

Simple Sequence Repeat Marker Development and Mapping Targeted to Previously Unmapped Regions of the Strawberry Genome Sequence

Daniel J. Sargent,* Paulina Kuchta, Elena Lopez Girona, Hailong Zhang, Thomas M. Davis, Jean-Marc Celton, Annalisa Marchese, Malgorzata Korbin, Kevin M. Folta, Vladimir Shulaev, and David W. Simpson

Abstract

The genome sequence of the woodland strawberry (*Fragaria vesca* L.) is an important resource providing a reference for comparative genomics studies and future sequenced rosaceous species and has great utility as a model for the development of markers for mapping in the cultivated strawberry *Fragaria xananassa* Duchesne ex Rozier. A set of 152 microsatellite simple sequence repeat (SSR) primer pairs was developed and mapped, along with 42 previously published but unmapped SSRs, permitting the precise assignment of 28.2 Mbp of previously unanchored genome sequence scaffolds (13% of the *F. vesca* genome sequence). The original ordering of *F. vesca* sequence scaffolds was performed without a physical map, using predominantly SSR markers to order scaffolds via anchoring to a comprehensive linkage map. This report complements and expands resolution of the *Fragaria* spp. reference map and refines the scaffold ordering of the *F. vesca* genome sequence using newly devised tools. The results of this study provide two significant resources: (i) the concurrent validation of a substantial set of SSRs associated with these previously unmapped regions of the *Fragaria* spp. genome and (ii) the precise placement of previously orphaned genomic sequence. Together, these resources improve the resolution and completeness of the strawberry genome sequence, making it a better resource for downstream studies in *Fragaria* spp. and the family Rosaceae.

Fragaria is the most commercially important soft fruit genus, primarily due to the cultivation of the genetically complex octoploid species *F. xananassa* ($2n = 8x = 56$). In 2009, the world production of strawberries exceeded 4.1 million t and the crop was valued in excess of US\$4 billion (FAO, 2011). Due to its economic importance, *F. xananassa* has been the subject of much genetic research aimed at developing superior cultivars with enhanced disease resistance, fruit quality, and other characters, prompting the development of a number of molecular marker maps for this species (Rousseau-Gueutin et al., 2008; Sargent et al., 2009; van Dijk et al., 2010).

Simple sequence repeats (SSRs) have been the marker of choice in the genus *Fragaria* for linkage map development to date (Rousseau-Gueutin et al., 2008; Sargent et al., 2006; Sargent et al., 2009; Spigler et al., 2010; van Dijk et al., 2010) due to their abundance in the genome, their codominant and highly polymorphic nature, and their relative ease of development from enriched genomic libraries and expressed sequence tag (EST) collections (Celton et al., 2009; Sargent et al., 2003). The diploid *Fragaria* spp. reference map has been constructed using

D.J. Sargent, E. Lopez Girona, A. Marchese, and D. W. Simpson, East Malling Research, New Road, East Malling, Kent ME19 6BJ, UK; P. Kuchta and M. Korbin, Research Institute of Horticulture, Konstytucji 3 Maja 1/3, 96-100 Skierniewice, Poland; H. Zhang, The Glycomics Center, Univ. of New Hampshire, Durham, NH; T.M. Davis, Dep. of Biological Sciences, University of New Hampshire, Durham, NH; J.M. Celton, Biotechnology Dep., Univ. of the Western Cape, Private Bag X17, Bellville 7535, South Africa; K. M. Folta, Horticultural Sciences Dep., Univ. of Florida, Gainesville, FL; V. Shulaev, Dep. of Biological Sciences, Univ. of North Texas, 1155 Union Cir., Denton, TX. Received 12 May 2011. *Corresponding author (dan.sargent@emr.ac.uk).

Abbreviations: EMBL, European Molecular Biology Laboratory; EST, expressed sequence tag; FC, *Fragaria* pseudochromosome; FG, *Fragaria* linkage group; FVxFB, *Fragaria vesca* '815' × *Fragaria bucharica* '601'; FvH4, *Fragaria vesca* L. 'Hawaii 4'; PCR, polymerase chain reaction; SSR, simple sequence repeat; STS, sequence tagged site.

Published in The Plant Genome 4:165–177. Published 19 Aug. 2011.
doi: 10.3835/plantgenome2011.05.0014
© Crop Science Society of America
5585 Guilford Rd., Madison, WI 53711 USA
An open-access publication

All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher.

predominantly SSR markers, and a total of 272 have previously been mapped in the diploid *Fragaria* spp. reference linkage map *Fragaria vesca* '815' × *Fragaria bucharica* '601' (FV×FB) (Sargent et al., 2008). The reference map was developed from an interspecific F₂ population derived from a cross between *Fragaria vesca* L. and its closest diploid relative *Fragaria bucharica* Losinsk.; it spans seven linkage groups, corresponding to the basic, seven-member *Fragaria* spp. chromosome set and covers a total of 528.1 cM. This map contains a total of 422 sequence-characterized sequence tagged site (STS) markers mapped in the full progeny (Ruiz-Rojas et al., 2010) and a further 230 STS markers mapped using the bin-mapping progeny (Illa et al., 2011; Sargent et al., 2008).

The recent publication of the woodland strawberry (*F. vesca*) genome sequence was a milestone in plant biology (Shulaev et al., 2011). Not only was it one of the smallest plant genomes ever to be sequenced, but sequencing was performed using only short-read sequencing technologies. The complex polyploid genome of the cultivated strawberry genome precludes de novo assembly of sequences generated using current sequencing technologies, and so *F. vesca* ($2n = 2x = 14$), a close relative of *Fragaria × ananassa* Duchesne ex Rozier and widely regarded as the closest extant descendent of the octoploid's A genome donor (Davis et al., 2009), was chosen as a surrogate system for the development of a diploid reference genome sequence for the genus (Shulaev et al., 2011). The *F. vesca* 'Hawaii 4' (FvH4) genotype was chosen due to its amenability to genetic transformation, its self-compatibility and therefore highly homozygous genome, and its perpetually flowering "semperflorens" habit (Shulaev et al., 2011). *Fragaria vesca* 'Hawaii 4' was sequenced to 39x depth of coverage and the sequence is composed of a total of 219 Mbp of nucleotide data contained in more than 3200 genome sequence scaffolds.

Despite the large number of scaffolds, the majority of the genome sequence, 209.3 Mbp (96%), is contained in just 247 scaffolds of over 50 kbp in length (Shulaev et al., 2011). The genome was assembled into scaffolds and pseudochromosomes without a physical map, assigning the vast majority of contiguous sequences (198.1 Mbp of the sequence, contained in 204 sequence scaffolds) to a physical location on the *F. vesca* genome using molecular markers genetically mapped to the FV×FB linkage map (Shulaev et al., 2011). Of these, 131 scaffolds containing 169.5 Mbp of sequence have been anchored using the full FV×FB mapping progeny. The remaining 73 scaffolds were anchored using a selective mapping approach known as bin mapping (Sargent et al., 2008), the majority through sequencing of a reduced complexity genome scan of each of the six genotypes that comprise the selective mapping or "bin" set (Celton et al., 2010).

To complement and extend the existing SSR marker resources available for *Fragaria* spp., in particular for the development of saturated linkage maps, we developed a set of novel SSR markers from regions of the *F. vesca* genome that were unanchored and thus previously unmapped

or had been anchored only through bin mapping. We designed 152 novel primer pairs flanking selected polymorphic SSRs and, through mapping of these loci and a further 42 previously published SSRs, anchored a further 28.2 Mbp of scaffolded genome sequence to precise positions on the FV×FB reference map using the full mapping progeny of 76 individuals. Using these additional data we have improved the resolution and completeness of the anchoring of the *F. vesca* genome and have created a set of revised pseudochromosomes for FvH4.

