Genomics Approaches to Understanding Ripening Control and Fruit Quality in Tomato

J. Giovannoni
USDA-ARS and Boyce Thompson Institute for Plant Research
Cornell University campus
Ithaca, NY 14853
USA

Keywords: tomato, ethylene, fruit ripening mutants, microarrays, ESTs

Abstract
The maturation and ripening of fleshy fruits contributes a major component of human diets, nutrition and agricultural activity. While ripening brings about highly desirable changes in fruit character and chemistry in terms of flavor, appearance, texture and nutrition, the advanced stages of ripening lead to sub-optimal fruit quality and eventually post-harvest loss. Fruit biologists have studied numerous fruiting species with the intent of identifying strategies and technologies toward improving desirable ripening attributes while minimizing those with negative consequences. Tomato has emerged as a model for fleshy fruit ripening, in part due to its ease of use as a model system resulting from facilitating attributes including simple genetics, numerous characterized mutants, cross-fertile wild germplasm to promote genetic studies and routine transformation technology. In the last decade, the tomato system has been further complemented with molecular and genomic tools including dense genetic maps, large EST collections and the recently initiated genome sequencing effort. The isolation of genes corresponding to several previously described tomato ripening mutations has led to considerable advancement of a genetic regulatory model of fruit ripening while emerging genomics technologies promise exceptional opportunities for continued advancement of this field.

INTRODUCTION
Fleshy fruits including berries, stone, pome and numerous additional temperate and tropical fruits undergo a ripening process in which the biochemistry, physiology and structure of the organ are developmentally and biochemically altered to influence appearance, texture, flavor and aroma to attract seed dispersing organisms (reviewed in Seymour et al., 1993). While the specific regulatory and biochemical programs resulting in ripening phenomena vary among species, changes typically include: a) modification of visible pigmentation usually through alteration of chlorophyll, carotenoid, and/or flavonoid synthesis and accumulation; b) textural modification, often leading to softening, via alteration of cell turgor and cell wall metabolism and/or structure; c) modification of sugars, acids, and volatile chemistry leading to significant modifications in nutritional quality, flavor and aroma characteristics; and d) generally enhanced susceptibility to opportunistic pathogens (at least in part due to loss of cell wall structural integrity). In sum, these attributes define the ripening of most fleshy fruits and lead to enhanced fruit quality while simultaneously representing a production and distribution challenge in terms of increased post-harvest decay. While fruit species are classically defined physiologically on the basis of the presence (climacteric) or absence (non-climacteric) of elevated respiration and synthesis of the gaseous hormone ethylene at the onset of ripening (Lelievre et al., 1997), fruit displaying both ripening physiologies typically follow the general developmental changes outlined above.

Ripening is studied in numerous species though tomato has clearly developed as the most widely used model system for ripening due to the availability of extensive compatible germplasm, informative ripening mutants, comprehensive genetic maps and marker systems, efficient transformation and the recent development of genomics tools including large insert BAC libraries, extensive EST collections and an ongoing genome
sequencing effort (Reviewed in Adams-Phillips et al., 2004; Giovannoni, 2004). While the focus here is on the current state of tomato ripening research, especially as revealed through recent mutant analyses, it is important to recognize that other model systems for ripening are also important and are fairly widely used for research and practical purposes. For example, strawberry is a primary model for non-climacteric ripening (Aharoni and O’Connell, 2002) and peach, apple and citrus are developing as strong models for both ripening and genomics analysis of tree species. Peach is a target of increased genomics activity including development of substantial and broadly representative EST collections (http://www.genome.clemson.edu/gdr/ and http://genomics.msu.edu/fruitdb/analyses/nectarine.html) and has been a model for analysis of conserved ethylene response and ripening mechanisms in fruit tree species (Rasori et al., 2002; Rasori et al., 2003). While rapid advances are likely to occur in tomato, analysis of additional fruit crops is critical to identification of both novel and conserved mechanisms of ripening and associated post-harvest traits.

