**Introduction**

DNA polymorphisms and mutations, some of which confer phenotypic variations, can be detected by DNA marker analyses. The first DNA marker, the restriction fragment length polymorphism (RFLP), has been used for linkage analysis to determine the genomic positions responsible for Huntington’s disease in humans (Gusella et al. 1983). The RFLP marker technologies were then applied to analysis of plant genetics, particularly tomato and maize (Helentjaris et al. 1986), to construct genetic maps, which are essential tools for positional cloning and quantitative trait loci (QTL) analysis of genes of interest. Then, several types of DNA markers, e.g., random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), simple sequence repeat (SSR) or microsatellite, and single nucleotide polymorphism (SNP), were made available through advances in the technologies for DNA analysis (Phillips and Vasil 2001).

DNA markers are used not only in basic sciences but also in applied studies (Kumar 1999). For example, DNA markers linking to desirable loci are used in the selection of elite lines from breeding populations, a process known as marker-assisted selection (MAS). F₁ hybrids are used in the production of commercial varieties in several cereal and vegetable crops, because the F₁ hybrids sometimes exhibit hybrid vigor and heterosis. Therefore, DNA markers can also be applied for purity testing to investigate the heterogeneity of F₁ hybrids, which is a combination of different alleles from the parental lines. In addition, in the management of genetic resources and quality control of food products, DNA markers have been employed for identification of species, cultivars, and varieties.

Whole-genome sequencing in plants was first achieved in *Arabidopsis thaliana* (The Arabidopsis Genome Initiative 2000), followed by rice (International Rice Genome Sequencing Project 2005). Since those initial reports, the genomes of more than 50 plants have been sequenced (Michael and Jackson 2013). In addition, massive transcriptome analysis has been performed in several plants by using the next-generation sequencers (NGSs) (Hamilton and Buell 2012). This has enabled the development of large numbers of DNA markers for several plants in a relatively short time. Numerous databases have been made available with the genome and DNA marker information for these plant species, e.g. TAIR for *Arabidopsis thaliana* (http://

**Note added in proof:** A paper on faba bean (*Vicia faba*) entitled “Development of EST-SSR markers and construction of a linkage map in faba bean (*Vicia faba*)” by El-Rodeny et al. was published in *Breeding Science*. The information on the EST-SSRs and map were added to the Kazusa Marker DataBase.
Contents of the Kazusa Marker DataBase

The Kazusa Marker DataBase was constructed using Red Hat Enterprise Linux Server release 5.6 as the computer operating system. The MySQL (http://dev.mysql.com) system, which is a relational database management system, was employed for management of the database contents. Most of the contents of this database were written using html text format and Ruby on Rails (RoR), the open-source web framework. The Kazusa DNA Research Institute, for crops and other plant species to enrich the available molecular genetic information on agronomical plants.

Table 1. Types and numbers of DNA markers registered in the Kazusa Marker DataBase

| Binomial nomenclature | Common name | Marker type | Abbreviation of marker name | No. of markers | Reference |
|-----------------------|-------------|-------------|----------------------------|----------------|-----------|
| Solanum lycopersicum  | Tomato      | Genome-SSR  | TGS                        | 13,501         | Shirasawa et al. 2010a |
|                       |             | EST-SSR     | TES                        | 7,599          | Shirasawa et al. 2010a |
|                       |             | Genome-SNP  | 1,473,798                  |                | Shirasawa et al. 2013b |
|                       |             | EST-SNP     | 5,607                      |                | Shirasawa et al. 2010b |
|                       |             | Intron-SNP  | TEI                        | 169            | Shirasawa et al. 2010a |
| Capsicum annuum       | Capsicum    | EST-SSR     | CaES                       | 5,751          | Shirasawa et al. 2013c |
| Fragaria × ananassa   | Strawberry  | EST-SSR     | FAES                       | 603            | Isobe et al. 2013   |
|                       |             | SSR derived from F. vesca | FVES | 3,746 | Isobe et al. 2013   |
|                       |             | Transcriptome-SSR | FATS | 125  | Isobe et al. 2013   |
| Raphanus sativus      | Radish      | EST-SSR     | RRS                        | 3,811          | Shirasawa et al. 2011 |
| Lotus japonicus       |             | dCAPS       | BM, TM                     | 82             | Sato et al. 2001    |
| Glicine max           | Soybean     | EST-SSR     | GMES                       | 7,020          | Sato et al. 2004    |
| Arachis hypogaea      | Peanut      | EST-SSR     | AHS                        | 3,187          | Hisano et al. 2007  |
|                       |             | Genome-SSR  | AHGS                       | 6,706          | Shirasawa et al. 2012b |
|                       |             | Transposable element | AhTE | 1,039 | Shirasawa et al. 2012a, 2012b |
| Trifolium pratense    | Red clover  | SSR         | RCS, TPSSR                 | 7,262          | Isobe et al. 2009   |
| Trifolium repens      | White clover| EST-SSR     | WCS                        | 1,993          | Isobe et al. 2012   |
| Eucalyptus camaldulensis | Eucalyptus | Genome-SSR  | EcGAS                      | 4,656          | Hirakawa et al. 2011 |
|                       |             | EST-SSR     | EcES                       | 1,028          | Hirakawa et al. 2011 |
Table 2. Genetic linkage maps registered in the Kazusa Marker DataBase

