Invited Review

Advances of Whole Genome Sequencing in Strawberry with NGS Technologies

Sachiko Isobe*, Kenta Shirasawa and Hideki Hirakawa

Kazusa DNA Research Institute, Kisarazu 292-0818, Japan

Next generation sequencing (NGS) is one of the most impactful technologies to appear in the 21st century, and has already brought important changes to agriculture, especially in the field of breeding. Construction of a reference genome is key to the advancement of genomic studies, and therefore, de novo whole genome assembly has been performed in various plants, including strawberry. Strawberry (Fragaria × ananassa) is an allopolyploid species (2n = 8x = 56), which has four discriminable subgenomes. Because of its complex genome structure, de novo whole genome assembly in strawberry has been considered a difficult challenge. However, recent advances of NGS technologies have allowed the construction of chromosome-scale de novo whole genome assembly. In this manuscript, we review the recent advances in de novo whole genome sequencing in strawberry and other Fragaria species. The genome structure and domestication history in strawberry is one of the largest questions in genetic and genomic studies in strawberry. Therefore, the domestication history in strawberry is also be reviewed based on comparisons of genes and genome sequences across Fragaria species.

Key Words: Fragaria × ananassa, genus Fragaria, Next generation sequencing.

Introduction

Next generation sequencing (NGS) is one of the most impactful technologies to appear in the 21st century. At present, NGS technology has contributed more to medicine than to agriculture, but important changes to agriculture are clearly commencing, especially in the field of breeding. The term “next generation” is an umbrella term referring to various sequencing technologies related to the Sanger sequencing method (Yang et al., 2014). Briefly, there are two generations of these technologies, known as the 2nd and 3rd generations. The 2nd generation NGS technologies includes platforms such as Illumina (https://www.illumina.com/), Ion Torrent (https://www.thermofisher.com/) and MGI (https://en.mgitech.cn/), and generate large number of sequence reads with short length. Currently, several sequencers appeared in the beginning of 2nd generation were went out of production, such as Roche 454 (https://www.roche.com/) and Life Technologies SOLiD (https://www.thermofisher.com/). The 3rd generation NGS technologies, which include PacBio (https://www.pacb.com/) and Oxford Nanopore Technologies (ONT, https://nanoporetech.com/), generate long sequence reads with a medium number of reads (Cao et al., 2017). Because the two platforms sequencing a single molecule DNA, the 3rd generation NGS is also called as single molecule sequencing. The quality of de novo whole genome assembly has been improved along with the advances of NGS technologies from the 2nd to 3rd generation, though the 2nd generation technologies are still the main platforms for re-sequencing and gene expression analysis. In addition, chromosome-scale assembly has become less difficult through the assistance of other NGS technologies—namely, optical mapping with the Bionano platform (https://bionanogenomics.com/) and libraries constructed with long DNA fragments such as 10X Genomics (https://www.10xgenomics.com/) and Hi-C (Lieberman-Aiden et al., 2009).

Construction of a reference genome is key to the advancement of genomic studies, and therefore, since the appearance of NGS technologies, de novo whole genome assembly has been performed in various plant species. According to Chen et al. (2019), a total of 181 crop horticultural species had been sequenced by December 31, 2018. Nevertheless, de novo whole genome assembly in polyploidy species is still challenging, because homoeologous chromosomes often cause chimeric assembly. The first polyploid genome assembly was amphidiploid paleopolyploid in soybean (Schmutz et al., 2010), after which the assembly of polyploid genomes expanded to other species, including strawberry. Kyriakidou et al. (2018) listed 47 sequenced
polyploid genomes reported by May 2018. In polyploid species, genome sequencing in ancestor diploid species was often performed prior to the targeted polyploid genome sequencing to construct a reference, such as for banana (D’Hont et al., 2012), cotton (Wang et al., 2012), wheat (Ling et al., 2013), and peanut (Bertioli et al., 2016).