Materials and Methods

Fragaria vesca '815' × *Fragaria bucharica* '601' Diploid *Fragaria* Reference Population

Novel markers developed in this investigation were tested for polymorphism between the grandparental genotypes (*F. vesca* '815' and *F. bucharica* '601') of the F₂ diploid *Fragaria* spp. reference mapping population FV×FB (Sargent et al., 2006). Segregation data for markers that were polymorphic between the grandparental genotypes were generated in the full FV×FB progeny of 76 individuals.

Simple Sequence Repeat Marker Development

Previously unanchored FvH4 sequence scaffolds over 50 kbp in length were submitted individually to the SSR server at the Genome Database for Rosaceae (Jung et al., 2008) to determine the precise locations of di- and trinucleotide SSR loci contained within each scaffold sequence. Di- and trinucleotide SSRs were investigated as these had been shown to be the most polymorphic motifs in previous mapping investigations in *Fragaria* spp. Cutoff criteria for candidate SSR marker loci were di- and trinucleotides with a minimum repeat length of 12 and 6 repeats, respectively. Where no SSRs were discovered under these criteria, the repeat length of dinucleotides was reduced to eight repeats. Simple sequence repeats were considered for marker development when they occurred with a minimum of 200 bp of high-quality sequence data flanking either side of the repeat region. A maximum of four pairs of polymerase chain reaction (PCR) primers targeting SSR loci were designed and synthesized per scaffold. Where possible, SSRs were selected for marker development at evenly spaced intervals throughout the scaffold sequence. Primers were designed to amplify products between 100 and 350 bp in length. All primer pairs were designed with PRIMER 3 (Rozen and Skaletsky, 2000) to have an oligonucleotide melting temperature of 55 to 65°C (optimum 60°C), a primer length of 20 to 24 bp (optimum 22 bp), and a 2-bp GC clamp at the 5' end of each primer. Primers were synthesized by Integrated DNA Technologies Ltd (Leuven, Belgium). Primer pairs amplifying markers mapping to the FV×FB map were named with the prefix FvH4 followed by a four digit numerical identifier, that is, FvH40001.

Polymerase Chain Reaction Conditions and Product Visualization

Following the touchdown protocol described by Sargent et al. (2003), amplicons were generated from the

grandparental genotypes of the mapping population FV×FB for novel SSR markers developed in this investigation and SSRs previously published by Sargent et al. (2008) and Spigler et al. (2010) that had not been previously scored in the full FV×FB mapping progeny. Initially, PCR products were assessed for polymorphism following agarose gel electrophoresis, ethidium bromide staining, and visualization over ultraviolet light. Where possible, at least one marker identified as polymorphic between the grandparental genotypes per sequence scaffold was labeled on the forward primer with one of two fluorescent dyes: 6-FAM or HEX (IDT, Leuven, Belgium). Labeled products were then sized following fractionation by capillary electrophoresis through a 3100 genetic analyzer (Applied Biosystems, Warrington, UK) and the data generated were collected and analyzed using the GENESCAN and GENOTYPER (Applied Biosystems, Warrington, UK) software applications.

Based on compatible product sizes and fluorescent dye colors, markers (primer pairs) were grouped into sets of up to 16 and multiplex PCR was performed using the “Type-it” PCR mastermix (Qiagen, Crawley, UK) following the manufacturer’s recommendations, except that PCRs were performed in a final volume of 12.5 µL. Reactions were performed using the following PCR cycles: an initial denaturation step of 95°C for 5 min was followed by 28 cycles of 95°C for 30 sec, an annealing temperature of 55°C decreasing by 0.5°C per cycle until 50°C for 90 sec, and 72°C for 30 sec, followed by a 30 min final extension step at 68°C. Samples were analyzed by capillary electrophoresis as described above.

Data Analysis and Map Construction

Segregation data for all new markers were analyzed for cosegregation with previously mapped markers using the published data for the FV×FB mapping population (Ruiz-Rojas et al., 2010) to determine the map position of the markers and thus the physical positions on the FvH4 genome sequence of the sequence scaffolds from which these markers were derived. Data were analyzed using JOINMAP 4.0 (Van Ooijen, 2006) using the Kosambi mapping function. Linkage group construction was determined using a minimum logarithm of the odds (LOD) score threshold of 3.0, a recombination fraction threshold of 0.35, ripple value of 1.0, jump threshold of 3.0, and a triplet threshold of 5.0.

Creation of Pseudochromosomes

Pseudochromosomes were created from FvH4 sequence scaffolds over 50 kbp in length that had been anchored to the FV×FB reference map by markers developed in this investigation or using previously mapped markers. Sequence scaffolds were arranged in order based on their map positions on the FV×FB linkage map and, where possible (when at least two markers were mapped to the same scaffold with sufficient resolution to permit orientation), scaffolds were orientated in accordance with the linkage group on which they were located. Scaffolds were separated

in the pseudochromosomes by arbitrary gaps of 10,000 nucleotides ($(N)_{10k}$) to allow clear demarcation between the end of one scaffold and the beginning of the next within each pseudochromosome. *Fragaria* spp. pseudochromosomes were plotted alongside the FV×FB linkage groups using HARRY PLOTTER (Moretto et al., 2010).

Linkage Group and Pseudochromosome Nomenclature

For clarity, when referring to markers mapped to the FV×FB linkage map, linkage group nomenclature followed that of Vilanova et al. (2008): *Fragaria* spp. linkage groups (FG) FG1 through FG7. When referring to markers or scaffolds located to one of the FvH4 pseudochromosomes, nomenclature followed that of Shulaev et al. (2011): *Fragaria* spp. pseudochromosome (FC) FC1 through FC7.

Simple Sequence Repeat Identification in *Fragaria vesca* ‘Hawaii 4’ Sequence Scaffolds

A FASTA (Pearson, 1990) file containing the seven FvH4 pseudochromosomes created from FvH4 sequence scaffolds over 50 kbp in length that had been anchored to the FV×FB reference map was submitted to MSATCOMMANDER v0.8.2 for Windows (Faircloth, 2008) for SSR screening. All SSRs with a di- or trinucleotide motif at least 12 and 6 repeats, respectively, were identified and compiled, listing their repeat length, motif, and pseudochromosome position.

Results

Simple Sequence Repeat Design, Amplification, Polymorphism, and Mapping in *Fragaria vesca* ‘815’ × *Fragaria bucharica* ‘601’

In total, 296 primer pairs were designed from FvH4 genome sequence scaffolds. Of these, 152 primer pairs (51%) fulfilled three key criteria: (i) they generated single-locus polymorphisms between the grandparental genotypes of the FV×FB mapping population, (ii) they were targeted to previously unmapped or bin-mapped sequence scaffolds and were thus located in areas of the genome from which SSR markers had not previously been developed, and (iii) they mapped to one of the seven diploid *Fragaria* spp. linkage groups. One hundred twenty-one of the remaining 144 primer pairs generated products that were monomorphic between the parents of the FV×FB mapping population. Of the remaining primer pairs, 14 generated complex, or nonspecific, amplification profiles, six failed to amplify a product, and three generated products of over 1 kbp in length that could not be reliably genotyped in the progeny. Supplemental Table S1 lists the locus names, primer sequences, repeat motif, repeat number, and European Molecular Biology Laboratory (EMBL) reference numbers of the 152 mapped primer pairs. Additionally, a further 42 polymorphic SSRs that had been developed in previous investigations (Sargent et al., 2008; Spigler et al., 2010) and that had been identified in scaffolds over 50 kbp in length but had not previously been scored in the full FV×FB mapping population were also mapped.