| Binomial nomenclature | Common name  | Map name          | No. of linkage groups | No. of mapped loci | Total length (cM) | Mean marker density (cM/loci) | Reference |
|-----------------------|--------------|-------------------|-----------------------|--------------------|-------------------|-------------------------------|-----------|
| Solanum lycopersicum  | Tomato       | Tomato-EXPEN2000  | 12                    | 2,116              | 1,503             | 0.7                           | Shirasawa et al. 2010a |
|                       | AMF2         |                   | 12                    | 990                | 1,468             | 1.5                           | Shirasawa et al. 2010b |
|                       | MF2          |                   | 13                    | 637                | 1,230             | 2.0                           | Shirasawa et al. 2010b |
| Fragaria × ananassa   | Strawberry   | Integrated map    | 28                    | 1,861              | 1,967             | 1.1                           | Isobe et al. 2013   |
| Raphanus sativus      | Radish       | GHRI              | 9                     | 843                | 1,129             | 1.4                           | Shirasawa et al. 2011 |
| Lotus japonicus       | Soybean      | SK1               | 6                     | 1,155              | 421               | 0.4                           | Hayashi et al. 2001 |
| Glycine max           | Peanut       |                    | 20                    | 693                | 2,688             | 4.0                           | Hisano et al. 2007  |
| Arachis hypogaea      |              |                   | 21                    | 1,144              | 2,226             | 2.0                           | Shirasawa et al. 2012b|
|                       |              |                   | 19                    | 326                | 1,333             | 4.3                           | Shirasawa et al. 2012b|
|                       |              |                   | 10                    | 597                | 544               | 0.9                           | Shirasawa et al. 2013a|
|                       |              |                   | 10                    | 798                | 461               | 0.6                           | Shirasawa et al. 2013a|
|                       |              |                   | 20                    | 1,469              | 1,442             | 1.0                           | Shirasawa et al. 2013a|
|                       |              | Integrated map    | 20                    | 3,693              | 2,651             | 0.7                           | Shirasawa et al. 2013a|
| Trifolium pratense    | Red clover   | HR × R130         | 7                     | 1,714              | 834               | 0.5                           | Isobe et al. 2009   |

Usage instruction for the Kazusa Marker DataBase

The top page of the Kazusa Marker DataBase represents crops registered in this database. Users can click either “Images”, “Scientific names” in the table, or icons below the table to access pages of each crop, which include “Keyword Search”, “Marker List”, “Reference list”, “Linkage Map”, “Physical Map”, and “Markers on the Genome” depending on crops (see below section). Through the “Keyword Search”, marker names, sequence names used for marker designing, and descriptions in comment boxes can be searched. The “Marker List” contained marker types, e.g., genome-SSR and EST-SSR, as described in the below section and Table 1. By selecting the marker type, lists of the markers comprised of “Marker Name”, “Marker Type”, and primer sequences are available. Then, clicking marker names enables users to obtain all information on the markers, e.g., sequence name corresponding the markers with a hyperlink to public DNA sequence databases, map positions (if available), PCR fragment size estimated from the sequence, experimental conditions such as methods on PCR and detections, SSR motif and the repeated number (if SSRs), gel images (if available), and reference articles. The “Markers on the Genome” is available to presume physical genome positions of the markers, if the genome sequences of the crop itself or its relatives are released. Although bulk data download is not supported in the current version, it is available upon request to markerdb@kazusa.or.jp as well as to us.