TOMATO MUTANTS ELUCIDATE MECHANISMS OF DEVELOPMENTAL REGULATION OF FRUIT RIPENING

The primary focus of ripening regulatory research in recent years has emphasized the synthesis, perception and role of ethylene in climacteric ripening (Klee, 2004). However, non-climacteric fruits such as strawberry can ripen with little induction of ethylene nor little apparent need for this hormone in driving the ripening process. The nature of “developmental” mechanisms that precede and regulate ethylene synthesis have represented a black box, as has the degree and nature of shared or common regulatory systems among fruit displaying climacteric and non-climacteric fruit physiologies. The tomato rin (ripening-inhibitor) mutation represented an important tool shedding light onto such mechanisms in that the mature fruit of the homozygous mutant fail to produce climacteric ethylene or ripen. Application of exogenous ethylene does not elicit ripening of rin fruit but can induce expression of ethylene-responsive genes, suggesting a role in regulation of both ethylene synthesis and components of ripening which function independently of ethylene. The predicted ethylene-independent component of climacteric ripening suggested by rin may also represent candidate components of any conserved aspect of ripening bridging climacteric/non-climacteric fruit development distinctions. The mutant rin locus was isolated via a positional cloning strategy that entailed high-density genetic mapping followed by isolation of large insert tomato genomics clones spanning the genetic markers that flanked the rin locus. Use of one such clone as a hybridization probe against ripe fruit cDNA libraries resulted in the discovery of a MADS-box transcription factor at the mutant locus (Vrebalov et al., 2002). Indeed, the basis of the rin mutation was shown to be a small deletion on chromosome 5 which altered two adjacent MADS-box genes, the second responsible for the large sepal phenotype characteristic of the rin mutant (Fig. 1). This second mutation is referred to as macrocalyx (mc). RIN-MADS homologous genes expressed in ripening fruit were identified in strawberry (non-climacteric), melon and banana (a monocot) suggesting conservation of RIN function among a broad range of species (Vrebalov et al., 2002). A second transcription factor corresponding to the non-ripening (nor) mutation was also isolated via positional cloning. nor/nor fruit are phenotypically similar to rin/rin in their lack of ripening and it appears NOR function includes regulation of RIN gene expression (Giovannoni, 2004). Isolation of the rin and nor loci has resulted in identification of two ripening transcription factor genes and evidence of their evolutionary conservation among a variety of fruit species (Fig. 2).

Physiological, genetic and molecular characterization of additional ripening mutants suggested that the phenotypically similar and dominant Green-ripe (Gr) and Never-ripe 2 (Nr-2) loci are allelic. Both map to the same region of chromosome 1 and both display phenotypes consistent with a role in fruit-specific ethylene response (Barry et al., 2005). The molecular basis of fruit-specific components of ethylene signaling remain to be identified. Gr and Nr-2 may represent such a component (Fig. 2) and both
are targets of positional cloning efforts in our laboratory. Preliminary evidence revealed a solid candidate for the $Gr$ and $Nr-2$ loci – a gene that appears to be altered in DNA sequence in both mutants.

REGULATION OF CAROTENOID ACCUMULATION BY LIGHT: IMPACT ON FRUIT RIPENING AND NUTRIENT QUALITY

Plants regulate levels of antioxidant pigments including carotenoids and flavonoids in part as a light response to protect against photo-oxidation. Carotenoids constitute the major component of ripe tomato fruit pigments and provide health benefits through antioxidant activity and as precursors to essential vitamins. For example, $\beta$-Carotene, $\alpha$-carotene, and $\beta$-cryptoxanthin all act as precursors of vitamin A. Dietary carotenoids also act as antioxidants and thus can quench and inactivate free radicals in the human body, in turn reducing net oxidative stress that can otherwise damage biomolecules. The result of increased antioxidant activity is reduced risk for chronic diseases including heart disease, diabetes and cancer. Lycopene, the primary carotenoid and red pigment of ripe tomato fruit, functions as a very potent antioxidant, and has been linked to reduced risk of prostate cancer and reduced risk of cardio-vascular disease (Sesso et al., 2003).

Though less important in tomato, flavonoids have several health-promoting characteristics and are also regulated by light (Adams-Phillips et al., 2004). Flavonoids are reported to have antibacterial, anti-inflammatory, anti-mutagenic, and anti-carcinogenic properties. Tomato fruit contain small amounts of flavonoids limited to their epidermal tissues. To increase flavonoid levels in tomato, the maize transcription factor genes $LC$ and $CI$ were expressed in tomato plants whose resulting fruits accumulated high levels of the flavonol kaempferol and the flavanone naringenin suggesting the potential for modification of tomato fruit to accumulate large amounts of flavonoids in addition to carotenoids (Bovy et al., 2002).

A well-characterized tomato mutation known as "high-pigment" ($hp$) is a rare example of a genetic regulator of total carotenoid accumulation and has long been of interest to tomato breeders as a tool to elevate carotenoid and associated color and nutrient activities in fruit. Cultivars harboring the $hp$ mutation typically yield fruit with 1.5-2 times the total carotenoids of nearly isogenic controls. While numerous carotenoid mutants exist in tomato, most represent lesions in steps in the synthesis pathway and as such result in shifts in carotenoid profiles to certain metabolites. $hp$ fruit are elevated in all carotenoids but maintain ratios similar to normal controls suggesting a regulatory rather than biosynthetic function for the corresponding gene product. We have recently isolated the gene responsible for the tomato $hp$ mutation (Liu et al., 2004). The $HP$ gene has homology to the human Damaged DNA Binding Protein 1 (DDBP1) shown to participate in a protein complex active in repair of UV-damaged DNA. Reduced activity of this gene in tomato is also related to light responsiveness, in this case, however, resulting in hyper-sensitivity to light leading to the production of additional antioxidant carotenoids and flavonoids as a protective measure against the perceived increased threat of photo-oxidation. Isolation of the $hp$ gene suggests that additional light signaling genes defined in non-crop species such as Arabidopsis (Fig. 3) may represent tools for modification of fruit nutrient and quality traits in crops species.