Genome Sequence Scaffolds Anchored to the *Fragaria vesca* '815' × *Fragaria bucharica* '601' Reference Map

The 194 SSR loci genotyped in the full mapping population anchored a total of 93 genome sequence scaffolds to the FV×FB reference map. Sixty of these scaffolds had been previously located to FV×FB mapping bins using a selective mapping strategy (Celton et al., 2010) while 33 had not previously been located to the FV×FB map. Simple sequence repeat markers developed for one scaffold (scf0513139) mapped to two distinct locations: one on FG2 and one on FG5 of the FV×FB reference map. The sequence of scaffold scf0513139 was therefore split at nucleotide 563,780 between the two mapped SSRs in an area of low sequence coverage, and the two sections of the scaffold were renamed 513139a and 513139b following the convention of Shulaev et al. (2011).

Ninety-three FvH4 genome sequence scaffolds were anchored using markers mapped in the full FV×FB progeny in this investigation, adding a further 28.8 Mbp of sequence to precise locations on the *F. vesca* genome. These, along with those anchored using markers mapped in the full FV×FB progeny by Shulaev et al. (2011), brings the total number of anchored scaffolds (including the two scaffolds created by splitting scaffold scf0513139) to 222. These scaffolds are anchored using a total of 411 genetic markers (Fig. 1), 194 mapped in the present investigation and a further 217 previously mapped markers (Ruiz-Rojas et al., 2010), and contain 197,682,269 nucleotides (197.7 Mbp) of sequence including embedded gaps. Using only markers that were contained in genome sequence scaffolds and by removing markers where their positioning on the FV×FB was questionable—through the erroneous positioning of an otherwise robustly positioned sequence scaffold within a region of the genetic map, poor fit of the data for a single marker in relation to other markers on the linkage group, or large amounts of data missing for a marker—we have produced an updated version of the reference linkage map for *Fragaria* spp. in which the positioning of the markers and underlying sequence scaffolds is robust and reliable. The distribution of scaffolds per linkage group in the updated map was not even, with FG3 containing the highest number of scaffolds (47) and FG7 the lowest (14). The distribution of sequence was also uneven, the longest pseudochromosome being FC6 (39.5 Mbp), which was almost twice the length of FC1 (20.0 Mbp), while FC3 contained the greatest number of nucleotides per centimorgan of its equivalent linkage group (687 kbp) and FC7 contained the fewest (372 kbp). Table 1 lists the number of scaffolds, linkage groups, and equivalent pseudochromosome lengths along with the number of nucleotides per scaffold and per centimorgan of each of the linkage groups of the FV×FB map.

The 222 scaffolds mapped to the FV×FB reference map using the full mapping progeny now represent 90% of the FvH4 genome sequence scaffolds over 50 kbp in length. Figure 2 shows the physical positions of the 222 sequence

scaffolds in relation to the markers to which they are anchored on the FV×FB reference map while Supplemental Table S2 lists the 411 genetic markers mapped to the FV×FB reference map that have been identified in genome sequence scaffolds herein or previously by Shulaev et al. (2011), their map positions, and the FvH4 genome sequence scaffolds that they anchor, along with the scaffold sizes in base pairs. Figure 3 displays scaffolds anchored to FG2 of the FV×FB linkage map in this investigation using the full mapping progeny and those anchored in the investigation of Shulaev et al. (2011) in which some scaffolds were anchored using bin mapping, showing the increased precision in the positioning of scaffolds in this investigation. A total of 10% (24) of scaffolds over 50 kbp in length, collectively containing 11.9 Mbp of sequence, were not anchored to the FV×FB map with the full FV×FB mapping progeny using SSRs developed in this investigation; however, of the 11.9 Mbp, 56% (6.7 Mbp) was previously bin mapped, leaving just 5.2 Mbp (2.4% of the FvH4 genome sequence contained in scaffolds over 50 kbp) unanchored to the *Fragaria* spp. genome. In most cases (92%), when scaffolds were not anchored in this investigation it was due to difficulties in designing primers for amplification of single SSR loci or a lack of polymorphism in the SSRs designed; however, in a very small number of cases (8% unmapped scaffolds), it was due to the absence of SSR motifs of sufficient length around which to design primers.

Fragaria vesca 'Hawaii 4' Pseudochromosomes

The seven FvH4 pseudochromosomes (FC1–FC7) containing the 222 genome sequence scaffolds over 50 kbp in length that were anchored to the FV×FB linkage map using the full mapping progeny in this investigation or by Shulaev et al. (2011), along with an eighth pseudochromosome, FC0, containing the 24 unanchored genome sequence scaffolds over 50 kbp in length, have been deposited at the strawberry genome browser (Shulaev et al., 2011; Jung et al., 2005) and have been denoted v1.1.

Simple Sequence Repeat Identification in *Fragaria vesca* 'Hawaii 4' Pseudochromosomes

From the 222 genome sequence scaffolds over 50 kbp in length anchored to the FV×FB genetic map, a total of 10,071 SSRs were identified, 7321 (72.7%) of which were dinucleotide and 2744 (27.3%) of which were trinucleotide, with a repeat length equal to or in excess of 12 and 6, respectively. Simple sequence repeat distribution across the seven FvH4 pseudochromosomes was random with no clustering observed in relation to physical distance (data not shown). There was a strong correlation ($R^2 = 0.923$) between chromosome physical length and number of SSRs per chromosome, with an average of one SSR every 19,518 nucleotides. Simple sequence repeat motifs were present in the genome at different frequencies. Within the dinucleotides, the (AG)_n motif was present at the highest frequency (59%), followed by (AT)_n (37%) and (AC)_n (4%) (Fig. 4a). The (CG)_n motif was not found in the FvH4 genome sequence scaffolds with