Plant species registered in the Kazusa Marker DataBase

Tomato

Tomato (S. lycopersicum), an important fruit crop throughout the world and a model for fresh fruit research, is an autogamous diploid species (2n = 2x = 24) with a genome of 900 Mb, and its sequences have been published by a multinational project consortium (The Tomato Genome Consortium 2012). The database contains information on DNA markers, as well as genetic linkage and physical maps. The DNA markers include both SSR and SNP markers. The SSR markers, 7,599 EST-SSR (TES) and 13,501 genome-SSR (TGS) designed from EST and BAC-end sequences, respectively, were developed to construct an interspecific high-density genetic linkage map (Shirasawa et al. 2010a), on which totals of 648 TES and 634 TGS were mapped. In addition, 674 EST-derived intronic polymorphism markers (TEI) were developed and 151 TEI markers were mapped (Shirasawa et al. 2010a). The SNP markers were also developed from EST and genome sequence data. Each of the EST-derived SNPs was developed from the alignment data of EST sequences derived from at least two tomato lines. From this analysis, 5,607 SNPs were identified in 2,634 contigs, and 793 were mapped on the two genetic linkage maps based on intraspecific crossings (Shirasawa et al. 2010b). On the other hand, the genome-SNPs were discovered by the re-sequencing strategy (Shirasawa et al. 2013b), in which sequence reads for six tomato lines by the ABI-5500xl SOLiD (DRA accession numbers: DRA001017 to DRA001022) were mapped onto the tomato reference genome, SL2.40 (The Tomato Genome Consortium 2012). A total of 1,473,798 genome-SNPs were identified and 1,536 polymorphic sequence (CAPS) markers, in which 19 restriction enzymes are employed, are also available. The positions of the DNA markers developed in this study were identified on the published tomato genome (The Tomato Genome Consortium 2012). Then, the DNA markers and the positions of the various markers. For tomato, pepper, and radish, the DNA markers were mapped on the reference sequences themselves or their relatives.
predicted genes in the tomato genome were ordered in parallel based on the physical positions of the reference genome, from which the users can obtain information on the DNA markers and the predicted genes. This tool is very useful to search for DNA markers or loci in genes of interest.

In addition, we have established a portal website for tomato genomics, KatomicsDB (Shirasawa and Hirakawa 2013: http://www.kazusa.or.jp/tomato/), because our group provides not only information on the DNA markers and genetic maps as described above but also inferences of SNP effects on gene functions and sequence data of gene-rich regions in the tomato genome. The KatomicsDB contains links to the marker database described above, a functional regions in the tomato genome. The KatomicsDB contains links to the marker database described above, a functional SNP database (Hirakawa et al. 2013a: http://plant1.kazusa.or.jp/tomato/), and a database for genome sequences of selected BAC clone mixtures in gene-rich regions (http://www.kazusa.or.jp/tomato_sbm/).

Pepper
Capsicum spp., including C. annuum, C. baccatum, C. chinense, C. frutescens, and C. pubescens, belong to the Solanaceae family, and are widely cultivated for use as vegetables and spices. Like tomato, all these species are autogamous diploids (2n = 2x = 24), while the size of their genomes (~ 3.3 Gb: Moscone et al. 2003) is more than three times larger than that of the other members of Solanaceae, e.g., tomato, potato, and eggplant. The pepper marker database includes mainly Capsicum EST-SSR (CaES) information. A total of 5,751 CaES markers were designed from the 118,060 EST sequences for Capsicum annuum obtained from a public DNA database, GenBank (http://www.ncbi.nlm.nih.gov). The CaES markers were mapped on the tomato genome by in silico analysis based on sequence similarity search, which is recognized as a model for the Solanaceae, because the genome sequence of pepper has not been reported (at the time of writing). However, the genotype structures between tomato and pepper are conserved and exhibit a macrosynteny relationship (Wu et al. 2010). Therefore, the positions of the pepper DNA markers and genes on the pepper genome can be speculated by mapping them onto the tomato genome (Shirasawa et al. 2013c). As a result, the positions of 2,245 of the CaES markers were identified on the tomato genome. Among the 2,245 markers, 96 CaES markers were subjected to genotyping analysis of 192 Capsicum accessions, which have been stocked at the Kihara Institute for Biological Research of the Yokohama City University, Japan, to reveal their genetic diversity. The polymorphism information content (PIC) values and allele sizes for the 192 accessions are also available from this database. As additional markers, SNPs detected in the matK and rbcL genes coded in the chloroplast genome, which are known as “barcode” sequences for the identification of species (CBOL Plant Working Group 2009), are also available for the 192 accessions (Accession numbers: AB721552 to AB721935).

