DNA marker applications to molecular genetics and genomics in tomato

Kenta Shirasawa and Hideki Hirakawa*

Department of Plant Genome Research, Kazusa DNA Research Institute, 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan

Tomato is an important crop and regarded as an experimental model of the Solanaceae family and of fruiting plants in general. To enhance breeding efficiency and advance the field of genetics, tomato has been subjected to DNA marker studies as one of the earliest targets in plants. The developed DNA markers have been applied to the construction of genetic linkage maps and the resultant maps have contributed to quantitative trait locus (QTL) and gene mappings for agronomically important traits, as well as to comparative genomics of Solanaceae. The recently released whole genome sequences of tomato enable us to develop large numbers of DNA markers comparatively easily, and even promote new genotyping methods without DNA markers. In addition, databases for genomes, DNA markers, genetic linkage maps and other omics data, e.g., transcriptome, proteome, metabolome and phenome information, will provide useful information for molecular breeding in tomatoes. The use of DNA marker technologies in conjunction with new breeding techniques will promise to advance tomato breeding.

Key Words: DNA marker, genetic linkage map, genome, database, protein structure, tomato (*Solanum lycopersicum*).

Introduction

DNA markers have promoted genetics, genomics and breeding in a wide range of plant species, including tomato, through their use in the construction of high-density linkage maps, which are a useful tool for marker-assisted selection, association analysis and QTL analysis. However, the development of sufficient numbers of DNA markers to saturate the linkage maps has proven costly in terms of time, labor and financial resources, since the DNA markers have mainly been developed from randomly selected clones of genomic and cDNA libraries or PCR with random primers. Moreover, the construction of large-insert genomic libraries, the use of so-called chromosome walking to cover the candidate genomic regions and the sequencing of the selected clones have also required the use of map-based cloning.

The genome sequences of *Arabidopsis thaliana* (The Arabidopsis Genome Initiative 2000) and rice (International Rice Genome Sequencing Project 2005) have greatly assisted in the production of a large number of DNA markers for marker-assisted selection in the breeding of vegetable and cereal crops (Varshney et al. 2005). Because the nucleotide sequences as well as the gene order in the genomes are generally conserved between the model plants and crops (Tang et al. 2008), genomes of the model plants could allow us to estimate the number and the variation of genes in particular plant species. Recently, the genome sequences of various crops have been analyzed using a next-generation sequencers (NGS), a GS FLX+ System (Roche, Basel, Switzerland), HiSeq2000 (Illumina, San Diego, USA), or a SOLiD 5500xl system (Life Technologies, USA), which can produce up to 700 Mb, 600 Gb and 180 Gb of sequence data in a single experiment, respectively: apple (Velasco et al. 2010), banana (D’Hont et al. 2012), cacao (Argout et al. 2011), Chinese cabbage (*The Brassica rapa* Genome Sequencing Project Consortium 2011), cucumber (Huang et al. 2009), grape (The French–Italian Public Consortium for Grapevine Genome Characterization 2007), maize (Schnable et al. 2009), melon (Garcia-Mas et al. 2012), papaya (Ming et al. 2008), pigeonpea (Varshney et al. 2012), potato (*The Potato Genome Sequencing Consortium 2011*), sorghum (Paterson et al. 2009), soybean (Schmutz et al. 2010) and strawberry (Shulaev et al. 2011). Tomato is one of the major vegetable crops and regarded as a model for fruiting plant and Solanaceae relatives, and the genome sequencing has recently been completed using both NGSs and a fluorescent capillary sequencer with the Sanger method (The Tomato Genome Consortium 2012).

Genome-wide DNA polymorphism information can be obtained relatively easily by using NGSs for genome-scale genetic analyses such as genome-based breeding and genome-wide association studies. In this review, we...
summarize the studies of DNA markers developed for the genetics and molecular breeding in tomato and their applications, e.g., genetic linkage map, QTL and gene mappings, comparative genomics and functional annotations of DNA polymorphism. In addition, we introduce the databases for tomato genomics and genetics, and finally describe future perspectives of tomato breeding using the advanced DNA markers and genotyping technologies.

**DNA markers and genotyping platforms in tomato**

Molecular genetics based on DNA markers in tomato plants began with 57 restriction fragment length polymorphisms (RFLPs) (Bernatzky and Tanksley 1986). Since then the number of RFLP markers has increased to approximately 1000 for use in tomato genetics (Tanksley et al. 1992). However, a large amount of DNA as well as much expenditure of time and labor are required in the RFLP analysis based on the Southern-blotting method. Therefore, the RFLP markers have been replaced by PCR-based cleaved amplified polymorphic sequence (CAPS) markers (http://solgenomics.net), which are more convenient to handle than RFLP markers because their use requires less DNA and simpler laboratory experiments. DNA fingerprinting techniques, e.g., random amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP), are also conducted to develop DNA markers in tomato, because no sequence information is required, and because of the high polymorphism ratio due to multi-locus detection by single marker analysis (Saliba-Colombani et al. 2000).

Along with advances in genomic studies in plants, large amounts of sequence information, e.g., >200,000 expressed sequence tags (ESTs) and approximately 90,000 bacterial artificial chromosome (BAC)-ends, have been released for tomato species. Simple sequence repeat (SSR) markers can be rapidly and easily developed by using the sequence information derived from computational SSR-motif searches and primer designs for their flanking sequences (Fukuoka et al. 2005). SSR markers have advantages over the RFLP, CAPS, RAPD and AFLP markers due to multi-allelic detection, high-transferability across species, available tagged sequences and flexibility with various laboratory systems such as gel and capillary electrophoreses. In particular, capillary electrophoresis with a fluorescent fragment analyzer has major advantages for automatic analysis (allowing >2,000 samples/day in one analyzer) and for high-resolution analysis (distinguishing differences of only 1-bp length). Therefore, more than 20,000 SSR markers have been developed from EST and BAC-end sequences and used as genetic and genomic tools in tomato species (Ohyama et al. 2009, Shirasawa et al. 2010a).

