A Quantitative Genetic Basis for Leaf Morphology in a Set of Precisely Defined Tomato Introgression Lines

Daniel H. Chitwood, Ravi Kumar, Lauren R. Headland, Aashish Ranjan, Michael F. Covington, Yasunori Ichihashi, Daniel Fulop, José M. Jiménez-Gómez, Jie Peng, Julin N. Maloof, Daniel H. Chitwood, Ravi Kumar, Lauren R. Headland, Aashish Ranjan, Michael F. Covington, and Neelima R. Sinha

INTRODUCTION

The tomato clade (Solanum sect. Lycopersicon) provides unique opportunities to study natural variation. As one of the world’s most important crops, intense focus has been dedicated to the genetic analysis of fruit size, shape, and sugar content between domesticated tomato (Solanum lycopersicum) and wild relatives (Frary et al., 2000; Fridman et al., 2004; Xiao et al., 2008). One of the most distant wild relatives of domesticated tomato, Solanum pennelli, originated in the deserts of Peru. When comparing tomato to a relative that inhabits such an extreme environment, other phenotypic differences undoubtedly underlie the successes of these two species (Moyle, 2008; Chitwood et al., 2013; Koenig et al., 2013). Changes in drought resistance, disease resistance, and water use efficiency, among many other traits, have enabled these two species to thrive in their respective environments. As much as the fruits of these two species differ (Figure 1A), so do their vegetative organs, such as leaves, which exhibit pronounced differences in size, complexity, and morphology (Figure 1B).

Surprisingly, few studies have explicitly studied quantitative trait loci (QTL) regulating leaf traits, in any species. Such phenotypes are associated with water use efficiency and thermo-regulation, traits important to yield (Nicotra et al., 2011; Chitwood et al., 2012a). Studies examining leaf morphology are often limited to analyses of size, dimensions of length and width, and complexity (Jiang et al., 2000; Pérez-Pérez et al., 2002; Holtan and Hake, 2003; Frary et al., 2004). Recently, a genome-wide association study using the maize (Zea mays) nested association mapping population identified liguleless genes as regulators of upright leaf angles. In addition to leaf angle, leaf length and width are regulated by many loci of small effect with little epistasis (Tian et al., 2011). If the complement of genetic changes responsible for differences in leaf shape between species is to be fully understood, similar quantitative genetics approaches will be required in the future. Although length, width, and the dimensions of leaves are important, natural variation in leaf morphology is immense, and methods to quantify the entirety of shape variance are required to determine the full complement of genes regulating differences in populations (Langlade et al., 2005; Chitwood et al., 2012b, 2012c, 2012d). Ultimately, the morphology of leaves is determined at the cellular level, and the genetics underlying natural variation in cellular traits are only now beginning to be examined (Massonnet et al., 2011; Sterken et al., 2012).

Despite the disparate phenotypic differences and ecological habitats occupied by species in the tomato clade, most species are interfertile (Stevens and Rick, 1986), a property that has been...
Figure 1. Phenotypic Differences between IL Parents and RNA-Seq and RESCAN Data for Chromosome 2 ILs.

Phenotypic differences in the fruit (A) and leaves (B) between domesticated tomato (S. lycopersicum) and a wild relative (S. pennellii). Beyond obvious differences in the size, shape, and color of fruits are differences in metabolite content. Leaves between these species vary in size, complexity, and shape and non-cell-autonomously provide the majority of photosynthate to fruits. Shown are the S. pennellii introgression regions for ILs covering chromosome 2 as determined by two methods: RNA-Seq (C) and RESCAN (D). The depth of coverage (distance from midpoint on y axis) and genotype (color and direction on y axis) of each SNP/indel is plotted against chromosomal position (x axis). Polymorphisms that match S. pennellii are colored green and plotted in the top half of each IL panel, while polymorphisms matching cv M82 are plotted in magenta in the bottom halves. The coloring is on a continuum such that the color approaches black as a position’s genotype approaches heterozygosity. The y axis tick marks indicate depths of coverage ranging from 0 to 100 (C) or 0 to 20 (D). Subsequent to genotyping, introgression boundaries consistent between the RNA-Seq and RESCAN analyses were delineated. Using these breakpoints, S. pennellii and cv M82 regions are summarized by horizontal lines at the top and bottom of each IL panel, respectively.
exploited to create introgression lines (ILs) between wild tomato species and domesticated cultivars. A unique property of many such lines is that they contain a single, defined, introgressed genomic region from a wild species donor in an otherwise domesticated background (Eshed and Zamir, 1995; Liu and Zamir, 1999; Monforte and Tanksley, 2000; Canady et al., 2005). A set of such ILs that tile genomic segments from desert-adapted S. pennellii into domesticated S. lycopersicum cv M82 (Eshed and Zamir, 1995; Liu and Zamir, 1999) has been extensively phenotyped, amassing a plethora of QTL. The S. pennellii ILs have been used to map QTL for metabolites, enzymatic activity, yield, and fitness traits as well as the genetic basis of heterosis (Rousseaux et al., 2005; Schauer et al., 2006, 2008; Semel et al., 2006; Stevens et al., 2007; Steinhauser et al., 2011). Ultrahigh-density genotyping of the ILs is a first step toward understanding the whole plant relationships that underlie domesticated traits, but our knowledge will always be limited by phenotype (Chitwood and Sinha, 2013; Zamir, 2013).

Here, we precisely define the boundaries of the S. pennellii ILs at both the genomic and transcriptomic levels. Importantly, the combination of ultrahigh-density genotyping with the recently completed tomato genome allows the exact gene content of these ILs to be determined, aiding breeding efforts and the molecular characterization of QTL. Using precisely defined ILs, we undertake a comprehensive phenotyping of leaf traits, from the organ to cellular level. Measuring leaf shape, size, complexity, and serration traits, in addition to pavement cell morphology and stomatal density and patterning, we detect 1035 QTL. 826 toward the direction of S. pennellii and 209 transgressive, beyond the phenotype of the domesticated parent. We observe distinct, highly heritable aspects of leaf shape and show that leaf shape, serrat, and complexity can be genetically separated, contrary to previous findings from mutagenesis-based approaches. We additionally observe a relationship between pavement cell size and stomatal density, suggesting that modulation of cell size may be a mechanism to alter the spacing of stomata in natural populations. Finally, we analyze our phenotypes within the context of previously reported metabolic, enzymatic, and whole-plant phenotypes, finding an association between leaf complexity and shape with mono- and disaccharide levels in the fruit pericarp. RNA-Seq analysis of gene expression in the vegetative apices of the ILs demonstrates an association between the expression of developmental and photosynthetic pathways with this constellation of traits. The results suggest that leaf morphology can modulate photosynthetic efficiencies and/or that natural variation regulating the shape of leaves affects fruit morphology, which in turn affects the accumulation of fruit sugar. Our results improve upon an important, stable genetic resource and offer insights into not only the quantitative genetic basis of leaf shape, but also its phenomic context at a whole plant level.

RESULTS

Fine-Scale Genotyping of ILs

For fine-scale genotyping of the 76 ILs, we generated a database of polymorphisms between the domesticated tomato species S. lycopersicum cv M82 and its wild relative, S. pennellii (see Supplemental Data Set 1 online). Single nucleotide polymorphisms (SNPs) and indels between species were identified using RNA-Seq and reduced representation genomic sequencing, hereafter referred to as restriction enzyme sequence comparative analysis (RESCAN; Monson-Miller et al., 2012; Seymour et al., 2012). Taking both approaches together, we identified ~750,000 polymorphisms between cv M82 and S. pennellii (see Supplemental Table 1 online). Ninety-nine percent of the polymorphisms identified have at most a single gene separating them, saturating coverage at the level of genetic loci.