Strawberry
Strawberry (F. × ananassa) is a popular fruit cultivated throughout the world, and possesses a complex genome structure due to its octoploid nature (2n = 8x = 56) and its allogamous reproductive system. The genome size of strawberry is estimated to be 692 Mb (Hirakawa et al. 2013b). A wild diploid species, F. vesca, is one of the probable ancestral species, and 240 Mb of its genome has been sequenced (Shulaev et al. 2011). This database includes three types of SSR markers and an integrated map. The SSR markers were designed from EST sequences of not only F. × ananassa but also F. vesca, because a larger number of EST sequences for F. vesca were available from public DNA databases than for F. × ananassa. A total of 3,746 SSRs derived from ESTs for F. vesca, 603 SSRs derived from ESTs for F. × ananassa, and 125 SSRs derived from transcriptomes for F. × ananassa markers were developed and subjected to map constructions (Isobe et al. 2013). Three genetic linkage maps were established using three mapping populations, and integrated into a consensus map consisting of 28 linkage groups with 1,856 loci, the number of which corresponded to the haploid chromosome number of F. × ananassa. In addition to the map constructions, the SSR markers were employed for the genetic diversity analysis of 129 strawberry cultivars. A total of 45 SSR markers were determined to be sufficient to distinguish 129 F. × ananassa lines except for four lines.

Radish
Radish (R. sativus), or Japanese daikon, is an allogamous species due to its self-incompatibility system, and has a diploid genome (2n = 2x = 18), sizing of 526 Mb (Arunmuganathan and Earle 1991). The radish is a vegetable crop and a member of the Brassicaceae, to which the genera Arabidopsis and Brassica also belong, but the genomic research on radish has not been progressed as far as for members of the Brassicaceae. The daikon marker database includes mainly EST-SSR markers. A total of 3,800 radish EST-SSR markers (RSS) were developed from 26,606 EST sequences (Accession numbers: FY428055 to FY454660) (Shirasawa et al. 2011). Genetic linkage maps of 630 RSS markers and 213 previously reported markers were obtained from this database. Subsequent comparative analysis of the Raphanus map with the Arabidopsis and B. rapa genomes (The Arabidopsis Genome Initiative 2000, The Brassica rapa Genome Sequencing Project Consortium 2011) revealed the genomic synteny between the two species. Therefore, the radish DNA markers were in silico mapped on the genomes of Arabidopsis and B. rapa to speculate on the positions of the radish DNA markers and genes on the radish genome. This analysis revealed the positions of 3,234 and 3,730 SSR markers on the Arabidopsis and B. rapa genomes, respectively.

Lotus japonicus
L. japonicus is not a crop but is recognized as a model for legume crops and symbiosis research because of its
rapid life cycle, fixed genotypes due to autogamous reproduction, simple and compact genome (2n = 2x = 12, 472 Mb), and easy transformability (Handberg and Stougaard 1992). The marker database for Lotus japonicus consists of DNA markers and their linkage map. A total of 1,073 SSR and 82 derived CAPS (dCAPS) markers were developed by comparative analysis of the genome sequences from two L. japonicus strains, Miyakojima MG-20 and Gifu B-129 (Sato et al. 2001). A genetic linkage map of the SSR and dCAPS markers, which consisted of six linkage groups covering 1,155 cM in total, were generated by using an F2 mapping population derived from a cross between the MG-20 and B-129 (Hayashi et al. 2001). By using this linkage map as a reference, the genome sequences of the MG-20 were anchored to the chromosomes of L. japonicus (Sato et al. 2008: http://www.kazusa.or.jp/lotus/).

**Soybean**

Because soybean (G. max) is a major crop that is important for oil and protein production, its genome (2n = 2x = 40, genome size of 1.1 Gb) was sequenced in spite of the complexity of its paleopolyploidy (Schmutz et al. 2010). The database includes EST-SSR markers and a genetic linkage map. A total of 6,920 EST-SSR markers were developed from 63,676 publicly available non-redundant soybean ESTs from public databases (Dana-Farber Center Institute; http://compbio.dfci.harvard.edu/tgi/). Among them, 693 SSR marker loci were combined with 242 RFLP, genome-SSR, and phenotypic markers. The resultant maps consisting 20 linkage groups covered 2,700.3 cM in total length (Hisano et al. 2007). The transferability of the 686 mapped markers was investigated for 24 Glycine accessions. The EST-SSR markers were in silico mapped on the genome sequences to identify the positions of the EST-SSR markers on the soybean genome.