Because the tomato genome sequence has been released (The Tomato Genome Consortium 2012), single nucleotide polymorphism (SNP), which is the most abundant polymorphism in genome in general, has been discovered by a re-sequencing strategy. In the re-sequencing strategy, the sequences obtained from the whole-genome, from complexity-reduced genomes such as the restriction site associated DNA (RAD), or from transcribed sequences are remapped onto the reference genome or unigenes with mapping software, e.g., Bowtie2 (Langmead et al. 2012) or BWA (Li and Durbin 2009) for the HiSeq2000 (Illumina Inc., San Diego, USA) and the GS reference mapper (Roche Applied Science, Mannheim, Germany) or MIRA (Chevreux et al. 2004) for the 454 GS FLX+ system (Roche Applied Science). By comparing the cDNA sequences, between 2,000 and 63,000 SNP candidates have been found in several studies (Hamilton et al. 2012, Jimenez-Gomez and Maloof 2009, Labate and Baldo 2005, Shirasawa et al. 2010b, Yamamoto et al. 2005, Yang et al. 2004). However, these candidates found by computational approaches frequently contain false positives due to the errors in the processes of sequencing or mapping, and thus it is necessary to select an accurate SNP site deeply covered with high-quality fragments on both strands to eliminate the false positives.

On the other hand, several SNP genotyping methodologies have been developed for application to various objectives. Depending on the purpose and degree of the throughput of the SNP for analysis, genotyping methods can be selected as follows. A huge number of SNPs in a small number of samples can be detected by the re-sequencing strategy using NGs. Conversely, a small number of SNPs in a large number of samples, e.g., marker-assisted selection and cultivar identification, can be detected by TaqMan assay (Life Technologies), dot-blot SNP analysis (Shiokai et al. 2010, Shirasawa et al. 2006), the Tm-shift genotyping method (Fukuoka et al. 2008) and high-resolution melting analysis (Shirasawa et al. 2010a), because none of these methods require electrophoresis. In addition, a high-throughput SNP analysis in a large number of samples can be performed effectively by array-based assays as genotyping platforms, GoldenGate and Infinium (Illumina Inc., San Diego, USA) and applied to the construction of high-density genetic linkage maps and performance of genome-wide association studies (Hamilton et al. 2012, Hirakawa et al. 2013, Shirasawa et al. 2010b, Sim et al. 2012). The diversity arrays technology (DARt) platform, which is one of the other array-based methods, has been applied to develop bin-mapped polymorphic markers across the introgression lines (ILs) population of tomatoes (Van Schalkwyk et al. 2012).

**DNA markers for genetics, genomics and molecular breeding in tomato**

**Genetic linkage maps**

In tomato, the first genetic linkage map was constructed with mainly RFLP markers for an interspecific population derived from a cross between *S. lycopersicum* and *S. pennellii* (Bernatzky and Tanksley 1986). This map consists of 112 RFLP and isozyme loci and covers ca. 760 cm (Table 1). Then, several interspecific genetic linkage maps were generated with RFLPs incorporating CAPS, SSR and SNP.
Table 1. Genetic linkage maps developed in tomato and its relatives

| Cross combinations                  | Population types | No. of marker loci | Marker types          | Map length (cM) | References                                      |
|------------------------------------|------------------|--------------------|-----------------------|-----------------|------------------------------------------------|
| **Interspecies**                   |                  |                    |                       |                 |                                                 |
| S. lycopersicum ‘LA1500’ × S. pennelli ‘LA716’ | F₂ (n = 46)      | 112                | RFLP, Isozyme          | 760             | Bernatzky and Tanksley (1986)                   |
| S. lycopersicum ‘VF36-7m2a’ × S. pennelli ‘LA716’ | F₂ (n = 67)      | 1030               | RFLP, Isozyme          | 1276            | Tanksley et al. (1992)                         |
| S. lycopersicum ‘E6203’ × S. hirsutum ‘LA1777’ | BC₁ (n = 149)    | 135                | RFLP                  | 1356            | Bernacchi and Tanksley (1997)                  |
| S. lycopersicum ‘Sun 1642’ × S. pimpinellifolium ‘LA1589’ | F₂ (n = 100)     | 108                | RFLP                  | 1174            | van der Knaap and Tanksley (2001)              |
| S. lycopersicum ‘E6203’ × S. pimpinellifolium ‘LA1589’ | BC₂F₄ (n = 170)  | 127                | RFLP                  | 1282            | Doganlar et al. (2002a)                        |
| S. lycopersicum ‘Yellow Stuffer’ × S. pimpinellifolium ‘LA1589’ | F₂ (n = 200)     | 93                 | RFLP                  | 1076            | van der Knaap and Tanksley (2003)              |
| S. lycopersicum ‘LE777’ × S. chmielewskii ‘CH6047’ | F₂ (n = 149)     | 255                | AFLP, CAPS, SCAR, SSR  | 887             | Jimenez-Gomez et al. (2007)                    |
| S. lycopersicum ‘Rio Grande’ × S. pimpinellifolium ‘LA1589’ | F₂ (n = 94)      | 97                 | CAPS, RFLP, SSR       | 1174            | Gonzalo and van der Knaap (2008)               |
| S. lycopersicum ‘Sausage’ × S. pimpinellifolium ‘LA1589’ | F₂ (n = 106)     | 96                 | CAPS, RFLP, SSR       | 1072            | Gonzalo and van der Knaap (2008)               |
| S. lycopersicum ‘LA925’ × S. pennelli ‘LA716’ | F₂ (n = 83)      | 2116               | SSR, SNP              | 1503            | Shirasawa et al. (2010a)                       |
| S. lycopersicum ‘LA925’ × S. pennelli ‘LA716’ | F₂ (n = 79)      | 3503               | SNP                   | 1670            | Sim et al. (2012)                              |
| S. lycopersicum ‘Moneymaker’ × S. pennelli ‘LA716’ | F₂ (n = 160)     | 3687               | SNP                   | 1155            | Sim et al. (2012)                              |
| S. lycopersicum ‘Moneymaker’ × S. pimpinellifolium ‘LA121’ | F₂ (n = 183)     | 4491               | SNP                   | 1049            | Sim et al. (2012)                              |
| **Intraspecies**                   |                  |                    |                       |                 |                                                 |
| S. lycopersicum ‘Leovil’ × S. lycopersicum var. cerasiforme ‘Cervil’ | F₁ (n = 153)     | 377                | AFLP, RAPD, RFLP      | 965             | Saliba-Colombani et al. (2000)                 |
| S. lycopersicum ‘Ailsa Craig’ × S. lycopersicum ‘Micro-Tom’ | F₂ (n = 120)     | 989                | SNP, SSR              | 1468            | Shirasawa et al. (2010b)                       |
| S. lycopersicum ‘M82’ × S. lycopersicum ‘Micro-Tom’ | F₂ (n = 135)     | 637                | SNP                   | 1423            | Shirasawa et al. (2010b)                       |