Cultivated strawberry (Fragaria × ananassa) is an octoploid species (2n = 8X = 56), originated in the 18th century by artificial crosses between two octoploid species, F. chiloensis and F. virginiana (Darrow, 1966). The genome structure of strawberry was initially considered to be allo-autopolyploidy with AABBBBCC (Federova, 1946) or AA AAA’A’BBBB models (Senanayake and Bringhurst, 1967) based on cytological studies. An allopolyploidy model (AAA’A’BBBB’B’) was then proposed based on isozyme segregation (Bringhurst, 1990) and later supported by the segregation patterns of DNA markers (Isobe et al., 2013; Kunihisa, 2011). Genus Fragaria contains 25 species comprising 13 diploids, 5 tetraploids, 1 hexaploid, 5 octoploids, and 1 decaploid (Hummer et al., 2009; Staudt, 2009). Although discussions regarding the ancestors in strawberry are ongoing, F. vesca and F. inumae are currently considered the most likely diploid progenitors (Rousseau-Gueutin et al., 2009). Therefore, the F. vesca genome was first sequenced as a reference in genetic and genomic analysis (Shulaev et al., 2011). In addition, recent advances in NGS technologies have led to the successful construction of an octoploid strawberry genome. In this manuscript, we review the recent advances in de novo whole genome sequencing in strawberry and other Fragaria species.

**De novo whole genome assembly in F. vesca**

De novo whole genome sequencing in F. vesca was reported by Shulaev et al. (2011) as a first reference genome in genus Fragaria. The genomes of a fourth-generation inbred line, ‘Hawaii 4’, were sequenced by Roche 454, Illumina Solexa and Life Technologies SOLiD with 39x coverage, resulting in the generation of 3,263 scaffolds (fvesca_v1.0_scaffolds) (Table 1). The total and N50 lengths of the scaffolds were 201.9 Mb and 1.36 Mb, respectively. The assembled scaffolds were aligned on a diploid Fragaria reference map, and 204 scaffolds were successfully anchored with a total length of 198.1 Mb as pseudomolecules (fvesca_v1.0_pseudo). Meanwhile, gene prediction was performed on the 3,263 scaffolds, and a total of 34,809 gene models were generated with an average length of 1,160 bp. A chloroplast genome was also constructed and 78 protein, 30 tRNA and 4 rRNA genes on the genome were predicted. The genome sequences and predicted genes are available in the Genome Database for Rosaceae (GDR, https://www.rosaceae.org/species/fragaria/all) (Jung et al., 2019).

The predicted genes were then re-annotated with transcript sequences derived from 25 different tissues with two biological replicates in the F. vesca cultivar ‘YW5AF7’ (Darwish et al., 2015). As a result, a total of 32,831 gene models were predicted with a total coding length of 38.8 Mb. The set of revised sequences was named version (v) 1.1. Genome assembly was also revised based on F. vesca linkage maps (Tennesen et al., 2013, 2014) constructed based on targeted sequence capture, which enabled mapping of 200 bp sequences including polymorphic sites. The v1.1 scaffold sequences were anchored via captured sequences located on the linkage maps, and conflict v1.1 scaffolds were broken and re-ordered according to the linkage maps. The revised version of F. vesca genome was designated v2.0.a1 (Tennesen et al., 2014).

The latest version of F. vesca genome at present (December 2019) is v4.0.a1 (Edger et al., 2018). Scaffolds were re-constructed with genomes of ‘Hawaii 4’ sequenced by PacBio. Chromosome-scale assembly was performed by hybrid assembly with Bionano optical mapping. The number and total length of the assembled scaffolds are 31 and 220.8 Mb, respectively, with an N50 length of 36.1 Mb. Of the 31 scaffolds, 9 cover 99.8% of the assembly, and these 9 scaffolds are anchored to 7 pseudomolecules by Bionano optical maps. A total of 24.96 Mb sequences were newly generated on the v4.0.a1 genome. Improved gene prediction and annotation analysis led to identification of an additional 1,496 genes. The pseudomolecule sequences are available on the GDR database.