**Peanut**

Peanut (A. hypogaea), or groundnut, is an autogamous allotetraploid (2n = 4x = 40) legume species with a genome of approximately 2.8 Gb (Arunuganathan and Earle 1991). It is used for food and oil production, and its probable ancestral species have been identified as A. duranensis and A. ipaënsis. The database consists of information on DNA markers and genetic linkage maps. As the DNA markers, a total of 6,706 genome-SSR (AHGS), 3,187 EST-SSR (AHS), and 1,039 transposon insertion length polymorphism markers (AhTE) have been developed from the sequence data collected from SSR-enriched genomic libraries (accession numbers: DH964238 to DH968256) (Shirasawa et al. 2012b), cDNA libraries (accession numbers: FS960760 to FS988327) (Koilkonda et al. 2012), and transposon-enriched genomic libraries (accession numbers: DE998420 to DE998823 and DH968257 to DH968767) (Shirasawa et al. 2012a, 2012b), respectively. The genome- and EST-SSR and transposon markers were subjected to constructions of five genetic linkage maps in Arachis, SKF2 and NYF2 for cultivated peanut (A. hypogaea), AF5 and BF6 for wild diploid relatives (A. duranensis, A. stenosperma, A. ipaënsis, and A. duranensis), and TF6 or A. hypogaea and an artificial amphidiploid (A. ipaënsis × A. duranensis)46 (Shirasawa et al. 2013a). In addition, the five genetic linkage maps were integrated with 11 published maps from other research groups under collaborations between Japan, Brazil, India, France, the US, and China (Shirasawa et al. 2013a). The EST-SSR markers were employed for the genetic diversity analysis of peanut accessions, including 17 Japanese, 4 American, 2 Indian, and 1 Chinese cultivated lines as well as 6 wild relatives (Koilkonda et al. 2012).

**Red clover**

Red clover (T. pratense) is an allogamous diploid legume (2n = 2x = 14, genome size of 468 Mb: Arumuganathan and Earle 1991) that is cultivated as a forage crop. The database for red clover consists of information on the DNA markers, RFLP and SSR markers, and genetic linkage map. The RFLP markers were developed to construct a genetic linkage map in red clover. The resultant map contains 157 RFLP markers and covers 535.7 cM in total (Isobe et al. 2003). Subsequently, 7,262 SSR markers were developed from 26,356 EST sequences (Accession numbers: BB902456 to BB928811), and employed to generate a genetic linkage map consisting of 1,434 marker loci covering 868.7 cM in total (Sato et al. 2005). Finally, additional new linkage maps together with the developed genetic linkage maps were integrated into the consensus map with 1,804 marker loci covering 836.6 cM in total (Isobe et al. 2009). The resultant genetic linkage map, i.e., HR × R130, is available from this database.

**White clover**

White clover (T. repens) is an allogamous allotetraploid legume (2n = 4x = 32, genome size of 999 Mb: Arumuganathan and Earle 1991) widely cultivated as a forage crop. We generated the white clover linkage maps using SSR markers in order to conduct comparative genomics analyses among legume species (Isobe et al. 2012). In this database, a total of 1,993 primers are available for the EST-derived SSR markers. A total of 15,214 EST sequences used for primer construction are also available through the accession numbers FY454661 to FY469874.

**Eucalyptus**

E. camaldulensis is a diploid species (2n = 2x = 22, genome size of 650 Mb) that is used in the pulp industry. Therefore, the genome sequencing of E. camaldulensis and development of markers have been performed to survey the genetic information and accelerate the process of molecular breeding (Hirakawa et al. 2011: http://www.kazusa.or.jp/eucalypt/). The eucalyptus marker database consists of information on 4,656 genome- and 1,028 EST-SSR markers, which were developed from the sequence data of the transcriptome and genome of E. camaldulensis, respectively.
The SSR markers were employed for the genetic diversity analysis of six *Eucalyptus* species, i.e., *E. camaldulensis*, *E. globules*, *E. grandis*, *E. nitens* and *E. urophylla*. The PIC values based on this analysis are also available from this database.

**Marker densities in the genetic and physical maps**

Marker density, or mean distances between any neighboring marker intervals, would be important information for gene mapping studies of a map-based cloning strategy or genome-wide association studies (GWAS), and for MAS in breeding programs. On the one hand, as for the linkage maps registered in the database, the marker densities between any neighboring loci were varied from 0.4 cM in *L. japonicus* to 4.3 cM in peanut, and 1.5 cM in average over the 14 genetic maps of the six species (Table 2). On the other hand, as for the physical maps among the ten species, tomato had the highest dense marker loci due to the massive SNP data from the re-sequencing analysis (Table 3). The markers were estimated to locate in every 600 bp interval in the tomato genome. In the remaining nine species, mean physical intervals of any neighboring two markers were ranging from 64 kb in red clover to 573 kb in *Capsicum* (Table 3). While availability of whole genome sequence data at present were limited to tomato (The Tomato Genome Consortium 2012), strawberry (Hirakawa et al. 2013b), *L. japonicus* (Sato et al. 2008), soybean (Schmutz et al. 2010), and eucalyptus (Hirakawa et al. 2011) among the plant species registered in the Kazusa Marker DataBase at the time of writing, genome sequences for the other species would be determined in near future with the cooperation of the NGSs as summarized in Genomes OnLine Database (http://www.genomesonline.org). The whole genome sequence data identify the physical genome positions of the DNA markers, and provide useful information for the gene mapping studies as well as MAS.