Markers (Bernacchi and Tanksley 1997, Doganlar et al. 2002a, Fulton et al. 2002, Gonzalo and van der Knaap 2008, Grandillo and Tanksley 1996, Jimenez-Gomez et al. 2007, Shirasawa et al. 2010a, Sim et al. 2012, Tanksley et al. 1992, 1996, van der Knaap and Tanksley 2001, 2003) as summarized in the SOL Genomics Network (SGN) (Mueller et al. 2005, http://solgenomics.net). The numbers of mapped loci ranged from 93 to 4,491 and covered 887 to 1,670 cM (Table 1). An intraspecific map, which is considered more useful for breeding than interspecific maps, has also been constructed with SSR and SNP markers using a population derived from a cross between the tomato cultivars “Micro-Tom” and either “Ailsa Craig” or “M82” (Table 1; Shirasawa et al. 2010b). In addition, a total of 7054 non-redundant SNPs between Micro-Tom and other cultivars have been genotyped by the array technologies (Hamilton et al. 2012, Shirasawa et al. 2010b) and these SNPs were mapped onto the tomato genome (Hirakawa et al. 2013). Currently, sequencing of the whole genome of Micro-Tom is underway (Aoki et al. 2011). These maps, SNPs and genome sequences will provide opportunities for map-based cloning of genes responsible for Micro-Tom-derived mutant lines provided from the National BioResource Project (NBRP) for the tomato (Saito et al. 2011: http://tomatoma.nbrp.jp).

Mapping of QTLs and genes for agronomical traits

In the map-based cloning strategy, to carry out subsequent regional fine mapping following the genome-wide linkage mapping, the introgression lines (ILs) composed of 76 lines, which together cover the entire genome of the donor parent, S. pennelli “LA716,” in the background of the recurrent parent, S. lycopersicum “M82,” have been developed by a method of marker-assisted backcrossings (Eshed and Zamir 1994). The ILs are available from the Tomato Genetics Resource Center (TGRC: http://tgrc.ucdavis.edu) and the NBRP for tomato (Saito et al. 2011, http://tomatoma.nbrp.jp). The developed genetic resources, e.g., genetic linkage maps and introgression lines, have been used for...
identification of agronomically important genes for disease resistance (CF-2: Dixon et al. 1996, CF-9: Jones et al. 1994, I2: Ori et al. 1997, Mi: Milligan et al. 1998, Pto: Martin et al. 1993, Sw-5: Brommonschenkel et al. 2000, Tm-1: Meshi et al. 1988, Tm-2: Meshi et al. 1989, Ve: Kawchuk et al. 2001, and reviewed in Flooad and Panthee 2012), fruit characteristics (Brix9-2-5: Fridman et al. 2000, FAS: Cong et al. 2008, FW2.2: Frary et al. 2000, LC: Munos et al. 2011, OVATE: Liu et al. 2002, SUN: Xiao et al. 2008, U: Powell et al. 2012), hybrid vigor (SFT: Krieger et al. 2010), plant architectures (D: Bishop et al. 1996, SP: Pnueli et al. 1998) and several traits summarized in The Tomato Genome Consortium (2012). Of course, DNA markers as well as the cloned genes themselves could be used for marker-assisted selection in breeding (Labate et al. 2007). Furthermore, associations between genotypes and phenotypes have been revealed: the allele distribution of FAS and LC for fruit locule number and flat shape and OVATE and SUN for elongated shape in tomato cultivars is strongly associated with fruit shape diversity (Fridman et al. 2011); and mutant alleles of D for dwarfism and SP for determinate plant height are observed in dwarf tomato and processing cultivars (Fig. 1).

**Comparative genomics**

Genetic linkage maps contribute to not only QTL and gene mapping but also comparative genomics, which has a significant impact on the fields of plant genetics and genomics. Because tomato is recognized as a representative experimental model of the Solanaceae family, comparative maps have been developed by connecting orthologous markers as anchors between the pairs of the plant species, such as tomato and potato (Tanksley et al. 1992), tomato and eggplant (Doganlar et al. 2002b, Fukuoka et al. 2012, Wu et al. 2009b), tomato and pepper (Livingstone et al. 1999, Prince et al. 1993, Tanksley et al. 1988, Wu et al. 2009a) and tomato and tobacco (Wu et al. 2010) and chromosome segments conserved in Solanaceae have been identified (Wu and Tanksley 2010). These results were confirmed by the comparative analysis of genome sequences between tomato and potato species (The Tomato Genome Consortium 2012). The syntenic relationship among Solanaceae family members will contribute to the transfer of knowledge obtained from studies of tomato to other Solanaceae crops with respect to genetics, genomics and molecular breeding.

**Functional annotation of DNA polymorphisms**

The functional marker that is responsible for gene of protein functions could be effectively applied to molecular breeding such as marker-assisted selection. SNPs can be used as functional markers, because they have the potential to link to the gene functions. The SNPs can be classified into the following groups according to the locations on the genome sequence: cSNPs (SNPs in coding sequences causing amino-acid substitutions), sSNP (SNPs in coding sequences not causing amino-acid substitutions), iSNPs (SNPs in intron regions), rSNPs (SNPs in regulatory regions), uSNPs (SNPs in untranslated regions) and gSNPs (SNPs in intergenic regions). The cSNPs may directly effect protein function, if they are located on a catalytic site, while the uSNPs and rSNPs may effect gene expressions. These SNPs probably effect alternations of gene functions and are used for functional SNPs. In the field of human genomics, the SNP variations have been studied by comparing individual genome sequences (Altshuler et al. 2000, International HapMap 3 Consortium 2010, Li et al. 2009). The HapMap project (International HapMap Consortium 2003) has provided a huge number of genome-wide SNPs collected from several populations through the Single Nucleotide Polymorphism database (dbSNP) at the National Center for Biotechnology Information (NCBI) and the F-SNP (Functional SNP) databases (Lee and Shatkay 2008). In the field of plant genomics, on the other hand, genome-wide functional SNPs...
have not been analyzed sufficiently.

To speculate the functional effects of the eSNPs in the tomato genome, the SNP locations in catalytic sites have been identified by using the three-dimensional structure of proteins constructed by homology modeling (Hirakawa et al. 2013). The binding clefts as catalytic site in proteins can be predicted by the F Pocket (Guilloux et al. 2009) or MetaPocket programs (Zhang et al. 2011). The amino acid residues important for catalytic activity have been predicted by calculating the protein-substrate affinity using binding simulation with the Autodock (Morris et al. 1998) or ASEdock of MOE software package (Kumar et al. 2011). Together with the information on the catalytic sites, the genes having the functional SNPs have been annotated by similarity searches against the KOG (Tatusov et al. 2003), KEGG (Ogata et al. 1999), NR in NCBI (http://www.ncbi.nlm.nih.gov), TAIR10 (Garcia-Hernandez et al. 2002) and PDB (Berman et al. 2000; http://www.pdb.org) databases and domain searches against the Pfam database (Punta et al. 2012). According to these analyses, the genes with the functional SNPs would be more applicable for molecular breeding than DNA markers linking to target genes, because the target genes might be lost by crossing over between the loci for the marker and the target genes in conventional marker-assisted selection (Shirasawa et al. 2004).