GENOMIC TOOLS FOR TOMATO AND THE SOLANACEAE

Interspecific sexual compatibility among cultivated tomato ($Solanum lycopersicum$) and a number of wild species has facilitated the development of high density genetic maps of tomato based on DNA marker loci (Tanksley et al., 1992). Using a set of highly conserved and single-copy DNA markers, researchers at several institutions have been able to develop integrated Solanaceae genetic maps that have facilitated both breeding and gene isolation in addition to localization of numerous loci corresponding to QTL and mutant loci (summarized at http://www.sgn.cornell.edu/). Extensive tomato germplasm collections including wild species, introgression lines and nearly isogenic mutants are also available through stock centers (Table I).
In addition to this excellent genetic resource, a large collection of tomato ESTs has been developed, with the heaviest concentration focused on fruit development and ripening (Van der Hoeven et al., 1992). The resulting EST collection is available to the research community and a subset has additionally been used to develop a public cDNA-based microarray. The tomato array includes over 8,600 sequence verified unigenes while raw and processed expression data resulting from replicated analyses of a time course of tomato fruit development is available on the web at the Tomato Expression Database (Alba et al., 2004, 2005). Comparative expression profiling of a normal and nearly isogenic ethylene insensitive cultivar revealed novel insights into ripening control including a role for ethylene in regulating gene expression of most carotenoid biosynthetic enzymes and provides evidence for a metabolic shift in amino acid synthesis toward methionine (a precursor to ethylene) at the onset of ripening (Alba et al., 2005). In addition, the fact that the tomato EST collection was developed from cDNA libraries that were not normalized or subtracted allowed for digital analysis of tomato gene expression based on EST prevalence (Fei et al., 2004). A key feature of such in silico analysis methods is the ability to compare gene expression (i.e. relative EST abundance) for homologous genes across diverse taxa. For example, analysis of highly homologous ESTs shared between tomato and Arabidopsis has facilitated identification of likely functionally equivalent homologs among otherwise similar family members (Fei et al., 2004). In a comparison of grape and tomato EST abundance, a number of genes induced during ripening and conserved between these species manifesting substantially different ripening physiology could be identified and included putative regulatory factors (Fei et al., 2004). The development of EST resources in additional fruit species will facilitate comparative genomics of fruit development and ripening and should prove useful in identifying additional conserved regulatory motifs. As the Solanaceae family includes important vegetable crop plants with highly conserved gene content it is feasible to utilize the tomato microarray resource for expression profiling of other members of this family (Moore et al., 2005).

INTERNATIONAL TOMATO GENOME EFFORT

Sequencing the tomato genome is the cornerstone of a larger international effort: “The International Solanaceae Genome Project” also known as SOL. The long-term goal of SOL is to establish a network of information, resources and scientists to tackle two of the most significant questions in plant biology and agriculture: 1) How can a common set of genes/proteins give rise to a wide range of morphologically and ecologically distinct organisms that occupy our planet? 2) How can a deeper understanding of the genetic basis of plant diversity be harnessed to meet the needs of society in an environmentally-friendly and sustainable manner? The Solanaceae also include the most important plant taxa in terms of vegetable crops. Fruit and tubers are major contributors of vitamins, fiber, carbohydrates, and phyto-nutrient compounds in our 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 contributing factors to human mortality.

Immediate application of the tomato genome sequence to other solanaceous species will be possible because the tomato genome is connected to these other species by comparative genetic maps and the level of microsynteny appears to be well conserved with respect to gene content and order (S. Tanksley, pers. commun.). Additionally, because the Solanaceae represent a distinct and divergent sector of flowering plants, distant from Arabidopsis, Medicago and rice (plant species for which genome sequence is available), the tomato genome sequence will provide a rich resource for investigating the forces of genome evolution over long periods of evolutionary time.