RESCAN and RESCAN data for the 76 ILs and detected SNPs were used to genotype each IL across the entire genome (data available at www-plb.ucdavis.edu/Labs/sinha/TomatoGenome/Resources.htm). A graphical summary of the S. pennellii introgressions for all the ILs shows that they tile over nearly the entire tomato genome (see Supplemental Figure 1 and Supplemental Table 2 online).

For each IL, we plotted the genotype of polymorphisms for all relevant chromosomes (ILs with introgressions on chromosome 2 are shown in Figure 1; see Supplemental Figures 2 and 3 online for all other chromosomes). The RNA-Seq and RESCAN-based genotyping results are consistent with and complement one another. Our RNA-Seq–based genotyping (Figure 1C) has a higher relative depth of coverage, aiding polymorphism identification, although a smaller portion of the genome was sequenced. RESCAN-based genotyping (Figure 1D), which yields a more even distribution of polymorphisms and includes nongenic regions.

Noncontiguous Introgressions and Bins

Ultrahigh-density genotyping revealed that seven ILs have multiple introgressions. The majority of the additional introgressions are on the same chromosome as the primary introgression (for example, IL2-1-1 and IL2-3; Figure 1; see Supplemental Figure 4 online); however, we found that IL9-3-1 has an ~100-kb introgression from S. pennellii at the top of chromosome 12 (see Supplemental Figures 2L and 5 online).

ILs harboring multiple introgressions have important implications for genetic mapping. Unique overlapping regions between introgressions define smaller intervals than the ILs, termed “bins.” The unique combinations of ILs that define a bin can dissect a QTL into considerably smaller intervals than the ILs themselves (Liu and Zamir, 1999; see Supplemental Data Set 2 online). Importantly, because of the additional introgressions, we changed the nomenclature of our bins compared with the original bins. As much of the literature uses the old bin designations and cannot be changed retroactively, we name our bins with a “d-” prefix (as in d-5E), denoting “Davis, CA.” The moniker is critical, as the bin names between the old and new systems do not correspond. Whereas previously the 76 S. pennellii ILs defined 107 bins, precisely defined IL boundaries reveal 112 bins (see Supplemental Figures 4 and 5 online). The majority of bins harbor <500 annotated genes (median = 177 genes, mean = 295.03 genes; see Supplemental Figure 6 and Supplemental Data Set 3 online). Increased bin numbers are caused by ILs with multiple introgressions where previously only one had been detected and slightly different boundaries between borders of ILs that were thought to be shared. In some instances, bins are noncontiguous, especially in the case...
of ILs harboring multiple introgressions (for example, d-2G follows the IL2-3 split; see Supplemental Figure 4 online) and when a smaller introgression lies completely within a larger introgression (for example, d-10A is divided by IL10-1-1; see Supplemental Figure 5 online). Precise knowledge of IL boundaries allows the gene content of bins to be known with near certainty. A list of annotated genes within the newly defined bins is provided (see Supplemental Data Set 4 online).

Heritability and Detected QTL

Having precisely defined the introgression boundaries, we next used this resource to determine the genetic basis underlying natural variation in leaf form, measuring a comprehensive suite of leaf traits in the ILs. A number of the traits are derived from outline analyses of terminal (“Term”) and distal lateral (“Lat”) leaflets from field-grown ILs. These include size (TermLFSize and LatLFSize) and measures of serration/lobing (LftCirc and LftSolid). To globally measure the differences in shape between leaflets, we used an elliptical Fourier descriptor (EFD) approach followed by a principal component analysis (Iwata et al., 1998; Iwata and Ukai, 2002). Over 11,000 leaflets were measured, including the terminal and left and right distal lateral leaflets, at a pseudoreplication of 15 leaflets per individual and replication of 10 individuals per IL (raw photos available at www-plb.ucdavis.edu/Labs/sinha/TomatoGenome/Resources.html). The principal components resulting from the EFD analysis describe intuitive, distinct aspects of shape segregating in the ILs, which we treat as traits (TermPC1-5 and LatPC1-5; Figure 2; Chitwood et al., 2012b, 2012c, 2012d). Together, the five principal components considered in this article explain ~75% of all leaf shape variance observed.

Additionally, we provide measures of leaflet length-to-width ratio (LftAR and LftRound), leaf complexity (CompPri, CompSec, CompInt, CompRachis, and CompAll), pavement cell size (CotPaveArea and CotPaveCnt), pavement cell morphology (CotPaveAR, CotPaveRound, CotPaveSolid, and CotPaveCirc), stomatal density (CotStom, LfAdStom, and LfAbsStom), stomatal patterning (CotStomInd), and flowering time (FlowTime) in the ILs (see Supplemental Figure 7 and Supplemental Data Sets 5 and 6 online). An extensive list detailing traits, what they represent, and how they are statistically modeled is provided for reference (see Supplemental Data Set 7 online; trait information has been deposited at Phenom-Networks, www.phenome-networks.com). Correlation analysis and hierarchical clustering suggest that our traits fall into three main categories. The first category includes leaf size and leaf complexity measures, the second pavement cell-related traits and stomatal density, and the third leaflet length-to-width ratio, serration, and shape (see Supplemental Figure 8 online).

The traits we measure vary widely in their broad-sense heritabilities. Previous reports analyzing measures of leaf length and width in maize (Tian et al., 2011) and shape analysis in Antirrhinum majus (Langlade et al., 2005) suggested a high genetic component to the variance in leaf morphology in these species. Our results are similar; the highest broad-sense heritability values are observed in measures of leaflet serration and length-to-width ratio (see Supplemental Figure 9 online). Specifically, two distinct shape components, PC1 and PC4, from the EFD analysis of leaflet shape are highly heritable, and as discussed later, show unique interactions with metabolite and yield-associated traits. Leaf complexity and size, as well as leaf stomatal density, exhibit intermediate heritability, whereas most of the cellular traits we measured have low heritability. This low heritability may not reflect a small genetic component to these phenotypes, but rather the high variance in these measurements due to the inability to account for micro-patterning across the entirety of the leaf surface, and, as we show, robust QTL for such traits can still be determined.

In total, we detect over 1000 QTL at a significance level of <0.05. 829 of these QTL are in the direction toward S. pennelli. Because the directionality of traits was determined by comparisons with S. pennelli grown in the greenhouse (because of the difficulty growing this species under Davis, CA field conditions), whether QTL are transgressive beyond S. pennelli cannot be determined. However, 209 QTL are transgressive beyond cv M82 (see Supplemental Figure 10 online). Of particular note is the numerous QTL that reduce leaf complexity and induce shape changes toward that of S. pennelli. The results suggest that large portions of the genome contribute to natural variation in leaf complexity and shape and that these polygenic contributions can act additively (although a limited role for epistasis in the ILs cannot be discounted).