**Databases for molecular genetics and genomics in tomato**

The SGN (Mueller et al. 2005; http://solgenomics.net) is recognized as one of the databases for molecular genetics and genomics in Solanaceae, e.g., tomato, potato, pepper, eggplant, tobacco, petunia and so on. This database provides information about not only maps and markers but also mass data for genomes, sequences and expression patterns of genes, metabolite pathways, phenotypes and QTLs. Furthermore, the database provides information on the news, events and publications related to Solanaceae and links to the external databases providing Solanaceae genetic information, e.g., data on the genome, ESTs, markers, QTLs and mutants.

Among the tomato databases for DNA markers, the Solanaceae Coordinated Agricultural Project (SoICAP; http://solcap.msu.edu), which focuses on translating genomic advances to tomato and potato breeding, has provided 62,576 SNPs and experimentally validated data on 96 SNPs for 85 tomato cultivars (Hamilton et al. 2012). The Tomato Mapping Resource Database (http://www.tomatomap.net) releases genotyping data of 52 indel, 102 RFLP, 205 SNP and 94 SSR markers for 102 tomato lines including 9 wild species. In the Tomato SNPs database (http://www.plb.ucdavis.edu/labs/malool/tomatosnp/), 12,568 and 5,004 SNPs detected by *in silico* analyses between *S. lycopersicum* and *S. habrochaites* and between *S. lycopersicum* and *S. pennellii* are available, of which 220 and 196 SNPs have been experimentally verified, respectively. The National Center of Biotechnology Information (NCBI) has established a database for SNPs known as dbSNP (http://www.ncbi.nlm.nih.gov/snp/), from which 376 tomato SNPs have been published. The VegMarks (http://vegmarks.nivot.affrc.go.jp) provides genetic linkage maps and genotyping data for 270 SNPs between *S. lycopersicum* “LA925” and *S. pennellii* “LA716” and 148 SSR markers for 10 lines. The MiBASE (http://www.pgb.kazusa.or.jp/mibase/), a database specific to Micro-Tom resources, also provides 1935 SNP candidates between Micro-Tom and either five lines and 409 EST-SSRs found by *in silico* analyses.

A portal website for tomato genomics, Kazusa Tomato Genomics Database (KaTomicsDB: http://www.kazusa.or.jp/tomato/) consisting of the following two databases, has been released. The first is the Tomato Marker Database (http://marker.kazusa.or.jp/tomato/), which mainly provides information on 8,297 SNP and 21,100 SSR markers, i.e., primer sequences and DNA fragments including marker loci, genetic linkage maps of the provided DNA markers and genotyping data of the SNPs for 42 lines (Hirakawa et al. 2013, Shirasawa et al. 2010a, 2010b). Moreover, most of the markers have been mapped on the tomato genome by similarity searches, and ordered with the predicted genes. The second is the Tomato Functional SNP Database (http://plant1.kazusa.or.jp/tomato/), which provides the genes with SNPs annotated by similarity searches against the databases of KOG (Tatusov et al. 2003), KEGG (Ogata et al. 1999), NR in NCBI (http://www.ncbi.nlm.nih.gov), TAIR10 (Garcia-Hernandez et al. 2002) and PDB (Berman et al. 2000; http://www.pdb.org). In addition, the web site allows visitors to browse the locations of SNPs on the three-dimensional structure built by homology modeling.

In addition to the databases described above, various databases for tomato genetics and genomics have been released (Table 2). In Table 2, these databases are roughly classified into seven categories, i.e., genome, DNA marker, ESTs, gene expression, metabolome, plant materials and portal sites. To integrate these databases, the Plant Genome DataBase Japan (PGDBj: http://pgdbj.jp) has been established through the National Bioscience Database Center (NBDC), Japan. In this database, over 50 plant species, including crops, fruits, trees and vegetables, have been registered. This database provides genome maps integrating the DNA markers, genetic linkage maps and QTLs collected from the databases and related articles.

**Future perspectives of DNA markers in tomato genetics, genomics and breeding**

The whole genome sequence of tomato has recently been released (Tomato Genome Consortium 2012). In this project, the genome sequences were linked to the 12 tomato chromosomes with two BAC-based physical maps and anchored/ oriented using a high-density genetic map, introgression line mapping and BAC fluorescence *in situ* hybridization (FISH). As the genetic map, an interspecific map, Tomato-EXPEN 2000 (Fulton et al. 2002), was selected because of
Shirasawa and Hirakawa

the highest-density SSR, CAPS and RFLP marker loci (2,116 loci covering 1,503 cM) (Shirasawa et al. 2010a). Taking these results together, the consortium has released 12 pseudomolecules consisting of 760 Mb of the predicted genome size of 960 Mb, which will be used as a reference tomato genome for development of DNA markers.

Table 2. Databases for tomato genetics, genomics and breeding available from the World Wide Web

