TOMATOMA: A Novel Tomato Mutant Database Distributing Micro-Tom Mutant Collections

Takeshi Saito1,4, Tohru Ariizumi1,4, Yoshihiro Okabe1, Erika Asamizu1, Kyoko Hiwasa-Tanase1, Naoya Fukuda1, Tsuyoshi Mizoguchi1, Yukiko Yamazaki2, Koh Aoki3 and Hiroshi Ezura1,*

1Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, 305-8572 Japan
2National Institute of Genetics, Yata 1111, Mishima, 411-8540 Japan
3Kazusa DNA Research Institute, 2-6-7 Kazusa-Kamatari, Kisarazu, 292-0818 Japan
4These authors contributed equally to this work

*Corresponding author: E-mail, ezura@gene.tsukuba.ac.jp; Fax, +81-29-853-7734
(Received December 15, 2010; Accepted January 8, 2011)

The tomato is an excellent model for studies of plants bearing berry-type fruits and for experimental studies of the Solanaceae family of plants due to its conserved genetic organization. In this study, a comprehensive mutant tomato population was generated in the background of Micro-Tom, a dwarf, rapid-growth variety. In this and previous studies, a population was generated in the background of Micro-Tom, including 8,598 and 6,422 M2 mutagenized lines produced by ethylmethane sulfonate (EMS) mutagenesis and γ-ray irradiation, and this study developed and investigated these M2 plants for alteration of visible phenotypes. A total of 9,183 independent M2 families comprising 91,830 M2 plants were inspected for phenotypic alteration, and 1,048 individual mutants were isolated. Subsequently, the observed mutant phenotypes were classified into 15 major categories and 48 subcategories. Overall, 1,819 phenotypic categories were found in 1,048 mutants. Of these mutants, 549 were pleiotropic, whereas 499 were non-pleiotropic. Multiple different mutant alleles per locus were found in the mutant libraries, suggesting that the mutagenized populations were nearly saturated. Additionally, genetic analysis of backcrosses indicated the successful inheritance of the mutations in BC1F2 populations, confirming the reproducibility in the morphological phenotyping of the M3 plants.

To integrate and manage the visible phenotypes of mutants and other associated data, we developed the in silico database TOMATOMA, a relational system interfacing modules between mutant line names and phenotypic categories. TOMATOMA is a freely accessible database, and these mutant recourses are available through the TOMATOMA web site (http://tomatoma.nbrp.jp/index.jsp).

Keywords: Database • In silico • Micro-Tom • Mutagenesis • Mutant • Tomato.

Abbreviations: ARF, auxin response factor; DAE, days after emasculation; EMS, ethylmethane sulfonate; EST, expressed sequence tag; IL, introgression line; LD, lethal dose; MTA, material transfer agreement; NBRP, National BioResource Project; SGN, Solanaceae genomics network; SNP, single nucleotide polymorphism; TILLING, targeting induced local lesions in genomes; TGRC, Tomato Genetics Resource Center; WT, wild type.

Introduction

The tomato (Solanum lycopersicum) is an important crop in the fresh vegetable and food processing industry, with one of the highest productions for edible crops in the world. The tomato is also an excellent model plant for studying Solanaceae species, since it has a relatively small genomic size (950 Mb) and shares the same haploid chromosome number and conserved genome organization (i.e. a high level of gene synteny) with the other Solanaceous plants (Rick and Yoder 1988, Hille et al. 1989, Tanksley 2004). Additionally, the tomato bears fleshy berry-type fruits and is extensively used as a target for studying fruit development and fruit ripening/maturity, as well as metabolite analysis (Emmanuel and Levy 2002, Carrari and Fernie 2006, Egea et al. 2010, Mochida and Shinozaki 2010, Pineda et al. 2010, Sun et al. 2010, Yin et al. 2010).

Furthermore, studies of abiotic and biotic responses have been widely carried out in the tomato (Matsui et al. 2010, Rellan-Alvarez et al. 2010, Rivero et al. 2010, Uehara et al. 2010). The tomato has been selected as a core model plant for accelerating genomic studies in the Solanaceae family, and its genome is being sequenced by The International Solanaceae Genomics Project (SOL) (Mueller et al. 2005a, Mueller et al. 2005b, Mueller et al. 2009).

The majority of the tomato genome has now been released and is best represented at the Solanaceae Genomics Network web site (SGN, http://solgenomics.net/). Moreover, draft genome sequences of some wild relatives, large numbers of tomato expressed sequence tags (ESTs; 320,000 clones) and
the full-length cDNA sequence (13,722 unigene clones) have been released at the SGN, MiBASE (http://www.pgb.kazusa.or.jp/mibase/) and KaFTom (http://www.pgb.kazusa.or.jp/kaftom/) databases (Aoki et al. 2010, Ozaki et al. 2010, Bombarely et al. 2011). To render the genome sequence useful, the International Tomato Annotation Group (ITAG) has been engaging in gene annotation and had identified 7,464 protein coding genes as of 2009 (Mueller et al. 2009). Due to the great advances in the genome sequence project, it seems that now is the perfect time to exploit the genome information for exploring gene functions controlling important traits of the tomato and that the mutagen-induced mutant population could be a powerful tool for accelerating tomato functional genomics (Kuromori et al. 2009).

