Analysis of Sequence, Map Position, and Gene Expression Reveals Conserved Essential Genes for Iron Uptake in Arabidopsis and Tomato

Petra Bauer*, Thomas Thiel, Marco Klatte, Zsolt Bereczky, Tzvetina Brumbarova, Rüdiger Hell§, and Ivo Grosse
Institute of Plant Genetics and Crop Plant Research, D–06466 Gatersleben, Germany

Arabidopsis (Arabidopsis thaliana) and tomato (Lycopersicon esculentum) show similar physiological responses to iron deficiency, suggesting that homologous genes are involved. Essential gene functions are generally considered to be carried out by orthologs that have remained conserved in sequence and map position in evolutionarily related species. This assumption has not yet been proven for plant genomes that underwent large genome rearrangements. We addressed this question in an attempt to deduce functional gene pairs for iron reduction, iron transport, and iron regulation between Arabidopsis and tomato. Iron uptake processes are essential for plant growth. We investigated iron uptake gene pairs from tomato and Arabidopsis, namely sequence, conserved gene content of the regions containing iron uptake homologs based on conserved orthologous set marker analysis, gene expression patterns, and, in two cases, genetic data. Compared to tomato, the Arabidopsis genome revealed more and larger gene families coding for the iron uptake functions. The number of possible homologous pairs was reduced if functional expression data were taken into account in addition to sequence and map position. We predict novel homologous as well as partially redundant functions of ferric reductase-like and iron-regulated transporter-like genes in Arabidopsis and tomato. Arabidopsis nicotianamine synthase genes encode a partially redundant family. In this study, Arabidopsis gene redundancy generally reflected the presumed genome duplication structure. In some cases, statistical analysis of conserved gene regions between tomato and Arabidopsis suggested a common evolutionary origin. Although involvement of conserved genes in iron uptake was found, these essential genes seem to be of paralogous rather than orthologous origin in tomato and Arabidopsis.

Arabidopsis (Arabidopsis thaliana) serves as the reference for dicot genome analysis regarding gene sequence, gene number, and gene function. In the minimal genome concept, it is assumed that orthologous gene functions transmitted in a lineage-dependent vertical manner in related organisms should be conserved by sequence and function if they are essential (Mushegian, 1999). Using systematic gene knockout approaches and genome sequence comparisons, it was calculated that Bacillus subtilis may have about 250 to 300 essential genes (Kobayashi et al., 2003). Ninety-six percent of essential genes are conserved in bacteria and encode proteins with known functions in cell metabolism and energetics, information processing, and cell growth and division (Kobayashi et al., 2003).

Yeast (Saccharomyces cerevisiae) may have about 1,000 essential genes (Winzeler et al., 1999; Giaever et al., 2002). In plants, the number of conserved and essential plant genes still remains to be experimentally determined. It is unclear whether searching for conserved sequence homologs in expressed sequence tag (EST) and genomic databases of diverse plant species may provide clues to essential gene functions. A potential obstacle for such homolog searches in plants is the fact that plant genomes underwent large evolutionary genomic rearrangements involving polyploidization, chromosome rearrangement, and partial gene loss (Gaut et al., 2000; Simillion et al., 2002). The analysis of the full genome sequence suggested that the Arabidopsis genome contained duplicated blocks corresponding to ancient and recent genome duplication events (Arabidopsis Genome Initiative, 2000; Blanc et al., 2000, 2003; Vision et al., 2000; Simillion et al., 2002; Bowers et al., 2003; Ermolaeva et al., 2003). These genome duplication events led to horizontal multiplication of homologous genes known as paralogs. The presence of paralogous genes hampers straightforward genome colinearity studies between diverged species of different families.

Full genomic sequence of the five Arabidopsis chromosomes and skeletons of genomic and EST markers mapped onto the chromosomes of other dicot species are currently available for assessing homologous gene functions (Fulton et al., 2002; Gebhardt et al.,

---

1 This work was supported by Deutsche Forschungsgemeinschaft grants in the Emmy Noether program (Ba1610/3–1 to P.B.) and in the Arabidopsis Functional Genomics Network program (Ba1610/4–1 to P.B. and R.H.), and by the Bundesministerium für Bildung und Forschung to the Bioinformatics Center Gatersleben-Halle (grant to I.G.).

2 Present address: Heidelberg Institute for Plant Sciences, University of Heidelberg, Im Neuenheimer Feld 360, D-69120 Heidelberg, Germany.

* Corresponding author; e-mail bauer@ipk-gatersleben.de; fax 49–39482–5139.

[w] The on-line version of this article contains Web-only data.

Article-publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.104.047233.
the ferric chelate reductase (FRO) gene

Essential gene functions involved in iron assimilation

Bauer and Bereczky, 2003; Curie and Briat, 2003). Both tomato and Arabidopsis mobilize iron by re-

results in leaf chlorosis and severe growth retardation.

Stephan, 2003). Lack of appropriate iron mobilization

sufficient iron from their environment (Hell and

of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,

iron deficiency in tomato and Arabidopsis and in-

Ni content of any organism. In plants, iron deficiency is caused

iron uptake is a strictly essential mechanism for growth

reactions and biological processes in all organisms,
Table 1. Amino acid sequence comparisons of tomato and Arabidopsis homologs and alignment results using BLASTP at http://www.arabidopsis.org/Blast

Arabidopsis iron uptake homologs were only represented and considered for further analysis if they showed E-values similar to the best hits. The only exceptions to this were three FRO proteins marked by **, which were retained due to their map position. For the purpose of highlighting tomato-Arabidopsis conserved gene clusters in Figure 3, the iron uptake genes were marked by a color code (compare Table I with Fig. 3 and Supplemental Table I). Arabidopsis iron uptake homologs located within a distance of ±400 genes (about ±2 Mb) were marked by the same color (e.g. At4g19680-IRT2 and At4g19690-IRT1). Homologous Arabidopsis iron uptake genes that were located in different regions of the genome (at a distance of more than 400 genes) were highlighted by different colors. The best hits were highlighted in dark blue, thereafter in light blue, dark green, light green, brown-orange, and light orange.