**De novo whole genome assembly in cultivated strawberry (F. × ananassa)**

De novo whole genome assembly was first reported by Hirakawa et al. (2014). Genome sequences of the Japanese strawberry variety ‘Reikou’ were obtained by Illumina GAIIx, HiSeq, and Roche 454. This was a pioneering work in polyploid genome sequencing, and algorithms for the discrimination of multiple heterozygous genomes were not well established when the work started. Therefore, two types of assembly, namely FAN_r1.1 and FANhybrid_r1.2, were constructed (Table 1). FAN_r1.1 is an assembly sequence generated from Illumina reads that represents the entire genome sequence in ‘Reikou’. FAN_r1.1 consists of 625,966 sequences, comprising a total of 697.8 Mb with an N50 length of 2,201 bp. The genome size of ‘Reikou’ was estimated as 692 Mb based on kmer frequency analysis with Illumina reads in the Hirakawa et al. (2014) manuscript. However, our later kmer frequency analysis with additional Illumina reads suggested that the genome size of ‘Reikou’ was 811 Mb (not published). The total length of FAN_r1.1 is almost the same as the estimated genome size based on kmer frequency analysis (811 Mb), suggesting that the assembled genome covers the whole genome of ‘Reikou’. FANhybrid_r1.2 was generated as a virtual reference...
Genome (haploid genome) that integrates the genome sequences of homoeologous chromosomes by eliminating heterozygous bases. Both Illumina and Roche 454 sequences were used for FANhybrid_r1.2. The number of sequences in FANhybrid_r1.2 is 211,588, with a total and an N50 length of 173.2 Mb and 5,137 bp, respectively. Although the genome sequences created in this study were fragmented, the first assembled strawberry genome sequences contributed to subsequent molecular genetics and gene expression analyses (Landi et al., 2017; Wada et al., 2017). The two assembled genome sequences are available on Strawberry GARDEN (http://strawberry-garden.kazusa.or.jp/, Fig. 1) and GDR. Gene sequences and marker information are also included in the Strawberry GARDEN. The available functions in Plant GARDEN are keyword search, Blast search and extraction of sequences. Genome and gene sequences are displayed on JBrowse.

### Table 1. Statistics of assembled Fragaria genomes.

| Species | F. vesca | F. × ananassa | F. iinumae | F. nubicola | F. orientalis |
|---------|----------|---------------|------------|-------------|--------------|
| Sequence name | fvesca_v1.0_scaffolds | FAN_r1.1 | FII_r1.1 | FNI_r1.1 | FOR_r1.1 |
| Assembled line | Hawaii 4 | Reikou | 221 | 265.6 | — |
| Estimated genome size (Mb) | ~240 | 811 | 208 | 208 | 208 |
| Assembly level | Scaffolds | Scaffolds | Scaffolds | Scaffolds | Scaffolds |
| Number of sequences | 3,263 | 625,966 | 117,822 | 117,822 | 3,309 |
| Total length (Mb) | 201.9 | 211,588 | 215,024 | 215,024 | 206.4 |
| N50 length (bp) | 1,361,426 | 30,775,793 | 210,780 | 210,780 | 1,275 |
| Sequencing platforms | Roche 454, Illumina Solexa, Life Technologies SOLiD | fvesca_v1.0_scaffolds with a linkage map | Illumina GAIx, HiSeq, Roche 454 | Illumina GAIx, HiSeq, Roche 454 | Illumina HiSeq, 10X Genomics, linkage map |
| Reference | Shulaev et al. (2011) | — | Hhirakawa et al. (2014) | Hhirakawa et al. (2014) | Hhirakawa et al. (2014) |

* Germplasm provided by Prof. Yanagi Tomohiro of Kagawa University, Japan.
of all reads against the genomes of strawberry (FAN_r1.1) and other organisms, including humans, bacteria, and nematodes, in order to estimate the ratio of contamination in each library. As a result, 144 of the 288 libraries were confirmed to include more than 50% of the strawberry sequences. Genome and gene sequences are displayed on JBrowse.