Menda et al. (2004) developed a comprehensive isogenic tomato mutant population in the genetic background of cultivar (cv.) M82. A total of 13,000 M2 families, which were derived from ethylmethane sulfonate (EMS) and fast-neutron mutagenesis, were visually phenotyped in the field, and 3,417 independent mutants were classified into a morphological catalog on the website 'Genes That Make Tomatoes' (http://zamir.sgn.cornell.edu/mutants/). Recently, new mutant collections were generated in the genetic backgrounds of the cv. Red Setter and Tpaadasu using an EMS mutagenesis approach (Gady et al. 2009, Minoia et al. 2010). These isogenic mutant resources are useful for dissecting the mechanisms underlying mutant phenotypes, and such mutagenized populations are also being used to develop targeting induced local lesions in genomes (TILLING) platforms, which represent a high-throughput reverse genetic strategy to screen for point mutations in specific regions of targeted genes (McCallum et al. 2000a, McCallum et al. 2000b, Colbert et al. 2001, Minoia et al. 2010). During the last decade, several mutant populations in the genetic background of the miniature cv. Micro-Tom, which were derived from EMS (Meissner et al. 1997), fast-neutron bombardment (David-Schwartz et al. 2001), activation tagging (Mathews et al. 2003) and Activator/Dissociation (Ac/Ds) transposon tagging (Meissner et al. 2000), have been developed, although these resources are not currently available to the scientific community.

The C. M. Rick Tomato Genetics Resource Center (TGRC) at the University of California, Davis (http://tgrc.ucdavis.edu/) plays an important role as a seed stock center of various tomato genotypes, providing phenotypic information on 1,023 monogenic mutants at 625 putative genetic loci. These stocks come from several sources, including spontaneous and induced mutants and natural variants from the edible tomato and wild relatives, including many members of the Solanum genus. The TGRC also provides 1,160 lines of wild species and 1,023 monogenic mutants for a total of 2,303 unique lines. These seed materials are available upon request following the completion of a material transfer agreement (MTA) contract.

In Japan, as a part of the National BioResource Project (NBRP) funded by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan (Yamazaki et al. 2010), results in the cultivation of Micro-Tom in an indoor growth room.

**Results**

**Morphological appearance of Micro-Tom**

Fig. 1 shows the plant behavior of the miniature tomato cv. Micro-Tom compared with that of cv. M82, the background variety in the saturated mutant libraries used by Menda et al. (2004). Most cultivated tomato varieties, including M82, have certain drawbacks when grown in limited space due to their large size (approximately 1 m in height in the adult stage) and a relatively long life cycle (90–110 d from seed germination to fruit maturation) (Meissner et al. 1997, Emmanuel and Levy 2002). In contrast, Micro-Tom exhibits dwarfism (approximately 10–20 cm height) and a rapid life cycle, with fruit maturity occurring 70–90 d after sowing. Micro-Tom can be grown indoors.
at a high density of up to 1,325 plants m$^{-2}$ (Scott and Harbaugh 1989), which is ideal for indoor cultivation in most plant biology laboratories (Fig. 1C). Additionally, a highly efficient Agrobacterium-mediated transformation method has been established for this cultivar (Dan et al. 2006, Sun et al. 2006), which is indispensable for investigating gene function. Moreover, Micro-Tom is capable of both intraspecific and interspecific cross-pollination with most cultivated tomato varieties and wild relatives, allowing for the transfer of mutations into commercially available varieties and the production of F$_2$ hybrid populations for positional candidate gene cloning. Furthermore, various types of genetic and molecular tools have been developed in Micro-Tom, including metabolite profile annotation and full-length cDNA collections (Iijima et al. 2008, Aoki et al. 2010). These beneficial traits make Micro-Tom an attractive model cultivar for studying tomato plants, and, thus, it was chosen as the core genetic background to develop comprehensive mutant populations in this study.

**Frequency of mutation in EMS-mutagenized populations**

To optimize EMS mutagenesis, Micro-Tom seeds were treated with different EMS concentrations (0.3, 0.5, 1.0 or 1.5%) and for different incubation times (12, 16 or 48 h incubation) from 2004 to 2009 (Supplementary Table S1). For each EMS treatment, the lethal dose (LD), rate of surviving M$_1$ plants and rate of fertile M$_1$ plants were determined. The rate of surviving M$_1$ plants was probably dependent on the concentrations of the EMS treatments (0.3–1.5%), ranging from 34.7 to 94.1%, with the LD ranging from 2 to 63. The rate of fertile M$_1$ plants ranged from 19.0 to 87.9%.

In total, 8,598 of 22,500 mutagenized M$_1$ plants were fertile, and M$_2$ seeds were harvested from each of the single M$_1$ plants (Fig. 2). Then, approximately 10 M$_2$ seeds produced from the same M$_1$ plant were sown as an M$_2$ family, and M$_3$ seeds were harvested from each M$_2$ family. By November 2010, 6,483 M$_2$ families (~65,000 M$_2$ seeds) had been sown in the greenhouse. The germination rates of the M$_2$ families appeared to vary from 32.6 to 78.0%, depending on the severity of the EMS treatments (Supplementary Table S2). A total of 5,267 M$_2$ families produced M$_3$ seeds, whereas 1,216 M$_2$ families did not produce M$_3$ seeds (Supplementary Table S2).

In the process of propagating M$_2$ seeds, individual M$_2$ plants were initially inspected for visible phenotypic changes (Fig. 2). Overall, 865 individual mutants were isolated. The rates of mutant recovery in the 1.0% EMS treatment were 12.0% (EMS-3), 18.0% (EMS-7) and 17.9% (EMS-8) (Supplementary Table S2). A relatively lower rate of mutant recovery was found in lower EMS treatments (for EMS-1, -2, -4 and -6, the rate was from 5.6 to 13.6%). In contrast, the highest rate of mutant recovery was found for the 1.5% EMS treatment (EMS-10, 30%); the rate of fertility in the M$_1$ plants was 6.5% (Supplementary Table S1), and the rate of germination in the M$_2$ plants was 44.9% (Supplementary Table S2). Such an extremely low rate of fertility was not efficient in the mutagen approach employed here, and thus we concluded that the 1.0% EMS treatment represented the most suitable conditions for Micro-Tom EMS mutagenesis in this study.

