museum, London and the Universidad Central de Venezuela). Future research will include collaboration in a worldwide project whose aim is the creation of a comprehensive inventory of *Solanum* (headed by Lyhn Bohs, Sandra Knaap and Michael Knee and recently funded by the National Science Foundation, NSF (USA).

Breeding. Traditional breeding for lulo, tree tomato and potato. Relevant research advances includes the release of a high vitamin C Lulo (*S. quitoense*), the release of a lulo cultivar named "lulo la selva" (*S. quitoense* x *S. hirtum*) showing resistance to nematodes, and the field identification of inter-specific *Cyphomandra betacea* materials tolerant to anthracnose (CORPOICA)

Tissue culture and genetic transformation. *in vitro* multiplication, plant regeneration and genetic transformation of potato, lulo and tree tomato (Universidad Nacional de Colombia, CIB, Universidad Catolica de Oriente).

Integrated pest management. Formulation of biopesticides -- use of BT toxins for pest control in potato and tomato (CORPOICA, Universidad Nacional de Colombia, CIB).

Funding of Solanaceae research

Maintenance of the germplasm banks of plants, animals and microorganisms is funded by the Colombian Ministry of Agriculture and CORPOICA. Other research is funded by the Instituto Colombiano para el Desarrollo de la Ciencia y la Tecnologia (COLCIENCIAS), CEVIPAPA, the Bolsa Nacional Agropecuria, and the Universities. However, funds for genomic / molecular studies are needed, particularly for species with potential international commercialization (*C. betacea*, *S. quitoense* and *P. peruviana*). Genomics research conducted on these species will be of great importance for diversity and adaptation studies proposed in the SOL project. Likewise, information derived from the sequencing of the tomato genome and from functional and comparative genomics with other Solanaceae species as proposed in the SOL project will be of great importance for developing a genomics platform for these Andean species.

Topics of coffee research

CENICAFE is the Research Center of the National Coffee Growers' Federation of Colombia, an organization that unifies coffee producers through-out the country and brings economic and social development to the coffee zone. Since its creation in 1927, it has funded scientific research, agricultural extension, and diversification programs to support the coffee industry, as well as, physical infrastructure (road building, electrification), education, health, and other services that provide economic and social development for the 4 million people in the coffee growing area. The Federation founded CENICAFE in 1938 with the mission to generate scientific knowledge and technologies to help sustainability of coffee production in Colombia while preserving natural resources in the coffee regions (soil, water, flora, and fauna). CENICAFE's research covers the complete range of coffee production: from genetic studies aimed at development of new coffee varieties right through marketing.

In 1991, CENICAFE initiated a collaborative project with Cornell University to train a multidisciplinary team of Colombian scientists in Entomology, Plant Breeding, Plant Pathology and Molecular Biology. This project was funded by the National Coffee Growers Federation of Colombia and has generated the expertise that will now enable CENICAFE to work as a full research partner in coffee genomics with Cornell University to tailor these efforts towards applications to coffee improvement.

Coffee genomics initiative. CENICAFE is currently preserving and characterizing one of the largest *Coffea arabica* germplasm collections in the world with more than a thousand accessions (several from Ethiopia, a center of diversity for coffee). Our genomics studies will be oriented toward exploring and characterizing genetic diversity in the genus *Coffea,* in terms of insect and pathogen resistance, quality and yield to use it for CENICAFE's breeding efforts in *C. arabica*. CENICAFE has invested several years in the development of populations for mapping purposes and QTL analysis. The parental genotypes for the crosses were selected based on the diversity studies performed by Moncada (2000). Analysis of the heritable characteristics of quantitative and qualitative type on populations derived with that objective, will allow us to obtain information on the number, chromosomal location, effect and genetic interactions of loci that control the expression of those characters.

We are currently funded by the Colombian Ministry of Agriculture to develop a coffee genetic map. Our initial objective is to screen 500 SSR and map those that are polymorphic onto one or more mapping populations. The isolation and characterization of microsatellite markers was initiated in Susan McCouch's laboratory during the joint project between Cornell and CENICAFE. Additional SSR markers are being developed at CENICAFE. Evaluation of SSR polymorphism in the parents is being conducted at CENICAFE to obtain the most informative markers for population analysis. In addition, 1000 COS markers, developed in collaboration with Steve Tanksley, will be screened and those that are polymorphic will be mapped. The map, including the identity and the order of the markers, will be published. The map will lay the foundation for identifying QTLs associated with traits of interest for coffee improvement and will represent a basic framework for saturation with additional markers in the future. This will be a first step in the process of identifying genes of agronomic importance. Genes of particular interest are those involved in resistance to insects and diseases, in the flowering process, yield, and in cup quality, among others.

The specific markers chosen for map construction will provide interspecies links and also provide tools that are easy and inexpensive to use for large scale breeding applications. Both SSRs and COS represent highly informative PCR-based markers that will be efficient for use in both basic genetics research and breeding applications.

6) EUROPEAN UNION (EU)

In 2000 Commissioner Busquin launched the European Research Area - ERA – policy aimed at creating a comprehensive set of measures directed at promoting the integration of European research. One of the major objectives has been to identify areas of strategic interest in which integration is to be pursued with high priority. A model for this philosophy has been the steps taken towards the creation of a European Plant Genomics network, under the ERA-NET programme, and the establishment of a Technology Platform for Plant Genomics and Biotechnology. The former initiative is comprised of national science policy makers, representatives of the major funding agencies including scientists and experts. The objective will be to harmonize funding strategies and mechanisms, develop a framework for cross border collaborative research and establish a long-term common European policy for plant genomics. The Plant Genomics ERA-NET is becoming operational on January 1, 2004.

The development of 2025 Plant Genomics Platform is aimed at bringing together all the stakeholders (science, industry, society) to develop a 20 year vision for plant genomics and biotechnology in Europe. This vision will include economic and industrial priorities as well as Europe's role in the global context in a manner that will allow the establishment of a strategic research agenda to meet long-term goals. This Plant Genomics Platform constitutes an important step towards a coherent European strategy for plant genomics and will provide a firm basis for long term funding.

The European Solanaceae Genome Initiative is aimed to bring together a broad scientific base for Solanaceae research. Europe is blessed with a strong and dynamic research community with more than 100 groups that are directly involved in Solanaceae research. The well-structured national programs are reflected in the countries write-ups with excellent complementarity between the groups and already established research alliances. Europe's Solanaceae research community is highly motivated to contribute to the SOL project particularly when the vision of the international project is highly consistent with the views developed in Europe. The plan is to submit an Integrated Solanaceae Project application for funding in FP6. In tomato the European countries are already preparing grant applications to sequence 5 chromosomes. For potato a large number of BACs will be sequenced (The

Netherlands, Germany) as a part of an already existing European collaboration. Other topics that are of interest include physical mapping, SNP marker development, EST sequencing, map-based cloning and QTL cloning, microarray development, metabolic profiling technologies, TILLING and a bioinformatics effort that will benefit from the solid European expertise in genome annotation and comparative genomics.

The FP6 funding opportunities for a Solanaceae genome project are in Priority 5 – "Food quality and safety". This will provide complementary funding for the sequencing of the tomato genome as a first step to support the exploitation of the genome resource to improve the nutritional quality of Solanaceae vegetables. The funding from the European Commission will strengthen the integration of the national programs and consolidate Europe's contribution to the International Solanaceae Genome Project. Europe is highly motivated to participate in the international SOL project and to contribute significantly to the tomato genome sequencing effort as well as follow up biology along the proposed line of research of system biology of diversity and adaptation. **A grant proposal integrating the concepts described above will be submitted to the European Commission in 2004 under the European FP 6 program and will be coordinated by Chris Bowler from Laboratory of Molecular Plant Biology Stazione Zoologica, ITALY and ENS/CNRS FRANCE.** This initiative is part of a broader drive to integrate Europe's plant biology research where plant genomics has been recognized as a high priority technology area.