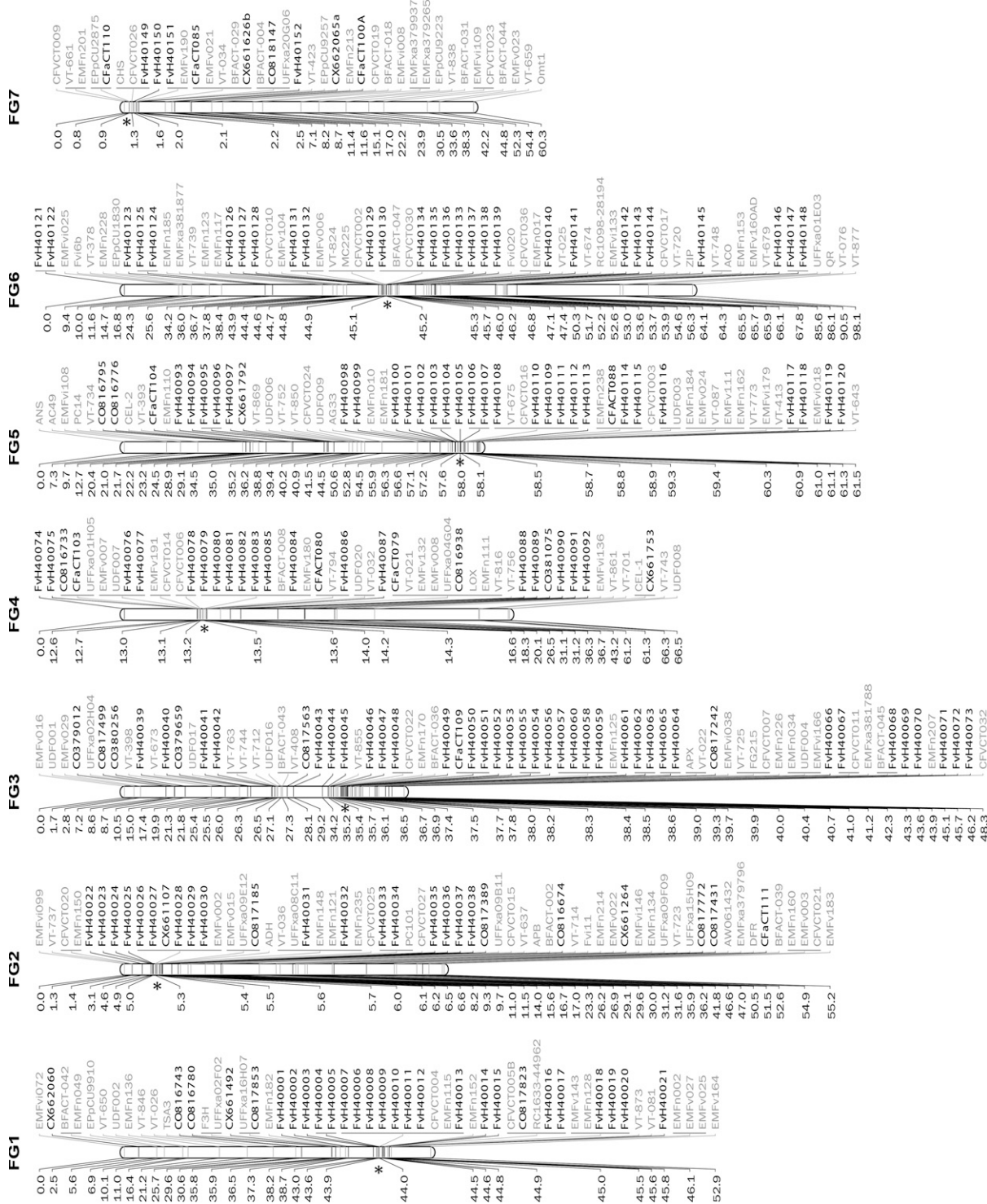


Figure 1. A genetic linkage map of the *Fragaria vesca* '815' x *Fragaria bucharica* '601' (FvxFB) mapping progeny of 76 seedlings showing the map positions of the 411 sequence characterized genetic markers used to anchor the scaffolds over 50 kbp in length of the *Fragaria vesca* L. 'Hawaii 4' (FvH4) genome sequence assembly. Novel markers mapped for the first time in the full progeny in this investigation are given in bold and genetic distances are given in cM. Asterisks denote the putative positions of centromeric regions on each linkage group. FG, *Fragaria* linkage group.

Table 1. *Fragaria vesca* L. 'Hawaii 4' (FvH4) pseudochromosome characteristics. The number of sequence scaffolds anchored to each *Fragaria* pseudochromosome giving the lengths of *Fragaria vesca* '815' × *Fragaria bucharica* '601' (FV×FB) linkage groups that they were anchored to and their equivalent pseudochromosome lengths along with the number of nucleotides per scaffold and per centimorgan of each of the linkage groups of the FV×FB map.

Linkage group	No. of scaffolds	Pseudochromosome (FC) length (bp)	Linkage group length (cM)	Megabase pairs per scaffold	Base pairs per centimorgan
FG1 [†]	31	19,956,132	52.9	0.64	377,371
FG2	29	24,527,021	55.2	0.85	444,467
FG3	47	33,188,684	48.3	0.71	687,449
FG4	29	29,352,125	66.5	1.01	441,492
FG5	31	27,567,315	61.5	0.89	448,577
FG6	41	39,546,558	98.1	0.97	402,961
FG7	14	22,432,496	60.3	1.60	372,151
Total	222	196,570,331	442.8		

[†]FC, *Fragaria* pseudochromosome.

*FG, *Fragaria* linkage group.

a repeat number $n \geq 12$. Within the trinucleotides, the (AAG)_n motif was present at the highest frequency (33%), followed by (AAT)_n (21%) and (ATG)_n (10%), the other seven trinucleotide repeat motifs making up the remaining 36% (Fig. 4b).

Discussion

In this investigation we have developed and mapped a set of 152 novel polymorphic SSR markers from previously unmapped regions of the diploid *Fragaria* spp. genome, which will thus be useful for the continued development of genetic linkage maps in strawberry. Additionally, the mapping of these novel SSR markers has improved the precision of the anchoring of the FvH4 genome sequence assembly (Shulaev et al., 2011) through the location of a further 28.2 Mbp of sequence data to the diploid *Fragaria* spp. reference map FV×FB using the full mapping progeny (Ruiz-Rojas et al., 2010).

Due to their inherent transferability, codominant inheritance, and generally high levels of polymorphisms, SSRs have found enormous utility as genetic markers and have been used extensively as markers in numerous plant species including the development of reference maps for many rosaceous genera (Aranzana et al., 2003; Celton et al., 2009; Gisbert et al., 2009; Graham et al., 2004; Yamamoto et al., 2007). The SSR markers developed in this investigation mapped to discrete positions on the diploid *Fragaria* spp. reference map (Ruiz-Rojas et al., 2010) and permitted the precise anchoring of a further 28.2 Mbp *Fragaria* spp. genome sequence data than had previously been anchored (Shulaev et al., 2011). In total, 90% of the genome sequence scaffolds over 50 kbp in length have now been anchored using STS markers in the full FV×FB progeny and a total of 97.6% are anchored either using a conventional- or bin-mapping strategy. The sequencing of the FvH4 genome was not performed as a means to an end, and it was envisaged that the resultant sequence would be utilized a generic tool to assist the entire *Fragaria* spp. research community. Bin mapping, as performed by Shulaev et al. (2011), is a highly effective strategy for the placement of a large number of markers onto a linkage map with minimum experimental

time and cost. However, the technique lacks precision and, while scaffolds anchored in this way were placed accurately in mapping bins on the FV×FB genetic map, the ordering of markers within the bins cannot be determined. Figure 3 displays the positions of scaffolds anchored in this investigation through the mapping of markers in the full FV×FB progeny compared to the placement of scaffolds using bin mapping by Shulaev et al. (2011). Mapping anchor markers in the full FV×FB progeny has increased the precision with which the scaffolds were anchored and therefore the utility of the anchored sequences for further research activities.

Thus, the pseudochromosomes presented in this investigation will improve the utility and generic value of the FvH4 sequence as a reference tool for genomics studies within *Fragaria* spp. and between plant genera. Marker systems exploiting single nucleotide polymorphisms, such as cleaved amplified polymorphic sequence (CAPS) markers (Ruiz-Rojas et al., 2010) and high resolution melting analyses, have found utility in other species including apple (*Malus pumila* Mill.), potato (*Solanum tuberosum* L.), and almond [*Prunus dulcis* (Mill.) D. A. Webb] (Chagné et al., 2008; De Koeper et al., 2010; Wu et al., 2010), but such markers cannot be readily applied to the cultivated strawberry and its wild progenitors due to the complex polyploid nature of their genomes. The SSRs developed and utilized for scaffold anchoring in this investigation will be of great utility for the genetic characterization of and the continued development of genetic linkage maps for the octoploid cultivated strawberry *F. ×ananassa* (Sargent et al., 2009; van Dijk et al., 2010).