**M$_2$ plant phenotyping**

The M$_2$ plants were inspected for alteration in visible phenotypes at many developmental stages from seed germination to fruit maturation (Fig. 2). The production of 6,422 lines of M$_2$ seeds by $\gamma$-ray irradiation was previously described (Matsukura et al. 2007), and these M$_2$ plants were also phenotyped in this study. All of the visible phenotypes were classified into 15 major phenotypic categories and 48 subcategories (Table 1). The categories used to describe the phenotypes were derived

---

**Fig. 2** Flow chart for constructing mutagenized populations. WT Micro-Tom seeds (M$_0$) were treated with mutagen, and the resulting M$_1$ seeds were sown in the greenhouse to harvest M$_2$ seeds from single M$_1$ plants. Ten M$_2$ seeds were sown and grown as an M$_2$ family, and whole M$_1$ seeds were harvested from the same M$_1$ family. In this process, each M$_2$ plant was initially inspected for visible phenotypic alterations, and the M$_3$ homozygous seeds from each mutant were also obtained. Individual mutant images and their phenotypic categories were registered in TOMATOMA.
from systematic phenotyping of the mutant populations, as reported previously (Menda et al. 2004). By November 2010, 6,483 and 2,700 M2 families (9,183 M2 families in total) were phenotyped from EMS-mutagenized and \( \gamma \)-ray-irradiated M2 populations, respectively, and 865 and 183 individual mutants were identified (1,048 mutants in total). Of the 865 and 183 analyzed M2 plants, 442 and 107 M2 plants were found to show more than one phenotypic category (pleiotropic) in the EMS-mutagenized and \( \gamma \)-ray-irradiated populations (Fig. 3). In contrast, 423 EMS-derived and 76 \( \gamma \)-ray-derived M2 plants were found to be non-pleiotropic mutants showing only one phenotypic category. In total, 1,819 phenotypic categories were found in 1,048 mutants (Table 1).

Fig. 4 shows the 15 major categories of defined visible phenotypes and the number of phenotypic categories found in the mutant populations. The most abundant category was plant size (26.8%), followed by plant habits (26.6%), leaf morphology (12.9%) and leaf color (10.3%) in the EMS population (Fig. 4A). In the \( \gamma \)-ray-irradiated population, the most abundant category was plant size (23.8%), followed by leaf morphology (13.5%), plant habits (11.8%), leaf color (11.2%) and seeds (10.6%) (Fig. 4B). In both populations, phenotypic categories associated with the vegetative organs were dominant. In fact, four major categories related to vegetative organs (plant size, plant habit, leaf morphology and leaf color)

### Table 1

<table>
<thead>
<tr>
<th>Major category</th>
<th>Subcategory</th>
<th>No. of mutant lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seeds</td>
<td>Germination</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Seedling lethality</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Slow germination</td>
<td>11</td>
</tr>
<tr>
<td>Plant size</td>
<td>Extremely small</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Small plant</td>
<td>316</td>
</tr>
<tr>
<td></td>
<td>Large plant</td>
<td>55</td>
</tr>
<tr>
<td>Plant habit</td>
<td>Internode length</td>
<td>169</td>
</tr>
<tr>
<td></td>
<td>Branching</td>
<td>216</td>
</tr>
<tr>
<td></td>
<td>Aborted growth</td>
<td>8</td>
</tr>
<tr>
<td>Leaf morphology</td>
<td>Leaf width</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Leaf size</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Leaf texture</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>Other leaf morphology</td>
<td>102</td>
</tr>
<tr>
<td>Leaf color</td>
<td>Purple leaf</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Yellow leaf</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Yellow-green leaf</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Dull green/gray leaf</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Variegation</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>White leaf</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Other leaf color</td>
<td>39</td>
</tr>
<tr>
<td>Flowering timing</td>
<td>Early</td>
<td>1</td>
</tr>
<tr>
<td>Inflorescence structure</td>
<td>Many flowers</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Few flowers</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Inflorescence structure not formed</td>
<td>2</td>
</tr>
<tr>
<td>Flower morphology</td>
<td>Flower homeotic mutation</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Flower organ size</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Flower organ width</td>
<td>11</td>
</tr>
<tr>
<td>Flower color</td>
<td>White flower</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Pale yellow flower</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Strong yellow flower</td>
<td>1</td>
</tr>
<tr>
<td>Fruit size</td>
<td>Small fruit</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Large fruit</td>
<td>12</td>
</tr>
<tr>
<td>Fruit morphology</td>
<td>Long fruit</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Rounded fruit</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td>1,479</td>
</tr>
</tbody>
</table>

The number of phenotypic categories scored in M2 plants is shown. Note that the total number of phenotypic categories exceeds the total number of mutants because a mutant sometimes presented more than one phenotype (pleiotropy).