7) FRANCE

Centers of *Solonaceae* research in France

Research Institutions
INRA Plant Physiology & Biotechnology, Bordeaux
INRA Genetics and Breeding of Fruit & Vegetables, Avignon
INRA Genomics and Biotechnology of Fruit, Toulouse
INRA Horticultural Plant Growing Systems, Avignon
INRA Safety and Quality of Plant Products, Avignon
INRA Cell wall Polysaccharides laboratory, Nantes
CNRS Plasts and Cell Differenciation Laboratory, Grenoble
INRA Plant Pathology, Antibes
INRA Plant Pathology, Bordeaux
INRA Plant Genomic, Evry
Universities
University of Bordeaux
University of Toulouse, National Polytechnic Institute
University of Grenoble

Sources of Support

Research is supported by INRA institutional resources, by grants of the French Ministries of Research and Agriculture, by Genoplante programme and by EU programmes (APOPHYS, DREAM)

Major Topics of Research

- *Production of genetic resources, mapping*. Mapping and genotyping research activities on tomato, pepper, potato and eggplant are centred at INRA-Avignon (M. Causse, V. Lefebvre, MC Daunay), partly in interaction with breeding companies. A pepper BAC library was constructed at INRA-Avignon and INRA-Evry (Genoplante). Generation of EMS mutants at INRA-Bordeaux (C. Rothan).

- *Development and implementation of genomics analytic platforms*. INRA-Toulouse (M. Bouzayen) is the source centre for tomato microarrays production and distribution and is in charge of sequencing of the unigene set of tomato ESTs. INRA-Bordeaux (C. Rothan) is setting facilities for implementation of TILLING strategy on the tomato model. Tomato proteome is characterized at INRA-Avignon (M Causse).

- *Metabolomic platform.* Metabolic profiling is analysed through NMR and HPLC at INRA-Bordeaux.

Major areas of biological interest are

- *Fruit development and fruit ripening.* The control of fruit development and ripening is addressed through the characterisation of hormone-dependent regulation of gene transcription (INRA-Toulouse). Early stages of fruit development are studied through investigation of the control of cell division and cell enlargement (INRA-Bordeaux). Fruit composition in relation to consumer perception is studied at INRA-Avignon
- *Metabolism, transport, and sink-source interactions.* This research is mainly conducted at INRA-Bordeaux on the tomato.
- *Secondary metabolism.* Research on carotenoid biosynthesis in relation with plast differenciation are developed at CNRS/University Fournier, Grenoble.
- *Fruit cell wall structure and modifications.* Cell wall genes involved in fruit firmness are studied at INRA Avignon and modifications of cell wall structure during fruit development are addressed through biochemical and biophysics approaches at INRA Nantes.
- *Disease Resistance in Solanaceae.* Among the major pathogens studied in several SOL species are potyviruses (INRA-Avignon and Bordeaux), nematodes (INRA-Avignon and Antibes) and Phytophthora (INRA-Avignon)

Potential Contribution of France to the Tomato Sequencing Project

Decision on whether France they will make a major contribution to the tomato genome sequencing project is still under discussion. A proposal will be submitted to CNS (National Centre for Sequencing) in 2004 (likely March/April). An EU funding to support contribution to the tomato sequencing will also be sought and a proposal will be submitted in 2004.

Following discussion with INRA national authorities, it is likely that France will strongly contribute to the functional genomics program through the generation of tomato generic resources for high throughput reverse genetics (EMS, TILLING, proteomics).

8) GERMANY

Centers of *Solanaceae* research in Germany

Research Institutions
Max Planck Institute for Plant Breeding Research (MPIZ), Cologne
Max Planck Institute for Molecular Plant Physiology (MPIMP), Golm
Max Planck Institute for Chemical Ecology (MPICE), Jena
Institute for Plant Breeding and Crop Plant Research (IPK), Gatersleben
Institute for Plant Biochemistry (IPB), Halle

Universities
University of Halle
University of Cologne
Technical University Munich/Weihenstephan (TU Weihenstephan)
University of Hannover

Breeding Companies

Germany has several, strong potato breeding companies, which are organized in the Gesellschaft für Pflanzenzüchtung e.V. (GFP), Arbeitsgruppe 13 (Kartoffeln).

Agrochemical companies

Plantech (Bayer Crop Science)

Sources of Support

Research is supported by grants from the German Ministry for Research and Education (BMBF, e.g. the GABI programme), by the Deutsche Forschungsgemeinschaft (DFG, e.g. a “Sonderforschungsbereich” SFB in Halle), the European Union (e.g. the EuDicot and APOPHYS projects) and by institutional resources of the Max Planck society and the two Leibniz Institutes (IPB and IPK).

Major Topics of Research

- *Production and molecular characterization of genetic resources, mapping of agronomic traits, marker development, analysis of gene activation in response to biotic and abiotic stress*. This research focuses on potato. It is centered at MPIZ-Cologne and at the TU Weihenstephan, partly in interaction with breeding companies.

- *Positional cloning, physical and comparative mapping and genomic sequencing in potato and tomato.* This type of research is carried out at MPIZ-Cologne.

- *Development and implementation of genomics analytic platforms*. This occurs in particular at MPIMP-Golm (expression arrays, metabolomics) and at IPK-Gatersleben and IPB-Halle (metabolomics, in the latter case, especially secondary metabolism).

Major areas of biological interest are:

- *Genome analysis*. Research on the potato genome started 1986 at MPIZ-Cologne (C. Gebhardt) with the construction of molecular maps and proceeded with the localization of genetic factors controlling resistance to pathogens and tuber traits. This work led to the molecular cloning of the *R1* gene for resistance to late blight and the *Gro1* gene for resistance to the root cyst nematode *Globodera rostochiensis*. Major interests are (1) the molecular basis of QTL for quantitative resistance and QTL for traits related to carbohydrate metabolism (tuber starch and sugar content, cold sweetening), (2) population genetics (LD mapping, haplotyping, natural variation) of breeding pools and wild potato species, (3) comparative function maps.

- *Metabolism, transport, and sink-source interactions.* This research was initiated in the early 1990's, profiting from the possibility for routine transformation of Solanaceae species. Initial work was often in tobacco, but quickly potato and tomato became the major species used. The work includes detailed analysis by reverse genetics of many major pathways in central carbon and nitrogen metabolism, also analysis of transport protein families. A particular area of research was the pathway of starch synthesis, resulting in the ability to radically alter the structure and properties of starch. Major centers for this research include the MPIMP-Golm, IPK-Gatersleben and the Universities of Cologne (Prof. Flügge) and Hohenheim (Prof. Von Wiren; The biotech company Plantech - now part of Bayer Crop Science) was a product of these research activities. Recent developments include studies of hormonal signals and of bacterial pathogenic factors that influence sink growth and cell functions (MPIZ-Cologne, IPB-Halle, University of Halle, IPK-Gatersleben).

- *Secondary metabolism.* The IPB-Halle in particular has a long-standing expertise in the analysis of phenylpropanoid and alkaloid metabolism in *Solanaceae* spps.

- *Biotic and abiotic stress.* This on the one hand encompasses the mapping of resistance genes (MPIZ-Cologne, University of Hannover, TU Weihenstephan), and on the other hand the analysis of stress responses with genetic, biochemical and genomics tools (in particular at the IPB-Halle, the University of Halle, the University of Frankfurt (Prof. Nover), TU Weihenstephan (Prof. Wenzel).

- *Chemical Ecology.* Research on Solanaceae spps. plays a central role in the molecular analysis of plant-insect interactions at the MPICE-Jena, in particular in the work of Profs Gershenzon and Baldwin.

- *Plant architecture*. The origin of axillary meristems, shoot branching and apical dominance is studied at the molecular level in tomato at MPIZ-Cologne (K. Theres)

Potential Contribution of Germany to the Tomato Sequencing Project

At this time, it is unlikely that funds are available in Germany to provide significant direct support to a Solanaceae sequencing program (see below).

The major potential contribution of Germany to a potential tomato sequencing program would be to provide strong support for follow-up activities to ensure effective use of the information contained from the tomato genome sequence. In particular:

- In providing resources and programs to allow transfer of the results from tomato into potato breeding programs. Tomato breeding and cultivation is not a major activity in Germany, whereas potato is an important crop. It is also an important crop in some of the new members to the EU (e.g., Poland).

- In providing centers with a strong analytic infrastructure to allow analysis of genetic diversity in potato, tomato and other Solanaceae. In addition to basic research, this directly addresses topics related to the nutritional and health value of crops, and therefore provides a strong support for breeding programs. These are already being used, for example, in an in-depth analysis of tomato introgression lines to identify loci that produce major shifts in metabolism. The possibility of attracting large funding for a systematic study of metabolite levels in several crops including potato and tomato, in order to provide a database of food composition, a background for assessment of substantial equivalence, and a background for breeders to use in breeding in the future.