**Future directions**

Until now, the Kazusa Marker Database includes information on DNA markers, genetic linkage maps, and physical maps for 10 plant species comprised of mainly crops. Because our research groups have been working on more than 25 plant species, the contents of this database will increase when we publish papers on each project. Databases for the DNA markers have been globally established in each crop species, institutes, and countries, a situation which is considered to be undesirable for users. To overcome the problem, an integrated database of the plant genome-related information, i.e., PGDBj (http://pgdbj.jp), has been established (Asamizu et al. 2013), which includes parts of the marker and map information registered in the Kazusa Marker DataBase. In addition, we are planning to provide graphical views of the marker positions on genome sequences or linkage maps by using GBrowse (Stein et al. 2002) or CMap (Fang et al. 2003) from the Kazusa Marker DataBase. The user-friendly interfaces will accelerate comparative analysis of QTL and GWAS loci across the plant species, which will also contribute to gene isolation and molecular breeding.

**Acknowledgements**

We thank Ms. Mitsuyo Kohara for her technical assistance. This work was supported by the KAKENHI Grant-in-Aid for Scientific Research (C) (24510286), Japan Society for the Promotion of Science; and the Kazusa DNA Research Institute Foundation.

**Literature Cited**

Arumuganathan, K. and E.D. Earle (1991) Nuclear DNA content of some important plant. Plant Mol. Biol. Rep. 9: 208–218.

Asamizu, E., H. Ichihara, A. Nakaya, Y. Nakamura, H. Hirakawa, T. Ishii, T. Tamura, K. Fukami-Kobayashi, Y. Nakajima and S. Tabata (2014) Plant Genome DataBase Japan (PGDBj): a portal website for the integration of plant genome-related databases. Plant Cell Physiol. 55: e8.

CBOL Plant Working Group (2009) A DNA barcode for land plants. Proc. Natl. Acad. Sci. USA 106: 12794–12797.

Fang, Z., M. Polacco, S. Chen, S. Schroeder, D. Hancock, H. Sanchez and E. Coe (2003) CMap: the comparative genetic map viewer. Bioinformatics 19: 416–417.

Gusella, J.F., N.S. Wexler, P.M. Conneally, S.L. Naylor, M.A. Anderson, R.E. Tanzi, P.C. Watkins, K. Ottina, M.R. Wallace, A.Y. Sakaguchi et al. (1983) A polymorphic DNA marker genetically linked to Huntington’s disease. Nature 306: 234–238.

Hamilton, J.P. and C.R. Buell (2012) Advances in plant genome

### Table 3. The estimated genome sizes and marker densities of the plant species in the Kazusa Marker DataBase

| Binomial nomenclature               | Common name            | Total no. of markers | The estimated genome size (Mb) | Mean marker density (kb/loci) |
|-------------------------------------|------------------------|----------------------|-------------------------------|------------------------------|
| *Solanum lycopersicum*              | Tomato                 | 1,500,674            | 900                           | 0.6                          |
| *Capsicum annuum*                   | Capsicum               | 5,751                | 3,300                         | 573.8                        |
| *Fragaria × ananassa*               | Strawberry             | 4,474                | 692                           | 154.7                        |
| *Raphanus sativus*                  | Radish                 | 3,811                | 526                           | 138.0                        |
| *Lotus japonicus*                   |                        | 1,155                | 472                           | 408.7                        |
| *Glycine max*                       | Soybean                | 7,020                | 1,115                         | 158.8                        |
| *Arachis hypogaea*                  | Peanut                  | 10,932               | 2,813                         | 257.3                        |
| *Trifolium pratense*                | Red clover             | 7,262                | 468                           | 64.4                         |
| *Trifolium repens*                  | White clover           | 1,993                | 999                           | 501.3                        |
| *Eucalyptus camaldulensis*          | Eucaly                | 5,684                | 650                           | 114.4                        |