QTL Regulating Leaf Shape, Complexity, and Serration

No IL comes close to approximating the shape of a S. pennelli leaflet, the longest axis of which lies perpendicular to cv M82
(see Supplemental Figure 11 online). *S. pennellii* leaflets are also more orbicular and lack the distinct deltoid tip and lobing of a domesticated tomato leaflet. Because the ILs tile the tomato genome and a majority of ILs possess a significant shape QTL (see Supplemental Figure 10 online), this reinforces the idea that leaflet shape is highly polygenic with an important additive component.

The largest contributing loci to leaflet shape in the IL population include ILs 4-3 and 5-4, which are much wider than cv M82, and ILs 2-1 and 9-1-2, which are transgressively narrower than cv M82 leaflets (see Supplemental Figure 11A online). A close comparison of the averaged leaflet outlines of the two widest and two narrowest ILs against each other reveals that these ILs alter their shape in distinct ways (see Supplemental Figure 11B online). For example, the tip of IL5-4 remains much more distinct than that in IL4-3, while increasing its wideness at the base of the leaflet. Similarly, IL2-1 and 9-1-2 vary distinctly in the degree of constriction at their proximal ends. These four ILs represent four different extremes in PC1-PC4 space. Although a simplification, PC1 explains shape variance relating to overall length-to-width ratio changes, whereas PC4 tends to explain variance relating more to the distinctness in shape between ILs, such as the distribution of laminar outgrowth along the proximal-distal axis and cordate bulges at the base of leaflets (Figure 2).

Extensive developmental and mutagenesis-based approaches in model systems suggest that a suite of leaf morphology features, including serration, complexity, and laminar outgrowth, are under similar genetic regulation, including the activities of auxin and KNOX, CUC, and TCP family members (Barkoulas et al., 2007). How do these leaf features behave within the context of segregating natural variants? At least one IL, IL4-3, possesses significant QTL in the *S. pennellii* direction for all traits measured, including significantly decreased length-to-width ratio (influenced by laminar outgrowth), decreased serration and lobing (as measured by circularity and solidity), and decreased leaf complexity counts. Nonetheless, serration and shape are genetically separable. For example, IL5-4 is significantly wider than cv M82 but is transgressively more serrated and has increased leaf complexity (Figure 3). Looking at representative leaflets, it becomes apparent that measurements of serration and lobing versus shape impinge upon each other to some degree; for example, the increased lobing in IL5-4 creates proximal lobes that likely contribute toward its increased width. The varying combinations of leaflet shape, serration, and complexity are also present in the transgressively thinner ILs 2-1 and 9-1-2. That these leaf morphology traits segregate independently from each other suggests that either unique genes contribute to natural variation in these features or that the spatiotemporal regulation of known factors is modulated independently from each other.

### QTL Regulating Cellular Phenotypes

Natural variation in organ shape must arise during development from differences in the patterning, division, and expansion of cells. Additionally, cellular features are important for adaptations to abiotic conditions, such as the patterning and response of stomata to the xerophytic conditions found in the native habitat of *S. pennellii* (Heichel and Anagnostakis, 1978). To measure natural variation at the cellular level in the ILs, we analyzed the size of pavement cells, their shape characteristics, and stomatal density and patterning on the adaxial side of cotyledons. We additionally measured the density of stomata on the adaxial and abaxial sides of mature, field-grown leaves.

One IL in particular, IL10-3, consistently exhibits extreme phenotypes for a number of cellular features (see Supplemental Figure 12 online). Like *S. pennellii*, IL10-3 has a significantly lower stomatal density on the adaxial side of cotyledons and true leaves, and it also possesses the largest pavement cell size measured in the ILs. Beyond the developmental interest of adaxial stomatal density, IL10-3 additionally possesses significantly reduced stomatal density on the abaxial side of true leaves, an important adaptive trait considering the desert habitat of *S. pennellii* (see Supplemental Data Sets 5 and 6 online). Interestingly, the stomatal index of IL10-3 is not significantly different from cv M82 (“CotStomInd”; see Supplemental Data Sets 5 and 6 online), suggesting that the number of stomata per pavement cell is not the major cause of these phenotypes and that the larger pavement cell size pushes the stomata away from each other, reducing their density. The significant correlations between pavement cell size and stomatal density (see Supplemental Figures 8 and 13 online) may represent an evolutionary mechanism to modulate stomatal spacing.

### Bin Mapping and Gene Candidates

Bin mapping can be a qualitative endeavor. If one IL possesses a significant phenotypic difference from cv M82 and an overlapping IL does not, then a QTL interval can be narrowed by exclusion. Similarly, a shared region between ILs with similar phenotypes can be used to delimit a QTL by parsimony. Applying exclusion and parsimony can be difficult when dealing with QTLs, though. For example, is an IL not significant enough to use to exclude a region, or conversely is a QTL significant enough to apply parsimony? To help solve this predicament, for each bin, we use a marginal regression approach, regressing bin genotypes of individuals against their trait values (see Supplemental Data Sets 5 and 6 online). In this manner, the number of stomata per pavement cell is not the major cause of these phenotypes and that the larger pavement cell size pushes the stomata away from each other, reducing their density. The significant correlations between pavement cell size and stomatal density (see Supplemental Figures 8 and 13 online) may represent an evolutionary mechanism to modulate stomatal spacing.
Moreover, SP5G cosegregates for the \textit{pht5.4} QTL regulating plant height and is tightly linked to the \textit{OB-SCURAVENOSA} locus, and it has been suggested that the tight linkage of these two loci contributes to the coincident features of compact plant habit and chloroplast-obscured venation in processing tomato varieties (Jones et al., 2007). Similarly, we observe leaf phenotypes in IL5-4 as well, including wider leaves (see Supplemental Figure 11 online) with more serration and complexity (Figure 3). Given previously demonstrated connections between flowering time pathways and leaf morphology.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure3.pdf}
\caption{Leaflet Shape, Serration, and Leaf Complexity Are Genetically Distinct Components.}
\begin{description}
\item[(A)] Representative leaflets from ILs with significant shape QTL. Given are the direction and significance of length-to-width ratio change relative to cv M82 (AR), serration/lobing (circularity), and leaf complexity. Note that despite considerable accumulated genetic evidence suggesting otherwise, these features do not follow each other, and different ILs exhibit different combinations of these traits.
\item[(B)] Graphs demonstrating the breaking between AR, circularity (Circ.), and complexity (Comp.). In each graph, AR is on the \textit{x} axis for comparison showing IL values for circularity and complexity on the \textit{y} axis. Colors indicate significance of trait deviations for the \textit{y} axis.
\end{description}
\end{figure}
(Willmann and Poethig, 2011), it is not unreasonable to suspect SP5G as a modulator of leaf shape in addition to flowering time.

Not all bins are as small as the aforementioned, but knowing the gene content of an interval can provide a list of potential candidates regulating traits. For example, bin d-9B is inferred to possess a QTL that increases leaflet aspect ratio (AR; see Supplemental Figure 14B and LatPC1 in Supplemental Figure 31 online). An ARF16 homolog, a modulator of leaflet width and hyponasty in Arabidopsis, lies in this interval (Liu et al., 2011). Similarly, d-8F is an interval regulating leaflet AR as well (see Supplemental Figure 14C online). Within this bin lies a homolog of FILAMENTOUS FLOWER/GRAMINIFOLIA, a regulator of laminar outgrowth, as well as an ATHB-2 homolog, which in addition to its role in the shade avoidance response, also modulates leaf width (Siegfried et al., 1999; Steinidler et al., 1999; Golz et al., 2004; Eshed et al., 2004). As adaxial-abaxial polarity determinants modulate leaf complexity in tomato (Kim et al., 2003a, 2003b), the inclusion of LEUNIG and KANADI2 homologs (Kerstetter et al., 2001; Cnops et al., 2004; Stahle et al., 2009) on the d-8A interval, which regulates leaf complexity, is interesting (see Supplemental Figure 14D online).