Deadlines For Submission Of Genomics Submissions

The framework and aims of the second phase of the GABI programme have already been implemented and the reviewing process for projects in the second phase is completed, and does not include this aim. The DFG is not really able to support activities on this scale, but concentrates instead on funding focused biological projects.

However, in parallel with attempts to raise international support for Solanaceae research and more specifically for tomato genome sequencing, it will be possible in Germany to seek sources of money internally for the development and application of these tools.

9) ISRAEL

Centers of *Solanaceae* research in Israel

Research Institutions
Agricultural Research Organization (ARO), The Volcani Center
MIGAL Research Institute
The Weizmann Institute of Science
The Technion Israel Institute of Technology

Universities
Tel Aviv University
The Hebrew University of Jerusalem
Ben-Gurion University of the Negev

Sources of Support

The research is supported in part by national foundations (Ministry of Science, Ministry of Agriculture, Israeli Science Foundation) and by bi-national foundations (BARD, GIF, DIP, BSF).

Major tools of Research

1. *Introduction of exotic germ plasm, marker assisted breeding and cloning of important agronomical traits as well as analysis of synteny.* This research, which focuses on tomato and pepper, is performed at the ARO, the Hebrew University and the Weizmann Institute of Science.

2. *Development of EMS, Fast Neutron Ds transposon insertion population in tomato.* This research is performed at the ARO, the Hebrew University and the Weizmann Institute of Science.

3. *Development and implementation of genetic and genomics platforms.* This research is performed at most institutes and universities.

4. Development and implementation of trans-activation system in Tomato and tobacco. This research is performed at the Weizmann Institute (Yuval Eshed).

5. *Genetic and molecular characterization of carotenoid biosynthesis*: The groups of Zamir and Hirschberg have identified 110 single-gene mutations in *L. esculentum* CV M82, which are impaired in carotenoid biosynthesis or accumulation in fruits, flowers and leaves. Biochemical and molecular characterization of these mutants will provide unique insights into isoprenoid biosynthesis and its regulation.

6. *Transcriptional profiling by cDNA microarray.* A microarray of more than 2,000 fruit-specific cDNAs has been established at The Hebrew University. This array, together with arrays obtained from Cornell, is being used to monitor global gene expression in fruits and flowers of various tomato mutants.

Major areas of biological interest

a. *Molecular biology and metabolism of fruits.* This research is preformed on the model plant tobacco as well as on tomato and pepper. The studies include isolation and analysis of promoters and genes expressed in early stages of fruit development, identification of genes, regulatory elements and enzymes that modulate various traits such as sugars, amino acids, secondary metabolism (vitamins, flavonoids, carotenoids, aroma), hormones, cuticle development, fruit size and shape, pungency, color, fragrance, aroma, texture, amino acids. There is great emphasis in the research on the carotenoids biosynthesis pathway were a number of the key genes were isolated and characterized. Major centers for this research include the ARO, MIGAL Research Institute and The Weizmann Institute of Science and the Hebrew University of Jerusalem.

b. *Biotic and abiotic stresses.* The research on plant pathogen interactions includes pathogen stresses as well as salt draught stresses. Methodologies include identification of genes associated with disease-resistance, signal transduction of stress responses including reactive oxygen species response, genes and networks controlling stress responses in relation to growth and development. Breeding programs, including development of molecular markers, includes tomatoes for resistance to TYLCV resistance, pepper for resistance to TMV, CMV, TSWV and powdery mildew, root knot nematodes. The research on abiotic stresses includes the role of small heat shock proteins in the protection against extreme temperature stresses (tomato), the genetics of water loss and drought stress (tomato and pepper) and the functions of aquaporins in draught and salt stresses (tomato). Major centers for this research include the ARO, The Weizmann Institute of Science, the Hebrew University of Jerusalem and Ben-Gurion University.

c. Control of plant physiology and development. This research includes phloem genomics, as well as molecular, genetic and genomics studies on he role and function of specific xanthophylls in photosynthesis, organ morphogenesis, leaf polarity, polar morphology, lateral expansion of plant organs, the control of the transition from vegetative growth to flowering, flowering time as well as flower and fruit development, is studied using genetic and molecular tools, such as mutant analysis, and identification of important genes. Expression patterns and functions of crucial genes are characterized. In addition, circadian clock and its relation to flowering time are studied. Mutant analysis, genetic mapping and molecular analysis is used to study the genetics of development in tomato, in particular meristem function, plant architecture and leaf shape. GA signal transduction and response are studied in petunia. Major centers for this research include the Weizmann Institute of Science, the Hebrew University of Jerusalem and the Technion Institute of Technology, Haifa.

Resources

1. cDNA libraries of tomato flower and young fruit, subtracted cDNA from parthenocarpic fruit. cDNA library of pepper fruit, genomic library of *L. hirsutum*. Pepper mapping populations: *C. frutescens* introgression lines (under construction), *C. chinense* introgression lines (some chromosomes), F2 intra-specific *C. annuum,* BC2 inter-specific *C. annuum x C. frutescens* (The ARO).

2. -EMS population: ~1000 families in Micro-Tom.
- Insertion lines with a Ds transposon: ~3000 lines out of which only 800 were recently propagated, the rest of the seeds is ~8 years old. (Avi Levy, The Weizmann Institute).
- 6000 EMS and 8000 Fast Neutron mutants in the variety M82 (The Hebrew University - 'The genes that make tomatoes' a saturated phenotypic database http://zamir.sgn.cornell.edu/mutants/). Seventy six *L. pennellii* introgression lines in M82 are distributed by the Tomato Genetics Resource Center in Davis, CA. <http://tgrc.ucdavis.edu/>

10) ITALY

Centers of *Solanaceae* research in Italy

Research Institutions
National Agency for New technologies, Energy and the Environment (ENEA), Biotechnology Unit (Rome, Matera)
National Research Council (CNR), Institute of Plant Genetics (Naples, Bari)
Stazione Zoologica (SZN), Naples

Universities
University of Viterbo
University of Naples
University of Bari
University of Siena
University of Florence

Companies
Metapontum Agrobios
Farmen SnC
Biogen SrL

Sources of Support

The research is supported by the Italian Ministry of Research and the Italian Ministry of Agriculture.

Major topics of Research

- *EST sequencing and re-sequencing*. This is mainly performed at ENEA and by a group coordinated by C. Bowler. The goal is to re-, within June 2004, 12.000 Cornell ESTs.

- *Transcriptional profiling* Several groups are using arrays bought from TIGR (potato, Univ. of Naples) and Cornell (tomato, ENEA/ CNR Naples/ SZN). SZN has developped a "boutique" macroarray and ENEA a "boutique" microarray, comprising mainly genes for light signal transduction/carotenoid biosynthesis. Several groups are using differential display and cDNA-AFLP as profiling tools.

- *Mutant isolation/characterization*: The group of G.P. Soressi (Univ. of Viterbo) has a large collection of mutants for fruit colour/quality and plant architecture.

- *Virus-induced gene silencing* It is being used at ENEA as a functional genomics tool (M. Tavazza/ G. Giuliano) using PVX and blue light photoreceptors. Constructs are now being developed in TRV.

- *Bioinformatics:* Tomato promoter analysis using bioinformatic tools is being developed at the University of Florence (M. Buiatti).

Major areas of biological interest are

- *Carotenoid biosynthesis/fruit quality in tomato*. Transcriptional regulation/metabolic engineering studies are conducted at ENEA (C. Rosati/G. Giuliano. Some metabolic engineering is also conducted at Metapontum Agrobios (F. Cellini). QTLs for fruit carotenoid content are being mapped in Naples (S. Grandillo/L. Monti). A large mutant collection is available at the University of Viterbo (A. Mazzucato-G.P. Soressi)

- *Photomorphogenesis in tomato.* The group of C. Bowler (Stazione Zoologica di Napoli) focuses mainly on phytochrome signal transduction, while the ENEA group (G. Perrotta/G. Giuliano) focuses mainly on cryptochromes.

- *Biotic and abiotic stress.* Main stresses under study comprise fungi (oidium), for which a monogenic recessive resistance has been mapped at the University of Bari (L. Ricciardi); viruses (TYLC, CMV) for which transgenic resistant lines have been developed at ENEA (M. Tavazza) and Metapontum Agrobios (F. Cellini); insects, for which genetic engineering is ongoing at the University of Naples (L. Monti); salinity/drought, for which mutant studies/transcriptional profiling are ongoing at Naples (S. Grillo/L. Monti) and ENEA Trisaia (A. Maggio); cold and soft rot in potato, for which interpsecific hybridization is ongoing at the University of Naples (L. Frusciante).