### Table 1 (continued)

<table>
<thead>
<tr>
<th>Major category</th>
<th>Subcategory</th>
<th>No. of mutant lines</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fruit color</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Fruit ripening</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Sterility</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Disease and stress</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Disease and stress response</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td><strong>1,819</strong></td>
<td>340</td>
</tr>
</tbody>
</table>
represented >60% of the total phenotypes. These trends were similarly observed in the EMS-mutagenized populations in the M82 and Red Setter genetic backgrounds (Menda et al. 2004, Mionia et al. 2010). Although the distributions of the phenotypic categories were similar between the EMS-mutagenized and γ-ray-irradiated populations, the frequency of subcategories related to seedling lethality (9/340) and slow germination (27/340) was greater in the γ-ray-irradiated populations than in the EMS-mutagenized populations (14/1,479 and 11/1,479, respectively) (Table 1); thus, the phenotypic category of ‘Seeds’ represented the fifth largest group in the γ-ray-irradiated lines (Fig. 4B).

Mutants similar to previously reported mutants

Supplementary Fig. S1 presents the appearance of the mutants in which the visible phenotypes were similar to those of previously reported mutants. For example, TOMJPG0114 produced simple leaves with single leaflets, which was similar to entire or indole-3-acetic acid 9 (iaa9) (Wang et al. 2005). TOMJPG0226 had elongated leaflets without serration, similar to lanceolate (Ori et al. 2007). TOMJPG3699 exhibited leaf margins that curled adaxially, similar to flaccio (wilfy) (Bowman et al. 1984), and TOMJPE6000 showed incomplete Chl deficiency, being white initially and later becoming green, as was observed in ghost (Scolnik et al. 1987). TOMJPG5663 set pink colored fruits, similar to y (Adato et al. 2009, Ballester et al. 2010), and TOMJPE6152 had light colored petals and set orange colored fruits, as in tangerine (Isaacson et al. 2002). Furthermore, TOMJPG4290 exhibited inflorescences with a single flower, similar to uniflora (Dielen et al. 2004), and TOMJPE6414 exhibited inflorescences that were exceptionally large and excessively branched, similar to anantha (Lippman et al. 2008). TOMJPG2941 had greatly condensed, wrinkled, dark green leaves and foreshortened internodes, similar to dumpy (Koka et al. 2000), while TOMJPE1832 set fruits with an increased number of locules, similar to fascinated (Cong et al. 2008). TOMJPG1331 set elongated fruits, similar to sun (Xiao et al. 2008), and TOMJPG2614 set pear-shaped fruits, similar to ovate (Liu et al. 2002). It is possible

Fig. 4 Classification of visible mutant phenotypes by 15 major phenotypic categories. (A) The 15 major categories and the number of phenotypic categories found in M2 plants of EMS-mutagenized populations. (B) The 15 major categories and the number of phenotypic categories found in M2 plants of γ-ray-irradiated populations.
that these mutants represent new alleles of corresponding mutations, and further analysis will be essential for clarifying this possibility. In this study, we provide evidence that TOMJPG0114 represents a new allele of the entire/iaa9 locus (see below).

**Leaf structure mutants**

Supplementary Fig. S2 shows representative mutants related to leaf color and morphology. TOMJPE6398 developed pale green leaves, and TOMJPE6352 and TOMJPE6472 developed variegated leaves. TOMJPE5236 and TOMJPE6397 produced lesions such as pointed leaves. TOMJPE5278 produced wrinkled leaves, and TOMJPE6380 produced abaxially curled, small leaves. Additionally, mutants with large serrated leaves (TOMJPE6586), rounded margin leaves (TOMJPE6653), glossy leaves (TOMJPE8506 and TOMJPG1450) and adaxially curled leaves (TOMJPG2156) were isolated.

**Flower mutants**

Supplementary Fig. S3 shows representative mutants related to flower color and morphology. For example, TOMJPE6468 produced red anthers. Additionally, mutants with a short calyx (TOMJPE6371), opened anthers (TOMJPE6425), adaxially curled petals (TOMJPE6505), very narrow petals (TOMJPE6360 and TOMJPE8910) and twisted petals (TOMJPE8534) were isolated.

**Fruit mutants**

Fig. 5 shows several representative mutants that were defective in normal fruit development. TOMJPG1949 bore fruits with projections, while TOMJPG1236 and TOMJPG1700 produced fruits with scarred surfaces. TOMJPG2075 and TOMJPE8505 produced chapped fruits, while TOMJPG2082 set fruits with spots on the epidermis. Moreover, mutants with differentiation of flower- or fruit-like structures in succession from the fruits were screened (TOMJPG1572, TOMJPG3104 and TOMJPE8533).

**Mutants related to plant size and habit**

Supplementary Fig. S4 shows representative mutants related to plant size and habits. For example, extremely small mutants were obtained (TOMJPE6517). Additionally, long internode mutants (TOMJPG1890 and TOMJPG2385) and short internode mutants (TOMJPE6646) were found. Furthermore, mutants in which the lateral shoots grow thickly (TOMJPE8949) and mutants in which short and curled trichomes were formed (TOMJPE6574) were isolated. TOMJPG1366, TOMJPG2574 and TOMJPG4093 showed deformed overall plant development.

**Phenotype verification**

To test the reproducibility of phenotyping in the M2 plants, we examined whether the mutant phenotypes were inherited in
the M3 plants. In this experiment, 202 individual M3 mutant plants were re-evaluated for visible variation. This analysis indicated that 122 of these 202 M3 plants were perfectly reproducible phenotypically, while 17 mutants did not perfectly segregate for the described phenotypes. In contrast, 63 mutants did not show any described phenotypes. Thus, re-examination of the mutant phenotypes seems to be necessary to verify the mutant phenotypes. Moreover, among the mutants in which the M3 phenotype was verified, 18 individual mutants were backcrossed to wild-type (WT) plants, and their genetic segregations in BC1F2 populations were examined. This genetic analysis indicated that the distribution of phenotypic categories in the examined mutants was stably inherited in the BC1F2 plants (Table 2), confirming the successful inheritance of the mutations. Most of the tested mutants appeared to carry monogenic recessive mutations (Table 2), except for TOMJPG2637, which most probably carried a dominant or a semi dominant mutation (Supplementary Table S3).