To resolve the complexity of genome sequencing in polyploid species, Yanagi et al. (2017) proposed a micromanipulating approach for genome sequencing of single somatic chromosomes. Single chromosomes in strawberry ‘Reikou’ were dissected under a light microscope with a glass needle, and DNA from the 288 dissected chromosomes were individually amplified for genome sequencing by the Illumina platform. Because the amount of DNA in a single chromosome was extremely small, other genomes in air often contaminated the process of library construction. Therefore, each of the 1000 reads in the libraries were sampled for a BLAST search against the genomes of strawberry (FAN_r1.1) and other organisms, including humans, bacteria, and nematodes, in order to estimate the ratio of contamination in each library. As a result, 144 of the 288 libraries were confirmed to include more than 50% of the strawberry reads in the libraries. According to the BLAST search of all reads against the F. vesca genome (v2.0a1), the 144 libraries were classified into the 7 pseudomolecules on the F. vesca genome. Although there are still a few issues to be solved, such as sequence bias caused by whole genome amplification by PCR, the approach is expected to contribute to elucidation of the structure of the complex strawberry genome.

Chromosome-scale strawberry genome assembly was then reported by Edgar et al. (2019b) based on scaffolds constructed by NovoMAGIC 3 (NRGene, https://www.nrgene.com/solutions/denovomagic/). NovoMAGIC is a genome assembly method only available as a commercial service provided by NRGene. The assembly is done based on Illumina reads with paired end (PE), mate pair (MP), and 10X Genomics libraries, and just recently PacBio reads have also been included. NovoMAGIC has an excellent function that discriminates multiple heterozygous genomes in polyploid species, and hence it has been used for polyploid genome assembly in crops such as wheat (Avni et al., 2017). The genome of a strawberry variety, ‘Camarosa’, was sequenced by Illumina, 10X Genomics and PacBio platforms with a total of 65x coverage for DenovoMAGIC 3 assembly, and Hi-C clustering was performed to generate chromosome-scale scaffolds. The assembled genome sequences (Camarosa_Genome_v1.0.a1) consisted of 28 pseudomolecules with a total and N50 lengths of 805.5 Mb and 27.8 Mb, respectively. The genome size of ‘Camarosa’ was estimated as 813 Mb by flow cytometry analysis, and the assembled genome sequences covered 99% of the genome. A total of 108,087 protein coding genes in 341.4 Mb were predicted on the Camarosa_Genome_v1.0.a1 genome. The sequences are available on the GDR database.

Chromosome-scale scaffolds were also constructed for the ‘Reikou’ genome (Table 1, in preparation). Assembly was performed by DenovoMAGIC 3 with Illumina PE, MP, and 10X Genomics reads. The total length of the assembled sequences was 1,406 Mb, and consisted of 32,715 sequences and an N50 length of 3.9 Mb. Because the scaffolds were generated for all haplotype genomes in ‘Reikou’, so-called “phased genome” in the DenovoMAGIC analysis, the length was almost double the estimated genome size. The assembled genome was designated as FAN_r2.3. The scaffolds were aligned on a high density linkage map constructed by an S1 mapping population of ‘Reikou’ (Nagano et al., 2017) with an IStraw 90K Axiom® Array (Bassil et al., 2015) and SSR markers (Isobe et al., 2013). The ‘Reikou’ linkage map consisted of 31 linkage groups, and a total of 62 (31 × 2 haplotypes) pseudomolecules were developed with a total length of 1,733.5 Mb. The FAN_r2.3 sequence is available on the download site in Strawberry GARDEN.

**De novo whole genome assembly in other Fragaria spp.**

The genomes of three diploids, F. iinumae, F. nipponica, F. nubicola, and one tetraploid, F. orientalis, were sequenced by Illumina HiSeq (Hirakawa et al., 2014). De novo assemblies were performed by SOAPdenovo v1.05 and GapCloser 1.10, and generated scaffolds with a total length of 200–206 Mb in diploid species, and 214 Mb in tetraploid F. orientalis (Table 1). The assembled genome sequences are available in the Strawberry GARDEN and GDR databases.

Chromosome-scale genome assembly in F. iinumae was recently reported by Edgar et al. (2019a). PacBio reads were obtained with 172x coverage and assembled by FALCON v0.3.0. The generated scaffolds were polished with Illumina reads by Pilon and then anchored to a F. iinumae map. As a result, eight chromosome-scale pseudomolecules were generated. The total length of the assembled sequences was 265.6 Mb, and consisted
of eight scaffolds and an N50 length of 33.98 Mb. The assembled genome sequences are available in the GDR databases.