Identifying causal genes for QTL obviously requires fine-mapping, but the delimitation of the exact gene content in bins (see Supplemental Data Set 4 online) provides a powerful starting place to begin studies of natural variation in tomato.

**Associations between Leaf Morphology and Fruit Sugar Metabolism**

One of the advantages of a true-breeding genetic resource, such as the S. pennelli ILS, is the ability to meta-analyze phenotypic data sets with a common genetic basis (Zamir, 2013). Perhaps the most interesting relationship between phenotypes that has been established by such studies (Schauer et al., 2006, 2008; Steinhauser et al., 2011; Toubiana et al., 2012) is a prominent negative correlation between harvest index (the ratio of fruit yield to overall biomass) and metabolite levels in the pericarp. Generally, the more biomass of a plant dedicated to fruit production, the lower the metabolite concentrations in the fruit. The antagonism between metabolite levels and harvest index obviously bodes badly for breeding efforts to increase both of these critical traits simultaneously. However, the relationship is understandable, especially if viewed from the perspective of limited resources, nutrient allocation, and effects of metabolite dilution at the whole-plant level. Critical to the understanding of these whole plant relationships is detailed knowledge of not only nutrient sinks (i.e., fruits, seeds, and flowers) but the ultimate source (leaves).

To better understand the role that leaves play in these relationships, we performed a correlation analysis of leaf traits with the existing phenomics database (Phenom-Networks, www.phenome-networks.com; Figure 4; see Supplemental Figures 48 and 49 online). We divide all traits analyzed into five major groups, defined by the studies from which they are reported and the phenotype that they measure. Largely, traits belonging to a group describe related phenotypes, but for consistency they are first defined by the study from which they originate, the authors’ terminology in those studies, and most importantly their class designations in the Phenom-Networks database. “MET” traits are derived from Schauer et al. (2006, 2008) and measure metabolite levels in the fruit pericarp. “MOR” traits (for “metabolism,” using the nomenclature of Schauer et al. [2006, 2008] and Phenom-Networks) include both yield-related traits and explicit morphological measurements of fruits and flowers. The term “morphology” is used loosely and in contrast with the “metabolite” traits also measured by Schauer et al. (2006, 2008). For example, “MOR” traits include fruit Brix, earliness (the ratio of red fruit yield to total fruit yield), and plant weight, even though these are not strictly morphological features. “ENZ” traits, derived from Steinhauer et al. (2011), measure enzymatic activities in the fruit pericarp, and “SEED” traits, reported by Toubiana et al. (2012), measure metabolite levels in seeds. Traits described in this article were termed “DEV” traits because of their relevance to leaf development. MET, MOR, ENZ, and SEED traits (represented in blue, magenta, yellow, and orange in figures, respectively) are described in Supplemental Data Set 8 online and their values and correlations with other traits provided in Supplemental Data Sets 9 to 11 online. DEV values are provided in Supplemental Data Set 5 online and described in Supplemental Data Set 7 online.

Hierarchical clustering of the mean z-scores for traits reveals a strong negative relationship of harvest index and yield-associated (MOR) traits with metabolite levels (MET; dotted box, upper right-hand corner of Figure 4A), demonstrating the robustness of this previously described relationship (Schauer et al., 2006, 2008; Toubiana et al., 2012). The metabolites exhibiting the strongest negative correlation with harvest index and MOR traits are related to nitrogen and amino acid metabolism in both the fruit and seed (for a close-up of the trait identities in Figure 4A, see Supplemental Figure 50 online; amino acids indicated by asterisk).

Leaf traits (DEV; black) cluster exclusively outside of the aforementioned complex of phenotypes (i.e., the strong negative relationship between harvest index and metabolites; Figure 4A). As DEV traits were not considered in previous analyses, the traits from other classes that cluster with DEV traits are informative as to the importance of leaf traits as correlates of metabolism and yield. For example, a small group of highly heritable DEV traits explaining the length-to-width ratio of leaflets (LftAR, LftRound, LatPC1, and TermPC1) cocluster and are significantly correlated with MOR traits related to length-to-width ratio in fruit and seeds (Figure 4B, red box in Figure 4A). Length-to-width ratio is the major source of shape variance (>40%) in field-grown leaflets (PC1; Figure 2). Such correlation suggests that leaf shape is not independent from the genetic basis of morphology in disparate organs, with implications for the independent modulation of organ shapes during evolution. As we discuss below, changes in either leaf or fruit morphology can explain correlations we observed with fruit sugar, demonstrating the difficulty of organ specific breeding efforts.

Brix and pericarp levels of Suc, Gic, Fru, Gal, mannose, and trehalose (small dotted box in lower right-hand corner in Figure 4A; see Supplemental Figure 50 online) cluster away from the previously described constellation of harvest index and metabolite antagonisms. Also included in this cluster are plant weight and earliness. This suggests a more prominent relationship between carbon metabolism and leaves than the previously found connection between nitrogen metabolism and harvest index. A more detailed analysis confirms the special relationship
FIGURE 4. Relationship between Leaf Morphology and Previously Measured IL Traits.

(A) Hierarchical clustering of leaf traits with previously studied traits. DEV (black), leaf development traits from this study; MOR (magenta), whole-plant, yield, and reproductive morphological traits as described by Schauer et al. (2006, 2008); MET (blue), metabolic traits described in the same studies; ENZ (yellow), enzymatic activities, as measured by Steinhauser et al. (2011); SEED (orange), seed metabolites, as described by Toubiana et al. (2012). Hierarchical clustering is based on absolute correlation values, with red denoting negative Pearson correlation coefficients and yellow positive. The top half of the plot shows significant correlations (p < 0.05) between traits after global multiple test adjustment, indicated in black. Trait identities are indicated as a marginal rug plot along the sides of the graph. The large group of highly correlated traits (in the top right-hand corner, indicated by the dotted line) is consistent with previous reports of negative correlation between MOR traits (including harvest index, [HI], indicated by an arrow) with fruit metabolite levels. DEV traits cluster away from this previously described relationship and closely associate with Brix, plant weight, and mono- and disaccharide levels, indicated by the dotted lined box toward the bottom of the graph. A more detailed view of the hierarchical clustering is found in Supplemental Figure 50 online.
between leaf complexity and shape with sugar metabolism in the fruit (see Supplemental Figure 48 online), and jackknifing suggests that the correlations are robust and not artifacts resulting from undue influence of a few ILs (see Supplemental Figure 51 and Supplemental Data Set 12 online). Additionally, if only correlation between leaf development traits with traits from other classes are considered, then not only do Brix and sugar levels in the fruit exhibit the highest connectivity with leaf complexity and shape, but they are among the most significant correlations (see Supplemental Figure 49 online).