The SSR markers were employed for the genetic diversity analysis of six *Eucalyptus* species, i.e., *E. camaldulensis*, *E. globules*, *E. grandis*, *E. nitens* and *E. urophylla*. The PIC values based on this analysis are also available from this database.
sequencing. Plant J. 70: 177–190.
Handberg, K. and J. Stougaard (1992) Lotus japonicus, an autogamous, diploid legume species for classical and molecular genetics. Plant J. 2: 487–496.
Hayashi, M., A. Miyahara, S. Sato, T. Kato, M. Yoshikawa, M. Taketa, M. Hayashi, A. Pedrosa, R. Onda, H. Imaizumi-Anraku et al. (2001) Construction of a genetic linkage map of the model legume Lotus japonicus using an intraspecific F_2 population. DNA Res. 8: 301–310.
Helenjaris, T., M. Slocum, S. Wright, A. Schaefer and J. Nienhuis (1986) Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. Theor. Appl. Genet. 72: 761–769.
Hirakawa, H., Y. Nakamura, T. Kaneko, S. Isobe, H. Sakai, T. Kato, T. Hibino, S. Sasamoto, A. Watanabe, M. Yamada et al. (2011) Survey of the genetic information carried in the genome of Eucalyptus camaldulensis. Plant Biotechnol. 28: 471–480.
Hirakawa, H., K. Shirasawa, A. Ohyama, H. Fukuoka, K. Aoki, C. Rothan, S. Sato, S. Isobe and S. Tabata (2013a) Genome-wide SNP genotyping to infer the effects on gene functions in tomato. DNA Res. 20: 221–233.
Hirakawa, H., K. Shirasawa, S. Kosugi, K. Tashiro, S. Nakayama, M. Yamada, M. Kohara, A. Watanabe, Y. Kishida, T. Fujishiro et al. (2013b) Dissection of the octoploid strawberry genome by deep-sequencing of the genomes of Fragaria species. DNA Res. 2: 169–181.
Hisano, H., S. Sato, S. Isobe, S. Sasamoto, T. Wada, A. Matsuno, T. Fujishiro, M. Yamada, S. Nakayama, Y. Nakamura et al. (2007) Characterization of the soybean genome using EST-derived microsatellite markers. DNA Res. 14: 271–281.
International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436: 793–800.
Isobe, S., I. Klimenko, S. Ivashuta, M. Gau and N.N. Kozlov (2003) First AFLP linkage map of red clover (Trifolium pratense L.) based on cDNA probes and its transferability to other red clover germplasms. Theor. Appl. Genet. 108: 105–112.
Isobe, S., R. Kölliker, H. Hisano, S. Sasamoto, T. Wada, I. Klimenko, K. Okumura and S. Tabata (2009) Construction of a consensus linkage map for red clover (Trifolium pratense L.). BMC Plant Biol. 9: 57.
Isobe, S.N., H. Hisano, S. Sato, H. Hirakawa, K. Okumura, K. Shirasawa, S. Sasamoto, A. Watanabe, T. Wada, Y. Kishida et al. (2012) Comparative genetic mapping and discovery of linkage disequilibrium across linkage groups in white clover (Trifolium repens L.). G3 2: 607–617.
Isobe, S.N., H. Hirakawa, S. Sato, F. Maeda, M. Ishikawa, T. Mori, Y. Yamamoto, K. Shirasawa, M. Kimura, M. Fukami et al. (2013) Construction of an integrated high density simple sequence repeat linkage map in cultivated strawberry (Fragaria × ananassa) and its applicability. DNA Res. 20: 79–92.
Koffer, R., C. Schlotterer and T. Lepley (2007) SciRoKo: a new tool for whole genome microsatellite search and investigation. Bioinformatics 23: 1683–1685.
Koikonda, P., S. Sato, S. Tabata, K. Shirasawa, H. Hirakawa, H. Sakai, S. Sasamoto, A. Watanabe, T. Wada, Y. Kishida et al. (2012) Large-scale development of expressed sequence tag-derived simple sequence repeat markers and diversity analysis in Arabidopsis spp. Mol. Breed. 30: 125–138.
Kumar, L.S. (1999) DNA markers in plant improvement: an overview. Biotechnol. Adv. 17: 143–182.
Michael, T.P. and S. Jackson (2013) The first 50 plant genomes. Plant Genome 6: doi: 10.3835/plantgenome2013.03.0001in.
Moscone, E.A., M. Barani, I. Ebert, J. Greilhuber, F. Ehrendorfer and A.T. Hunziker (2003) Analysis of nuclear DNA content in Capsicum (Solanaceae) by flow cytometry and Feulgen densitometry. Ann. Bot. 92: 21–29.
Phillips, R.L. and I.K. Vasil (2001) DNA-based markers in plants, 2nd ed. Kluwer Academic Publishers, Dordrecht, p. 512.
Rozen, S. and H. Skaletsky (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132: 365–386.