**Clarification of domestication history in strawberry with a genomic approach**

The genome structure and domestication history are among the major unresolved issues in genetic and genomic studies in strawberry. Tennessen et al. (2014) discussed the evolutionary origins of strawberry based on dense linkage maps constructed with targeted capture sequences by POLiMAPS. By comparing strawberry and diploid *Fragaria* sequences, they concluded that one of the four subgenomes in strawberry originated from *F. vesca*, another was from *F. iinumae*, and the remaining two subgenomes were from an unknown ancestor close to *F. iinumae*. Their findings agreed with the results of a study by Hirakawa et al. (2014), which compared the strawberry genome (FANhybrid_r1.2) with the genomes of five *Fragaria* spp., *F. vesca*, *F. iinumae*, *F. nipponica*, *F. nubicola*, and *F. orientalis*. Based on the SNP-haplotype comparison, Sargent et al. (2016) also proposed a similar model, A-A, b-b, X-X, in which A, b, and X represented *F. vesca*-like, *F. iinumae*-like, and an unknown donor close to the *F. iinumae* genome, respectively.

Meanwhile, Yang and Davis (2017) performed phylogenetic analysis in *Fragaria* spp. using the Fluidigm Access Array system and Roche 454 platform by comparing 24 single or low copy nuclear genes in 16 *Fragaria* species. They hypothesized that genomes in octoploid species were contributed by at least four diploid species (*F. vesca*, *F. iinumae*, *F. bucharica*, *F. viridis*) and one unknown allele donor.

After construction of a chromosome-scale reference genome, Edger et al. (2019b) performed *de novo* transcriptome assembly in every described diploid *Fragaria* species. Phylogenetic analysis based on 19,302 nuclear genes with strawberry and other *Fragaria* species suggested that *F. iinumae* and *F. nipponica*, which are endemic species in Japan, were progenitors. Edger and colleagues hypothesized that the two species (*F. iinumae* and *F. nipponica*) donated genomes for generating five tetraploid species in China. The third candidate diploid ancestor is *F. viridis*, which distributes in Europe and Asia, and partially overlaps its distribution with a hexaploid *F. moschata* in East Eurasia. It was assumed that *F. moschata* originated from a hybridization between *F. viridis* and a tetraploid derived from *F. iinumae* and *F. nipponica*. The fourth candidate ancestor is *F. vesca* subsp. *bracteata*, which is an endemic species in North America. Among these four candidate ancestors, diploid species in hand, it was hypothesized that octoploid species were generated in North America by integration of a hexaploid generated in East Eurasia and an *F. vesca* subsp. *bracteata*. Edger et al. (2019b) also explored the potential existence of a “dominant” subgenome, which would have greater gene contents and higher gene expression, based on the abundance and distribution of transposable elements (TE), and concluded that a dominant subgenome was contributed by the *F. vesca* progenitor.

**Conclusion**

Only a decade ago, the construction of a chromosome-scale reference genome in strawberry was only a dream. Now that dream is a reality, and we have new basic information that can be used to answer a host of questions in this crop. Edger et al. (2019a) put forward a hypothesis in regard to the origin of strawberry, but further studies will be needed to verify their version of the domestication history. For example, Liston et al. (2019) immediately argued against the hypothesis based on the results of a chromosome-scale phylogenomic analysis and phylogenetic analysis based on an *F. moschata* linkage map. They averred that Edger et al. (2019b) drew the wrong conclusion from their gene-scale phylogenetic analysis, and proposed that while *F. vesca* and *F. iinumae* were possible ancestors in strawberry, *F. nipponica*, *F. viridis*, and *F. moschata* were not. However, shortly after this refutation, Edger et al. (2019a) countered with the construction of a chromosome-scale reference genome in *F. iinumae* for chromosome-scale comparison in multiple *Fragaria* species. NGS technology will continue to fuel such conversation, and to generate further chromosome-scale assembled genomes and accurate gene prediction. It is expected that the development of high quality assembled genomes will in turn impact evolutional studies as well as genomics, molecular genetics and breeding in strawberry.