Association between Photosynthetic Gene Expression and Leaf Morphology

Leaf morphology may correlate with sugar accumulation in the fruit due to a variety of factors, indirectly and directly involving leaves. (1) The coregulation of fruit and leaf morphology through similar gene regulatory networks (especially considering that fruits are modified leaves) may lead to shape changes in the fruit that affect the accumulation of sugars. (2) Natural variation in leaf morphology may affect photosynthetic efficiencies through physiological parameters (Nicotra et al., 2011; Chitwood et al., 2012a). The latter hypothesis is particularly intriguing considering that >80% of sugars in the fruit are produced directly by photosynthesis in leaves and subsequently translocated through the phloem (Heatherington et al., 1998; Lytovchenko et al., 2011). Nonetheless, fruit photosynthesis has been demonstrated to significantly affect the accumulation of fruit sugars (Powell et al., 2012), and the role of fruit morphology in this process remains to be more fully explored.

To explore these hypotheses, we correlated the gene expression levels in the 76 ILs, as measured in the vegetative apex using RNA-Seq, against other IL trait values (Figure 5A; see Supplemental Data Sets 13 to 15 online). After hierarchically clustering traits and the expression profiles of those genes significantly correlated with at least one trait (after multiple test correction; see Supplemental Data Set 16 online), a distinct group (indicated by an asterisk in Figure 5A) contained genes with numerous correlations to leaf development (DEV) traits. In addition to leaf complexity, LfTcirc/LfTSolid (measures of serration), and PC1/LfTAR (length-to-width ratio), fruit Gic is represented among the traits significantly correlated with this group of genes (see Supplemental Figure 52 online).

As might be predicted, potential regulators of leaf morphology are present in this group of genes, including ARF3/ETTIN, ARF4, AGO1, SAW1, BELL1, PIN5, and GRF7 homologs (see Supplemental Data Set 17 online), which regulate laminar outgrowth, patterning, indeterminacy, and cell expansion. Support for an intimate association between the genetic coregulation of leaf and fruit morphology with fruit sugar is apparent in genes such as AUXIN-RESPONSE FACTOR4 (ARF4), which modulate not only leaf shape through auxin and adaxial-abaxial pathways, but also the morphology of the fruit (Jones et al., 2002; Yifhar et al., 2012; Sagar et al., 2013).

A Gene Ontology enrichment analysis for genes within this group reveals numerous significantly enriched categories related to photosynthesis (Figures 5A and 5B; see Supplemental Data Set 18 online). It is unlikely that leaf complexity and shape modulate the levels of photosynthetic genes via changes in overall blade area given the negative correlation between leaf complexity and leaflet area (see Supplemental Figure 13 online). Rather, the correlation of photosynthetic gene expression with leaf development traits may reflect the influence of leaf morphology on photosynthetic efficiency physiologically. It is also possible that developmental gene regulatory networks impinge upon photosynthesis pathways more directly, independent of leaf shape. ARF4 again provides a striking example: Not only does this gene regulate both leaf and fruit morphology (Yifhar et al., 2012), but it also regulates the accumulation of chloroplasts and the greening of fruits (Jones et al., 2002), which were recently shown to affect sugar levels (Sagar et al., 2013). Considering the relationship between leaf and fruit shape, a fuller understanding of the causative factors underlying sugar accumulation will require a similar gene expression analysis in fruits to that performed here in vegetative apices.

DISCUSSION

Although breeding traits from wild relatives into domesticated lines is important in and of itself, the knowledge of the identities of genes regulating these traits can be incredibly powerful and explanatory (Frary et al., 2000; Fridman et al., 2004; Xiao et al., 2008; Kimura et al., 2008; Li and Chetelat, 2010). In this study, the high-density genotyping we perform at the genetic and transcriptomic levels elaborates upon previous genetic maps of the S. pennellii ILs and provides outstanding resolution of the recombination breakpoints that define their introgressions. In many instances, the increased resolution resulting from our genotyping yielded insights into the bin structure and the complement of genes harbored by the ILs. Precise knowledge of the genetic content of each IL is a prerequisite for positional cloning of QTL, reverse genetics, and genetic genomics. Together with the recently completed tomato genome (Tomato Genome Consortium, 2012), the means to begin understanding the genetic basis of natural variation in the tomato complex are coming into place (Moyle, 2008; Ranjan et al., 2012).

Despite the exhaustive study of harvest and fruit-related phenotypes, little has been done to study leaves in the S. pennellii ILs. Indeed, approaches utilizing natural variation to study leaf
development are rare in any species (Jiang et al., 2000; Holtan and Hake, 2003; Kimura et al., 2008; Tian et al., 2011). Even rarer are studies of variation in leaf morphology by a quantitative means capable of describing total shape variance (Iwata et al., 1998; Iwata and Ukai, 2002; Langlade et al., 2005; Chitwood et al., 2012b, 2012c, 2012d). Our analyses reveal that leaf morphology is highly heritable and that the *S. pennellii* ILs are a valuable resource to study natural variation in leaves (Figure 2; see Supplemental Figure 9 online). Moreover, the drastic, but superficially simple shape differences in leaves between *S. pennellii* and *S. lycopersicum* cv M82 are regulated through a complex, polygenic genetic basis (see Supplemental Figure 10 online). For example, in addition to the simple regulation of length-to-width ratio, different QTL can impart distinct shapes by which leaflets modulate their width (see Supplemental Figure 11 online). Additionally, we demonstrate that leaflet serration and leaf complexity segregate independently from shape characteristics and even each other (Figure 3). This observation goes against the prevailing wisdom that leaf morphology is regulated by common genetic elements (Barkoulas et al., 2007) and suggests that only a small fraction of the genes regulating the tremendous variation of leaf morphology in natural populations has been discovered.

What mechanisms could explain the disparate association between leaf morphology and sugar levels in the fruit? A purely developmental connection between leaves and fruit is likely part of the explanation. Among other correlations we detect is a relationship between traits describing the length-to-width ratio of leaflets with traits explaining similar dimensions of fruit and seeds (Figure 4B). Considering the extensive analysis of fruit size and morphology in tomato (Frary et al., 2000; Cong et al., 2008; Xiao et al., 2008), it will be interesting to analyze genetic perturbations modulating fruit phenotypes for their effect on vegetative development and vice versa (Wu et al., 2011). Another possibility is that leaf development is altered by genetic changes in overall metabolism (Hackel et al., 2006) or that manipulation of carbon metabolism itself induces morphological changes in leaves (Tsi et al., 1997; Geigenberger et al., 2004; Lawson et al., 2006; Rains and Paul, 2006).

It is also possible that leaf morphology affects sugar metabolism in the fruit. Not only are leaves the ultimate source of most photoassimilates, but overwhelming evidence suggests that sugars are apoplastically unloaded from the phloem into the fruit (Fridman et al., 2004; Baxter et al., 2005; Hackel et al., 2006; Zanor et al., 2009), suggesting a directionality to the correlations we observe. The traits that correlate with fruit sugar levels are not related to ratios of biomass to yield (Do et al., 2010) or the size of leaves, but rather, sensu stricto, leaf shape and complexity. That the expression profiles of photosynthetic genes correlate with leaf morphology traits (Figures 5A and 5B) only bolsters the idea that leaf shape can modulate fruit sugar levels via a photosynthetic mechanism. Similar analyses to those presented here of photosynthetic gene expression in IL fruits are required to better resolve the roles of vegetative and reproductive organs in sugar accumulation.