<table>
<thead>
<tr>
<th>Iron Uptake Tomato Gene Product</th>
<th>Arabidopsis Protein with Sequence Similarity</th>
<th>E-Values</th>
<th>Score (bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LeIRT1</td>
<td>At5g19680 (IRT1)</td>
<td>1e-117</td>
<td>417</td>
</tr>
<tr>
<td>Metal transporter</td>
<td>At5g19680 (IRT2)</td>
<td>1e-106</td>
<td>382</td>
</tr>
<tr>
<td>At1g31260 (ZIP10)</td>
<td>1e-106</td>
<td>381</td>
<td></td>
</tr>
<tr>
<td>At5g45105 (ZIP8)</td>
<td>2e-098</td>
<td>356</td>
<td></td>
</tr>
<tr>
<td>At2g44032 (ZIP9)</td>
<td>4e-093</td>
<td>335</td>
<td></td>
</tr>
<tr>
<td>LeIRT2</td>
<td>At5g19680 (IRT1)</td>
<td>1e-114</td>
<td>407</td>
</tr>
<tr>
<td>Metal transporter</td>
<td>At1g31260 (ZIP10)</td>
<td>1e-112</td>
<td>402</td>
</tr>
<tr>
<td>At4g19680 (IRT2)</td>
<td>1e-104</td>
<td>376</td>
<td></td>
</tr>
<tr>
<td>At2g44032 (ZIP9)</td>
<td>3e-091</td>
<td>332</td>
<td></td>
</tr>
<tr>
<td>At5g45105 (ZIP8)</td>
<td>5e-089</td>
<td>324</td>
<td></td>
</tr>
<tr>
<td>LeNRAPl</td>
<td>At1g80830 (NRAMP1)</td>
<td>1e-160</td>
<td>561</td>
</tr>
<tr>
<td>Metal transporter</td>
<td>At1g67330 (NRAMP4)</td>
<td>1e-085</td>
<td>314</td>
</tr>
<tr>
<td>At1g47240 (NRAMP2)</td>
<td>1e-083</td>
<td>307</td>
<td></td>
</tr>
<tr>
<td>At2g23150 (NRAMP3)</td>
<td>2e-083</td>
<td>307</td>
<td></td>
</tr>
<tr>
<td>At4g18790 (NRAMP5)</td>
<td>3e-083</td>
<td>305</td>
<td></td>
</tr>
<tr>
<td>At1g15960 (NRAMP6)</td>
<td>3e-074</td>
<td>276</td>
<td></td>
</tr>
<tr>
<td>LeNRAPl</td>
<td>At2g23150 (NRAMP3)</td>
<td>0</td>
<td>768</td>
</tr>
<tr>
<td>Metal transporter</td>
<td>At5g67330 (NRAMP4)</td>
<td>0</td>
<td>740</td>
</tr>
<tr>
<td>At1g47240 (NRAMP2)</td>
<td>0</td>
<td>738</td>
<td></td>
</tr>
<tr>
<td>At4g18790 (NRAMP5)</td>
<td>0</td>
<td>665</td>
<td></td>
</tr>
<tr>
<td>At1g80830 (NRAMP1)</td>
<td>3e-088</td>
<td>323</td>
<td></td>
</tr>
<tr>
<td>At1g15960 (NRAMP6)</td>
<td>2e-037</td>
<td>154</td>
<td></td>
</tr>
<tr>
<td>LeFRO-TC-124302</td>
<td>At5g23990</td>
<td>1e-119</td>
<td>423</td>
</tr>
<tr>
<td>Iron reductase</td>
<td>At5g23990</td>
<td>1e-118</td>
<td>421</td>
</tr>
<tr>
<td>At1g01580 (FRO2)</td>
<td>1e-105</td>
<td>377</td>
<td></td>
</tr>
<tr>
<td>At1g01590 (FRO1)</td>
<td>2e-096</td>
<td>349</td>
<td></td>
</tr>
<tr>
<td>At1g23020 (FRO3)</td>
<td>5e-094</td>
<td>341</td>
<td></td>
</tr>
<tr>
<td>At5g50160**</td>
<td>7e-043</td>
<td>171</td>
<td></td>
</tr>
<tr>
<td>At5g49730**</td>
<td>8e-042</td>
<td>168</td>
<td></td>
</tr>
<tr>
<td>At5g49740**</td>
<td>7e-040</td>
<td>161</td>
<td></td>
</tr>
<tr>
<td>LeFRO-TC-129233</td>
<td>At1g01580 (FRO2)</td>
<td>1e-101</td>
<td>366</td>
</tr>
<tr>
<td>Iron reductase</td>
<td>At1g01590 (FRO1)</td>
<td>1e-096</td>
<td>350</td>
</tr>
<tr>
<td>At5g23990</td>
<td>7e-095</td>
<td>344</td>
<td></td>
</tr>
<tr>
<td>At1g01590 (FRO1)</td>
<td>1e-093</td>
<td>340</td>
<td></td>
</tr>
<tr>
<td>At1g23020 (FRO3)</td>
<td>1e-092</td>
<td>336</td>
<td></td>
</tr>
<tr>
<td>At5g50160**</td>
<td>3e-046</td>
<td>182</td>
<td></td>
</tr>
<tr>
<td>At5g49730**</td>
<td>2e-040</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>At5g49740**</td>
<td>4e-040</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>LeFRO1</td>
<td>At1g01580 (FRO2)</td>
<td>0</td>
<td>819</td>
</tr>
<tr>
<td>Iron reductase</td>
<td>At1g01590 (FRO1)</td>
<td>0</td>
<td>732</td>
</tr>
<tr>
<td>At1g23020</td>
<td>0</td>
<td>732</td>
<td></td>
</tr>
<tr>
<td>At5g23990</td>
<td>0</td>
<td>708</td>
<td></td>
</tr>
<tr>
<td>At1g23020</td>
<td>3e-177</td>
<td>618</td>
<td></td>
</tr>
<tr>
<td>At5g50160</td>
<td>1e-097</td>
<td>354</td>
<td></td>
</tr>
<tr>
<td>At5g49740</td>
<td>1e-094</td>
<td>344</td>
<td></td>
</tr>
<tr>
<td>At5g49770</td>
<td>2e-092</td>
<td>337</td>
<td></td>
</tr>
<tr>
<td>LeNAS</td>
<td>At1g60240 (NAS3)</td>
<td>1e-110</td>
<td>394</td>
</tr>
<tr>
<td>Nicotianamine synthase</td>
<td>At1g54630 (NAS4)</td>
<td>1e-105</td>
<td>379</td>
</tr>
<tr>
<td>At1g04905 (NAS1)</td>
<td>1e-104</td>
<td>374</td>
<td></td>
</tr>
<tr>
<td>At5g56808 (NAS2)</td>
<td>1e-100</td>
<td>360</td>
<td></td>
</tr>
<tr>
<td>LeFER</td>
<td>At5g23150 (FLU, MHLH29)</td>
<td>1e-057</td>
<td>220</td>
</tr>
<tr>
<td>MHLH protein</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
could be aligned with AtNRAMP2 (phylogenetic tree not shown).

AtFRO2 is required for iron reduction upon low iron supply (Robinson et al., 1999). Three LeFRO sequences and four Arabidopsis FRO sequences had significant similarity among each other and to AtFRO2. The LeFRO-TC124302 sequence was most similar to Arabidopsis FRO homologs At5g23980 and At5g23990. LeFRO1 and LeFRO-TC129233 were most related to AtFRO2 and AtFRO1. The three FRO-like proteins—At5g50160, At5g49730, and At5g49740—were not as related in sequence. Their genes were only retained because of their mapping location, as will be shown later.

NAS is essential for iron homeostasis in tomato. Four different Arabidopsis NAS sequences were in the database (see also Suzuki et al., 2001; Becher et al., 2004), which were more related to each other rather than to the tomato LeNAS (CHLORONERVA; Fig. 1B). Arabidopsis NAS sequences could be divided into two subgroups, encoded by genes on chromosomes 1 and 5, respectively.

A single Arabidopsis protein that we named FER-like regulator of iron uptake (AtFRU; At2g28160, bHLH029; Heim et al., 2003) corresponded to LeFER (see also Jakoby et al., 2004). All other predicted bHLH domain protein sequences from either Arabidopsis or tomato differed significantly outside the bHLH domain (data not shown).

In summary, it was possible to predict a unique homologous Arabidopsis-tomato gene pair in only a single case, namely that of LeFER-AtFRU. For the other four gene functions Arabidopsis had more and larger gene families encoding these functions than tomato.

Mapping of Iron Uptake Genes in Tomato

A further criterion for gene homology between two species is location in a colinear region of the two genomes. For map position comparison, it was necessary first to map the tomato genes.