As an additional contribution to *Fragaria* spp. genomic information resources needed for mapping and other marker-dependent applications in strawberry, we have also surveyed the content and distribution of SSR repeats within the FvH4 genome sequence. As with other plant species, such as *Arabidopsis thaliana* (L.) Heynh., rice (*Oryza sativa* L.), almond (*Prunus dulcis*), peach [*Prunus persica* (L.) Batsch], and rose (*Rosa* spp.) (Jung et al., 2005; McCouch et al., 2002), the most abundant repeat motifs for di- and trinucleotide SSRs in the FvH4 genome sequence were (AG)_n and (AAG)_n, respectively. The distribution of repeat motifs for SSRs throughout the

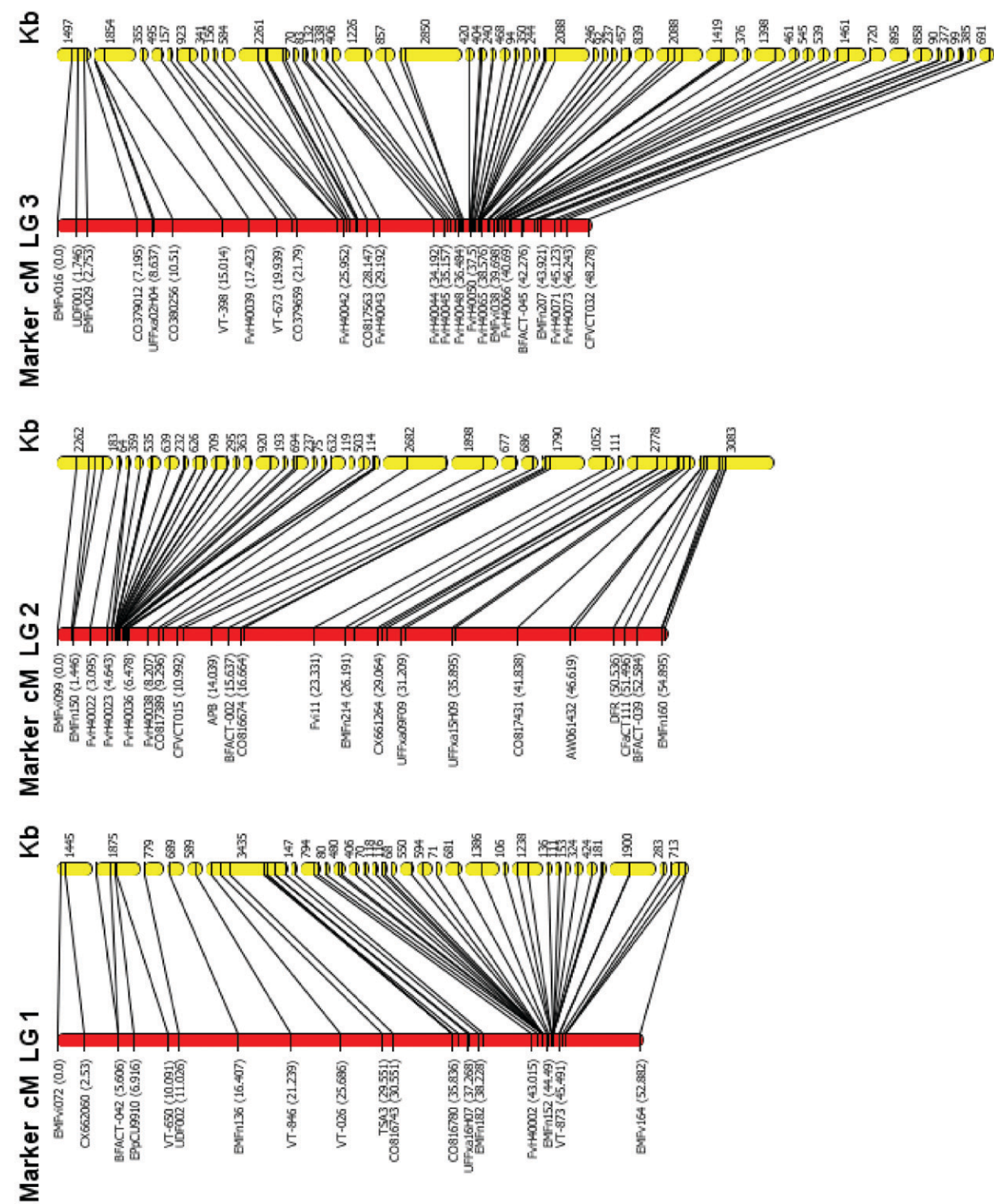


Figure 2. (cont'd on next page) The genetic locations of the *Fragaria vesca* L. 'Hawaii 4' (FvH4) genome sequence scaffolds. The physical positions of the 222 sequence scaffolds greater than 50 kbp in length are shown in relation to the markers to which they are anchored on the *Fragaria vesca* '815' × *Fragaria bucharica* '601' (Fv×FB) reference map. Physical distances are given in kilobase pairs and genetic distances are given in centimorgans. LG, linkage group.

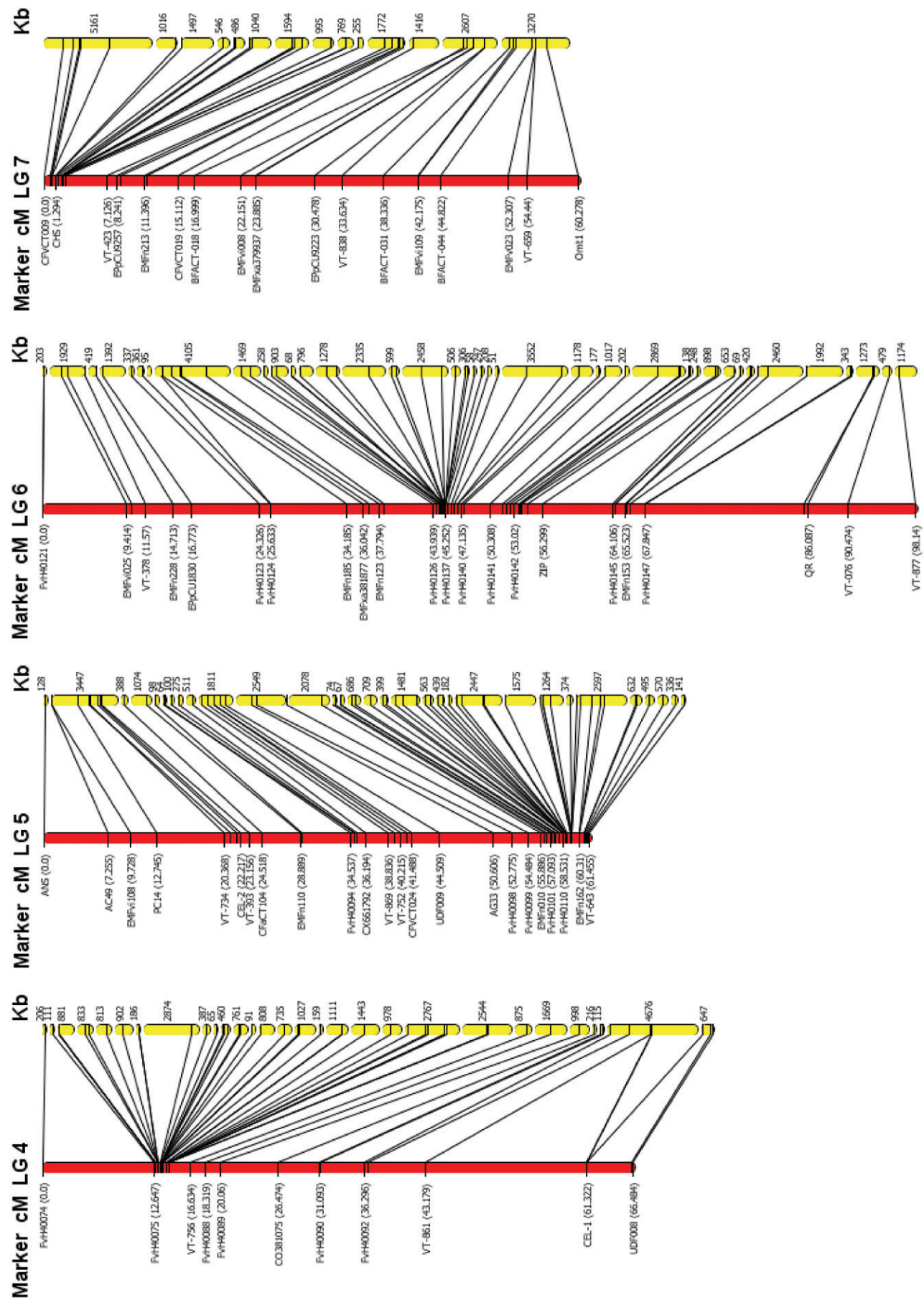


Figure 2. Continued.