Regardless of the mechanism, our results highlight an often overlooked fact: Leaves, as the major source of photoassimilate in the fruit (Heatherington et al., 1998; Lytovchenko et al., 2011; Powell et al., 2012) and as organs with a shared developmental

---

**Figure 5.** An Association between Photosynthetic Gene Expression and Leaf Morphology.

(A) Hierarchical clustering of traits and gene expression profiles in the vegetative apex measured across the 76 ILs. Gene expression profiles across ILs were regressed against traits and only those genes with at least one significant correlation with a trait were considered. Colors indicate significant correlation after multiple test adjustment with a gene expression profile and the class to which the correlated trait belongs (DEV, black/white; MOR, magenta; MET, blue; ENZ, yellow; and SEED, orange). One cluster of genes (indicated by an asterisk) significantly correlate with numerous DEV traits related to leaf development.

(B) Gene Ontology enrichment analysis of the gene group with an asterisk reveals numerous significant categories related to photosynthesis.
foundations with carpels, are an important consideration for any breeding effort. Substantial theoretical and empirical evidence has accumulated that leaf shape and size play major roles in water use efficiency and thermoregulation (Parkhurst and Loucks, 1972; Givnish and Vermeij, 1976; Poorter et al., 2010; Nicotra et al., 2011; Chitwood et al., 2012a). Additionally, the relationship in the fossil record between temperature and precipitation with leaf serration and size further supports a functional significance of leaf shape (Bailey and Sinnott, 1915; Wolfe, 1971; Greenwood, 1992; Wilf et al., 1998). The environment imposed by agriculture upon domesticated species, including tomato, is radically different from that encountered by their wild ancestors. If the morphology of leaves matters with respect to photosynthetic efficiency, then leaf size and shape may have been bred as much as the crop, even when the crop is not the leaf, as in tomato. Just as leaf angle has been responsible for most yield increases in maize over the past century (Duvick, 2005; Tian et al., 2011), we propose that leaves may have affected the tomato fruit through developmental and photosynthetic mechanisms.

METHODS

Plant Materials, Growth Conditions, and Experimental Design

Second-generation Solanum pennellii ILs (Eshed and Zamir, 1995; Liu and Zamir, 1999) and Solanum lycopersicum cv M82 seeds were obtained from the Tomato Genetics Resource Center (University of California, Davis) and Dani Zamir (Hebrew University, Rehovot, Israel). For information about the seed stocks used for different lines and experiments, see Supplemental Data Set 19 online.

In mid April, seed were washed in 50% bleach for ~2 min, rinsed, and placed onto water-soaked paper towels in Phytatray ( Sigma-Aldrich) in preparation for field planting. Seeds were placed in darkness for 3 d before moving to a 16:8 light cycle in growth chambers for 4 d. Seedlings were then transplanted into 5 x 10 subdivided trays (11 x 22 inches) in Sunshine Mix soil (Sun Gro) in a greenhouse. Importantly, seedlings were transplanted in trays in the same randomized block design used in the field: Not only did this assist field transplanting, but it allowed cellular trait measurements to be taken in the lath house. Twenty-one days after plating, seedlings were then transferred to a lath house (early May). In both the greenhouse and lath house, seedlings were vigorously top watered and allowed to completely dry between waterings to harden for the field.

Thirty-five days after plating, seedlings were hand transplanted to the field (late May). Transplanted seedlings were initially sprinkler watered followed by ditch irrigated. Ditch irrigation was used as needed throughout the season. ILs and cv M82 were arranged in a block design with 10 replicates. Each block consisted of two rows. Arrangement within each block was randomized.

Anthesis began in early June, and field measures of leaf morphology were taken in early July from mature leaves. Harvest of fruit began late August and continued until mid September.

All traits (except flowering time), whether cellular traits measured in the lath house or field measurements of mature leaves, were studied in a 2010 field in Davis, CA. Additionally, leaf complexity traits were measured in a 2011 field in Davis, CA and the results incorporated together with 2010 data in statistical models. Flowering time was measured exclusively from the 2011 field season.

For the RNA-Seq-based genotyping and expression analysis experiments, seeds of the ILs and two parents were washed in 50% bleach for ~2 min. Afterwards, seeds were placed in darkness for 3 d before moving to a 16:8 light cycle in growth chambers for 5 d. Seedlings were then transplanted into 2 x 5 pots per tray in Sunshine Mix soil (Sun Gro). For each replicate, six seedlings of each parent or IL were planted per pot. The 76 ILs (and two replicates each of cv M82 and S. pennellii) were divided into four cohorts of 20 randomly assigned genotypes. These cohorts sampled different shelves and regions of shelves across four temporal replicates in a Latin square design to account for positional effects on growth. Within a cohort’s assigned space for each of four temporal replicates, pots were randomly distributed. The seedlings were harvested 5 d after transplanting (13 d of growth in total). Cotyledons and mature leaves > 1 cm in total length were excluded, and remaining tissues (including the shoot apical meristem) above the midpoint of the hypocotyl were pooled, for all individuals in a pot, into 2-mL microcentrifuge tubes and immediately frozen in liquid nitrogen.

RNA-Seq Library Preparation

mRNA isolation and RNA-Seq library preparation were performed from 80 samples at a time using a high-throughput RNA-Seq protocol (Kumar et al., 2012). The prepared libraries were sequenced in pools of 12 for replicates 1 and 2 (one lane each) and in pools of 80 for replicates 3 and 4 (seven lanes) at the UC Davis Genome Centre Expression Analysis Core using the HiSeq2000 platform (illumina).

Preprocessing RNA-Seq and RESCAN Sequence Data

Preprocessing of reads involved removal of low quality reads (phred score < 20), trimming of low-quality bases from the 3' ends of the reads, and removal of adapter contamination using custom Perl scripts. The quality-filtered reads were sorted into individual libraries based on barcodes and then barcodes were trimmed using the Fastx toolkit.

RNA-Seq Mapping

RNA-Seq reads were initially mapped to the Heinz reference genome using BWA (parameters: -e 15 -i 10 -k 1 -l 125 -n 0.05; Li and Durbin, 2009). Nonuniquely mapped reads and reads with a mapping quality < 20 were discarded. Unmapped reads were subsequently remapped using TopHat (parameters: -m 1 -g 1 -l 1-segment-length 22 I = 5000-library-type fr-unstranded--solexa1.3-quals--butterfly-search; Trapnell et al., 2009). Mapping and intermediate processing (including sorting, filtering, and duplicate removal using samtools [Li et al., 2009] and Picard [http://picard.sourceforge.net/]) were automated using a Perl script available at http://github.com/mfcovington/RNaseq_mapping.

To remove reads originating from repeat-rich genomic regions, RESCAN sequencing reads were initially mapped to the Sol Genomic Network’s tomato repeat database using BWA (BWA parameters: -e 15 -i 10 -k 1 -l 25 -n 0.05) (we created the fasta file for this from the gff3 file available at ftp://ftp.sgn.cornell.edu/organisms/Solanum_lycopersicum/ annotation/ITAG2.3_release/ITAG2.3_repeats.gff3). Reads not mapped to the repeat database were extracted using bam2fastq program (http://www.hudsonalpha.org/gsi/software/bam2fastq.php). Subsequently, these repeat-filtered reads were mapped to the Heinz reference genome using the same BWA parameters. Samtools (with the '–bq 1' option) was used to retain the reads that mapped uniquely to the reference genome.