LeNAS and LeFER map positions were known to be on tomato chromosomes 1 and 6, respectively (Ling et al., 1999, 2002). Here, we mapped IRT1 and IRT2, FRO, and NRAMP homologs by restriction fragment length polymorphism analysis onto the tomato genome ("Materials and Methods;" Fig. 2). We found that LeIRT1 and LeIRT2 were mapped on chromosome 2 between CT255 and TG554. LeNRAMP1 was localized on chromosome 11 between TG47 and TG400. LeNRAMP3 was localized on chromosome 2 between CT59 and TG154. LeFRO-TC124302 could be mapped to chromosome 3 between CT171 and TG525. LeFRO-TC129233 and LeFRO1 were both localized in between TG460 and CT165 on chromosome 1. All genes were mapped as codominant markers with the exception of LeFRO-TC124302, which was mapped as dominant marker in the L. esculentum background. Since a homolog of LeFRO-TC124302 also could not be found in the potato EST database, we hypothesize that this gene is lacking from the potato and Lycopersicon pennellii genomes.

Analysis of Conserved Gene Content in Arabidopsis and Tomato Chromosomal Regions Containing Iron Uptake Genes

Chromosomal regions harboring iron uptake genes of Arabidopsis and tomato genomes were analyzed for their level of conserved genes. For this analysis, we selected COS markers that were mapped by Fulton et al. (2002) within an approximately 10-cM distance of tomato iron uptake genes and for which we could identify the corresponding Arabidopsis homologs by sequence similarity searches ("Materials and Methods;" Supplemental Table I, available at www.plantphysiol.org). Among the Arabidopsis COS markers, we identified those that were located up to 400 genes upstream or downstream of the respective iron uptake
genes (approximately ±2 Mb) and defined them as being located in a similar region (“Materials and Methods,” Supplemental Table I; Fig. 3). For illustration, COS markers and iron uptake genes that are homologous and colocalize in the tomato and Arabidopsis genomes were highlighted by the same color (Table I; Supplemental Table I; Fig. 3). We found that 8 tomato COS markers out of 24 were located in the region of LeIRT1/LeIRT2 and recognized at least one Arabidopsis region containing either AtIRT1/AtIRT2, AtZIP8, or AtZIP10 (Fig. 3A). Two of these tomato COS markers were homologous to multiple Arabidopsis COS markers that were located in two and three IRT/ZIP regions, respectively. Therefore, we could identify three Arabidopsis regions harboring IRT/ZIP genes that shared several conserved gene sequences with the LeIRT1/LeIRT2 region in tomato.

Although only 10 COS markers were available for the LeNRAMP1 region, we could identify 2 that were located near AtNRAMP1 (Fig. 3B). Single LeNRAMP1-neighboring COS markers identified regions of AtNRAMP4, AtNRAMP5, and AtNRAMP6. Three COS markers were located near AtNRAMP3. The LeNRAMP3 region showed similarity to the region of AtNRAMP3 (3 COS markers out of 26), and, to a lesser extent, to that of AtNRAMP4 (two COS markers) and AtNRAMP5 (one COS marker; Fig. 3C).

The tomato region of LeFRO-TC124302 showed similar gene sequences to the Arabidopsis regions around the FRO-like homologs At5g23990/At5g23980 (4 COS markers out of 35), AtFRO2/AtFRO1 (one COS marker), and At5g50160/At5g49730/At5g49740 (seven COS markers; Fig. 3D). The LeFRO1/LeFRO-TC129233 region showed similarity to the regions of AtFRO2/AtFRO1 (3 COS markers out of 21) as well as At5g23980/At5g23990 (one COS marker; Fig. 3E). No indication for conserved genes was found for the region of Arabidopsis FRO3.

For the LeNAS regions we identified four corresponding regions in Arabidopsis (twice 2 and twice 3 COS markers out of 29; Fig. 3F), indicating that all four Arabidopsis NAS regions were related.

Only a single COS marker out of 21 recognized the LeFER and AtFRU regions (Fig. 3G). With the exception of the LeFER/AtFRU regions, we could thus determine that 20% to 50% of the analyzed tomato COS markers recognized Arabidopsis genes located in clusters in the vicinity of the respective iron uptake genes.

We realized that clusters of conserved COS markers could also be found between regions that did not appear related by the presence of homologous iron uptake genes. To analyze this point further, we searched for matches between COS markers of 6 unrelated iron uptake regions (42 comparisons of non-homologous regions and the 7 comparisons of homologous regions as control; Table II). We found that the number of matching COS markers was highest when comparing Arabidopsis and tomato regions with
corresponding iron uptake genes (three to five conserved genes; only exception was LeFER/AtFRU). Only in the case of LeFER was the highest level of gene conservation observed with a different region. Only 4 out of 42 comparisons of regions with nonhomologous iron uptake genes showed similar levels of conserved gene content as the regions with corresponding iron uptake genes (three and four conserved COS markers), namely LeFRO-TC124302-At1g09240 (AtNAS3), LeFER-At5g23990 (FRO-like), and LeNRAMP1-AtFRU. In 19 comparisons with nonhomologous iron uptake gene regions, no matching COS marker was detected. In 12 cases, a single matching COS marker was found. In six cases, two matching COS markers were found. These observations indicate that the level of conserved gene sequences between homologous iron uptake gene regions generally tends to be higher between homologous iron uptake gene regions than between regions with unrelated iron uptake genes.
To analyze the significance of these findings, we determined the probability of the occurrence of tomato-Arabidopsis homolog clusters at random within 2-Mb intervals in the Arabidopsis genome. Similar results were obtained for 4-Mb (±2 Mb) windows (data not shown). We calculated P-values for the occurrence of clustering of the selected Arabidopsis gene homologs within ±1-Mb windows by taking into account the number of neighbors recognizing tomato homologs in the iron uptake region, the number of homologous genes of the same tomato region in the entire Arabidopsis genome, and the number of all genes in the Arabidopsis genome ("Materials and Methods"). Small P-values (P < 0.05) indicate that the number of homologous Arabidopsis neighbors recognized by the corresponding tomato region exceeded significantly the number of neighbors according to random clustering within the ±1-Mb window. We found that 7 out of 31 Arabidopsis iron uptake genes analyzed (22.6%) had P-values less than 0.05, and 43 out of 279 noniron uptake genes analyzed (15%) showed clustering with P-values less than 0.05 (Supplemental Table II). The cumulative distribution of P-values for iron uptake genes indicated that overall iron uptake genes were significantly more clustered than noniron uptake genes (data not shown). Three Arabidopsis homologs of *LeIRT1/LeIRT2* had P-values less than 0.015, namely At1g31260, At5g45105, and At4g19690/At4g19680 (Table III). Moreover, 9 out of 11 COS homologs that we suspected to be located in these regions (compare Supplemental Table I and Fig. 3) had P-values less than 0.02 (Table III). The 2 out of 11 remaining markers had P-values of about 0.06, just slightly above the cutoff value of 0.05 below which we considered results to be statistically significant (Table III). Thus, *LeIRT1* maps together with four other tomato markers from the ±10-cM region surrounding the *LeIRT1* locus to a single ±1-Mb window on the Arabidopsis genome region containing At1g31260 together with 419 genes. This means that in this Arabidopsis region the homologs of tomato markers of the *LeIRT1/LeIRT2* region are highly enriched, which is reflected by the average density of about one of these markers every 83 genes (5/419). We would expect only about one of these markers every 500 genes (52/26,404) if the genes were randomly distributed. The significantly similar gene content of the *IRT/ZIP* regions of tomato and Arabidopsis suggests a common evolutionary origin. An equally significant conservation of clustered genes was found for *LeFRO-TC124302/At5g49740/At5g50160 (Table III). Significant clustering was obtained for *LeFRO1/LeFRO-TC129233/At1g01580/At1g01590* as well as *LeNAS/At1g09240* (Table III). P-values between 0.05 and 0.075 were found for the regions of *LeNRAMP1/At2g23150* and *LeFRO-TC124302/At1g01580/At1g01590* (Table III). All other comparisons of iron uptake regions between tomato and Arabidopsis showed no significant clustering, with P-values higher than 0.1 (Table III). However, we speculate that, for some of the *NRAMP, FRO, and NAS* gene regions, P-values would decrease if more COS markers were available for the analysis of these regions.