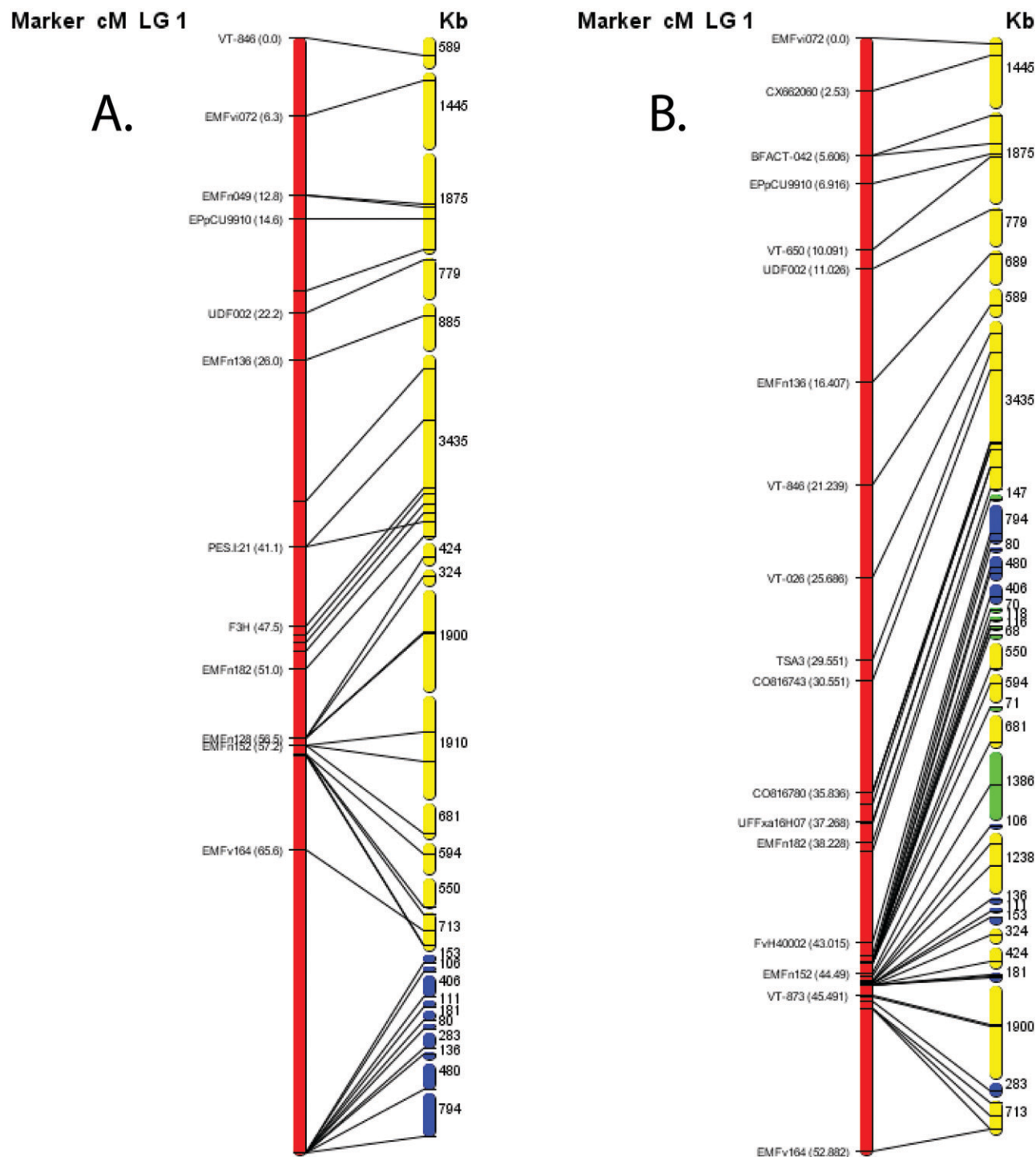
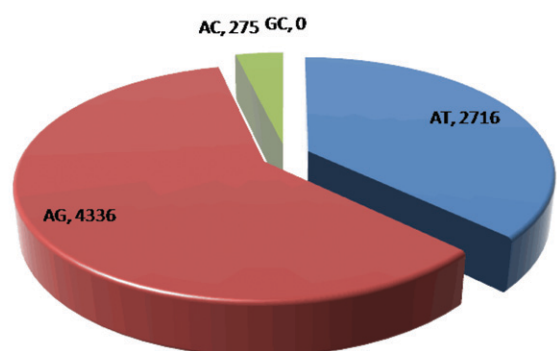


Figure 3. A comparison of scaffold positions anchored by bin mapping and by markers mapped in the full *Fragaria vesca* '815' × *Fragaria bucharica* '601' (FVxvFB) on the *Fragaria* linkage group (FG) 2 of the FVxvFB reference map showing A) sequence scaffolds mapped using bin mapping (blue scaffolds) by Shulaev et al. (2011) and B) the positions of the same scaffolds (shown in blue) anchored using markers mapped in the full progeny in this investigation. Scaffolds shown in green were anchored for the first time in this investigation and were not present on the map of Shulaev et al. (2011). LG, linkage group.

whole FvH4 genome sequence was similar to that found in ESTs sequenced from *F. ×ananassa* for dinucleotide repeats but differed for trinucleotide repeats, with the (AAT)_n and (ATG)_n repeats present in much lower frequencies in the ESTs sequenced and analyzed by (Folta et al., 2005) using similar criteria to define the presence

of di- and trinucleotide SSRs. In two more recent studies, however, (Zhang and Deng, 2010) found a similar percentage of (AAG)_n trinucleotide repeats in SSRs found within the 17,565 *Fragaria* spp. ESTs deposited in the EMBL nucleotide sequence repository, with other motifs present at lower frequencies, and (Bombarely et al., 2010)

A. Dinucleotide repeats



B. Trinucleotide repeats

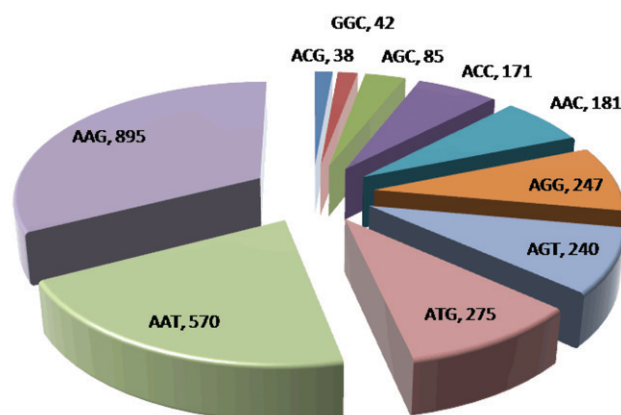


Figure 4. Simple sequence repeat (SSR) motifs identified in the *Fragaria vesca* L. 'Hawaii 4' (FvH4) pseudochromosomes showing frequency distributions of SSR repeat motifs within the FvH4 genome sequence scaffolds over 50 kbp in length. A) shows distribution of dinucleotide repeats and B) shows distribution of trinucleotide repeats.

reported similar proportions of both di- and trinucleotide repeats within a collection of 4500 *F. ×ananassa* ESTs to this investigation, suggesting that SSR repeat motifs in *Fragaria* spp. are distributed in similar frequencies within the coding and noncoding portions of the genome.