- *Profiling of local ecotypes.* This is being performed using a combination of traditional (biochemical and sensory profiling) and molecular techniques (transcriptional profiling, development of molecular markers) at several institutes, including CNR, ENEA and SZN.

- *Reproductive biology.* The main centers are at the University of Siena (M. Cresti, biology of pollen), University of Viterbo (G.P. Soressi/A. Mazzucato, parthenocarpy), and University of Naples (L. Frusciante, meiotic mutants in potato).

Potential Contribution of Italy to the Tomato Sequencing Project

Although it is unclear, at the present moment, if Italy will contribute to the genomic sequencing effort, plans are being put in place for applying for funds for the sequencing of the euchromatic portion of a chromosome (approx 20 Mb). The actual sequencing would be performed at the University of Padua and at ENEA.

Deadlines For Submission

A first deadline will be at the end of 2003/beginning of 2004, with further deadlines in 2004.

11) JAPAN

Places of Solanaceae *research in Japan*

Research Institutions

Kazusa DNA Research Institute, Chiba
National Institute of Vegetable and Tea Science (NIVTS), Mie
Chiba Prefectural Agriculture Research Center, Chiba
National Institute of Genetics, Mishima

RIKEN, Yokohama

Universities
Tsukuba University
University of Tokyo
Tohoku University
Hokkaidou University
Tokyo University of Agriculture and Technology
Nagoya University

Companies in Japan with commercial interests in Solanaceae
Sakata Seed Corporation
TAKII & Co., LTD.
SUNTRY Ltd.
KAGOME CO.,Ltd.
Nippon Del Monte Corporation

Sources of Support

In Japan, research is supported by grants from Ministry of Education, Culture, Sports, Science and Technology (MEXT)_the Japan Society for the Promotion of Science (JSPS)_Ministry of Agriculture, Forestry and Fisheries (MAFF) and The New Energy and Industrial Technology Development Organization (NEDO). Unlike other institutions, Kazusa DNA Research Institute is supported by the grant from the Chiba local state government.

Major Topics of Research

- EST, full-length cDNA sequencing

37,000 ESTs from 5’ end of cDNA clones prepared from fruits and leaves of a dwarf tomato, MicroTom were sequenced. PCR products of independent 11,000 EST clones were spotted on Nylon filters for macroarray analysis.
Full-length cDNA libraries of MicroTom are ordered.

- Metabolomics
- Disease resistance
- Fruit development and ripening
- Plant development and hormone biology
- Tobacco BY2 cell biology
- Eggplant breeding

Potential Contribution of Japan to the Tomato Sequencing Project
We have not yet discussed about how we contribute to tomato genome sequencing itself. However, we will contribute an important part of the Tomato Sequencing Project, sequencing of tomato full-length cDNA clones, especially those specific to tomato. As shown in the full-length cDNA sequencing projects of Arabidopsis (RIKEN, Seki et al., *Science* 296: 141-145, 2002) and rice (MAFF, Kikuchi et al.,

Science 30: 376-379, 2003), to which Japanese research groups have contributed substantially, full-length cDNA sequence information is crucial to determine the transcribed and spliced sequences from the genome nucleotide sequences, especially those of previously unknown genes. We plan to sequence at least 1,000 full-length clones within one year. If we have a grant of full-length cDNA sequencing, we will increase the number of clones accordingly.

12) KOREA

Centers of *Solonaceae* Research in Korea

Research Institutions
Korea Research Institute of Bioscience and Biotechnology (KRIBB), Taejon
National Institute of Agricultural Biotechnology (NIAB), Suwon
National Horticultural Research Institute (NHRI), Suwon
National Alpine Agricultural Experiment Station, Pyungchang

Universities
Seoul National University, Seoul
Korea University, Seoul
Kyungbook National University, Taegu
Chungbook National University, Cheongjoo
Kangwon National University, Chuncheon

Breeding Companies
Nongwoo Bio. Co., Yeujoo
Heungnong Seed Co. (Seminis Inc.), Seoul
Joongang Seed Co. (Seminis Inc.), Seoul
more than 10 small seed companies are selling pepper and tomato seeds.

Sources of Support

The research could be supported by CFGC (Crop Functional Genomics Research Center, a 21st Frontier Research Program of Ministry of Science and Technology), Biogreen 21 program of Rural Development Administration and by institutional resources of each institute. Individual grants are also from KOSEF (Korea Science and Engineering Foundation).

Major Topics of Research

- *Molecular mapping and BAC library construction for pepper genome*. This research is performed in CPMGBR (Center for Plant Molecular Genetics and Breeding, funded by KOSEF) in Seoul National University (Dr. B.D. Kim) and focused on pepper. The research areas include molecular cloning of disease resistance and quality traits, pepper BAC library construction, development of various plant transformation vectors and detection of QTLs for agronomically important traits.

- *EST, Microarry and construction of platforms for functional genomics of pepper*. This occurs in particular at the Plant Genomics Lab. of KRIBB (Dr. D. Choi). The research topics include generation of pepper EST (currently around 30,000 EST), development of microarray using generated EST, development of bioinformatics tools and pepper expression profile DB. This project is currently funded by CFGC (Crop Functional Genomics Research Center) and PDRC (Plant Diversity Research Center) both are 21st century frontier research programs of ministry of science and technology.

Major areas of biological interest are

- *Plant-microbe interactions.* Four major labs in Korea are focused in this research. Functional genomics of pepper transcription factors which responsed to pathogen is performed using microarray, over-expression in model plants and VIGS techniques (Dr. D. Choi in KRIBB). Molecular mapping and cloning of hot pepper resistant locus for *Xanthomonas campestris* pv. *vesicatoria* and *Phythophthora capsici* are performed using molecular genetics approach (Dr. B.D. Kim in Seoul National University). Characterization of genes involved in systemic acquired resistance of pepper and studies on signaling network of disease resistant are also performed (Dr. K.H. Paek and Dr. B.K. Hwang in Korea University). Marker assisted selection for breeding of pepper multiple disease resistance are also performed (Dr. H.K. Park in Seoul National University). Development of multiple disease resistant pepper root stocks is performed using classical breeding techniques (Dr. B.S. Kim in Kyungpuk National University).

- *Functional genomics using VIGS.* Virus-induced gene silencing is a powerful tool for studying functions of a number of genes. VIGS of selected *Nicotiana benthamiana* genes for plant development and signaling is performed by Dr. H.S. Pai in KRIBB.

- *Metabolic engineering.* Generation of small scale EST and metabolic engineering of patato for lignans production are performed by Dr. J.H. Jeon in KRIBB.

- *Abiotic stress.* Functional characterizations of pepper genes involved in environmental stresses and studies on functional genomics of osmosensors in pepper are performed by funding from CFGC (Drs. C.B. Hong and Y.N. Hong in Seoul National University). Signal transduction studies on cold stressed pepper using cDNA microarray are also performed (Dr. H.B. Kwon in Sunmoon University).

- *Transformation technology development.* Pepper has extremely low frequency of transformation using currently available techniques. There are two government-funded research projects are performed by private company (Dr. C.H. Harn in Nongwoobio Co. and Dr. Y.S. Kim at Kumho Co.)

Potential Contribution of Korea to the Tomato Sequencing Project

As shown in this write-up, *Solanaceae* genomics researches in Korea are mainly focused on hot pepper plant. Even though *Solanaceous* plants are important crops in Korea (2.2 billion US$ per year), tomato is a minor crop among them. Subsequently, there is no research project on tomato genomics in Korea at this moment. Research scientists in *Solanaceae* genomics area could agree with the importance of tomato genome sequencing as a reference genome of *Solanaceae* plants but we still have long way to persuade the funding agency and plant research community in Korea.

Following is the potential contribution of Korea to the Solanaceae Genome Initiatives program in conditions of separate funding is available from our government:

- Participate in sequencing of one tomato chromosome (hopely chromosome No. 2 and up to 20 Mb).
- Parallel sequencing of portion of pepper chromosome (linkage group No. 2)
- Shotgun sequencing of euchromatin region of pepper chromosome (up to 50,000 random sequences for two years term).
- Extending of pepper EST to 100,000 (currently we have about 30,000 EST from 8 different cDNA libraries) using more disected cDNA libraries.