In summary, Arabidopsis and tomato genome regions with homologous iron uptake genes were generally characterized by the presence of multiple conserved genes. Since the Arabidopsis genome frequently contained more than one region with conserved gene content for any of the studied tomato iron uptake regions, it did not appear to be sufficient to determine a functional homology based on map position and sequence alone.

### Functional Analysis of Tomato-Arabidopsis Iron Uptake Genes

Within gene families, specific biological functions of gene family members are conferred by their specific expression patterns. Iron uptake genes should be expressed in the root and/or induced by iron deficiency, indicating a function in iron mobilization of external or internal iron. Here, we investigated which of the iron uptake homologs fulfilled these expression pattern criteria in tomato and Arabidopsis. Gene expression was surveyed by analyzing EST expression data available at http://www.tigr.org for tomato genes (Supplemental Table III), as well as experimental gene expression studies in tomato and Arabidopsis ("Materials and Methods;" Figs. 4 and 5).

### Table II. Number of COS marker matches between seven Arabidopsis and seven tomato regions containing iron uptake homologs, resulting in 42 comparisons of regions with nonhomologous iron uptake genes and seven control comparisons of regions with homologous iron uptake genes (in bold)

<table>
<thead>
<tr>
<th></th>
<th>A1g19690</th>
<th>A1g90830</th>
<th>A2g23150</th>
<th>A5g23990</th>
<th>A1g01580</th>
<th>A1g09240</th>
<th>A2g28160</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LeIRT1/LeIRT2</strong></td>
<td>4 + 1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><strong>LeNRAMP1</strong></td>
<td>1</td>
<td>2 + 1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>LeNRAMP3</strong></td>
<td>1</td>
<td>–</td>
<td>3 + 1</td>
<td>–</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>LeFRO-TC124302</td>
<td>4</td>
<td>–</td>
<td>–</td>
<td>3 + 1</td>
<td>–</td>
<td>3 + 1</td>
<td>1</td>
</tr>
<tr>
<td>LeFRO1/LeFRO-TC129233</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3 + 1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>LeNAS</td>
<td>–</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>–</td>
<td>3 + 1</td>
<td>–</td>
</tr>
<tr>
<td>LeFER</td>
<td>–</td>
<td>–</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1 + 1</td>
</tr>
</tbody>
</table>

Iron uptake homologs themselves are counted as + 1 in the controls. –, No COS marker matches were found.
Table III. P-values indicating the occurrence of tomato homolog clusters at random within ±1-Mb intervals in the Arabidopsis genome

P-values were calculated by Fisher's exact test ("Materials and Methods"). P-values less than 0.05 are considered to show significant nonrandom clustering. Tandemly repeated genes were counted as one. P-values for iron uptake genes are in bold.

<table>
<thead>
<tr>
<th>Tomato Region</th>
<th>Tomato Gene/Marker</th>
<th>Arabidopsis Gene</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LeIRT1/LeIRT2</td>
<td>LeIRT1, LeIRT2</td>
<td>At4g19680/At4g19690</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>T1668</td>
<td>At4g20410</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>T1698</td>
<td>At4g19880</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>CLEC27M9</td>
<td>At4g16250</td>
<td>0.607</td>
</tr>
<tr>
<td></td>
<td></td>
<td>At4g18130</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>CLPT1E8</td>
<td>At4g18970</td>
<td>0.014</td>
</tr>
<tr>
<td>LeIRT1, LeIRT2</td>
<td>A1g31260</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T1654</td>
<td>At1g29520</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>CLPT2K20</td>
<td>At1g31160</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>CLPT1E8</td>
<td>At1g29660/At1g29670</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>At1g33811</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>T984</td>
<td>At1g30580</td>
<td>0.001</td>
</tr>
<tr>
<td>LeIRT1, LeIRT2</td>
<td>A5g45105</td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T1105</td>
<td>A5g43280</td>
<td>0.062</td>
</tr>
<tr>
<td></td>
<td>T1698</td>
<td>A5g45020/At5g44990</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A5g4400</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>CLPT1E8</td>
<td>A5g45670</td>
<td>0.064</td>
</tr>
<tr>
<td>LeNRAMP1</td>
<td>LeNRAMP1</td>
<td>A1g80830</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>T161</td>
<td>A1g77470</td>
<td>0.353</td>
</tr>
<tr>
<td></td>
<td>T675</td>
<td>At1g77670</td>
<td>0.353</td>
</tr>
<tr>
<td>LeNRAMP1</td>
<td>A5g67330</td>
<td>0.229</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T1071</td>
<td>A5g65760</td>
<td>0.328</td>
</tr>
<tr>
<td>LeNRAMP1</td>
<td>A2g23150</td>
<td>0.334</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T1071</td>
<td>A2g24280</td>
<td>0.320</td>
</tr>
<tr>
<td></td>
<td>T1787</td>
<td>A2g27920</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>CLET10011</td>
<td>A2g27920</td>
<td>0.055</td>
</tr>
<tr>
<td>LeNRAMP1</td>
<td>A4g18790</td>
<td>0.337</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T1014</td>
<td>A4g21110</td>
<td>0.342</td>
</tr>
<tr>
<td>LeNRAMP1</td>
<td>A1g15960</td>
<td>0.377</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T1787</td>
<td>A1g1500</td>
<td>0.377</td>
</tr>
<tr>
<td>LeNRAMP3</td>
<td>LeNRAMP3</td>
<td>A2g23150</td>
<td>0.280</td>
</tr>
<tr>
<td></td>
<td>T1480</td>
<td>A2g25140</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>CTOB9L18</td>
<td>A2g23780</td>
<td>0.262</td>
</tr>
<tr>
<td></td>
<td>T1744</td>
<td>A2g26560</td>
<td>0.626</td>
</tr>
<tr>
<td>LeNRAMP3</td>
<td>A5g67330</td>
<td>0.144</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T6734</td>
<td>A5g67370</td>
<td>0.140</td>
</tr>
<tr>
<td></td>
<td>CLER17N11</td>
<td>A5g65700</td>
<td>0.276</td>
</tr>
<tr>
<td>LeNRAMP3</td>
<td>A4g18790</td>
<td>0.655</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T1003</td>
<td>A4g18370</td>
<td>0.659</td>
</tr>
<tr>
<td>LeFRO-TC124302</td>
<td>LeFRO-TC124302</td>
<td>A5g23990/At5g23980</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>T1540</td>
<td>A5g23940</td>
<td>0.074</td>
</tr>
<tr>
<td></td>
<td>T1424</td>
<td>A5g24460</td>
<td>0.069</td>
</tr>
<tr>
<td></td>
<td>T1429</td>
<td>A5g23670</td>
<td>0.075</td>
</tr>
<tr>
<td>LeFRO-TC124302</td>
<td>A1g01580/At1g01590</td>
<td>0.526</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CLEC37C6</td>
<td>A1g01780</td>
<td>0.542</td>
</tr>
<tr>
<td></td>
<td>T1659</td>
<td>A1g05350</td>
<td>1.000</td>
</tr>
<tr>
<td>LeFRO-TC124302</td>
<td>A5g49730/At5g49740</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A5g50160</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T1119</td>
<td>A5g53180</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>T1321</td>
<td>A5g52050</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>T1155</td>
<td>A5g51150</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>T648</td>
<td>A5g51540</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>T753</td>
<td>A5g52390</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>T1278</td>
<td>A5g50320</td>
<td>0.001</td>
</tr>
<tr>
<td>LeFro1/LeFro-129233</td>
<td>LeFro1/LeFro-129233</td>
<td>A1g01580/At1g01590</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>CLESI J</td>
<td>A1g01820</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>T1409</td>
<td>A1g01610</td>
<td>0.019</td>
</tr>
</tbody>
</table>