| Categories              | Database Names                                      | URLs                                      |
|-------------------------|-----------------------------------------------------|-------------------------------------------|
| Portal sites            |                                                     |                                           |
| Kazusa Tomato Genomics Database (KaTomicsDB) | http://www.kazusa.or.jp/tomato/                    |
| Plant Genome DataBase Japan (PGDBj)      | http://pgdbj.jp                                      |
| SOL Genomics Network (SGN)          | http://solgenomics.net/                             |
| eusol (eusol)              | http://www.eu-sol.net/                              |
| Lat-SOL network (Lat-SOL)         | http://enia.inta.gov.ar/lat-sol/                    |
| Plants Database            | http://plants.usda.gov/                             |
| PURDUE University         | http://www.hort.purdue.edu/hri/                      |
| Solanaceae Genomics Resource | http://solanaceae.plantbiology.msu.edu/              |
| Genome databases          |                                                     |                                           |
| PlantGDB                 | http://www.plantgdb.org/SiGDB/                      |
| A Tomato Integrated Database (TOMATOMICS) | http://bioinf.mind.meiji.ac.jp/tomatomics/         |
| International Solanaceae Genome Project | http://sol.kribb.re.kr/tomatogenome/                |
| Italian SOLAnaceae genomics resource (ISOL) | http://biosrv.cab.unina.it/isola/                   |
| Tomato SBM Database       | http://www.kazusa.or.jp/tomato_sbm/                 |
| DNA marker databases      |                                                     |                                           |
| Tomato Marker Database    | http://marker.kazusa.or.jp/tomato/                  |
| A DNA marker database for vegetables (VegMarks) | http://vegmarks.nivot.affrc.go.jp/                 |
| Tomato Mapping Resource Database | http://www.tomatomap.net/                           |
| Solanaceae Coordinated Agricultural Project (SoI CAP) | http://solcap.msu.edu/                             |
| dbSNP (Short Genetic Variations) (dbSNP)   | http://www.ncbi.nlm.nih.gov/snp/?term=Solanum+lycopersicum&SITE=NcbiHome&submit=Go|
| Tomato SNP                | http://www-plb.ucdavis.edu/labs/maloof/tomatosnp/   |
| EST databases             |                                                     |                                           |
| Micro-Tom Database (MiBASE) | http://www.kazusa.or.jp/jsol/microtom               |
| Kazusa Full-length Tomato cDNA Database (KafTom) | http://www.pgb.kazusa.or.jp/kaftom/               |
| Solanaceae EST Database (SoI EST) | http://biosrv.cab.unina.it/solestdb/               |
| A Comparative Omics Database for Plant Trichome (TrichOME) | http://www.planttrichome.org/trichomedb/esthyspecies_detail.jsp?species=Solanum%20lycopersicum |
| DFCI Tomato Gene Index (DFCI) | http://comphio.dfc.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=tomato |
| Gene expression databases |                                                     |                                           |
| PLEXdb (TomPLEX)          | http://www.plexdb.org/plex.php?database=Tomato     |
| TIGR Solanaceae Genomics Resource (SGED) | http://www.jcvi.org/potato/                      |
| Tomato Functional Genomics Database | http://ted.bti.cornell.edu/                         |
| Plant Transcription Factor Database (PlantTFDB) | http://plantfdb.cbi.pku.edu.cn/index.php?sp=Sly   |
| Metabolome databases      |                                                     |                                           |
| Kazusa Plant Pathway Viewer (KaPPA-View4 SOL) | http://kpv.kazusa.or.jp/kpv4-sol/               |
| MassBase                  | http://webs2.kazusa.or.jp/massbase/                 |
| Plant material databases  |                                                     |                                           |
| EU-SOL BredDB database (BreedDB) | https://www.eu-sol.wur.nl/                      |
| The ECPGR Tomato Database | http://documents.plant.wur.nl/cgn/pgr/tomato/     |
| Tomato Genetics Resource Center (TGRC) | http://tgrc.ucdavis.edu/                           |
| TOMATOMA                  | http://tomatoma.nlbpp.jp/                           |
| Tomato Mutant Database    | http://zamir.sgn.cornell.edu/mutants/               |
| Tomato Mutant DB (LycoTILL) | http://www.agrobios.it/tilling/                   |
| Others                   |                                                     |                                           |
| Tomato Functional SNP Database | http://plant1.kazusa.or.jp/tomato/               |
| GMO Detection method Database (GMDD) | http://gmdn.shgmo.org/event/view/113            |
| Solanaceae Source         | http://www.nhm.ac.uk/research-curation/research/projects/solanacearesource/ |
| The Tomato Genetics Cooperative | http://tgc.ifas.ufl.edu/                         |

The enhancing capacity of the NGSs will enable genotyping by sequencing (GBS), in which a huge number of SNPs can be genotyped by re-sequencing of multiple lines in a single experiment against reference genome sequences (Davey
In maize and rice, both of whose genome sequences have been determined (International Rice Genome Sequencing Project 2005, Schnable et al. 2009), the GBS strategy has been applied to construct the genetic linkage maps (Elshire et al. 2011, Xie et al. 2010). Even in plant species having larger and more complex genomes, the GBS has been carried out by sequencing of restriction site associated DNA to reduce the genome complexity (Rowe et al. 2011). In addition to the genetic mappings, the NGSS have also been applied to identify mutant genes directly. Sequencing of target regions with sequence capture technology is a straightforward strategy (Fu et al. 2010, Galvão et al. 2012).

In Arabidopsis, rice and legumes with small genome sizes, whole genome re-sequencing with genetic segregation analysis has already identified mutant genes (Abe et al. 2012, Schneeberger et al. 2009, Uchida et al. 2011). In tomato, such re-sequencing methods will be available under the advances of the sequencing technology. Together with the advanced genotyping methods, high-throughput phenotyping systems and genetic statistics for hundreds of thousands of segregation data are also required for the progression of tomato molecular genetics.

In the studies on the functional SNPs taking account of the protein structural information, the amino-acid residues important for protein activities could be revealed by using computational approaches, e.g., calculation of the affinities between enzymes and substrates by binding simulation (Ishikawa et al. 2010). This strategy has been greatly advanced in medical sciences, i.e., pharmacogenomics and pharmacogenetics, in which medical molecules are designed to fit genotypes of patients (Evans and Relling 1999). In plants, on the other hand, the virtually predicted desirable genotypes can be selected from natural populations and artiﬁcial mutants by SNP analysis. The screening systems of mutant genes, e.g., targeting induced local lesions in genomes (TILLING) (Okabe et al. 2011, 2012) and deep-sequencing of target genes (Rigola et al. 2009), have already been developed in tomato. Moreover, in the future, desirable genotypes can be created via new breeding techniques, e.g., site-directed mutagenesis by gene targeting with nucleases or oligonucleotides (Lusser et al. 2012), whereas these technologies have been applied to only maize, tobacco, rice and oilseed rape, but not yet in tomato. Tomato molecular breeding will proceed through the SNP analysis and the genome manipulations.

Acknowledgments

We thank Satoshi Tabata, Sachiko Isobe and Shusei Sato (Kazusa DNA Research Institute) for encouragement and critical reading of the manuscript. This work was supported by the Kazusa DNA Research Institute Foundation, the KAKENHI Grant-in-Aid for Scientific Research (C) (24510286), JSPS, Japan, the Genomics for Agricultural Innovation Foundation (DD-4010/SGE-1001), MAFF, Japan and the Integrated Database Project Foundation, MEXT, Japan.

Literature Cited

Abe,A., S.Kosugi, K.Yoshida, S.Natsume, H.Takagi, H.Kanzaki, H.Matsumura, K.Yoshida, C.Mitsuoka, M.Tamira et al. (2012) Genome sequencing reveals agronomically important loci in rice using MutMap. Nat. Biotechnol. 30: 174–178.