**Literature Cited**

Avni, R., M. Nave, O. Barad, K. Baruch, S. O. Twardziok, H. Gundlach, I. Hale, M. Mascher, M. Spannagl, K. Wiebe, K. W. Jordan, G. Golan, J. Deck, B. Ben-Zvi, G. Ben-Zvi, A. Himmelbach, R. P. MacLachlan, A. G. Sharpe, A. Fritz, R. Ben-David, H. Budak, T. Fahima, A. Korol, I. D. Faris, A. Hernandez, M. A. Mikel, A. A. Levy, B. Steffenson, M. Maccaferri, R. Tuberosa, L. Cattivelli, L. V. Faccioli, A. Cieriotti, K. Kashkush, M. Pourkieharianish, T. Komatsuda, T. Eilam, H. Sela, A. Sharon, N. Ohad, D. A. Chamovitz, K. F. X. Mayer, N. Stein, G. Ronen, Z. Peleg, C. J. Pozniak, E. D. Akhunov and A. Distelfeld. 2017. Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. Science 357: 93–97.

Bassil, N. V., T. M. Davis, H. Zhang, S. Ficklin, M. Mittmann, T. Webster, L. Mahoney, D. Wood, E. S. Alperin, U. R. Rosyara, H. Koehler-Vanc Putten, A. Monfort, D. J. Sargent, I. Amaya, B. Denoyes, L. Bianco, T. van Dijk, A. Pirani, A. Iezzoni, D. Main, C. Peace, Y. Yang, V. Whitaker, S. Verma, L. Bellon, F. Brew, R. Herrera and E. van de Weg. 2015. Development and preliminary evaluation of a 90 K Axiom® SNP array for the allo-octoploid cultivated strawberry *Fragaria × ananassa*. BMC Genomics 16: 155. DOI: 10.1186/s12864-015-1310-1.

Bertioli, D. J., S. B. Cannon, L. Froenicke, G. Huang, A. D.
Edger, P. P., R. VanBuren, M. A. Hardigan, M. Colle, M. R. McKain, R. D. Smith, S. J. Teresi, A. D. L. Nelson, C. M. Wai, E. I. Alger, K. A. Bird, A. E. Yocca, N. Jackson, R. Michelmore and P. Ozias-Akins. 2016. The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. Nat. Genet. 48: 438–446.

Brinthurst, R. S. 1990. Cytogenetics and evolution in American *Fragaria*. Hort. Sci. 106: 679–683.

Cao, Y., S. Fanning, S. Proos, K. Jordan and S. Srikumar. 2017. A berry genome. Nat. Genet. 51: 541–547.

Ashman. 2019. Revisiting the origin of octoploid strawberry. Nat. Genet. 52: 2–4.

*Hirakawa*, H., K. Shirasawa, S. Kosugi, K. Tashiro, S. Nakayama, M. Yamada, M. Kohara, A. Watanabe, Y. Kishida, T. Fujishiro, H. Tsurowka, C. Minami, S. Sasamoto, M. Kato, K. Nanri, A. Komaki, T. Yanagi, Q. Guoxin, F. Maeda, M. Ishikawa, S. Kuhara, S. Sato, S. Tabata and S. N. Isobe. 2014. Dissection of the octoploid strawberry genome by deep sequencing of the genomes of *Fragaria* species. DNA Res. 21: 169–181.

Hummer, K. E., P. Nathewet and T. Yanagi. 2009. Decaploidy in *Fragaria iturupensis* (Rosaceae). Am. J. Bot. 96: 713–716.

Isobe, S. N., H. Hirakawa, S. Sato, F. Maeda, M. Ishikawa, T. Mori, Y. Yamamoto, K. Shirasawa, M. Kimura, M. Fukami, F. Hashizume, T. Tsuji, S. Sasamoto, M. Kato, K. Nanri, H. Tsurowka, C. Minami, C. Takahashi, T. Wada, A. Ono, K. Kawashima, N. Nakazaki, Y. Kishida, M. Kohara, S. Nakayama, M. Yamada, T. Fujishiro, A. Watanabe and S. Tabata. 2013. Construction of an integrated high density simple sequence repeat linkage map in cultivated strawberry (*Fragaria × annanassa*) and its applicability. DNA Res. 20: 79–92.