(Table continues on following page.)
First, we analyzed expression of IRT/ZIP genes (Fig. 4A). LeIRT1 and LeIRT2 were previously shown to be expressed in the root, whereby LeIRT1 was iron regulated (Bereczky et al., 2003). LeIRT2 EST sequences were available in the EST database, but no data were available for LeIRT1. LeIRT2 EST sequences were all derived from root tissue (data not shown). In Arabidopsis (Fig. 4A), AtIRT1 and AtIRT2 were mainly expressed in roots and expression was induced upon iron deficiency, as expected (Eide et al., 1996; Vert et al., 2001). AtIRT2 appeared to be induced at a higher level than AtIRT1. The AtZIP8 expression pattern was similar to that of AtIRT2. AtZIP10 transcripts were detected in roots upon sufficient iron supply. AtZIP7 was mainly expressed in leaves and induced upon low iron supply.

Expression of NRAMP genes was not further analyzed here. LeNRAMP1 and AtNRAMP1 were previously found to be root specific and iron regulated (Curie et al., 2000; Bereczky et al., 2003). LeNRAMP3, AtNRAMP3, and AtNRAMP4 were iron regulated and also expressed in the leaf (Curie et al., 2000; Thomine et al., 2000; Bereczky et al., 2003). EST sequences were available for LeNRAMP3 and were derived from different tissues (data not shown).

EST sequences for LeFRO-TC124302 were found in roots and callus (data not shown). Experimental expression analysis showed that expression of LeFRO-TC124302 was not only root specific, but also slightly iron regulated and dependent on a functional LeFER gene in tomato (Fig. 5). On the other hand, transcripts for LeFRO-TC129233 were not detected experimentally in leaves and roots (data not shown). All EST sequences were derived from flower libraries, indicating that LeFRO-TC129233 was not involved in root iron uptake. As analyzed previously, LeFRO1 was expressed in roots and leaves, up-regulated by iron deficiency, and dependent on a functional LeFER gene (Li et al., 2004). Since LeFRO-TC124302 appeared to be specific to the L. esculentum genome, we tested whether LeFRO-TC124302 was essential for iron reduction in L. esculentum M82 introgression lines containing at the expected LeFRO-TC124302 locus an introgressed L. pennellii chromosomal fragment devoid of LeFRO-TC124302 (Eshed et al., 1992). We found that these introgression lines had similar levels of iron reductase activity as the control line M82 and grew in a very

Table III. (Continued from previous page.)

<table>
<thead>
<tr>
<th>Tomato Region</th>
<th>Tomato Gene/Marker</th>
<th>Arabidopsis Gene</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1327</td>
<td>At1g01300</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>LeFRO1/LeFRO-129233</td>
<td>At1g23020</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>T1173</td>
<td>At1g20200</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>LeNAS</td>
<td>At1g09240</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>T1118</td>
<td>At1g09830</td>
<td>0.069</td>
<td></td>
</tr>
<tr>
<td>T1687</td>
<td>At1g09430</td>
<td>0.016</td>
<td></td>
</tr>
<tr>
<td>CLET4E21</td>
<td>At1g06680</td>
<td>0.254</td>
<td></td>
</tr>
<tr>
<td>LeNAS</td>
<td>At1g56430</td>
<td>0.190</td>
<td></td>
</tr>
<tr>
<td>T1176</td>
<td>At1g56700</td>
<td>0.183</td>
<td></td>
</tr>
<tr>
<td>T1485</td>
<td>At1g54780</td>
<td>0.226</td>
<td></td>
</tr>
<tr>
<td>T1687</td>
<td>At1g60810</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>LeNAS</td>
<td>At1g04950</td>
<td>0.262</td>
<td></td>
</tr>
<tr>
<td>T141</td>
<td>At1g05010</td>
<td>0.261</td>
<td></td>
</tr>
<tr>
<td>T852</td>
<td>At1g07020</td>
<td>0.260</td>
<td></td>
</tr>
<tr>
<td>LeNAS</td>
<td>At1g56080</td>
<td>0.629</td>
<td></td>
</tr>
<tr>
<td>T1084</td>
<td>At1g27190</td>
<td>0.643</td>
<td></td>
</tr>
<tr>
<td>CTOC14118</td>
<td>At5g60910</td>
<td>1.000</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. Expression analysis of Arabidopsis iron uptake homologs in roots and leaves of plants grown upon low (−Fe) and sufficient iron supply (+Fe). A, IRT/ZIP genes; B, FRO genes; C, NAS genes; D, FRU gene. Expression analysis was performed by semiquantitative RT-PCR and Southern-blot analysis. Amplification of elongation factor cDNA EF1β-α (At5g19510) served as constitutive control.
similar way (data not shown). *LeFRO*-TC124302 was therefore not essential for iron uptake and iron reduction in *L. esculentum*, indicating that presumably its function was performed by the redundant *LeFRO1* gene. Arabidopsis *FRO*-like genes were all expressed in roots and leaves and tended to be induced upon iron deficiency (Fig. 4B). *AIFRO1*, *AIFRO2*, and At5g23990 shared a similar expression pattern. Expression of these genes was induced by iron deficiency in roots and leaves. Upon sufficient iron supply, *AIFRO1*, *AIFRO2*, and At5g23990 were expressed at a higher level in roots than in leaves. At5g23980 and *AIFRO3* were both expressed in leaves upon sufficient iron supply. *FRO3* was also expressed in roots under sufficient iron. Upon low iron supply, the two genes were induced in roots. *FRO3* was also induced in leaves upon iron deficiency, whereas expression of At5g23980 was at a similar level than in leaves upon sufficient iron supply. Thus, for all Arabidopsis *FRO* genes, the expression patterns may suggest a function in iron deficiency responses.

We found expression of Arabidopsis *NAS* genes in leaves and roots upon sufficient and low iron supply (Fig. 4C), supporting the tomato data that NAS activity was indeed required constantly (Ling et al., 1999). However, we observed that *AtNAS1* and *AtNAS2* were expressed in roots stronger than in leaves. Upon low iron supply, *AtNAS1* and *AtNAS2* were induced in the root. *AtNAS4* was induced by iron deficiency in roots and leaves. At sufficient iron supply, *AtNAS4* was mainly expressed in leaves. *AtNAS3* was mainly expressed in the leaf upon sufficient iron supply. Taking our expression data together indicates that Arabidopsis *NAS* genes are iron regulated. We identified Arabidopsis lines with T-DNA insertions in the coding regions of *AtNAS1*, *AtNAS2*, and *AtNAS4* (data not shown). None of these lines segregated for any discernible phenotypes, suggesting that *NAS* genes are at least partially redundant.