Polymorphism Identification for RNA-Seq

Polymorphisms between cv M82 and Heinz or S. pennellii and Heinz were identified using a set of Perl scripts (available at http://github.com/ mfcovington/snp_identification). These scripts identify potential SNPs/indels based on pileup data extracted from the sequence alignments. RNA-Seq reads that encroach upon introns can lead to the identification of false SNPs/indels if the portion of the read that protrudes into the intron is not long enough to be recognized by TopHat as containing an intron.
Polymorphism Identification for RESCAN

VarScan (v2.2.7), a software for variant detection from next generation sequencing data, was used to call SNPs for generating parental references (Koboldt et al., 2012). To this end, the pileup2snp command was used (parameters: --min-coverage 4–min-reads 2–min-avg-qual 20–min-var-freqation 0.9–p-value 0.05). Subsequently, the compare function of VarScan (with the “unique” option) was used to discard common SNPs between cv M82 and S. pennellii, compared with the Heinz genome. The S. pennellii SNP reference was further refined by removing SNPs represented in more than five of the 76 ILs, since the reads contributing to these SNPs are likely derived from unannotated repetitive regions. These steps yielded preliminary RESCAN polymorphism list.

Polymorphism Noise Reduction

The preliminary RNA-Seq and RESCAN polymorphism lists were filtered to remove spurious SNPs/indels. This was done by genotyping the parental lines that were used for initial polymorphism identification (cv M82 and S. pennellii) and removing any SNPs/indels that return an unexpected genotype (using a Perl script available at http://github.com/mfcovington/snp_identiﬁcation). This noise reduction step results in final RNA-Seq and RESCAN versions of the database of polymorphisms between cv M82 and S. pennellii (see Supplemental Data Set 1 online).

IL Genotyping and Plotting

IL sequence data was mapped to the Heinz reference genome. Mpileup information from each alignment was interrogated for every chromosomal position in the relevant version of the SNP database (i.e., RNA-Seq versus RESCAN) to determine the number of reads matching cv M82 versus S. pennellii (using a set of Perl scripts available at http://github.com/mfcovington/geno_plot). The genotype of every SNP/indel for all ILs can be found at www-pib.ucdavis.edu/Labs/sinha/TomatoGenome/Resources.htm.

These data were plotted with a ggplot2-based (Wickham, 2009) approach that conveys genotype and depth of coverage information for each polymorphism across the genome (using an R script available at http://github.com/mfcovington/geno_plot). The technique we developed to plot genotypes is also able to show heterozygous regions. Polymorphism positions that are S. pennellii or cv M82 for every sequencing read are shown as bright green or magenta, respectively. As the ratio of S. pennellii to cv M82 events out, the color of the data point approaches black. For example, our stock of IL1-12 is segregating and the introgressed region is clearly heterozygous for the pool we sequenced (see Supplemental Figures 2A and 3A online).

Trait Measurement

Cellular traits were measured from the adaxial side of cotyledons from plants in the lath house or from the adaxial and abaxial side of true leaves from the field. In all cases, dental impression (Provil Novo Light Standard Fast; Pearson Dental Supplies) was applied using an application gun and allowed to dry before archiving. Fingernail polish (Sally Hanson Double Duty) was applied to impressions, allowed to dry completely, removed from the impression, and floated on microscope slides with water. Water was removed and the nail polish remained affixed to the slide. Micrographs of samples were taken using a standard compound microscope. For each individual impression, two micrographs were taken to ensure representative measures. For each micrograph, four cotyledon pavement cells were traced using Bamboo Tablets (Wacom) in ImageJ (Abramoff et al., 2004) and the area and shape descriptors recorded. For stoma and epidermal cell counts, Bamboo Tablets were used to quickly place dots in ImageJ over the feature of interest, followed by custom macros that would count and record the number of features. Pseudoreplication was averaged.

Leaf complexity measurements were taken in the field. Pairs of measurers would measure two leaves per plant, including primary, secondary, and intercalary leaflet numbers. Leaf complexity was measured in both 2010 and 2011 field seasons. Pseudoreplication was averaged.

Leaf shape traits were derived from photographs. More than 11,000 leaflets were measured for this study. For each individual, five leaves were collected into plastic Ziploc bags and transported back to lab. For each leaf, the terminal and two distal lateral leaflets were dissected and arranged under nonreflective glass (a total of 15 leaflets per individual were measured). Olympus SP-500 UZ cameras were mounted on copy stands (Adorama 36-inch Deluxe Copy Stand) and controlled remotely by computer using Cam2Com software (Sabisk). Using custom ImageJ (Abramoff et al., 2004) macros, individual leaflets were extracted and named appropriately to denote individual and leaflet type (terminal, distal lateral left, and distal lateral right). Leaflet outlines were then batch processed in ImageJ to measure circularity, solidity, AR, and roundness.

Global analysis of leaflet shape was conducted using EFDs followed by principal component analysis using the program SHAPE (Iwata and Ukai, 2002). Object contours were extracted as chain-code. Chain-code was subsequently used to calculate normalized EFDs. Normalization was based upon manual orientation with respect to the proximal-distal axis of the leaflet. Principal component analysis was performed on the EFDs resulting from the first 20 harmonics of Fourier coefficients. Coefficients of EFDs were calculated at +2 and -2 standard deviations for each principal component and the respective contour shapes reconstructed from an inverse Fourier transformation. Principal components resulting from terminal and lateral leaflets were considered separately, and the remaining pseudoreplication was averaged.

Raw data of leaflet photos and micrographs of epidermal impressions can be found at the following database: www-pib.ucdavis.edu/Labs/sinha/TomatoGenome/Resources.htm. Trait information has also been deposited at Phenom-Networks (www.phenome-networks.com).

Statistical Modeling and QTL Analysis

Traits were modeled using mixed-effect linear models with the lme4 package (http://CRAN.R-project.org/package=lme4) in R (R Development Core Team, 2011). Before modeling, the distribution of the trait was checked to determine if it was normal, and if not, it was appropriately transformed. A thorough description of transformations applied to traits, whether model terms were treated as fixed or random, and the significance of terms is provided in Supplemental Data Set 7 online. Models were selected through a process of backward selection, in which two models differing by only the presence of a single term were compared to determine the significance of the term in explaining variance in the data.
The process was repeated for all terms (replacing the previously tested term and testing another), and at the end of the process the most non-significant term (using a P value threshold of 0.05) was removed from the model. This process was iterated until only significant terms remained in the model. We then performed a forward selection check of the resulting minimal model, adding terms previously removed back to the model and comparing to the minimal model to ensure that the non-significance of removed terms persists. That the distribution of residuals in the model was normal was verified. Model fitted values were used for subsequent analysis. P values from models for significant differences between ILs and cv M82 were extracted using the pvals.fnc function from the language R package (http://CRAN.R-project.org/package=languageR).

For bin analysis, each individual was factored as to whether it did or did not possess a particular bin. This process was repeated for each bin. Marginal regression was performed by fitting a linear model between each trait as a function of the presence for each bin. Resulting significance values for each bin with respect to a given trait were then multiple test adjusted using the Holm method to control the family-wise error rate at 0.05 level.