Simple sequence repeat distribution on the FV×FB genetic map was nonrandom, with clustering apparent on all seven linkage groups (Fig. 1) (Ruiz-Rojas et al., 2010; Sargent et al., 2006). The random physical distribution of SSR loci across the FvH4 pseudochromosomes suggests that the clustering of markers on the FV×FB linkage map is due to recombination suppression in specific regions of the genome. Such recombination suppression is apparent on linkage maps of other species, including papaya (*Carica papaya* L.) and barley (*Hordeum vulgare* L.), and has been shown to be located within centromeric regions (Chen et al., 2007; Wenzl et al., 2006). The direct effects of centromeric regions on the suppression of recombination have been demonstrated by (Lambie and Roeder, 1986) and thus we postulate that the regions of distinct recombination suppression

apparent on all seven FV×FB linkage groups may be the centromeric regions of the seven *Fragaria* spp. chromosomes. It cannot be ruled out, however, that the clustering, particularly where more than one distinct region was observed on a single linkage group as on FG3, might have been caused by other factors such as low homology between the interspecific chromosomes due to high sequence divergence between *F. vesca* and *F. bucharica* or genomic rearrangements in the evolution of the two species since they diverged from a common ancestor.

Using the 152 SSRs developed in this investigation and 42 previously published SSRs we have anchored 93 scaffolds that had not previously been located, to precise but, in the majority of cases, unorientated positions using the full FV×FB mapping progeny and created seven pseudochromosomes composed of 222 sequence scaffolds covering 197.7 Mbp. As greater numbers of plant genomes are sequenced (Kaul et al., 2000; Ming et al., 2008; Shulaev et al., 2011; Tuskan et al., 2006; Velasco et al., 2007, 2010; Vogel et al., 2010; Yu et al., 2002), it has become possible to perform whole genome synteny analyses at the DNA sequence level (Gar et al., 2011; Illa et al., 2011; Jung et al., 2010). Marker density on the FV×FB map presented is one marker every 1.09 cM and, thus, anchoring the majority of the FvH4 sequence using the full mapping progeny will permit such analyses to be performed for *Fragaria* spp. with far greater precision than could have been achieved if scaffolds had been located using bin mapping alone, where markers were assigned to 45 mapping bins with an average bin length of 12.6 cM (Sargent et al., 2008). The seven FvH4 pseudochromosomes presented in this report will also assist resequencing efforts and genomic investigations in the cultivated strawberry, *F. ×ananassa*, and its wild octoploid progenitors *Fragaria chiloensis* (L.) Mill. and *Fragaria virginiana* Mill.

Supplemental Information Available

Supplemental material is available free of charge at <http://www.crops.org/publications/tpg>.

Acknowledgments

We gratefully acknowledge the provision of HARRY PLOTTER by Marco Moretto (FEM-IASMA) and the assistance of Alastair Plant and Thomas Foreau in the development and testing of SSR primers identified using the GDR SSR miner. *Fragaria* genomics at East Malling Research is funded by the BBSRC and the East Malling Trust. This research was supported in part by USDA-CSREES National Research Initiative (NRI) Plant Genome Grant 2008-35300-04411 and New Hampshire Agricultural Experiment Station Project NH00535 (to TMD).

References

- Aranzana, M.J., A. Pineda, P. Cosson, E. Dirlewanger, J. Ascasibar, G. Cipriani, C.D. Ryder, R. Testolin, A. Abbot, G.J. King, A.F. Iezzoni, and P. Arús. 2003. A set of simple sequence repeat (SSR) markers covering the *Prunus* genome. *Theor. Appl. Genet.* 106:819–825.
- Bombarely, A., C. Merchante, F. Csukasi, E. Cruz-Rus, J.L. Caballero, N. Medina-Escobar, R. Blanco-Portales, M.A. Botella, J. Munoz-Blanco, J.F. Sanchez-Sevilla, and V. Valpuesta. 2010. Generation and analysis of ESTs from strawberry (*Fragaria ×ananassa*) fruits and evaluation of their utility in genetic and molecular studies. *BMC Genomics* 11:503. doi:10.1186/1471-2164-11-503