Next, to examine the level of saturation in our mutant populations, an allelism test was performed. In this analysis, the presence of multiple alleles per locus was investigated by crossing mutants sharing similar phenotypes. We chose five pairs of mutants sharing similar phenotypes and found that these mutant pairs were allelic, since the F1 plants from the mutant pairs showed the same mutant phenotypes (data not shown), confirming the presence of multiple alleles per locus.

**Isolation and characterization of new iaa9 mutant alleles**

To confirm further the level of saturation, we investigated whether multiple alleles per locus were recovered from our mutant populations by a forward genetic approach. In this experiment, we isolated two EMS-derived mutants (TOMJPE2811 and TOMJPE5405) and one γ-ray-derived mutant (TOMJPG0114) showing similar phenotypes. These three mutants produced leaves with reduced complexity and only one pair of lobed major lateral leaflets merged with the terminal leaflet, and all of these phenotypes are known to be caused by mutation or down-regulation in the entire/IAA9 locus (Wang et al. 2005, Wang et al. 2009), whereas WT tomato leaves are typically unipinnately compound leaves with several pairs of lobed major lateral leaflets (Supplementary Fig. S1A). IAA9 is a member of the Aux/IAA gene family and acts as a transcriptional repressor of the signaling pathway of the plant hormone auxin (Guilfoyle and Hagen 2007). Members of the Aux/IAA family are subject to proteolysis upon auxin treatment, while they bind to families of auxin response factor (ARF) transcription factors in the absence of auxin treatment, repressing auxin biological responses. Thus, the disappearance of Aux/IAA9 stimulates ARF binding to downstream genes, promoting auxin responses. Therefore, it is expected that disruption of IAA9 will trigger ubiquitous auxin responses.

Backcross experiments demonstrated that the genetic traits of TOMJPE5405 and TOMJPE0114 were monogenic recessive mutations (Table 2), which is consistent with the fact that the entire/iaa9 locus is controlled by monogenic recessive mutations (Wang et al. 2005). To examine whether any mutations were present in the IAA9 gene, the whole genomic DNA region containing the IAA9 coding region was amplified by PCR, and their sequences were determined in WT and three mutant plants. Pfam motif analysis (http://pfam.sanger.ac.uk/) revealed that the IAA9 protein has two functional motifs: an IAA protein motif (amino acids 67–341) and an IAA–ARF dimerization motif (amino acids 220–332). IAA9 contains four highly conserved domains (I, II, III and IV). All of these motifs and domains are considered important regions for full activity of the IAA9 protein as a transcription repressor (Wang et al. 2005). To date, two iaa9 mutant alleles (iaa9-1 and iaa9-2) have been isolated in the M82 and Ailsa Craig backgrounds (Zhang et al. 2007).

A single DNA substitution causing a T to A substitution was found at the 237th nucleotide when the translational start was represented as 1 in TOMJPE2811 (iaa9-3, Fig. 6A). This substitution caused the production of a stop codon, resulting in a polypeptide of only 79 amino acids (Fig. 6C). A single DNA substitution causing a G to A substitution was found at the 951st nucleotide position in TOMJPE5405 (iaa9-4). This substitution did not cause an amino acid substitution, but instead produced a truncated mRNA transcript with a 63 bp deletion in the fifth exon, most probably because the substitution was present at an exon–intron junction (Fig. 6B). This truncated mRNA transcript gave rise to a polypeptide with a loss of 21 amino acids in the fifth exon (amino acids 296–317), producing a 328 amino acid polypeptide (Fig. 6C). This deletion caused a partial loss of the IAA–ARF dimerization domain. Additionally, a 32 bp deletion was found in TOMJPC0114 (iaa9-5) at the 133rd nucleotide position, causing a frameshift and producing a stop codon at the 153rd nucleotide position, resulting in the

<table>
<thead>
<tr>
<th>Strain ID</th>
<th>BC1F2 segregation*</th>
<th>BC1F2 segregation*</th>
<th>χ² value*</th>
<th>Genetic trait*</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOMJPG0114</td>
<td>6:0</td>
<td>15:4</td>
<td>0.158</td>
<td>Monogenic recessive</td>
</tr>
<tr>
<td>TOMJPG5405</td>
<td>8:0</td>
<td>30:9</td>
<td>0.077</td>
<td>Monogenic recessive</td>
</tr>
<tr>
<td>TOMJPG0236</td>
<td>10:0</td>
<td>112:33</td>
<td>0.389</td>
<td>Monogenic recessive</td>
</tr>
<tr>
<td>TOMJPG789</td>
<td>7:0</td>
<td>91:26</td>
<td>0.481</td>
<td>Monogenic recessive</td>
</tr>
<tr>
<td>L170</td>
<td>4:0</td>
<td>103:22</td>
<td>3.651</td>
<td>Monogenic recessive</td>
</tr>
<tr>
<td>TOMJPG1949</td>
<td>4:0</td>
<td>17:4</td>
<td>0.397</td>
<td>Monogenic recessive</td>
</tr>
<tr>
<td>TOMJPG2547</td>
<td>7:0</td>
<td>120:35</td>
<td>0.483</td>
<td>Monogenic recessive</td>
</tr>
<tr>
<td>TOMJPG3104</td>
<td>6:0</td>
<td>14:4</td>
<td>0.074</td>
<td>Monogenic recessive</td>
</tr>
<tr>
<td>L33</td>
<td>2:0</td>
<td>25:7</td>
<td>0.166</td>
<td>Monogenic recessive</td>
</tr>
</tbody>
</table>

* The number of progeny exhibiting wild-type (WT) or mutant phenotypes from BC1F2 populations is shown.