*LeFER* is expressed in the root but not in the leaves and cotyledons (Ling et al., 2002). The tomato *BHLH* genes with weak sequence similarity to the tomato *LeFER* gene did not seem to be expressed in roots since all EST data were derived from flower tissue or callus (data not shown). *AIFRU* was mainly expressed in roots and hardly in leaves, whereby root expression was induced by iron deficiency (Fig. 4D). The expression pattern of *AIFRU* therefore suggests that the gene was indeed involved in iron deficiency responses in Arabidopsis. *AtBHLH021* was only expressed in flowering tissues (data not shown). Spliced transcripts for *BHLH* genes *AtBHLH022*, *AtBHLH090*, *AtBHLH035*, and *AtBHLH027* were not detected at all (data not shown).

Overall, expression data contributed to assigning homologous functions. In general, we found that gene family members from tomato differed more significantly in their expression patterns than did the homologous Arabidopsis gene family members, which tended to retain similar expression patterns.

**DISCUSSION**

In this study, we assigned functionally and structurally homologous gene functions involved in iron uptake between tomato and Arabidopsis. For this purpose, we based our studies not only on sequence comparisons but also took into account map position and functional expression data. Analyzing, in addition, levels of conserved gene content adjacent to the genes of interest allowed us to predict whether gene family members were redundant genes resulting from internal genome duplication events or nonredundant single-acting genes with distinct biological functions. We found that it was easier to determine the functional homologs in tomato using the Arabidopsis information than vice versa. Due to the lower gene complexity in tomato, we hypothesize that it might be generally convenient to utilize Arabidopsis genome information to predict tomato homologs.

**Gene Homology and Map Position Reflect Genome Evolution in Tomato and Arabidopsis**

Most comparative genome programs are based on finding orthologs by using the criterion sequence similarity (e.g. http://www.tigr.org/tdb/tgi/cgi/GO.html). In this study, sequence similarity alone was not a sufficient criterion for determining functional homologs for iron uptake between Arabidopsis and tomato. For a given gene, we frequently identified more than one related gene in the other species. The Arabidopsis genome in particular had a higher number of homologs compared to tomato. Extrapolating from our data would suggest that Arabidopsis may have more and larger gene families than tomato. Taking the number of most related *FER*-like, *NAS*, *IRT*, *NRAMP*, and *FRO* genes together would indicate that Arabidopsis had 21 genes (1 + 4 + 5 + 6 + 5), whereas tomato had only 10 genes for these functions (1 + 1 + 2 + 3 + 3). Van der Hoeven et al. (2002) calculated that only 50% of tomato genes were uncovered from the EST analysis. In our study, 7 of the 10 tomato genes were uncovered by EST analysis. Correcting our numbers to an estimated 14 tomato genes versus...
21 Arabidopsis genes would hence suggest that Arabidopsis had about 33% more genes than tomato for the families analyzed. This result was supported by the similar findings of Van der Hoeven et al. (2002) on the comparison of multigene family copy numbers between tomato EST data and Arabidopsis genome data.

The comparative mapping studies presented here showed that 20% to 50% of analyzed COS markers were located in the vicinity of corresponding iron uptake genes in tomato and in Arabidopsis (except for those of the LeFER region). In contrast, only 0% to 14% of these COS markers matched with Arabidopsis regions containing nonhomologous iron uptake genes. The clustering of conserved COS marker positions in between tomato and Arabidopsis was calculated to be significant for several of them, and so it can be excluded that all clusters occurred at random. For some of the compared Arabidopsis-tomato regions with P-values greater than 0.05, we speculate that an increase of the number of analyzed markers would decrease the P-values and make the clustering statistically significant. We avoid utilizing the term collinearity for these observations since the actual gene order was not conserved in all cases. Similar gene content suggests a common evolutionary origin of the corresponding Arabidopsis and tomato genome regions. The analysis also showed internally duplicated regions in the Arabidopsis genome. It seems likely that the reason for higher multigene family copy numbers in Arabidopsis was the genome structure. Some duplications uncovered here have been found in previous studies as internally duplicated regions in the Arabidopsis genome; for example, the duplication events involving chromosome 1, chromosome 4, and chromosome 5 containing IRT/ZIP gene regions as well as the chromosome 1 duplication involving NAS genes (Blanc et al., 2003). We also detected signs of duplications that had not been previously detected and might reflect hidden duplications (Simillion et al., 2002), such as the chromosome 5 duplication involving NAS genes. Due to similar gene content and sequence similarity, we predict that NAS genes reveal an older duplication event responsible for the duplication of NAS regions on chromosomes 1 and 5 as well as one or two more recent duplication events involving chromosomes 1 and 5. The tomato genome was presumably also partially duplicated, even though to a lower degree. LeIRT1/LeIRT2 appear as recent duplicates based on high sequence similarity. Both LeFRO1/LeFRO-TC129233 and LeFRO-TC124302 regions showed levels of gene conservation with the AtFRO2 region as well as the region containing the three related FRO genes At5g50160/At5g49730/At5g49740. The latter three Arabidopsis genes were related to tomato FRO genes in terms of gene content despite being more distantly related in sequence. These observations might suggest that either the genomic region was also duplicated in tomato or that the duplication occurred early in dicot evolution. We hypothesize that LeFRO1 and LeFRO-TC129233 are present in tandem duplication in the tomato genome.

Importance of Functional Expression Data for Assignment of Functional Homology and Redundancy in Iron Uptake

This study dealt mainly with essential gene functions required in the root and/or for iron regulation. We found that expression data greatly contributed to assigning appropriate and unique homologous gene pairs. For IRT and FRO sequences, it was indispensable to consider sequence, map position, and gene expression aspects together. For example, among the 15 Arabidopsis ZIP genes (Mäser et al., 2001), AtIRT1 and AtIRT2 as well as LeIRT1 and LeIRT2 are the most similar pairs, respectively, in each genome. AtIRT1 and LeIRT1 are likely encoding the essential and iron-regulated iron transporters in both species (Varotto et al., 2002; Vert et al., 2002; Bereczky et al., 2003). AtIRT2 is iron regulated but not essential (Varotto et al., 2002). In contrast, LeIRT2 is neither iron regulated nor dependent on the regulator FER and thus not likely involved in iron mobilization (Eckhardt et al., 2001; Bereczky et al., 2003). In addition, LeIRT1 and LeIRT2 are very closely related in sequence, suggesting that the two genes are recent tandem duplications. In contrast, AtIRT1 and AtIRT2, which are also located in tandem, are much more diverged in sequence as if the tandem duplication event was much older than that in tomato. We found that AtZIP8 shared a similar expression pattern with AtIRT2. AtZIP8 was most related in sequence to AtIRT1 and AtIRT2. Hence, it is a possibility that AtIRT2 function overlaps with a redundant function provided by AtZIP8. It seems, however, unlikely that, despite their tandem location with IRT1, AtIRT2 and LeIRT2 are functioning as orthologs. Similarly, our study suggests that LeIRT2 and AtZIP10 may share a common function.