Jung, S., T. Lee, C.-H. Cheng, K. Buble, P. Zheng, J. Yu, J. Humann, S. P. Picklin, K. Gasic, K. Scott, M. Frank, S. Ru, H. Hough, K. Evans, C. Peace, M. Olmstead, L. W. DeVetter, J. McFerson, M. Coe, J. L. Weizryn, M. E. Staton, A. G. Abbott and D. Main. 2019. 15 years of GDR: New data and functionality in the Genome Database for Rosaceae. Nucleic Acids Res. 47: D1137–D1145.

Kunihisa, M. 2011. Studies using DNA markers in *Fragaria × annanassa*: genetic analysis, genome structure, and cultivar identification. J. Japan. Soc. Hort. Sci. 80: 231–243.

Kyrtiadikou, M., H. H. Tai, N. L. Anglin, D. Ellis and M. V. Stromvik. 2018. Current strategies of polyploid plant genome sequence assembly. Front. Plant. Sci. 9: 1660. DOI: 10.3389/fpls.2018.01660.

Landi, L., R. M. De Miccolis Angelini, S. Pollastro, E. Feliziani, F. Faretra and G. Romanazzi. 2017. Global transcriptome analysis and identification of differentially expressed genes in strawberry after preharvest application of Benzothiadiazole and Chitosan. Front. Plant. Sci. 8: 235. DOI: 10.3389/fpls.2017.00235.

Lieberman-Aiden, E., N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Giri, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander and J. Dekker. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326: 289–293.

Ling, H. Q., S. Zhao, D. Liu, J. Wang, H. Sun, C. Zhang, H. Fan, D. Li, L. Dong, Y. Tao, C. Gao, H. Wu, Y. Li, Y. Cui, X. Guo, S. Zheng, B. Wang, K. Yu, Q. Liang, W. Yang, X. Lou, J. Chen, M. Feng, J. Jian, X. Zhang, G. Luo, Y. Jiang, J. Liu, Z. Wang, Y. Sha, B. Zhang, H. Wu, D. Tang, Q. Shen, P. Xue, S. Zou, X. Wang, L. Liu, F. Wang, Y. Yang, X. An, Z. Dong, K. Zhang, X. Zhang, M. C. Luo, J. Dvorak, Y. Tong, J. Wang, H. Yang, Z. Li, D. Wang, A. Zhang and J. Wang. 2013. Draft genome of the wheat A-genome progenitor *Triticum urartu*. Nature 496: 87–90.

Liston, A., N. Wei, J. A. Tennessen, J. Li, M. Dong and T.-L. Ashman. 2019. Revisiting the origin of octoploid strawberry. Nat. Genet. 52: 2–4.

Nagano, S., K. Shirasawa, H. Hirakawa, F. Maeda, M. Ishikawa and S. N. Isobe. 2017. Discrimination of candidate subgenome-specific loci by linkage map construction with an *S*1 population of octoploid strawberry (*Fragaria × ananassa*) and its applicability. DNA Res. 20: 79–92.

*Tomato* genome. Nat. Genet. 48: 438–446.
Rousseau-Gueutin, M., A. Gaston, A. Ainouche, M. L. Ainouche, K. Olbricht, G. Staudt, L. Richard and B. Denoyes-Rothan. 2009. Tracking the evolutionary history of polyploidy in *Fragaria* L. (strawberry): new insights from phylogenetic analyses of low-copy nuclear genes. Mol. Phylo. and Evol. 51: 515–530.

Sargent, D. J., Y. Yang, N. Surbanovski, L. Bianco, M. Buti, R. Velasco, L. Giongo and T. M. Davis. 2016. HaploSNP affinities and linkage map positions illuminate subgenome composition in the octoploid, cultivated strawberry (*Fragaria × ananassa*). Plant Sci. 242: 140–150.