Altshuler,D., V.J.Pollara, C.R.Cowles, W.J.Van Etten, J.Baldwin, L.Linton and E.S.Lander (2000) An SNP map of the human genome generated by reduced representation shotgun sequencing. Nature 407: 513–516.

Aoki,K., H.Nagasaki, S.Tatsumoto, Y.Minakuchi, E.Kaminuma, A.Toyoda, K.Suda, S.Kawamura, K.Igarashi and K.Yano (2011) Micro-Tom genome sequencing and genome-wide identification of DNA polymorphisms with Heinz 1706. In: Abstracts of 8th Solanaceae and 2nd Cucurbitaceae Genome Joint Conference, Kobe, Japan, p. 26.

Argout,X., J.Salse, J.M.Aury, M.J.Guilminan, G.Droc, J.Gouzy, M.Allegre, C.Chaparro, T.Legavre, S.N.Maximova et al. (2011) The genome of Theobroma cacao. Nat. Genet. 43: 101–108.

Berman,H.M., J.Westbrook, Z.Feng, G.Gilliland, T.N.Bhat, H. Weissig, J.N.Shindyavol and P.E.Bourne (2000) The protein data bank. Nucleic Acids Res. 28: 235–242.

Bernacchi,D. and S.D.Tanksley (1997) An interspecific backcross of Lycopersicon esculentum × L. hirsutum: linkage analysis and a QTL study of sexual compatibility factors and floral traits. Genetics 147: 861–877.

Bertazzi,J. and S.D.Tanksley (1986) Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics 112: 887–898.

Bishop,G.J. et al. (1996) The tomato Dwarf gene isolated by heterologous transposon tagging encodes the first member of a new cytochrome P450 family. Plant Cell 8: 959–969.

Brommonschenkel,S.H., A.Frary, A.Frary and S.D.Tanksley (2000) The broad-spectrum tspovirus resistance gene Siv-5 of tomato is a homolog of the root-knot nematode resistance gene Mi. Mol. Plant Microbe Interact. 13: 1130–1138.

Chevreux,B., T.Pfisterer, B.Drescher, A.J.Driesel, W.E. Müller, T. Wetter and S.Suahai (2004) Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Res. 14: 1147–1159.

Cong,B., L.S.Barrero and S.D.Tanksley (2008) Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. Nat. Genet. 40: 800–804.

Davey,J.W., P.A.Hohenlohe, P.D.Etter, J.Q.Boone, J.M.Catchen and M.L.Blaxter (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat. Rev. Genet. 12: 499–510.

D’Hont,A., F.Denoeud, J.M.Aury, F.C. Baurens, F.Carreol, O. Garsmeur, B.Noel, S.Bocs, G.Droc, M.Rouard et al. (2012) The banana (Musa acuminate) genome and the evolution of monocotyledonous plants. Nature 488: 213–217.

Dixon,M.S., D.A.Jones, J.S.Keddie, C.M.Thomas, K.Harrison and J.D.Jones (1996) The tomato Cf-2 disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell 84: 451–459.

Doganlar,S., A.Frary, H.M.Ku and S.D.Tanksley (2002a) Mapping quantitative trait loci in inbred backcross lines of Lycopersicon pimpinellifolium (LA1589). Genome 45: 1189–1202.

Doganlar,S., A.Frary, M.C.Daunay, R.N.Lester and S.D.Tanksley (2002b) A comparative genetic linkage map of eggplant (Solanum melongena) and its implications for genome evolution in the
functional genomics into rational therapeutics. Science 286: 487–89.

Galvão, V.C., J.K. Nordström, C. Lanz, P. Sulz, J. Mathieu, D. Posé, M. Schmid, D. Weigel and K. Schneeberger (2010) Repeat subtraction-mediated sequence capture from a complex genome. Plant J. 62: 898–909.

Guiloux, V.L., P. Schmidtke and P. Tuffery (2009) Fpocket: an open source platform for ligand pocket detection. BMC Bioinformatics 10: 168.

Hammond-Kosack, P.J. Balint-Kurti and J.D.Jones (1994) Isolation of the tomato disease resistance genes encode cell surface-like receptors. Proc. Natl. Acad. Sci. USA 91: 6995–6999.

Kawamoto, E.S., R.J., J.C. Tanksley (1996) QTL analysis of horticultural traits differentiating the cultivated tomato from the closely related species Lycopersicon pimpinellifolium. Theor. Appl. Genet. 92: 935–951.
DNA marker for genetics and genomics in tomato

with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760.
Li, R., Y. Li, X. Fang, H. Yang, J. Wang, K. Kristiansen and J. Wang (2009) SNP detection for massively parallel whole-genome re-sequencing. Genome Res. 19, 1124–1132.

Liu, J., J. Van Eck, B. Cong and S.D. Tanksley (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. Proc. Natl. Acad. Sci. USA 99: 13302–13306.

Livingstone, K.D., V.K. Lackney, J.R. Blauth, R. van Wijk and M.K. Jahn (1999) Genome mapping in capsicum and the evolution of genome structure in the solanaceae. Genetics 152: 1183–1202.

Lusser, M., C. Parisi, D. Plan and E. Rodriguez-Cerezo (2012) Deployment of new biotechnologies in plant breeding. Nat. Biotechnol. 30: 213–239.

Martin, G.B., S.H. Brommonschenkel, J. Chunwongse, A. Frary, M.W. Ganal, R. Spivey, T. Wu, E.D. Earle and S.D. Tanksley (1999) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science 262: 1432–1436.

Meshi, T., F. Motoyoshi, A. Adachi, Y. Watanabe, N. Takamatsu and Y. Okada (1998) Two concomitant base substitutions in the putative replicase genes of tobacco mosaic virus confer the ability to overcome the effects of a tomato resistance gene, Tm-1. EMBO J. 7: 1575–1581.

Meshi, T., F. Motoyoshi, T. Maeda, S. Yoshishowa, H. Watanabe and T. Okada (1989) Mutations in the tobacco mosaic virus 30-kD protein gene overcome Tm-2 resistance in tomato. Plant Cell 1: 515–522.

Milligan, S.B., J. Bodeau, J. Yaghoobi, I. Kaloshian, P. Zabel and V.M. Williamson (1998) The root knot nematode resistance gene Mi from tomato is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. Plant Cell 10: 1307–1319.

Ming, R., S. Hou, Y. Feng, Q. Yu, A. Donne-Laporte, J.H. Saw, P. Senin, W. Wang, B.V. Ly, K.L.T. Lewis et al. (2008) The draft genome of the transgenic tropical fruit tree papaya (Carica papaya Linnaeus). Nature 452: 991–996.