The most drastic example of why functional gene analysis data are needed for investigating homologous gene functions was provided by the FRO genes. Among the eight Arabidopsis FRO-like genes, AtFRO2 was unequivocally identified as the FRO gene in Arabidopsis due to genetic experiments (Robinson et al., 1999). From expression studies, we suggest that four other Arabidopsis FRO gene homologs could share additional functions in iron mobilization. In tomato, LeFRO1 and LeFRO-TC124302 may both be root FRO genes acting upon iron mobilization. Both genes are iron regulated in roots and dependent on FER (see also Li et al., 2004) The analysis of introgression L. esculentum lines devoid of LeFRO-TC124302 suggested that LeFRO-TC124302 was not essential and that a redundant iron reductase gene must exist in the Lycopersicon genome. This redundant function could well be provided by LeFRO1. In contrast, LeFRO-TC129233 was also similar in sequence to AtFRO2 and the genes were located in regions with conserved gene content. However, LeFRO-TC129233 was not
found to be expressed in roots or leaves, nor upon iron deficiency, suggesting that LeFRO-TC129233 may be involved as putative iron reduscate in a different biological process than LeFRO1 and LeFRO-TC124302.

The **chloronerva** (*nas*) tomato mutant is characterized by a distinct leaf chlorosis and root phenotype upon low iron supply (Scholz et al., 1992). A similar mutant has not been described in Arabidopsis. Our analysis of the four *NAS* genes in Arabidopsis indicated that *NAS* genes were duplicated at least twice. Interestingly, despite of a strong sequence conservation, we found differential regulation of *NAS* genes in iron deficiency responses, whereby three of the four *NAS* genes were up-regulated by low iron supply. We established previously that *NAS* was required for induction of LeFER-mediated *LeIRT1* and *LeNRAMP1* induction upon iron deficiency (Bereczky et al., 2003).

In accordance with these observations, iron regulation of *NAS* genes in Arabidopsis might suggest that nicotianamine could have a function in iron mobilization or signaling. Due to the high sequence conservation, *NAS* genes form a partially redundant gene family. Absence of single functional gene members such as observed in the metallophyte and zinc hyperaccumulator *Arabidopsis halleri* may lead to discrete dominant phenotypes—for instance in *A. halleri* zinc tolerance (Becher et al., 2004).

The Arabidopsis genome contains 162 predicted *BHLH* genes (Bailey et al., 2003). Sequence comparisons show that outside the conserved *BHLH* domain, the AtFRU protein is quite distinct from all other *BHLH* proteins. The genome regions of the two genes did not show apparent conservation of gene clusters. LeFER and AtFRU were both expressed in roots, whereby AtFRU was also expressed to a low level in leaves. AtFRU was induced upon iron deficiency. Recent genetic experiments suggest that LeFER and AtFRU serve a conserved biological function in iron regulation (Jakoby et al., 2004). Perhaps genes encoding regulatory components have lower evolutionary constraints on sequence conservation than genes encoding transporters and enzymes.

Arabidopsis and tomato show similar physiological responses to iron deficiency (strategy I), so that most likely iron regulation and iron uptake are conserved processes in these two species that involve conserved gene functions. If these gene functions were not performed by orthologs, at least they were expected to be compensated by paralogs. We predict from our analysis that essential homologous gene functions of iron uptake are indeed involved and conserved between Arabidopsis and tomato. However, detailed sequence and map position analysis indicated that these conserved genes are most likely of paralogous origin rather than of orthologous origin. Despite the similarities of genes and proteins involved in iron mobilization, tomato induces root morphological alterations, root hair proliferation, and transfer cell development as a response to iron deficiency (Schmidt, 1999; Schikora and Schmidt, 2001, 2002). These morphological alterations are supposed to aid iron uptake. In Arabidopsis, morphological alterations to iron deficiency are less pronounced, and it is possible that Arabidopsis relies instead on its duplicated gene functions.

### Table IV. Oligonucleotide primers used in RT-PCR experiments

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Number</th>
<th>5′ Primer 5′–3′</th>
<th>3′ Primer 5′–3′</th>
</tr>
</thead>
<tbody>
<tr>
<td>LeFRO-TC129233</td>
<td>TC129233</td>
<td>atgggggtataggggtacagag</td>
<td>tctgctcgtcatatgcctcttgt</td>
</tr>
<tr>
<td>LeFRO-TC124302</td>
<td>TC124302</td>
<td>tgtgacgctcaaggctgatcca</td>
<td>ggcctacaaacaaatctca</td>
</tr>
<tr>
<td>LeEF-1a</td>
<td>TC123773</td>
<td>cctctgggctgtcaactctgt</td>
<td>ctgggttgtagagcttactatc</td>
</tr>
<tr>
<td>AtZIP7</td>
<td>At2g04032</td>
<td>gctgacttcgagaagactct</td>
<td>atcaagttgagagcaagaaga</td>
</tr>
<tr>
<td>AtZIP8</td>
<td>At5g45105</td>
<td>atgcagagccgatccacag</td>
<td>ttctaaactgctgagataaag</td>
</tr>
<tr>
<td>AtZIP10</td>
<td>At1g31260</td>
<td>tgaacgtctgcttatgcat</td>
<td>catctattggagctcgcctct</td>
</tr>
<tr>
<td>AItRT1</td>
<td>At4g19690</td>
<td>gcacggtcgctcgcgggtgt</td>
<td>atccacacatgtatcgcct</td>
</tr>
<tr>
<td>AItRT2</td>
<td>At4g19680</td>
<td>gtaaagacactgtcgcagcag</td>
<td>tgcgcttcgaataaataacaaat</td>
</tr>
<tr>
<td>AITR0-like</td>
<td>At5g23990</td>
<td>gataagacctcaagagcagtta</td>
<td>caaacatatatgaagatgcgtta</td>
</tr>
<tr>
<td>AITR0-like</td>
<td>At5g23980</td>
<td>gataagatcctcaagaaatgta</td>
<td>aacacaagatgatagctcata</td>
</tr>
<tr>
<td>AITR01</td>
<td>At1g01590</td>
<td>ccaacaccttctcggtgatt</td>
<td>tgtgaaccacacatcgtgat</td>
</tr>
<tr>
<td>AITR02</td>
<td>At1g01580</td>
<td>tctcaacactctctcctctcatct</td>
<td>caaacatatgataaagatgcgtt</td>
</tr>
<tr>
<td>AINAS3</td>
<td>At1g09240</td>
<td>caacctcggttgattctctct</td>
<td>atagagaaagctcaagagact</td>
</tr>
<tr>
<td>AINAS4</td>
<td>At1g56430</td>
<td>agggagaggctcagatgtggt</td>
<td>acacagacttcttctgagtgt</td>
</tr>
<tr>
<td>AINAS1</td>
<td>At5g04950</td>
<td>ggggttaatgaggtgctcgg</td>
<td>agacagaatgaaagagacagt</td>
</tr>
<tr>
<td>AINAS2</td>
<td>At5g56080</td>
<td>ggcatcagctggacggtggt</td>
<td>ctctgacatccttcctcatac</td>
</tr>
<tr>
<td>AIFRU BHLH029</td>
<td>At2g28160</td>
<td>atggagaaagatgacgcacgc</td>
<td>tcaataagttgacagacgcgg</td>
</tr>
<tr>
<td>AIBHLH 090</td>
<td>At1g10610</td>
<td>atgatgatgagagagagagagagagagag</td>
<td>ttatactccttgccatgagctcgtat</td>
</tr>
<tr>
<td>AIBHLH 035</td>
<td>At5g57150</td>
<td>atggagatactctccacaaagaaatagc</td>
<td>ttgagagacagagagagagagagagctgga</td>
</tr>
<tr>
<td>AIBHLH 027</td>
<td>At4g29930</td>
<td>atggagaatctctccacaaagaaatagc</td>
<td>tcaaaaccataaagagctgatctcttcct</td>
</tr>
<tr>
<td>AIBHLH 022</td>
<td>At4g21330</td>
<td>atggagaggagagagagagagagagag</td>
<td>ttagatgtccttcctaatcagta</td>
</tr>
<tr>
<td>AIBHLH 021</td>
<td>At2g16910</td>
<td>atggacgctcagacggagagggtgc</td>
<td>taagttggtggagagagagagagagag</td>
</tr>
<tr>
<td>AIFRH-a</td>
<td>At5g19510</td>
<td>atggaggagagagagagagagagagag</td>
<td>aatactgagagagagagagagagagag</td>
</tr>
</tbody>
</table>
MATERIALS AND METHODS