* The number of progeny exhibiting WT or mutant phenotypes from BC1F2 populations is shown. WT and ‘mutant’ in parentheses indicate WT-like and mutant-like phenotypes, respectively.

* The values of χ² were calculated from the progeny in the BC1F2 populations.

* Genetic traits were estimated based on the χ² values.
loss of 305 functional amino acids and producing seven truncated polypeptides. These results suggested that the functional activity of the \textit{IAA9} was modified by DNA substitution or deletion in these mutants.

Reduced mRNA accumulation of the \textit{IAA9} gene triggers fruit development prior to pollination, or parthenocarpy, and the decreased level of \textit{IAA9} mRNA accumulation is likely to be associated with the increased level of severity of leaf morphology and parthenocarpic phenotypes (Wang et al. 2005). The severity of leaf morphology was unchanged among the three observed mutant alleles (Fig. 6D). To examine the levels of parthenocarpic fruit development in three \textit{iaa9} mutants. As a control, a pollinated WT plant is shown. DAE, day after emasculation.

Fig. 6 Isolation and characterization of new \textit{iaa9} mutant alleles. (A) Genomic structure of \textit{IAA9}, and the mutations found in three mutant alleles (\textit{iaa9-3}, \textit{iaa9-4} and \textit{iaa9-5}). Gray boxes indicate exons, and regions between the boxes indicate introns. The location and nature of mutations and are shown by arrows. A bent arrow indicates the translational start. (B) Structure of \textit{IAA9} mRNA transcripts in the WT and the \textit{iaa9-4} mutant. Numbers represent numbering of exons. (C) \textit{IAA9} protein structure in the WT and three \textit{iaa9} mutants. Numbers within exons represent the length of polypeptides. The \textit{iaa9-5} protein is truncated by seven amino acids, resulting in a 51 polypeptide length. (D) Leaf morphology of the WT and three \textit{iaa9} mutants in a 30-day-old plant. Bar = 10 cm. (E) Rate of parthenocarpic fruit production in the WT and three \textit{iaa9} mutants after emasculation. A t-test was used to determine a statistically significant difference. Different lower case letters represent different classes. The standard error (SE) is shown. A pollinated WT plant is shown \([WT(p)]\) as a control, \(n = 10\). (F) Time course analysis of parthenocarpic fruit development in three \textit{iaa9} mutants. As a control, a pollinated WT plant is shown. DAE, day after emasculation.
iaa9-5 and WT(p). These results suggested that iaa9-4 was a weaker mutant allele than iaa9-3 and iaa9-5. The delayed fruit expansion in the emasculated iaa9-4 plants resulted in delayed fruit maturation: parthenocarpic fruit ripening took >42 DAE in WT(p), iaa9-3 and iaa9-5 mutant plants, whereas it took >50 DAE in the iaa9-4 mutant (Fig. 6F). To compare the severity of phenotypes in ripe red fruits, the length and width of ripe red parthenocarpic fruits were determined. As shown in Supplementary Fig. S6A–C, there was no significant difference in size among the WT(p) and mutant plants. Additionally, parthenocarpic fruit weight was similar in the WT(p) and in the mutants (Supplementary Fig. S6D). However, a significant difference was found in the thickness of the pericarp: the pericarp thickness of parthenocarpic fruits in iaa9-3, iaa9-4 and iaa9-5 was 4.0 ± 0.1, 3.8 ± 0.1 and 3.9 ± 0.1 mm, respectively, while in the WT(p), it was 2.9 ± 0.1 mm, indicating that disruption of the IAA9 gene influenced pericarp development (Supplementary Fig. S6E, F). These results suggested that IAA9 was essential for regulating the precise timing triggering early fruit development, while the other factors probably determined final fruit size and weight. IAA9 was also important for normal pericarp development, although the thickness was not dependent on the severity of mutant alleles.

**Tomato mutant database TOMATOMA**

All of the visible phenotyping data, mutant images and other cultivation information (germination rate, mutation rate, first flowering date, first fruit harvesting date, flower to fruit period and total harvested seeds) were registered in a relational database, TOMATOMA (Fig. 7; http://tomatoma.nbrp.jp/index.jsp). This database can be searched according to strain ID or freely searched according to keywords from the detailed description of phenotypes, 15 major phenotypic categories and 48 subcategories (Fig. 7A). Detailed data, including the germination ratio, rate of mutant appearance in the M3 family and first flowering date, as well as photographs of the mutants, are shown for the page of each mutant (Fig. 7C). This database is publicly accessible through its web interface, and the mutant seeds are being distributed upon completion of an MTA contract between users and the University of Tsukuba. In addition to the mutant collections, TOMATOMA provides the tomato varieties and wild tomato relatives that are commonly used for experimental studies (Supplementary Table S5). Moreover, TOMATOMA provides M3 mutagenized seeds (up to 10 seeds per M3 family), and currently (November 2010) 2,236 EMS-mutagenized and 2,700 γ-ray-irradiated lines are available upon request. The populations of M3 seeds are also useful for conditional screening for mutants that survive under a wide range of different environmental conditions.