In best case, we can perform all of above. However, in worst case, we may only perform shotgun sequencing of euchromatin region of pepper chromosome and ESTs.

Deadlines for Submission of Grants Proposal

There are two possible funding sources in Korea for this project. One is BioGreen21 Program operated by Korean Rural Development Administration (RDA) and the deadline for the submission of grant proposal is February 2004. The other is Crop Functional Genomics Center (CFGC), the 21st century frontier research program operated by Ministry of Science and Technology (MOST) and the deadline for research proposal will be July 2004. However, the framework and aims of the both programs have already been implemented and we need discuss about the importance of the Solanaceae genome initiatives with them before we submit our proposal.

13) PERU

The situation of the Andes as the space for evolution of several genera and species of Solanaceae should be considered. In addition to Peru, Ecuador and Colombia might be included as participating countries to represent Andean diversity (and developing countries). These countries' national programs conduct conservation, evaluation and breeding research on Andean fruits (tree tomato, S. quitoense).

Peru/CIP- Potato

1. Germplasm: extensive ex-situ genebank of wild and cultivated potato species held and researched in trust at CIP under FAO auspices.
2. References: Potatoes of Peru, Ochoa; Potatoes of Bolivia, Ochoa- both books fully illustrated with habitats and suggestions for use of the gene pools
3. Institutions: Univ. Nacional Agraria (La Molina); Univ. Nacional (Cuzco)- hybrid populations and scientific capacity and genetic resources for crop improvement, but scarce economic resources

Peru- Tomato

1. Germplasm: 6 wild species are native; no *ex situ* genebanks are supported (desperately needed for building national capacity and for global research).
2. Institutions: No breeding/scientific program of any kind; only about two people in Peru would agree that they dedicated part time (20%) work to this group.

3. Status: The collection and documentation has finished the exploration phase. The challenge is to find the biological, logistical, and utilitarian reasons to continue collection/evaluation. Rick´s thesis of potentials from ecological observations used by E. Vallejos, in finding chilling injury ¨genes¨ in *Lycopersicon hirsutum* close to 3,000 msnm is a proper model.
4. Opportunities: promote field-work especially for systematic collection, conservation, evaluation and use. E.g. *L. parviflorum*- 3500 m. resistance to cold temperatures and insects; *L. pimpinellifolium*; *cercaeifolium* - direct exploitation for market and home use; use in breeding. Experienced individuals can help by reducing the chances of reinventing the wheel, providing complementary information that may be usefully integrated into current data bases (e.g. the correct spelling of collection sites). "Gray literature" papers should also be scanned for info.

Andean countries: Solanaceae fruits

Ecuador and Colombia's national programs conduct conservation, evaluation and breeding research on Andean fruits (tree tomato, S. quitoense).

CIP Research Interests:

1. Phenotypic, histological (cell level) and eventually genetic description of the Phytophthora - Solanum interaction via a systems approach - i.e., comparing pathogenicity and adaptation of the pathogen to diverse hosts, infection/resistance process, and virulence across host genera.
2. Functional characterization of allele series for host-pathogen interaction: viruses and fungi that affect multiple Sols.

Overview of CIP Activities in Genomics

A) CIP is currently applying sequence- expression-based strategies to characterize genes and regulatory systems related to disease resistance and carbohydrate metabolism in potato and sweetpotato, through a combination of in-house and collaborative arrangements. Gene function-based strategies are being explored and anticipated to be of much utility in the near future.

1. *Consensus and functional maps of disease resistance in potato* (late blight) (Collaborative with CU, CIAT, CRP-CU, UCB, UW, MPI)
   a. Construction of candidate gene collection and sequence databases for hypothesis testing from comparative biology (resistance and defense; heterologous source).
   b. Comparative QTL analysis, candidate gene mapping and cladistic analysis of resistance trait architecture (Solanaceae).
   c. Association genetics in genebank accessions and bred stocks
   d. Expression profiling (cDNA-AFLP) and function assignment toward identification of candidate genes for general resistance/defense.
   e. BAC library Construction from CIP-bred source of general resistance for characterization of clustered gene families (R genes, osmotin) and applications of structural genomics to gene discovery
2. *Search, visualization and analytical tools for applications of comparative biology to primary traits of potato and sweetpotato* (Collaborative with NCGR; CGIAR; CU)
   a. Participation in design of a comparative genomic map viewer on the basis of ISYS.
   b. Upgrading of potato genetic maps with function and position-based tomato-Arabidopsis COS (EST).

c. Development of software prototype to visualize genomic maps and overlay marker, genes/ESTs, similarities, and QTLs to produce customizable graphics with hyperlink functionality.
d. Establishment of web-site to manage in-house bioinformatics databases, software and documentation based on open source and free technology, providing interfaces to Perl and Java and running under Windows or Linux/Unix.
3. *Gene discovery and diversity analysis for starch and sugar metabolism in sweetpotato and potato* (Collaborative with AAFC, NCSU and ARCS)
a. Development of ESTs from storage and fibrous root core libraries
b. Expression profiling and function assignment toward identification of candidate genes for root bulking and starch synthesis.
c. Trans-species phylogeny and clustering of starch gene sequences/organization
d. Assessment of diversity and discovery of superior alleles in germplasm by SNP
e. Development of allele-specific detection systems (SNP)
f. Mapping and QTL analysis with candidate genes for carbohydrate metabolism and storage root development
g. Diversity assessment and candidate gene expression in post harvest challenged native Andean potatoes resistant to cold sweetening
4. *Assembly of candidate gene collection and genomic tools to characterize stress response in colored Andean potatoes* (Collaborative with VTI, USDA/ARS)
a. Biochemical, phenotypic and GxE analyses of Andean potato varieties for nutraceutical production and stress response
b. Amplification, purification and full-length sequencing of candidate potato ESTs for stress tolerance, pigment and nutrient production, for construction of a “stress chip” microarray.
c. Expression profiling (microarray) under stress challenges and integrated databases for phenotype, sequence, expression and mapping data concerning stress response in Andean potato.
d. Characterization of diversity for micronutrient contents (Fe, Zn, vitamins, etc.); GxE and genetics for use

B) Expected Outputs:

A Consortium or Challenge Program on Root and Tuber Crop Genomics or Solanaceae Genomics would greatly enhance rates and achievements in the improvement of mandate species. The availability of improved varieties better suited to local needs and requiring fewer external inputs, will improve opportunities for the poor through higher availability of diversity, contributing to poverty alleviation and protection of the environment. The following practical outputs relevant to the goals and objectives of the CGIAR are envisioned.

a. Significant, rapid amplification of the available genetic resource base for crop improvement through access to a wider range of genetic information, leading to improved opportunities for breeding and variety development.
b. More basic and practical precise knowledge about genetic structures of important traits across taxa, and tools to orient and facilitate breeding, especially of complex and lesser-known species.
c. Minimial duplication of public and private research efforts and accelerated progress through crop, disciplinary and institutional synergies.
d. Databases, molecular and software tools and new varieties in the public domain, important vehicles for transfer the benefits of science research, conservation and improvement efforts to resource poor farmers.

Potential contribution to the tomato sequencing project:

The major support would consist of logistical and research orientation for *in situ* experimentation and interdisciplinary collaboration toward the discovery of adaptive genes/ allele series within and across genera;
Can contribute well-characterized germplasm collections (potato); herbaria (web-accessible), and possible regional coordination of research of specific topics –e.g. Phytophthora-Solanum interactions

Full program to use genomic information and tools in breeding (potato; other species.); Education, technology transfer/outreach program in crop and social science related to agricultural development, environment and health.

Consortium member of Challenge Program "Unlocking Genetic Diversity in Crops for the Resource Poor" (Director R. Zeigler) and "HarvestPlus" (Biofortification of staple crops; Director Howie Bouis)

Facilities, logistics for international exchange of germplasm; alliance with competent authorities at national and regional levels (global).

Deadlines for submission of Genomics proposals.

The Challenge Program "Unlocking Genetic Diversity in Crops for the Resource Poor" will call for competitive proposals for applications of comparative biology in early 2004. Case study: abiotic stress; alternative topics: resistance; partitioning; nutrition; quality.