The tomato genome contains 950 Mb of DNA which is organized into 12 chromosomes (Arumuganathan and Earle, 1991). Unlike the chromosomes of maize or rice, in which heterochromatin and euchromatin are interspersed, the heterochromatin in tomato is concentrated around the centromeres. This pericentric heterochromatin is
largely devoid of genes, but constitutes approximately 75% of the DNA (Peterson et al., 1996; Van der Hoeven et al., 2002). In contrast, the distal portions of each tomato chromosome are comprised of largely contiguous stretches of gene-rich euchromatin which correspond to less than 25% of the DNA (Peterson et al., 1996). Rather than sequencing the entire tomato genome (950 Mb), the international participants of the tomato sequencing effort will sequence the approximately 220 Mb of euchromatin that contains the majority of genes (Table 1). Ten countries have been funded to date to sequence the tomato genome and the European Union has provided additional funds to insure completion. The project is estimated to be largely completed by 2008.

CONCLUSIONS
Recent advances in ripening research have led to discoveries concerning regulatory mechanisms preceding ethylene and which may represent evolutionary conserved ripening functions. The identification of homologous genes in other species and functional demonstration of ripening activities suggests these indeed represent conserved regulatory functions. The advent of genomics technologies including genome sequencing and their deployment in horticultural models such as tomato will likely accelerate discovery in ripening and other areas in the near future, providing new opportunities for both young and established researchers. The continued study of additional fruit systems will insure that insights developed in models such as tomato can be translated into a broad range of fruit crops for maximum benefit to producers and consumers.

ACKNOWLEDGEMENTS
Funding in the Giovannoni laboratory was provided by the National Science Foundation, USDA-NRI, USDA-IFAFS and the USDA Agricultural Research Service.

Literature Cited
Table 1. Public websites for information on tomato molecular biology and genomics.

<table>
<thead>
<tr>
<th>Site Name</th>
<th>Function</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solanaceae Genomics Network</td>
<td>maps, ESTs, sequencing, central access portal</td>
<td><a href="http://www.sgn.cornell.edu/">http://www.sgn.cornell.edu/</a></td>
</tr>
<tr>
<td>Tomato Gene Index (TIGR)</td>
<td>ESTs and their analysis</td>
<td><a href="http://www.tigr.org/">http://www.tigr.org/</a></td>
</tr>
<tr>
<td>C. Rick Tomato Genetics Resource Center</td>
<td>public germplasm center, mutant and wild stocks</td>
<td><a href="http://tgrc.ucdavis.edu/">http://tgrc.ucdavis.edu/</a></td>
</tr>
<tr>
<td>Genes that Make Tomatoes</td>
<td>mutant stocks</td>
<td><a href="http://zamir.sgn.cornell.edu/mutants/">http://zamir.sgn.cornell.edu/mutants/</a></td>
</tr>
<tr>
<td>Tomato Microarray Warehouse</td>
<td>public array data site, raw array data</td>
<td><a href="http://ted.bti.cornell.edu/cgi-bin/miame/home.cgi">http://ted.bti.cornell.edu/cgi-bin/miame/home.cgi</a></td>
</tr>
<tr>
<td>Tomato Expression Database</td>
<td>analyzed microarray expression data</td>
<td><a href="http://ted.bti.cornell.edu/array/interface/search/search.html">http://ted.bti.cornell.edu/array/interface/search/search.html</a></td>
</tr>
<tr>
<td>Digital Expression Database</td>
<td>electronic expression based on EST abundance</td>
<td><a href="http://ted.bti.cornell.edu/digital/interface/search/search.html">http://ted.bti.cornell.edu/digital/interface/search/search.html</a></td>
</tr>
<tr>
<td>Tomato Metabolite Database</td>
<td>correlative expression and metabolite analysis</td>
<td><a href="http://tomet.bti.cornell.edu/">http://tomet.bti.cornell.edu/</a></td>
</tr>
<tr>
<td>Center for Gene Expression Profiling</td>
<td>microarray distribution</td>
<td><a href="http://bti.cornell.edu/CGEP/CGEP.html">http://bti.cornell.edu/CGEP/CGEP.html</a></td>
</tr>
<tr>
<td>Tomato EST Orders</td>
<td>EST distribution, clones and libraries</td>
<td><a href="http://ted.bti.cornell.edu/order/index.htm">http://ted.bti.cornell.edu/order/index.htm</a></td>
</tr>
</tbody>
</table>
Fig. 1. Physical characterization of the *ripening-inhibitor* (*rin*) locus on tomato chromosome 5. Right angle arrows represent the transcription start sites of the MADS-RIN and MADS-MC genes whose coding regions are represented by filled and hatched chromosome segments, respectively. The black bar above the chromosome represents the span of the deletion found in the *rin* mutant. The distance separating the two genes is 2.6 kb.

Fig. 2. Regulatory model of fruit ripening defined via analysis of tomato genes.
Fig. 3. Known and predicted candidate genes impacting fruit carotenoid accumulation via light signal transduction.