**Discussion**

For optimization of EMS mutagenesis, it is important to assess parameters related to the efficient generation of mutant collections. To examine this, Micro-Tom seeds were treated with different EMS concentrations (0.3, 0.5, 0.7, 1.0 and 1.5%) and for different incubation times (12, 16 and 48 h), and the rates of germination, fertility and mutant recovery were calculated (Supplementary Tables S1, S2). Using a low EMS concentration with a long incubation time did not yield a significantly higher rate of mutant recovery (EMS-6, 11.5%) compared with the same EMS concentration with a shorter incubation time (EMS-4, 13.6%; EMS-5, 15.0%), suggesting that incubation duration was not the critical parameter. However, the highest EMS concentration used (1.5%) yielded the highest mutant recovery rate (EMS-10, 30%), while an extremely low germination rate was a bottleneck in this condition (Supplementary Tables S1, S2). For the 1.0% EMS concentration, a high rate of mutant recovery was reproducibly observed (i.e. EMS-7, 18.0% and EMS-8, 17.9%), representing valid conditions for the efficient production of mutant populations. Consistently, 1.0% EMS treatment was regarded as better than 0.7% EMS treatment in terms of providing efficient mutant isolation from the mutagenesis populations in the Red Setter (Minoia et al. 2010). Additionally, the rate of mutant recovery was apparently higher for EMS mutagenesis than for γ-ray irradiation (Fig. 3, Table 1), which could be due in part to the fact that γ-ray irradiation often causes large DNA deletions, resulting in increased infertility, gametophytic lethality and decreased germination capacity (Morita et al. 2009) (Fig. 4, Table 1). Thus, it is possible that the majority of mutants with severe DNA deletions did not survive.

In this study, comprehensive mutant populations were developed in the genetic background Micro-Tom. Construction of the mutant populations took >7 years, during which we successively inspected plants for alteration of visible mutant phenotypes based on 15 major phenotypic categories and 48 subcategories (Table 1, Supplementary Tables S1, S2). To our knowledge, TOMATOMA is the third interactive database to become accessible to the scientific community, providing individual mutant seeds in the same genetic background, following the development of ‘Genes that Make Tomatoes’ in M82 and ‘LycoTILL’ in Red Setter (Menda et al. 2004, Minoia et al. 2010). Although the total number of mutants (1,048 mutants) isolated from our mutagenesis approach was less than in the previously developed populations in the M82 genetic background (3,417 mutants; Menda et al., 2004), our mutant resources could be more manageable because the Micro-Tom has many advantages in its cultivation and biological features (i.e. rapid life cycle and small size; Fig. 1).

The population size necessary for conducting tomato saturation mutagenesis is estimated to be in the range of 10^9 individuals in the case of EMS or fast-neutron irradiation (Emmanuel and Levy 2002), suggesting that our populations were nearly saturated (Table 1, Supplementary Table S2). Additionally, allelism tests and the forward genetics approach demonstrated the presence of multiple alleles per locus (Fig. 6). The identification of multiple mutant alleles facilitates gene function analysis, since allelic series often show a wide range
of phenotypes associated with a single gene, making it possible to dissect the correlation between gene functional activity and given phenotypes. For example, it is probable that functional IAA9 protein activity is correlated with the degree of parthenocarpic phenotypes (Fig. 6).

The reproducibility of phenotypic categories was also confirmed by performing visible phenotyping of the M3 plants, and it was found that 122 of 202 M3 plants were phenotypically reproducible. Furthermore, 18 mutants were backcrossed to WT plants, and their successful phenotypic inheritance in BC1F2 populations was confirmed (Table 2), indicating that these mutants carried stable mutations. Additionally, nine other mutants probably carried recessive mutations (Supplementary Table S3). The confirmed genetic quality of these mutations indicates that they are sufficiently stable to be useful for functional analysis.

While the isolation of candidate genes from Micro-Tom mutants by a forward genetics approach could become a
main subject in future studies, intraspecific and interspecific crosses and resulting F1 hybrid production are possible between Micro-Tom and most tomato cultivated varieties, as well as many members of the Solanum genus, including S. pimpinellifolium and S. pennellii, both of whose draft sequences are available at SGN. Additionally, high-density molecular maps in tomatoes have been generated based on the F2 progeny of the cultivated tomato (S. lycopersicum) and many of its wild relatives (Foolad 2007), all of which are available at SGN. Furthermore, introgression line (IL) populations in which S. lycopersicum is used as the core background carrying a single introgressed chromosome segment from S. pennellii have been constructed (Eshed and Zamir 1994). This information and these ILs are useful for constructing F2 mapping populations for fine-mapping of mutants, since they provide a nearly isogenic, highly polymorphic resource for mutants (Menda et al. 2004). These ILs, in addition to S. pimpinellifolium and S. pennellii, are available at TGRC or TOMATOMA (Supplementary Table S4).

While positional cloning requires comprehensive DNA markers that are polymorphic between parental plants, large numbers of genetic markers have been developed based on the F2 progeny of interspecific crosses between S. lycopersicum and its wild relatives. SGN presents information about these DNA markers, including restriction fragment length polymorphism (RFLPs), simple sequence repeats (SSRs), cleavage amplified polymorphic sequence (CAPS), amplified fragment length polymorphism (AFLP), random amplified polymorphic DNA (RAPD), sequence characterized amplified region (SCAR) and single nucleotide polymorphism (SNP) data. Most of these markers have been mapped on the linkage map Tomato Expen-2000, which is currently composed of a total of 2,604 DNA markers, which are freely available at SGN (Fulton et al. 2002). Not only an interspecific but also an intraspecific genetic map has been reported for S. lycopersicum based on AFLP, RFLP and RAPD markers (Saliba-Colombani et al. 2000). Recently, an intraspecific genetic linkage map was constructed derived from crosses between Micro-Tom and either Ailsa Craig or M82, based on the SNP markers discovered from EST sequences (Shirasawa et al. 2010). It is likely that >600 SNP markers can be utilized for genotyping analysis in F2 progeny derived from Micro-Tom and other cultivated varieties, including Ailsa Craig or M82, and these genetic maps appear to cover most regions of the tomato genome, providing considerable possibilities for performing positional cloning.