Morris, G.M., D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew and A.J. Olson (1998) Automated docking using a Lamarckian genetic algorithm and empirical binding free energy function. J. Computational Chemistry 19: 1639–1662.

Mueller, L.A., T.H. Solow, N. Taylor, B. Skwarecki, R. Buels, J. Binns, C. Lin, M.H. Wright, R. Ahrens, Y. Wang et al. (2005) The SOLGenes Network: a comparative resource for Solanaceae biology and beyond. Plant Physiol. 138: 1310–1317.

Munos, S., N. Ranc, E. Botton, A. Berard, S. Rolland, P. Duffe, Y. Carretoro, M.C. Le Pailler, C. Delalande, M. Bouzyan et al. (2011) Increase in tomato locule number is controlled by two single-nucleotide polymorphisms located near WUSCHEL. Plant Physiol. 156: 2244–2254.

Ogata, H., S. Goto, K. Sato, W. Fujibuchi, H. Bono and M. Kanehisa (1999) KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 27: 29–34.

Ohyama, A., E. Asamizu, S. Negoro, K. Miyatake, H. Yamaguchi, S. Tabata and H. Fukuoka (2009) Characterization of tomato SSR markers developed using BAC-end and cDNA sequences from genome database. Mol. Breeding. 23: 685–691.

Okabe, Y., E. Asamizu, T. Ariizumi, K. Shirasawa, S. Tabata and H. Ezura (2012) Availability of Micro-Tom mutant library combined with TILLING in molecular breeding of tomato fruit shelf-life. Breed. Sci. 62: 202–208.

Okabe, Y., E. Asamizu, T. Saito, C. Matsukura, T. Ariizumi, C. Brés, C. Rothan, T. Mizoguchi and H. Ezura (2011) Tomato TILLING tech-
nology: development of a reverse genetics tool for the efficient iso-
lation of mutants from Micro-Tom mutant libraries. Plant Cell Physiol. 52: 1994–2005.

Ori, N., Y. Eshed, I. Panar, G. Presting, D. Aviv, S. Tanksley, D. Zamir and R. Fluhr (1997) The J2C family from the wilt disease resistance locus I2 belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. Plant Cell 9: 521–532.

Paterson, A.H., J.E. Bowers, R. Bruggmann, I. Dubchak, J. Grimwood, H. Gundlach, G. Haberer, U. Hellsten, T. Mitros, A. Poliakov et al. (2009) The Sorghum bicolor genome and the diversification of grasses. Nature 457: 551–556.

Pnueli, L., L. Carmel-Goren, D. Hareven, T. Gutfinger, J. Alvarez, M. Ganal, D. Zamir and E. Lifschitz (1998) The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. Development 125: 1979–1989.

Powell, A.L.T., C.V. Nguyen, T. Hill, K.L. Cheng, R. Figueroa-Balderas, H. Aktas, H. Ashrafi, C. Pons, R. Fernández-Muñoz, A. Vicente et al. (2012) Uniform ripening encodes a Golden-2-like transcription factor regulating tomato fruit chloroplast development. Science 336: 1711–1715.

Prince, J.P., E. Pochard and S.D. Tanksley (1993) Construction of a molecular linkage map of pepper and a comparison of synteny with tomato. Genome 36: 404–417.

Punta, M., P.C. Coghill, R.Y. Eberhardt, J. Mistry, J. Tate, C. Bourssnell, N. Pang, K. Forslund, G. Ceric, J. Clements et al. (2012) The Pfam protein families database. Nucleic Acids Res. 40: D290–D301.

Rigola, D., J. van Oeveren, A. Janssen, A. Bonne, H. Schneiders, H.J.A. van der Poel, N.J. van Oorsouw, R.C.J. Hogers, M.T.J. de Both and M.J.T. van Eijk (2009) High-throughput detection of induced mutations and natural variation using KeyPoint™ technology. PLoS ONE 4: e4761.

Rodriguez, G.R., S. Muños, C. Anderson, S.C. Sim, A. Michel, M. Causse, B.B. Gardener, D. Francis and E. van der Knaap (2011) Distribution of SUN, OVA, T, and FAS in the tomato germplasm and the relationship to fruit shape diversity. Plant Physiol. 156: 275–285.

Rowe, H.C., S. Renaut and A. Guggisberg (2011) RAD in the realm of next-generation sequencing technologies. Mol. Ecol. 20: 3499–3502.

Saito, T., T. Ariizumi, Y. Okabe, E. Asamizu, K. Hiwasa-Tanase, N. Fukuda, T. Mizoguchi, Y. Yamazaki, K. Aoki and H. Ezura (2011) TOMATOMA: a novel tomato mutant database distributing Micro-Tom mutant collections. Plant Cell Physiol. 52: 283–296.

Saliba-Colombani, V., M. Causse, L. Gervais and J. Philouze (2000) Efficiency of AFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome. Genome 43: 29–40.

Schnutz, J., S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D.L. Hyten, Q. Song, J.J. Thelen, J. Cheng et al. (2010) Genome sequence of the palaeopolyploid soybean. Nature 463: 178–183.

Schneebeger, K., S. Ossowski, C. Lanz, T. Juul, A.H. Petersen, K.L. Nielsen, J.E. Jorgensen, D. Weigel and S.U. Andersen (2009) SHOREmap: simultaneous mapping and mutation identification by deep sequencing. Nat. Methods 6: 550–551.

Shiokai, S., K. Shirasawa, Y. Sato and T. Nishio (2010) Improvement of the dot-blot-SNP technique for efficient and cost-effective genotyping. Mol. Breeding. 25: 179–185.
Shirasawa, K., L. Monna, S. Kishitani and T. Nishio (2004) Single nucleotide polymorphisms in randomly selected genes among *japonica* rice (*Oryza sativa L.*) varieties identified by PCR-RF-SSCP. DNA Res. 11: 275–283.

Shirasawa, K., S. Shiokai, M. Yamaguchi, S. Kishitani and T. Nishio (2006) Dot-blot-SNP analysis for practical plant breeding and cultivation identification in rice. Theor. Appl. Genet. 113: 147–155.

Shirasawa, K., E. Asamizu, H. Fukuoka, A. Ohyama, S. Sato, Y. Nakamura, S. Tabata, S. Sasamoto, T. Wada, Y. Kishida et al. (2010a) An interspecific linkage map of SSR and intronic polymorphism markers in tomato. Theor. Appl. Genet. 121: 731–739.