Sequence Analysis and Database Searches

Amino acid sequences of AtFRO2 (Robinson et al., 1999), AtIRT1 (Eide et al., 1996), LeIrt1 (Eckhardt et al., 2001), AtNRAMP1 (Curie et al., 2000), AtNRAMP3 (Thomine et al., 2000), LeNRAMP1 and LeNRAMP3 (Bereczky et al., 2003), LeNAS5 (CHLORONERVA; Ling et al., 1999), and LeFER (Ling et al., 2002) were used to search homologous sequences from tomato (Lycopersicon esculentum) and Arabidopsis (Arabidopsis thaliana) by BLASTP and BLASTX. Arabidopsis amino acid sequence homologs were identified using the BLAST program (http://www.Arabidopsis.org/Blast) and retrieved at http://www.sgn.cornell.edu/maps/tomato_Arabidopsis/synteny_map.html. tomato amino acid sequences were identified and retrieved using the BLAST program (http://tigrblast.tigr.org/tgi). Sequences were retained for analysis if they were most significantly related by comparable E-values. Multiple amino acid sequences were aligned using ClustalX and N-J Trees were generated at http://www-igbmc.u-strasbg.fr/BioInfo/ClustalX/Top.html. N-J Trees were represented using the NJPLOT (Perriére and Gouy, 1996). Functional expression data were available at http://www.tigr.org for tomato genes.

Mapping of Tomato Genes and Analysis of Conserved Gene Regions

Genomic tomato DNA fragments were mapped by restriction fragment length polymorphism analysis using 43 F2 individuals of an L. esculentum/Lycopersicon pennelli mapping population according to Tankely et al. (1992). Genetic maps were generated using the MAPMAKER program (Lander et al., 1987).

Tomato COS markers that mapped within a distance of ±10 cM of iron uptake genes were selected for analysis of conserved gene regions at http://www.sgn.cornell.edu/maps/tomato_Arabidopsis/synteny_map.html. The encoded amino acid sequences of the selected COS markers were used to BLAST for corresponding Arabidopsis amino acid sequences at http://www.Arabidopsis.org/Blast. Arabidopsis genes were used as COS markers if their encoded peptides gave expected values below e-20 in alignments with encoded peptides of tomato COS markers (Supplemental Table I). Arabidopsis COS markers located up to 400 genes from iron uptake genes (approximately ±2 Mb) were considered to be in that same chromosomal region. For example, genes in the region of At4g19690 would be expected to have gene locus numbers between At1g56900 and At4g25600.

Statistical analysis and calculation of P-values for random clustering of homologous genes between Arabidopsis and tomato were performed as follows: For each of the Arabidopsis genes of interest, an interval of ±1 Mb was considered in which the gene of interest was in the center. If K was the total number of Arabidopsis genes in the ±1-Mb interval, L was the number of Arabidopsis genes located in this interval and homologous to genes from the corresponding ±10-cM tomato region minus tandem duplicates, M was the total number of genes in the Arabidopsis genome, and N was the number of genes in the Arabidopsis genome homologous to genes from the ±10-cM tomato region minus tandem duplicates, then the probability pL of finding L genes by chance given K, M, and N was computed by Fisher’s exact test as the right tail of the hypergeometric distribution:

\[ p_L = \sum_{i=0}^{K} \binom{K}{i} \binom{M-K}{N-i} \binom{M}{N} \]

When determining K, L, M, and N, the gene in the center of the ±1-Mb interval was not counted. The statistical significance was computed as the probability that L or more than L genes could occur by chance in the ±1-Mb region if the numbers K, M, and N were fixed. The P-values were assigned to the genes in the center of the ±1-Mb intervals, and P-values less than 0.05 were considered significant for nonrandom clustering. For example, At1g31260 maps at position 11,175,540 bp on Arabidopsis chromosome 1. The ±1-Mb interval centered at position 11,175,540 contains 420 genes, including At1g31260 (K = 419). Out of those 419 genes, five are homologues of COS markers from the ±10-cM tomato region containing LeIrt1 (L = 5). In total, there are 26,405 Arabidopsis genes, including At1g31260 (M = 26,404). Out of those 26,404 genes, 51 are homologs of COS markers from the ±10-cM region of LeIRT1 (n = 51). The density of 5 genes out of 419 is 6-fold higher than the density of 51 genes out of 26,404, and the P-value is 0.001.

Plant Growth and Plant Material

Tomato plants used in RNA expression analysis were derived from the lines L. esculentum T32383 fer (for mutant phenotype) and L. esculentum Money-maker (wild type). Homozygous plants of the first-generation introgression lines from L. pennelli (LA 716) in the genetic background of the processing tomato variety M82 were propagated and utilized in this study (Eshk et al., 1992).

For tomato iron uptake studies, 12-d-old plants were grown in a hydroponic Hoagland medium containing 0.1 (low iron) or 10 μM (sufficient iron) FeNaEDTA for 1 week according to Bereczky et al. (2003). Iron reductase assays were performed on roots as described by Stephan and Prochazka (1989).

For expression studies in Arabidopsis, 2-week-old Arabidopsis Columbia plants grown on solid Hoagland medium in the presence of 10 μM FeNaEDTA were transferred to Hoagland medium containing 10 μM FeNaEDTA (sufficient iron) or no iron and 20 μM bathophenanthroline disulfonic acid (low iron) for 5 d.

nas1, nas2, and nas4 T-DNA insertion lines were identified by database searches at the SALK Institute Web site (http://signal.salk.edu) and ordered from the Arabidopsis Biological Resource Center.

Gene Expression Analysis

Semiquantitative reverse transcription (RT)-PCR analysis was performed according to Bereczky et al. (2003). Total RNA was extracted from tomato using the Purescript RNA Isolation kit (Genta Systems, Minneapolis) and from Arabidopsis using the Invisorb Spin Plant RNA Mini kit (Invitex, Berlin). From 0.1 to 2 μg RNA were reverse transcribed into cDNA using oligo(dT) primer and the RevertAid First Strand cDNA synthesis kit (Fermentas GmbH, St. Leon-Rot, Germany). cDNA was amplified using specific primers according to standard procedures (Table IV). The number of amplification cycles was determined experimentally so that the reaction was analyzed in the exponential phase. PCR fragments were separated by agarose gel electrophoresis, blotted, and hybridized according to standard procedures. Amplification of elongation factor cDNA was used as a constitutive control to normalize all samples.

ACKNOWLEDGMENTS

Tomato mapping filters were provided by Dr. M. Galan. Help with the MAPMAKER program by Dr. X. Huang and Dr. M. Röder is greatly acknowledged.

Received June 2, 2004; returned for revision July 7, 2004; accepted July 7, 2004.

LITERATURE CITED


Copyright © 2004 American Society of Plant Biologists. All rights reserved.


Cadmium and iron transport by members of a plant metal transporter family in Arabidopsis with homology to Nramp genes. Proc Natl Acad Sci USA 97: 4991–4996