To explore the possible use of mutagenized lines through a reverse genetics approach, we are developing a TILLING platform using the EMS-mutagenized populations generated in this study. By constructing systematic platforms combining EMS-mutagenized resources and TILLING, we expect to obtain desirable mutants through large-scale screening, as well as multiple alleles per locus, representing ideal resources for examining the effects of mutations on specific traits in conjunction with the progress of the tomato sequencing project. For improving the quality and yield of tomatoes, a deep understanding of the molecular bases underlying these complex traits is necessary. This challenge requires comprehensive and genetically high-quality populations of mutants as well as the availability of these resources at the research-community level to promote functional analyses of tomatoes. Therefore, our Micro-Tom mutant collections, which are available to the community, could play considerable roles in elucidating key mechanisms controlling important tomato traits.

### Materials and Methods

#### Production of mutant collections

A flow chart showing the process from mutagenesis to the establishment of mutant lines is shown in Fig. 2 and Supplementary Tables S1 and S2. EMS treatments carried out from 2004 to 2006 and γ-ray irradiation were described previously (Matsukura et al. 2007, Watanabe et al. 2007). In 2008 and 2009, three different doses (0.7, 1.0 and 1.5%) of EMS (Sigma-Aldrich) were applied to Micro-Tom seeds. For the EMS treatments, Micro-Tom seeds (M0 seeds) were soaked in distilled water for 8 h at room temperature, followed by incubation in 100 ml of a fleshy prepared EMS solution (for each batch of 1,000–3,000 seeds) under gentle stirring for 16 h. The EMS solution was removed, and the mutagenized seeds were washed by gentle shaking in 100 ml of distilled water for 4 h. This washing step was repeated three times. The M1 seeds were sown, and M1 plants were cultivated in a greenhouse to produce the M2 seeds. Then, approximately 10 M2 plants from the same M1 plant were grown in nursery pots (145 cm3 per pot) as a family, and whole M3 offspring seeds were harvested from the same family in bulk. Additionally, self-fertilized seeds were harvested from each independent mutant.

Each M2 plant was phenotyped morphologically according to 15 major categories and 48 subcategories defined previously (Menda et al. 2004). For each mutant, their visible phenotypes, images, germination rate, flowering time, source of mutagen and ontology were recorded in the database. Although all the mutant seeds that are listed in TOMATOMA will be distributed upon request, we may stop providing mutant lines that have a very limited number of seeds. These mutant seeds will again become available once sufficient numbers of seeds are produced.

#### Sequencing analysis

Genomic DNA was extracted from 3-week-old seedling plants using a DNeasy Mini extraction kit (QIAGEN). The whole IAA9 coding region was amplified using IAA9F and IAA9R primers (Supplementary Table S5), and PCR products were purified using a gel DNA extraction kit (QIAGEN). Subsequently, the genomic sequence of IAA9 was determined using extracted DNA and gene-specific primers, as listed in Supplementary Table S5.
Total RNA was extracted from mature leaves of 30-day-old plants using an RNaseasy Mini kit (QIAGEN). Genomic DNA contamination was removed using the RNase-free DNase Set (QIAGEN). cDNA was generated from 1μg of total RNA using the SuperScript III First-Strand Synthesis System (Invitrogen). cDNA was then used as a template for reverse transcription–PCR (RT–PCR) with specific primers to amplify the IAA9 coding region (Supplementary Table S5).

Emasculation assay

Seeds of WT Micro-Tom, and M1 homozygous seeds of TOMJPE2811 (iaa9-3), TOMJPE5405 (iaa9-4) and TOMJPG0114 (iaa9-5) were sown on wet filter paper under constant light for 48 h at 25°C to stimulate seed germination. Then 7-day-old seedlings were transplanted into soil and grown under conditions of 16 h light/8 h dark at 25°C. Emasculation was performed using 1-month-old plants before day 1 of flowering. All anthers were removed from closed flowers using forceps, and all plants were grown under the same conditions as described above. The number of fruits that developed from emasculated flowers was counted at 10 DAE. The time course of parthenocarpic fruit development was also collected as indicated in Fig. 6. As a control, open flowers of WT plants were vibrated for 10 s per day for 3 d to stimulate pollination.

Database system

TOMATOMA is a PostgreSQL-driven relational database. The web application is written using Java and Jsp. The software and the database are hosted on the Linux server [3 GHz Intel (R) Xeno (R) CPU S160, Redhat Linux ES4].

Supplementary data

Supplementary data are available at PCP online.

Funding

This work is supported by the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) [part of ‘The National BioResource Project’ (ADD20035)].

Acknowledgments

We thank T. Watanabe and Y. Takahashi at the National Institute of Genetics for assisting in the construction of the TOMATOMA database. We also thank all the people who helped in developing the mutant populations. We are grateful to Tsutomu Arie for providing several lines of S. lycopersicum, S. peruvianum, S. chilense and S. pimpinellifolium. We are also grateful to Dani Zamir for providing the ILs. We would like to thank all members of our lab for fruitful discussion on this project.

References