Sato, S., T. Kaneko, Y. Nakamura, E. Asamizuy, T. Kato and S. Tabata (2001) Structural analysis of a Lotus japonicus genome. I. Sequence features and mapping of fifty-six TAC clones which cover the 5.4 Mb regions of the genome. DNA Res. 8: 311–318.
Sato, S., S. Isobe, E. Asamizu, N. Ohimoto, R. Kataoka, Y. Nakamura, T. Kaneko, N. Sakurai, K. Okumura, I. Klimenko et al. (2005) Comprehensive structural analysis of the genome of red clover (Trifolium pratense L.). DNA Res. 12: 301–364.
Sato, S., Y. Nakamura, T. Kaneko, E. Asamizuy, T. Kato, M. Nakao, S. Sasamoto, A. Watanabe, A. Ono, K. Kawashima et al. (2008) Genome structure of the legume, Lotus japonicus. DNA Res. 15: 227–239.
Schmutz, 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.
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 intron 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.
Shirasawa, K., M. Oyama, H. Hirakawa, S. Sato, S. Tabata, T. Fujiokea, C. Kimizuka-Takagi, S. Sasamoto, A. Watanabe, M. Kato et al. (2011) An EST-SSR linkage map of Raphanus sativus and comparative genomics of the Brassicaceae. DNA Res. 18: 221–232.
Shirasawa, K., H. Hirakawa, S. Tabata, M. Hasegawa, H. Kiyoshima, S. Suzuki, S. Sasamoto, A. Watanabe, T. Fujishiro and S. Isobe (2012a) Characterization of active miniature inverted-repeat-transposable elements in the peanut genome. Theor. Appl. Genet. 124: 1429–1438.
Shirasawa, K., P. Koilkonka, K. Aoki, H. Hirakawa, S. Tabata, M. Watanabe, M. Hasegawa, H. Kiyoshima, S. Suzuki, C. Kuvata et al. (2012b) In silico polymorphism analysis for the development of simple sequence repeat and transposon markers and construction of linkage map in cultivated peanut. BMC Plant Biol. 12: 80.
Shirasawa, K., D.J. Bertolli, R.K. Varshney, M.C. Moretzsohn, S.C. Leal-Bertolli, M. Thudi, M.K. Pandey, J.F. Rami, D. Foncèka, M.V. Gowda et al. (2013a) Integrated consensus map of cultivated peanut and wild relatives reveals structures of the A and B genomes of Arachis and divergence of the legume genomes. DNA Res. 20: 173–184.
Shirasawa, K., H. Fukuoka, M. Hatsuanaa, Y. Kobayashiy, I. Kobayashiy, H. Hirakawa, S. Isobe and S. Tabata (2013b) Genome-wide association studies using single nucleotide polymorphism markers developed by re-sequencing of the genomes of cultivated tomato. DNA Res. 20: 593–603.
Shirasawa, K., K. Ishii, C. Kim, T. Ban, M. Suzuki, T. Ito, T. Muranaka, M. Kobayashi, N. Nagata, S. Isobe et al. (2013c) Development of Capsicum EST-SSR markers for species identification and in silico...
mapping onto the tomato genome sequence. Mol. Breed. 31: 101–
110.
Shirasawa, K. and H. Hirakawa (2013) DNA marker applications to
molecular genetics and genomics in tomato. Breed. Sci. 63: 21–30.
Shulaev, V., D.J. Sargent, R.N. Crowhurst, T.C. Mockler, O. Folkerts,
A.L. Delcher, P. Jaiswal, K. Mockaitis, A. Liston, S.P. Mane et al.
(2011) The genome of woodland strawberry (Fragaria vesca). Nat.
Genet. 43: 109–116.
Stein, L.D., C. Mungall, S. Shu, M. Caudy, M. Mangone, A. Day,
E. Nicherson, J.E. Stajich, T.W. Harris, A. Arva et al. (2002) The
generic genome browser: a building block for a model organism
system database. Genome Res. 12: 1599–1610.
Temnykh, S., G. DeClerck, A. Lukashova, L. Lipovich, S. Cartinhour
and S. McCouch (2001) Computational and experimental analysis
of microsatellites in rice (Oryza sativa L.): frequency, length vari-
tion, transposon associations, and genetic marker potential.
Genome Res. 11: 1441–1452.
The Arabidopsis Genome Initiative (2000) Analysis of the genome se-
quence 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 Tomato Genome Consortium (2012) The tomato genome se-
quency provides insights into fleshy fruit evolution. Nature 485:
635–641.
Thiel, T., W. Michalek, R.K. Varshney and A. Graner (2003) Exploiting
EST databases for the development and characterization of gene-
derived SSR-markers in barley (Hordeum vulgare L.). Theor. Appl.
Genet. 106: 411–422.
Wu, F. and S.D. Tanksley (2010) Chromosomal evolution in the plant
family Solanaceae. BMC Genomics 11: 182.