Shirasawa, K., S. Isobe, H. Hirakawa, E. Asamizu, H. Fukuoka, D. Just, C. Rothan, S. Sasamoto, T. Fujishiro, Y. Kishida et al. (2010b) SNP discovery and linkage map construction in cultivated tomato. DNA Res. 17: 381–391.

Shulaev, V., D.J. Sargent, R.N. Crowhurst, T.C. Mockler, O. Folkerts, A.L. Delcher, P. Jaiswal, K. Mockaitis, A. Liston, S.P. Mané et al. (2011) The genome of woodland strawberry (*Fragaria vesca*). Nat. Genet. 43: 109–116.

Sim, S.C., G. Durszewitz, J. Plieske, R. Wieseke, M.W. Ganal, A. Van Deyneze, J.P. Hamilton, C.R. Buell, M. Causse, S. Wijeratte et al. (2012) Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. PLOS One 7: e40563.

Tang, H., J.E. Bowers, X. Wang, R. Ming, M. Alam and A.H. Paterson (2008) Synteny and collinearity in plant genomes. Science 320: 486-488.

Tanksley, S.D., R. Berntatzky, N.L. Lapitan and J.P. Prince (1988) Conservation of gene repertoire but not gene order in pepper and potato. Proc. Natl. Acad. Sci. USA 85: 6419–6423.

Tanksley, S.D., M.W. Ganal, J.P. Prince, M.C. de Vicente, M.W. Bonierbale, P. Broun, T.M. Fulton, J.J. Giovannoni, S. Grandillo, G.B. Martin et al. (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132: 1141–1160.

Tanksley, S.D., S. Grandillo, T.M. Fulton, D. Zamir, Y. Eshed, V. Petiard, J. Lopez and T. Beck-Buun (1996) Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. Theor. Appl. Genet. 92: 213–224.

Tatusov, R.L., N.D. Fedorova, J.D. Jackson, A.R. Jacobs, B.Kiryutin, E.V. Koonin, D.M. Krylov, R. Mazumder, S.L. Melhedov, A.N. Nikoliskaya et al. (2003) The COG database: an updated version includes eukaryotes. Bioinformatics 4: 41–54.

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796–815.

The *Brassica rapa* Genome Sequencing Project Consortium (2011) The genome of the mesopolyploid crop species *Brassica rapa*. Nat. Genet. 43: 1035–1039.

The French–Italian Public Consortium for Grapevine Genome Characterization (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449: 463–467.

The Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. Nature 475: 189–195.

The Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485: 635–641.

Uchida, N., T. Sakamoto, T. Kurata and M. Tasaka (2011) Identification of EMS-induced causal mutations in a non-reference *Arabidopsis thaliana* accession by whole genome sequencing. Plant Cell Physiol. 52: 716–722.

van der Knaap, E. and S.D. Tanksley (2001) Identification and characterization of a novel locus controlling early fruit development in tomato. Theor. Appl. Genet. 103: 353–358.

van der Knaap, E. and S.D. Tanksley (2003) The making of a bell pepper-shaped tomato fruit: identification of loci controlling fruit morphology in Yellow Stuffer tomato. Theor. Appl. Genet. 107: 139–147.

Van Schalkwyk, A.P., W. Wenzl, S. Smit, R. Lopez-Cobollo, A. Kilian, G. Bishop, C. Hefer and D.K. Berger (2012) Bin mapping of tomato diversity array (DArT) markers to genomic regions of *Solanum lycopersicum × Solanum pennellii* introgression lines. Theor. Appl. Genet. 124: 947–956.

Varshney, R.K., A. Graner and M.E. Sorrells (2005) Genomics-assisted breeding for crop improvement. Trends Plant Sci. 10: 621–630.

Varshney, R.K., W. Chen, Y. Li, A.K. Bharti, R.K. Saxena, J.A. Schlueter, M.T.A. Donoghue, S. Azam, G. Fan, A.M. Whaley et al. (2012) Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. Nat. Biotechnol. 30: 83–89.

Velasco, R., A. Zharkikh, J. Affourtit, A. Dhingra, A. Cestaro, A. Kalyanaraman, P. Fontana, S.K. Bhatnagar, M. Troggio, D. Pruss et al. (2010) The genome of the domesticated apple (*Malus × domestica* Borkh.). Nat. Genet. 42: 833–839.

Wu, F., N.T. Eannetta, Y. Xu, R. Durrett, M. Mazourek, M.M. Jahn and S.D. Tanksley (2009a) A COSII genetic map of the pepper genome provides a detailed picture of synteny with tomato and new insights into recent chromosome evolution in the genus *Capsicum*. Theor. Appl. Genet. 118: 1279–1293.

Wu, F., N.T. Eannetta, Y. Xu and S.D. Tanksley (2009b) A detailed synteny map of the eggplant genome based on conserved ortholog set II (COSII) markers. Theor. Appl. Genet. 118: 927–935.

Wu, F., N.T. Eannetta, Y. Xu, J. Plieske, M. Ganal, C. Pozzi, N. Bakher and S.D. Tanksley (2010) COSII genetic maps of two diploid *Nicotiana* species provide a detailed picture of synteny with tomato and insights into chromosome evolution in tetraploid *N. tabacum*. Theor. Appl. Genet. 120: 809–827.

Wu, F. and S.D. Tanksley (2010) Chromosomal evolution in the plant family Solanaceae. BMC Genomics 11: 182.

Xiao, H., N. Jiang, E. Schaffner, E.J. Stockinger and E. van der Knaap (2008) A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. Science 319: 1527–1530.

Xie, W., Q. Feng, H. Yu, X. Huang, Q. Zhao, Y. Xing, S. Yu, B. Han and Q. Zhang (2010) Parent-independent genotyping for constructing an ultra-high-density linkage map based on population sequencing. Proc. Natl. Acad. Sci. USA 107: 10578–10583.

Yamamoto, N., T. Tsugane, M. Watanabe, K. Yano, F. Maeda, C. Kuwata, M. Torki, Y. Ban, S. Nishimura and D. Shibata (2005) Expressed sequence tags from the laboratory-grown miniature tomato (*Lycopersicon esculentum*) cultivar Micro-Tom and mining for single nucleotide polymorphisms in *Lycopersicon esculentum* by computer aided analysis of expressed sequence tags. Mol. Breed. 14: 21–34.

Zhang, Z., Y. Li, B. Lin, M. Schroeder and B. Huang (2011) Identification of cavities on protein surface using multiple computational approaches for drug binding site prediction. Bioinformatics 27: 2083–2088.