Schmutz, J., S. B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D. L. Hyten, Q. Song, J. J. Thelen and J. Cheng. 2010. Genome sequence of the paleopolyploid soybean. Nature 463: 178–183.

Senanayake, Y. D. A. and R. S. Bringhurst. 1967. Origin of *Fragaria* Polyploids. I. Cytological Analysis. American J. of Botany 54: 221–228.

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, P. Burns, T. M. Davis, J. P. Slovin, N. Bassil, R. P. Hellens, C. Evans, T. Harkins, C. Kodira, B. Desany, O. R. Crasta, R. V. Jensen, A. C. Allan, T. P. Michael, J. C. Setubal, J. M. Celton, D. J. Rees, K. P. Williams, S. H. Holt, J. J. Ruiz Rojas, M. Chatterjee, B. Liu, H. Silva, L. Meisel, A. Adato, S. A. Filipchin, M. Troggio, R. Viola, T. L. Ashman, H. Wang, P. Dharmawardhana, J. Elser, R. Raja, H. D. Priest, D. W. Bryant, Jr., S. E. Fox, S. A. Givan, L. J. Wilhelm, S. Nathani, A. Christoffels, D. Y. Salama, J. Carter, E. Lopez Girona, A. Zdepski, W. Wang, R. A. Kerstetter, W. Schwab, S. S. Korban, J. Davik, A. Monfort, B. Denoyes-Rothan, P. Arus, R. Mittler, B. Flinn, A. Aharoni, J. L. Bennetzen, S. L. Salzberg, A. W. Dickerman, R. Velasco, M. Borodovsky, R. E. Veilleux and K. M. Folta. 2011. The genome of woodland strawberry (*Fragaria vesca*). Nat. Genet. 43: 109–116.

Staudt, G. 2009. Strawberry biogeography, genetics and systematics. Acta Hortic. 842: 71–84.

Tennessen, J. A., R. Govindaraju, T. L. Ashman and A. Liston. 2014. Evolutionary origins and dynamics of octoploid strawberry subgenomes revealed by dense targeted capture linkage maps. Genome Biol. Evol. 6: 3295–3313.

Tennessen, J. A., R. Govindaraju, A. Liston and T. L. Ashman. 2013. Targeted sequence capture provides insight into genome structure and genetics of male sterility in a gynodioecious diploid strawberry, *Fragaria vesca* ssp. bracteata (Rosaceae). G3 3: 1341–1351.

Wada, T., K. Oku, S. Nagano, S. Isobe, H. Suzuki, M. Mori, K. Takata, C. Hirata, K. Shimomura, M. Tsubone, T. Katayama, K. Hirashima, Y. Uchimura, H. Ikegami, T. Suyoshi, K. I. Obu, T. Hayashida and Y. Shibato. 2017. Development and characterization of a strawberry MAGIC population derived from crosses with six strawberry cultivars. Breed. Sci. 67: 370–381.

Wang, K., Z. Wang, F. Li, W. Ye, J. Wang, G. Song, Z. Yue, L. Cong, H. Shang, S. Zhu, C. Zou, Q. Li, Y. Yuan, C. Lu, H. Wei, C. Gou, Z. Zheng, Y. Yin, X. Zhang, K. Liu, B. Wang, C. Song, N. Shi, R. J. Kohel, R. G. Percy, J. Z. Yu, Y. X. Zhu, J. Wang and S. Yu. 2012. The draft genome of a diploid cotton *Gossypium raimondii*. Nat. Genet. 44: 1098–1103.

Yanagi, T., K. Shirasawa, M. Terachi and S. Isobe. 2017. Sequence analysis of cultivated strawberry (*Fragaria × ananassa Duch.*) using microdissected single somatic chromosomes. Plant Methods 13: 91. DOI: 10.1186/s13007-017-0237-8.

Yang, Y. and T. M. Davis. 2017. A new perspective on polyploid *Fragaria* (Strawberry) genome composition based on large-scale, multi-locus phylogenetic analysis. Genome Biol. Evol. 9: 3433–3448.

Yang, Y., B. Xie and J. Yan. 2014. Application of next-generation sequencing technology in forensic science. Genomics Proteomics Bioinformatics 12: 190–197.