14) SPAIN

Centers of *Solanaceae* research in Spain

Research Institutions
Centro Nacional de Biotecnología (CNB), Madrid
Centro de Edafología y Biología Aplicada del Segura (CEBAS), Murcia
Estación Experimental La Mayora (EELM), Málaga
Estación Experimental del Zaidín (EEZ), Granada
Instituto de Biología Molecular y Celular de Plantas (IBMCP), Valencia
Servicio de Investigación Agroalimentaria, Zaragoza

Universities
Universidad de Almería
Universidad de Barcelona
Universidad de Málaga
Universidad Politécnica de Cartagena
Universidad Politécnica de Valencia

Breeding Companies
Spain has strong seed and production breeding companies. The are mostly located in Almería and Murcia.

Sources of Support

The research is supported by the Ministry of Science and Technology (MCYT), by Genome Spain (Dependent of the MCYT and Ministry of Health), by the Instituto Nacional de Investigaciones Agrarias (INIA, MCYT) and by regional goverments.

Topics of Research

Due to the importance of production in Spain of several Solanaceae crops there are many topics covered. They are mainly based on small projects that are related to many biological aspects such as development, flowering, architecture, fruit formation and ripening, resistance to biotic and abiotic stresses, hormonal control of tuber formation, etc.

Recently, many labs have joined in the search for funds in order to generate genomic tools in Solanaceae. Tomato is one of the most important horticultural crops and Spain is the first producing and exporting country of tomato for fresh market. We pretend to develop genomic tools required for functional analysis of economically important genes. This proposal (GEnSOL) will support the following objectives: **(1)** Generation of cDNA libraries of complete transcripts (tomato), two hybrid system libraries (tomato), ESTs (tomato, pepper and potato) and DNA chips (tomato and pepper). **(2)** Gene expression profiles if tomato and pepper. **(3)** Set the conditions for proteomic analysis of tomato chromoplast. **(4)** Generation of tomato mutants by EMS mutagenesis. **(5)** Functional genomic analysis of tomato using genetic reverse techniques (VIGS, RNAi and TILLLING (Targeting Induced Local Lesions In Genomes). **(6)** Genetic analysis of natural variability of tomato and pepper.

The genomic tools developed in this project will be used for genome-wide analysis, mainly for functional characterization of genes expressed during fruit development, ripening and during responses of tomato plants to pathogen infection. Most importantly, this project represent the best way to collaborate in an international consortium (USA and European groups) dedicated to the genomic analysis of tomato, the most economically important horticultural crop in our country.

Potential Contribution of Spain to the Tomato Sequencing Project

At this time, the major opportunity to provide significant direct support to a Solanaceae sequencing in Spain comes from Genoma España and regional governments. For this purpose some members of the Spanish genome initiative will have a meeting with the persons responsible in Genoma España and legal representative of regional governments.

Another potential contribution of Spain to a tomato sequencing program would be to provide support for follow-up activities to ensure effective use of the information contained in the genome sequence. Thus, there are several genomics projects in Solanaceae currently under review. GenSOL. As previously mentioned, has been submitted to the Spanish Initiative in Genomics and Proteomics (McyT).. A second Solanaceae genomic Project BAInSOL is under review by the Genoma Spaña – Genome Canada where Solanaceae research as a priority area. The project addresses the question on how environmental stresses (biotic and abiotic) affect the sink strength in potato tubers/ tomato fruits and uses genomic tools (transcriptomics, proteomics and metabolomics) and reverse genetics on targeted genes. Up to 8 trilateral projects involving up to 15 different Spanish groups have been submitted to evaluation to McyT-GenoPlante-GABI initiative on genomics. All these projects using genomic tools and natural variability in different members of Solanaceae.

Deadlines For Submission Of Genomics Submissions

The possibility of submission of genomics proposal is closed as such. However the mission of Genoma España is to promote research in *genomics* and proteomics. Its aim is that Spain be an active partner in the search for new knowledge and the development of new applications. Because they look private investors and have autonomy as public investors there is not deadline as such.

15) TAIWAN

Centers or Institutes of *Solanaceae* research in Taiwan

TW-SOL-tomato is an initial joint team of graduate institute of Plant Biology in Taiwan University, Asian Vegetable Research and Development Center (AVRDC) and Known-You seed Co. LTD to focus on tomato studies. TW-SOL will be broadened in the coming years with Solanaceae research scattered in Academia Sinica, governmental institutes, universities and companies.

Sources of Support

Research is supported by Council of Agriculture (COA), National Science Council (NSC) and Asian Vegetable Research and Development Center (AVRDC).

Major topics of research

As for systems approach to diversity and adaptation, Taiwan locates in both tropic and sub-tropic areas and has diverse environments and species. Tomato occasionally incurs poor growth and fruit setting under hot-dry or hot-wet season in the tropics and sub-tropics. Therefore, AVRDC (http://www.avrdc.org/) has been breeding and collecting unique and substantial tomato germplasm, especially of "heat-tolerance" and "multiple diseases resistance" over the past 30 years. The consequent research is related.

1. *Production of genetic resources and mapping.* The germplasm of disease resistance, heat tolerance and other major horticultural traits are continuously collected by breeding program. Mapping of targeted genes follows.
2. *Improvement of fruit setting and production under high temperature.* Along with many tropical diseases, high temperature is the major bottleneck impeding fruit setting and summer production. We will focus on mining and studying “heat-stable” resistant genes against diseases such as bacterial wilt, bacterial spot, *fusarium* wilt and those caused by root-knot nematodes, and several viruses based on heat-tolerant tomato lines.
3. *Functional genomic analysis in tomato under heat-tolerance* Profiling the regulation of gene expression by heat-shock treatment in tomato is studied, especially those of lycopene related genes, genes involving heat-shock protein synthesis and heat-stable *R* genes.
4. *Comparative analysis* The heat-tolerance related comparative studies within Solanaceae family are proceeding.

Potential contribution of Taiwan to tomato sequencing project

Although Taiwan has finished sequencing rice chromosome 5, it is unlikely that funds are available in Taiwan to support tomato genome sequencing at this time. Nevertheless, any sequencing project related to heat-tolerant genotype and follow-up research activities will be strongly supported.

Deadlines for submission of genomics proposals
Council of Agriculture (COA) now guarantees the funding of these projects for this year. The 10-year-run TW-SOL proposal in the international SOL project is under discussion in National Science Council (NSC) and Council of Agriculture (COA), though both have been supporting ongoing relevant research for years.

16) THE NETHERLANDS

Centres of Solanaceae research in The Netherlands

- **Universities:**

Wageningen University
University of Amsterdam
University of Nijmegen

- **Research Institutes:**

Plant Research International
Agrotechnology and Food Innovations (formerly ATO)

- **Breeding companies:**

Potato: Agrico Research BV, Averis Seeds BV, HZPC Holland BV, C. Meijer BV, Van Rijn BV
Tomato: ENZA seeds, De Ruiter Seeds, Nickerson-Zwaan BV, Rijk Zwaan, SVS Holland BV, Syngenta Seeds BV

- Potato processing industry: AVEBE
- Genomics company: Keygene NV
- The Centre for BioSystems Genomics (CBSG; www.cbsg.nl) is the Netherlands network of plant genomics research. It links plant genomics research groups from Wageningen University, Utrecht University, University of Amsterdam and Catholic University of Nijmegen, the research institutes Plant Research International and Agrotechnology and Food Innovations, and the major Dutch agro-industrial parties in potato and tomato breeding and processing, in a scientific research programme with a strong market focus. The research has a strong focus on potato and tomato.

Sources of support

The research is supported in part by the Ministry of Economic Affairs and Ministry of Agriculture, Nature and Food Quality, the Netherlands Organisation for Scientific Research (NWO), The Dutch Technology Foundation STW. The Centre for BioSystems Genomics that focuses on genomics research is financed by the Netherlands Genomics Initiative, which is also part of NWO. In addition there are bilateral grants between the knowledge institutes and industry sometimes in collaboration with STW.

Topics of research

Biological research:

- **Plant-Pathogen Interaction:**

potato: Wageningen University and Plant Research International
tomato: Wageningen University, University of Amsterdam

- **Plant Quality:**

potato: Wageningen University

tomato: Wageningen University, Plant Research International, Agrotechnology and Food Innovations
petunia: University of Amsterdam, University of Nijmegen

- **Plant Genome:**

potato: Wageningen University, Plant Research International, Keygene NV
tomato: Wageningen University, Plant Research International, Keygene NV

Platforms:

- Genetics resources:

Centre for Genetic Resources, The Netherlands, Wageningen University, University of Nijmegen

- Mapping agronomics traits:

Wageningen University, Plant Research International, Keygene NV

- Genome Sequencing:

Greenomics, Plant Research International

- Micro-arrays, Proteomics and Metabolomics:

Plant Research International

The Netherlands contribution to the Solanaceae Genome Sequencing Initiative

Tomato:

- CBSG funds research to integrate Keygene's genetic AFLP map of tomato and the physical map consisting of BAC clones build by Rod Wing's group.
- CBSG funds sequencing a tilling path of BAC clones forming the top of chromosome 6
- CBSG funds sequencing of BAC clones located on chromosome 6 in regions involved in resistance and quality traits.
- NWO and CBSG fund bioinformatics research focused on tomato.