Meta-Analysis of Traits, Hierarchical Clustering, and Network Analysis

Traits from other studies used in meta-analysis were downloaded from www.phenome-networks.com. The studies from which traits are derived and whether or not they are included in this study are detailed in Supplemental Data Set 8 online. Only those traits for which data were collected for >60 ILs were considered, so as to not unduly bias results. In reality, this means that 67 ILs were measured for the trait with the fewest recorded values used in this study. For each trait in a data set, data were z-score normalized, and z-scores were averaged across replicates (see Supplemental Data Set 9 online). A correlation matrix (Pearson) was then created between all traits, both those measured in this study and those from others (see Supplemental Data Set 10 online). Significance values for correlations were determined and the false discovery rate controlled using the Benjamini and Hochberg method (see Supplemental Data Set 11 online; Benjamini and Hochberg, 1995). Subsequently, subsets of the correlation matrix would be analyzed, for example, only those correlations between a DEV trait with a trait of another class that are significant out of the 4,513,704 correlations tested. The 12,501 significant correlations between gene expression profiles and traits represented 3951 genes. These genes were hierarchically clustered based on their expression profile across ILs using the hclust function from the stats package in R, clustering by the absolute value of the Pearson correlation coefficient using Ward’s minimum variance method. Independently clustered genes and traits were then plotted against each other as a matrix shown in Figure 5A. Clusters of genes were then analyzed for enrichment of Gene Ontology terms at a 0.05 false discovery rate cutoff (goseq Bioconductor package; Young et al., 2010).

Accession Numbers

Sequence data from this article are presented in the supplemental data. All supplemental materials from this article are deposited in the DRYAD repository: http://dx.doi.org/10.5061/dryad.m5v5.

Supplemental Data

The following materials are available in the online version of this article.

**Supplemental Figure 1.** S. pennellii Introgression Summary for All 76 ILs.

**Supplemental Figure 2.** Rnaseq-Based Genotyping of All Chromosomes.

**Supplemental Figure 3.** RESCAN-Based Genotyping of All Chromosomes.

**Supplemental Figure 4.** Map of S. pennellii Introgression Lines, Chromosomes 1-6.

**Supplemental Figure 5.** Map of S. pennellii Introgression Lines, Chromosomes 7-12.

**Supplemental Figure 6.** Distribution of Genes per Bin.

**Supplemental Figure 7.** Z-Score Values of ILs Relative to cv M82.

**Supplemental Figure 8.** Correlation between Leaf Developmental Traits.

**Supplemental Figure 9.** Broad-Sense Heritability for Leaf Developmental Traits.

**Supplemental Figure 10.** Detected Leaf Development QTLs.

**Supplemental Figure 11.** Leaflet Shape QTL.

**Supplemental Figure 12.** IL10-3 Exhibits QTL Affecting Pavement Cell Size and Stomatal Density.

**Supplemental Figure 13.** Developmentally Insightful Correlations between Leaf Developmental Traits.

**Supplemental Figure 14.** Bin Mapping and Gene Correlations.

**Supplemental Figure 15.** Bin Mapping Result Legend.
Supplemental Figure 16. Bin Mapping Results for CompAll.
Supplemental Figure 17. Bin Mapping Results for CompInt.
Supplemental Figure 18. Bin Mapping Results for CompPri.
Supplemental Figure 19. Bin Mapping Results for CompRachis.
Supplemental Figure 20. Bin Mapping Results for CompSec.
Supplemental Figure 21. Bin Mapping Results for CotPaveAR.
Supplemental Figure 22. Bin Mapping Results for CotPaveArea.
Supplemental Figure 23. Bin Mapping Results for CotPaveCirc.
Supplemental Figure 24. Bin Mapping Results for CotPaveCnt.
Supplemental Figure 25. Bin Mapping Results for CotPaveRound.
Supplemental Figure 26. Bin Mapping Results for CotPaveSolid.
Supplemental Figure 27. Bin Mapping Results for CotStom.
Supplemental Figure 28. Bin Mapping Results for CotStomInd.
Supplemental Figure 29. Bin Mapping Results for FlowTime.
Supplemental Figure 30. Bin Mapping Results for LatLfSize.
Supplemental Figure 31. Bin Mapping Results for LatPC1.
Supplemental Figure 32. Bin Mapping Results for LatPC2.
Supplemental Figure 33. Bin Mapping Results for LatPC3.
Supplemental Figure 34. Bin Mapping Results for LatPC4.
Supplemental Figure 35. Bin Mapping Results for LatPC5.
Supplemental Figure 36. Bin Mapping Results for LiAPStom.
Supplemental Figure 37. Bin Mapping Results for LiASStom.
Supplemental Figure 38. Bin Mapping Results for LiStomRatio.
Supplemental Figure 39. Bin Mapping Results for LiAR.
Supplemental Figure 40. Bin Mapping Results for LiCirc.
Supplemental Figure 41. Bin Mapping Results for LiRound.
Supplemental Figure 42. Bin Mapping Results for LiSolid.
Supplemental Figure 43. Bin Mapping Results for TermLfSize.
Supplemental Figure 44. Bin Mapping Results for TermPC1.
Supplemental Figure 45. Bin Mapping Results for TermPC3.
Supplemental Figure 46. Bin Mapping Results for TermPC4.
Supplemental Figure 47. Bin Mapping Results for TermPC5.
Supplemental Figure 48. Significant Correlations between Leaf Complexity and Shape with Sugar Levels.
Supplemental Figure 49. Network Analysis Reveals a Relationship between Leaf Complexity and Shape with Sugars, Brix, and Biomass.
Supplemental Figure 50. Hierarchical Clustering of Traits Analyzed in This Study.
Supplemental Figure 51. Jackknifing Results Indicate Stable Correlations between Leaf Complexity and Shape with Sugar Metabolism and Yield Traits.
Supplemental Figure 52. Traits Significantly Correlated with a Distinct Cluster of Genes.
Supplemental Table 1. Summary of cv M82 Versus S. pennellii SNP/Indel Distribution for RNaseq and RESCAN Analyses.
Supplemental Table 2. Chromosomal Positions of the S. pennellii Introgression Boundaries for All 78 ILs.

ACKNOWLEDGMENTS

We thank numerous undergraduate researchers (University of California, Davis) for their contributions, including Tandis Arani, Andy Buck, Jessica Cruz, Michael Davis, Nora Downs, Tommy Hatcher, Max Mumbach, Dan Naylor, Nataly Raymundo, Katy Rush, Paradee Thammapichai, Thinh Thiem, Alaha Wahab, Jennifer Weil, Choua Yang, and Sharon Zimmerman. We thank Jim Jackson (University of California, Davis) for field care, and Dani Zamir (Hebrew University, Rehovot, Israel) and the Tomato Genetics Resource Center (University of California, Davis) for gifts of germplasm. D.H.C. is a fellow of the Life Sciences Research Foundation funded through the Gordon and Betty Moore Foundation. This work is supported through a National Science Foundation grant (IOS-0820854) awarded to N.R.S., J.N.M., and J.P.

AUTHOR CONTRIBUTIONS


A Quantitative Genetic Basis for Leaf Morphology in a Set of Precisely Defined Tomato Introgression Lines

Daniel H. Chitwood, Ravi Kumar, Lauren R. Headland, Aashish Ranjan, Michael F. Covington, Yasunori Ichihashi, Daniel Fulop, José M. Jiménez-Gómez, Jie Peng, Julin N. Maloof and Neelima R. Sinha

*Plant Cell* 2013;25;2465-2481; originally published online July 19, 2013;
DOI 10.1105/tpc.113.112391

This information is current as of July 14, 2015