Potato:

- CBSG funds sequencing of 200-250 BAC clones located in region of the genome accommodating loci involved in resistance and quality traits.
- The Dutch Technology Foundation, STW and EU Framework Programme 5 fund the construction of a physical BAC map of potato that is expected to be released in summer, 2004.
- A ultra high density map of potato (>10.000 markers) was constructed with amongst others EU funding, and this genetic map is being integrated with the physical map.
- NWO and CBSG fund bioinformatics research focused on potato.

Deadlines for submission of genomics research project applications

The Netherlands Genomics Initiative (NGI) started funding the Centre for BioSystems Genomics in 2003. The projects involved in Solanaceous genome research have started.

In parallel NGI and CBSG have formulated a proposal to sequence the potato genome and are searching for funding such a project by (inter)- national agencies.
Also grant proposals are formulated by the CBSG partners to raise additional funds for Solanaceae research. Granting agencies may be NWO, Ministries, Technology Foundation STW, EU and industry.

17) TURKEY

Centers of Solanaceae Research in Turkey

Research Institutions
Citrus and Greenhouse Research Institute-Antalya
Agean Agricultural Research Institute
Blacksea Agricultural Research Institute

Universities
Çukurova University-Adana
University of Trakya-Edirne
University of Suleyman Demirel-Isparta
Akdeniz University-Antalya
University of Ankara-Ankara
University of Gaziosmanpasa
University of Harran-Urfa
Mustafa Kemal University-Hatay (MKU)
Ege University-Izmir (EGE)
Uludag University-Bursa
Yeditepe University-Istanbul
Izmir Institute of Technology (YTE)-Izmir

Sources of Support

Research is supported by grants from the State Planning Organization (DPT), Natıonal Science of Foundation of Turkey (TUBA), the Scientific and Technical Research Council of Turkey (TUBITAK), University Research Foundation and, to a limited extend, by the private sector.

Main Topics of Research

*Germplasm screening and adaptation studies for tomato, pepper, eggplant and potato (by many research institute and universities)
*Development of genetic resources and mapping in tomato, pepper and eggplant (by _YTE)
*Analysis of eggplant genome and comparative mappinf with other solanaceae species (By IYTE)
*Determination of genetic diversity in eggplant and its wild relatives (by IYTE)
*Genetic control of biotic stress tolerance including viral, fungal, bacterial, nematode and insect resistance in tomato, pepper and eggplant (by MKU, IYTE, YED_TEPE, EGE)
*Genetic control of abiotic stress tolerance including drought and salinity in eggplant and tomato (by IYTE, EGE)
*Analysis of secondary metabolites in eggplant (by IYTE)
*Analysis of antioxidant content in pepper, tomato, eggplant (by IYTE)

Potential Contribution of Turkey to the Tomato Sequencing Project

National funds are not available for large scale sequencing projects in Turkey. However, it may be possible to obtain national or international grants for a smaller project targeted at eggplant. Eggplant is one of the most important vegetable crops in Turkey which ranks third in worldwide production. This project will focus on refining RFLP based molecular genetic linkage map by the addition of AFLPs, ESTs and microsatellite markers mapped in tomato. In addition, methyl filtered BACs will be sequenced to determine synteny between eggplant and tomato genomes.

Datelines for Submissions

An EU and national support for eggplant genome project will be pursued in 2004.

18) UNITED KINGDOM

Places of *Solanaceae* research in UK

Research Institutions

Horticulture Research International	(HRI)
John Innes Centre	(JIC)
Kew	
Natural History Museum	(NHM)
Sainsbury Laboratory	(SL)
Scottish Crops Research Institute	(SCRI)

Universities

Many independent groups in various institutions (see list below)
Aberystwyth
Bangor
Glasgow
Lancaster
London
Nottingham
Oxford
UEA
Sheffield
Southampton
York

Companies in the UK with commercial interests in Solanaceae

ATC
Unilever
Syngenta

Industry related organisations

The British Tomato Growers Association

British Potato Council

Sources of Support

In England and Wales research is mainly supported either by Biotechnology and Biological Sciences Research Council (BBSRC) or Department for Environment, Food and Rural Affairs (defra). Horticultural Development Council (HDC), fund development of near market research for UK horticulture. In Scotland the Scottish Office fund research via Scottish Executive Environment and Rural Affairs Department (SEERAD).

Topics of Research

- *Disease resistance and VIGS technology*
- *Fruit development and ripening*
- *Plant development and hormone biology*
- *Taxonomy and comparative genomics*
- *Key independent research areas*

Disease resistance and VIGS technology
The UK is world leading in Solanaceous research in disease resistance and has many pioneering laboratories working on Virus Induced Gene Silencing. Groups working in these areas are focused in the JIC and SL research institutions with many other important laboratories present in other research institutes and universities.

Fruit development and ripening
Historically work on fruit ripening using tomato has been of central importance in UK's Solanaceae research community. Much of the initial work in this area has been focused at Nottingham University, which continues to be a centre of excellence along with other significant groupings including those at Warwick-HRI, those in London and those in industry.

Plant development and hormone biology
There are many UK based research groups taking advantage of the unique features of Solanaceous sp in studying plant growth and development. This includes both study of meristems, leaf development and hormonal control of plant development. There are many world class groups working in this research at UK institutions and universities.

Taxonomy and comparative genomics
Of fundamental importance in obtaining the tomato genome sequences will be to utilise this knowledge in comparative genomics. This will be extremely useful to those working in the world-class centres at Kew and NHM on taxonomy and to other researchers studying comparative genomics.

Key independent research areas
One of UK's strengths in Solanacesous research are the numerous groups working on specific unique features of solanaceous sp. This is highlighted in the numerous research groups shown

below. Of those research areas not already highlighted are those focusing on metabolomics and those working towards enhancing the production of key metabolites with potential health benefits.

Potential Contribution of UK to the Tomato Sequencing Project

Current discussions with BBSRC and defra are encouraging with the UK aiming to sequence ~25Mb of the gene rich regions of Chr4. We also intend to annotate this sequence and generate a mirrored portal for web based access to this data.

Deadlines For Submission Of Genomics Submissions
BBSRC next deadline is January 19 2004.

19) UNITED STATES

Centers of *Solonaceae* research in the US

Research Institutions
USDA-ARS, Berkeley. CA
USDA-ARS, Ithaca, NY
Boyce Thompson Institute for Plant Research, Ithaca, NY

Universities
University of California, Berkeley
University of California, Davis
University of Florida, Gainesville
University of Utah, Salt Lake City
University of Arizona, Tuscon
University of Wisconsin, Madison
Colorado State University
Cornell University, Ithaca, NY
New Mexico State University
North Carolina State University
Ohio State University

(Note: the above lists of research institutions and universities are not inclusive of all laboratories involved in *Solanaceae* research in the US. Researchers working on or with *Solanaceae* can be found at most every US university with a major plant/agricultural research program. Indicated are those institutions with programs emphasizing *Solanaceae* genomics and/or *Solanaceae* genome utilization.)

Seed Companies and Processors
A number of important seed companies and processors (some based in the US and some not) are either directly involved in *Solanaceae* breeding or have an interest in *Solanaceae* improvement in the US and include:
Seminis Seed

Heinz
Hunt-Wesson
Lipton Foods
Uniliever
Frito-Lay

Agrochemical and biotechnology companies
Syngenta
Monsanto
Exelixis
Mendel

Sources of Support

The majority of Solanceae genomics research in the US is supported by the NSF (National Science Foundation) -Plant Genome and USDA-NRI Plant Genome programs. Additional support is provided by the Department of Energy (DOE), the National Institutes of Health (NIH) and industry including those companies mentioned above.

Major Topics of Research

- *Development of genetic resources.*
- *Analysis of Solanaceae genomes and their evolution.*
- *Genomic tool kit development - ESTs genomic libraries and physical maps.*
- *Expression, protein and metabolite profiling.*
- *Bioinformatics*

Major areas of biological interest are:

- *Genome organization and evolution*
- *Pathogen response and resistance*
- *Abiotic stress - e.g. temperature, water, nutrient, heavy metal*
- *Primary and secondary metabolism*
- *Chemical ecology*
- *Hormone signaling and response*
- *Fruit development and ripening*
- *Development and architecture*

Potential Contribution of the US to the Tomato Sequencing Project

The diversity of ongoing Solanaceae genomics programs position the US well for substantial contributions to tomato genome sequencing. 90,000 tomato BACs have been fingerprinted (Arizona) and are currently being anchored to the genetic map (Cornell/USDA). ESTs collections have been developed for tomato (Cornell/BTI/USDA) and potato (Berkeley) which will facilitate genome annotation. Comparative maps of potato, eggplant, and pepper anchored to tomato (Cornell) will facilitate extension of genome information to the broader Solanaceae, as will ongoing potato physical mapping efforts (Berkeley). Cytogenetic analysis of Solanaceae genomes (Colorado) and analysis of organelle genome sequence integration into the nuclear genome (Rutgers) will further enhance analysis of genome organization and refine sequencing strategy and data analysis. Bioinformatics (Cornell, NC State) and evolutionary genetics (Utah) specialists are in place to assist with genome sequence and cross-species comparative analysis and breeders are poised to implement downstream sequence information into DNA marker systems (Davis, New Mexico, Cornell, Florida, Texas, Ohio).

The US group plans to submit a proposal to the NSF-Plant Genome Program in July of 2004 to request funding for sequencing of six tomato chromosomes or approximately, one half of the tomato genome, and corresponding appropriate informatics support. The US group has some preferences for target chromosomes but specific chromosomes to be sequenced will be determined once we have ascertained the interests of our international collaborators to minimize overlap and insure that those able to do only one chromosome can target the one most likely to result in funding from their respective granting agency..

As indicated by numerous international collaborators at the November 3, 2003 workshop in Washington, DC, funding in their respective countries is likely but contingent on successful efforts by other countries to insure that in the end their respective investments will be realized by a full tomato gene-space sequence. A significant effort undertaken in the US will insure the highest probability of funding for our international collaborators.

Deadlines For Submission Of Genomics Proposals

NSF-Plant Genome Program	January 24, 2004
USDA-NRI Plant Genome Program	January 9, 2004

8) References

1 Carroll SB (2003) Genetics and the making of *Homo sapiens*. Nature 422: 849-857.
2 Paran I and Zamir D (2003) Quantitative traits in plants: beyond the QTL. Trends in Genetics 19: 303-306.
3 Chase MW et al. 1993 Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene rbcL. Ann. Missouri Bot. Gard. 80: 528-580.
4 Paterson AH et al. (2000) Comparative genomics of plant chromosomes. Plant Cell 12: 1523-1539.
5 Bennetzen J (2002) Opening the door to comparative plant biology. Science 296: 60-63.
6 Van der Hoeven R, C Ronning, J Giovannoni, G Martin, SD Tanksley (2002) Deductions about the number, organization and evolution of genes in the tomato genome based on analysis of a large EST collection and selective genomic sequencing. Plant Cell 14: 1441-1456
7 Sweetlove LJ, Last RL, Fernie AR (2003). Predictive metabolic engineering: a Goal for Systems Biology. Plant Physiol 132: 420-425.
8 Hirschberg J (2001) Carotenoid biosynthesis in flowering plants. Curr Opin Plant Biol. 4: 210-218.
9 Pichersky E, Gang DR. (2000) Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. Trends Plant Sci. 2000: 439-45.
10 Giovannoni, J. (2001) Molecular biology of fruit maturation and ripening. Annu Rev Plant Physiol Plant Mol Biol 52: 725-749.
11 Menendez CM et al (2002) Cold sweetening in diploid potato: mapping quantitative trait loci and candidate genes. Genetics 162:1423-34.
12 Pedley KF, Martin GB (2003) Molecular basis of Pto-mediated resistance to bacterial speck disease in tomato. Annu Rev Phytopathol 2003:215-43.
13 Gebhardt C, Valkonen JP (2001) Organization of genes controlling disease resistance in the potato genome. Annu Rev Phytopathol. 2001: 79-102.
14 Dinesh-Kumar SP et al (2003) Virus-induced gene silencing. Methods Mol Biol 236: 287-94.
15 Tanksley SD and McCouch SR (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277: 1063-1066.
16 Koornneef M and Stam P (2001) Changing paradigms in plant breeding. Plant Physiol 125: 156-159.
17 Zamir D (2001) Improving plant breeding with exotic genetic libraries. Nature Review Genetics 2: 983-989.
18 Cong B et al. (2002) Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. Proc. Natl. Acad. Sci. USA 99: 13606-13611.
19 Bertin N et al (2003) Do genetic make-up and growth manipulation affect tomato fruit size by cell number, or cell size and DNA endoreduplication? Ann Bot 92: 415-24.
20 Kitano H (2002) Computational systems biology. Nature 420:206-10.
21 Minorsky PV (2003) Achieving the in Silico Plant. Systems Biology and the Future of Plant Biological Research. Plant Physiol 132: 404-409.
22 Jansen RC (2003) Studying complex biological systems using multifactorial perturbation. Nat Rev Genet 4 :145-51.
23 Pennisi E (2003) Systems biology. Tracing life's circuitry. Science. 302 :1646-9.
24 Ge H, Walhout AJ, Vidal M (2003) Integrating 'omic' information: a bridge between genomics and systems biology. Trends Genet 19 :551-60.
25 Mueller LA, Zhang P, Rhee SY (2003) AraCyc: a biochemical pathway database for Arabidopsis. Plant Physiol. 2003 132: 453-60.

26 Quiros CF (2003) Genomics and Breeding of Vegetable Crops http://veghome.ucdavis.edu/classes/vc221/syllb221.html.

27 Tanksley SD et al. (1988) Conservation of gene repertoire but not gene order in pepper and tomato. Proc. Natl. Acad. Sci. 85: 6419-6423.

28 Livingstone KD et al (1999) Genome mapping in capsicum and the evolution of genome structure in the Solanaceae. Genetics 152:1183-202

29 Doganlar S et al. (2002) A comparative genetic linkage map of eggplant (*Solanum melongena*) and its implications for genome evolution in the Solanaceae. Genetics 161: 1697-711.

30 Eshed Y and Zamir D (1995) An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141, 1147-1162.

31 Monforte AJ and Tanksley SD (2000) Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in a *L. esculentum* genetic background: A tool for gene mapping and gene discovery. GENOME 43: 803-813.

32 Ji Y and Chetelat RT (2003) Homoeologous pairing and recombination in *Solanum lycopersicoides* monosomic addition and substitution lines of tomato. Theor Appl Genet 106(6): 979-89.

33 Emmanuel E and Levy AA (2002) Tomato mutants as tools for functional genomics. Curr. Opin Plant Biol 5: 112-117.

34 Mathews H et al. (2003) Activation tagging in tomato identifies a transcriptional regulator of anthocyanin biosynthesis, modification, and transport. Plant Cell 15, 1689-1703.

35 Chen Q, (2003) Construction of two BAC libraries from the wild Mexican diploid potato, *Solanum pinnatisectum*, and the identification of clones near the late blight and Colorado potato beetle resistance loci. Theor Appl Genet In press.

36 Brugmans B et al. (2002) A novel method for the construction of genome wide transcriptome maps. Plant J. 2002 31 :211-22.

37 Vandenbussche M et al. (2003). Toward the analysis of the petunia MADS box gene family by reverse and forward transposon insertion mutagenesis approaches: B, C, and D floral organ identity functions require SEPALLATA-like MADS box genes in petunia. Plant Cell 15 :2680-93.

38 Bohs L and R Olmstead (1997) Phylogenetic relationships in *Solanum* (Solanaceae) based on *ndhF* sequences. Sys Bot 22:5–17

39 Heiser C and Anderson, G (1999) "New" solanums. In: J. Janick (ed.), Perspectives on new crops and new uses. ASHS Press, Alexandria, VA.

40 Patiño, V. 1963. Plantas cultivadas y animales domesticas en America equinoccial. Tomo I. Frutales. Imprenta Departmental. Cali, Colombia.