- Celton, J.M., A. Christoffels, D.J. Sargent, X. Xu, and D.J.G. Rees. 2010. Genome-wide SNP identification by high-throughput sequencing and selective mapping allows sequence assembly positioning using a framework genetic linkage map. *BMC Biol.* 8:155. doi:10.1186/1741-7007-8-155
- Celton, J.M., D.S. Tustin, D. Chagne, and S.E. Gardiner. 2009. Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from *Malus* ESTs and *Pyrus* genomic sequences. *Tree Genet. Genomes* 5:93–107. doi:10.1007/s11295-008-0171-z
- Chagné, D., K. Gasic, R.N. Crowhurst, Y. Han, H.C. Bassett, D.R. Bowatte, T.J. Lawrence, E.H.A. Rikkerink, S.E. Gardiner, and S.S. Korban. 2008. Development of a set of SNP markers present in expressed genes of the apple. *Genomics* 92:353–358. doi:10.1016/j.ygeno.2008.07.008
- Chen, C.X., Q.Y. Yu, S.B. Hou, Y.J. Li, M. Eustice, R.L. Skelton, O. Veatch, R.E. Herdes, L. Diebold, J. Saw, Y. Feng, W.B. Qian, L. Bynum, L. Wang, P.H. Moore, R.E. Paull, M. Alam, and R. Ming. 2007. Construction of a sequence-tagged high-density genetic map of papaya for comparative structural and evolutionary genomics in brassicales. *Genetics* 177:2481–2491. doi:10.1534/genetics.107.081463
- Davis, T.M., M.E. Shields, Q. Zhang, E.G. Poulsen, K.M. Folta, J.L. Benetzen, and P. San Miguel. 2009. The strawberry genome is coming into view. *Acta Hort* 842:533–536.
- De Koeijer, D., K. Douglass, A. Murphy, S. Whitney, L. Nolan, Y. Song, and W. De Jong. 2010. Application of high-resolution DNA melting for genotyping and variant scanning of diploid and autotetraploid potato. *Mol. Breed.* 25:67–90. doi:10.1007/s11032-009-9309-4
- Faircloth, B.C. 2008. MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. *Mol. Ecol. Resour.* 8:92–94. doi:10.1111/j.1471-8286.2007.01884.x
- FAO. 2011. Time-series and cross sectional data relating to food and agriculture for some 200 countries. Available at <http://faostat.fao.org/> (verified 10 Aug. 2011). FAO, Rome, Italy.
- Folta, K.M., M. Staton, P.J. Stewart, S. Jung, D.H. Bies, C. Jesdurai, and D. Main. 2005. Expressed sequence tags (ESTs) and simple sequence repeat (SSR) markers from octoploid strawberry (*Fragaria × ananassa*). *BMC Plant Biol.* 5:12. doi:10.1186/1471-2229-5-12
- Gar, O., D.J. Sargent, C.J. Tsai, T. Pleban, G. Shalev, Dh. Byrne, and D. Zamir. 2011. An autotetraploid linkage map of rose (*Rosa hybrida*) validated using the strawberry (*Fragaria vesca*) genome sequence. *PLoS ONE* 6:e20463. doi:10.1371/journal.pone.0020463
- Gisbert, A.D., J. Martinez-Calvo, G. Llacer, M. Badenes, and C. Romero. 2009. Development of two loci at *Eriobotrya japonica* (Thunb.) Lindl.] linkage maps based on AFLPs and SSR markers from different *Rosaceae* species. *Mol. Breed.* 23:523–538. doi:10.1007/s11032-008-9253-8
- Graham, J., K. Smith, K. MacKenzie, L. Jorgenson, C. Hackett, and W. Powell. 2004. The construction of a genetic linkage map of red raspberry (*Rubus idaeus* subsp. *idaeus*) based on AFLPs, genomic-SSR and EST-SSR markers. *Theor. Appl. Genet.* 109:740–749. doi:10.1007/s00122-004-1687-8
- Illa, E., D.J. Sargent, E. Lopez Girona, J. Bushakra, A. Cestaro, R. Crowhurst, M. Pindo, A. Cabrera, E. Van der Knapp, A. Iezzoni, S. Gardiner, R. Velasco, P. Arus, D. Chagné, and M. Troggio. 2011. Comparative analysis of rosaceous genomes and the reconstruction of a putative ancestral genome for the family. *BMC Evol. Biol.* 11:9. doi:10.1186/1471-2148-11-9
- Jung, S., A. Abbott, C. Jesdurai, J. Tomkins, and D. Main. 2005. Frequency, type, distribution and annotation of simple sequence repeats in *Rosaceae* ESTs. *Funct. Integr. Genomics* 5:136–143. doi:10.1007/s10142-005-0139-0
- Jung, S., I. Cho, B. Sosinski, A.G. Abbott, and D. Main. 2010. Comparative genomic sequence analysis of strawberry and other rosids reveals significant microsynteny. *BMC Res. Notes* 3:168. doi:10.1186/1756-0500-3-168
- Jung, S., M. Staton, T. Lee, A. Blenda, R. Svancara, A. Abbott, and D. Main. 2008. GDR (genome database for Rosaceae): Integrated web-database for *Rosaceae* genomics and genetics data. *Nucleic Acids Res.* 36:D1034–D1040. doi:10.1093/nar/gkm803
- Kaul, S., H.L. Koo, J. Jenkins, M. Rizzo, T. Rooney, L.J. Tallon, T. Feldblyum, W. Nierman, M.I. Benito, X.Y. Lin, C.D. Town, J.C. Venter, C.M. Fraser, S. Tabata, Y. Nakamura, T. Kaneko, S. Sato, E. Asamizu, T. Kato, H. Kotani, S. Sasamoto, J.R. Ecker, A. Theologis, N.A. Federspiel, C.J. Palm, B.I. Osborne, P. Shinn, A.B. Conway, V.S. Vysotskaia, K. Dewar, L. Conn, C.A. Lenz, C.J. Kim, N.F. Hansen, S.X. Liu, E. Buehler, H. Altafi, H. Sakano, P. Dunn, B. Lam, P.K. Pham, Q. Chao, M. Nguyen, G.X. Yu, H.M. Chen, A. Southwick, J.M. Lee, M. Miranda, M.J. Toriumi, R.W. Davis, R. Wambutt, G. Murphy, A. Dusterhoft, W. Stiekema, T. Pohl, K.D. Entian, N. Terry, G. Volckaert, M. Salanoubat, N. Choise, M. Rieger, W. Ansoorge, M. Unseld, B. Fartmann, G. Valle, F. Artiguenave, J. Weisenbach, F. Quetier, R.K. Wilson, M. de la Bastide, M. Sekhon, E. Huang, L. Spiegel, L. Gnoj, K. Pepin, J. Murray, D. Johnson, K. Habermann, N. Dedhia, L. Parnell, R. Preston, L. Hillier, E. Chen, M. Marra, R. Martienssen, W.R. McCombie, K. Mayer, O. White, M. Bevan, K. Lemcke, T.H. Creasy, C. Bielke, B. Haas, D. Haase, R. Maiti, S. Rudd, J. Peterson, H. Schoof, D. Frishman, B. Morgenstern, P. Zaccaria, M. Ermolaeva, M. Pertea, J. Quackenbush, N. Volkovskiy, D.Y. Wu, T.M. Lowe, S.L. Salzberg, H.W. Mewes, S. Rounsley, D. Bush, S. Subramaniam, I. Levin, S. Norris, R. Schmidt A. Acarkan, I. Bancroft, F. Quetier, A. Brennicke, J.A. Eisen, T. Bureau, B.A. Legault, Q.H. Le, N. Agrawal, Z. Yu, R. Martienssen, G.P. Copenhagen, S. Luo, C.S. Pikaard, D. Preuss, I.T. Paulsen, M. Sussman, A.B. Britt, D.A. Selinger, R. Pandey, D.W. Mount, V.L. Chandler, R.A. Jorgensen, C. Pikaard, G. Juergens, E.M. Meyerowitz, A. Theologis, J. Dangl, J.D.G. Jones, M. Chen, J. Chory, M.C. Somerville, and Gen I. Ar. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796–815. doi:10.1038/35048692
- Lambie, E.J., and G.S. Roeder. 1986. Repression of meiotic crossing over by a centromere (CEN3) in *Saccharomyces cerevisiae*. *Genetics* 114:769–789.
- McCouch, S.R., L. Teytelman, Y.B. Xu, K.B. Lobos, K. Clare, M. Walton, B.Y. Fu, R. Maghirang, Z.K. Li, Y.Z. Xing, Q.F. Zhang, I. Kono, M. Yano, R. Fjellstrom, G. DeClerck, D. Schneider, S. Cartinhour, D. Ware, and L. Stein. 2002. Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Res.* 9:199–207. doi:10.1093/dnares/9.6.199
- Ming, R., S.B. Hou, Y. Feng, Q.Y. Yu, A. Dionne-Laporte, J.H. Saw, P. Senin, W. Wang, B.V. Ly, K.L.T. Lewis, S.L. Salzberg, L. Feng, M.R. Jones, R.L. Skelton, J.E. Murray, C.X. Chen, W.B. Qian, J.G. Shen, P. Du, M. Eustice, E. Tong, H.B. Tang, E. Lyons, R.E. Paull, T.P. Michael, K. Wall, D.W. Rice, H. Albert, M.L. Wang, Y.J. Zhu, M. Schatz, N. Nagarajan, R.A. Acob, P.Z. Guan, A. Blas, C.M. Wai, C.M. Ackerman, Y. Ren, C. Liu, J.M. Wang, J.P. Wang, J.K. Na, E.V. Shakhov, B. Haas, J. Thimmapuram, D. Nelson, X.Y. Wang, J.E. Bowers, A.R. Gschwend, A.L. Delcher, R. Singh, J.Y. Suzuki, S. Tripathi, K. Neupane, H.R. Wei, B. Irikura, M. Paidi, N. Jiang, W.L. Zhang, G. Presting, A. Windsor, R. Navajas-Perez, M.J. Torres, F.A. Feltus, B. Porter, Y.J. Li, A.M. Burroughs, M.C. Luo, L. Liu, D.A. Christopher, S.M. Mount, P.H. Moore, T. Sugimura, J.M. Jiang, M.A. Schuler, V. Friedman, T. Mitchell-Olds, D.E. Shippen, C.W. dePamphilis, J.D. Palmer, M. Freeling, A.H. Paterson, D. Gonsalves, L. Wang, and M. Alam. 2008. The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature* 452:991–997. doi:10.1038/nature06856
- Moretto, M., A. Cestaro, M. Troggio, F. Costa, and R. Velasco. 2010. Harry Plotter: A user friendly program to visualize genome and genetic map features. *In* ECCB10: 9th European Conf. Computational Biology, Ghent, Belgium. 26–29 Sept. 2010. SIB Swiss Institute of Bioinformatics, Basel, Switzerland.
- Pearson, W.R. 1990. Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.* 183:63–98. doi:10.1016/0076-6879(90)83007-V
- Rousseau-Gueutin, M., E. Lerceteau-Kohler, L. Barrot, D.J. Sargent, A. Monfort, D. Simpson, P. Arus, G. Guerin, and B. Denoyes-Rothan. 2008. Comparative genetic mapping between octoploid and diploid *Fragaria* species reveals a high level of colinearity between their genomes and the essentially disomic behavior of the culti-

- vated octoploid strawberry. *Genetics* 179:2045–2060. doi:10.1534/genetics.107.083840
- Rozen, S., and H. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 132:365–386.
- Ruiz-Rojas, J.J., D.J. Sargent, V. Shulaev, A.W. Dickerman, J. Pattison, S.H. Holt, A. Ciordia, and R.E. Veilleux. 2010. SNP discovery and genetic mapping of T-DNA insertional mutants in *Fragaria vesca* L. *Theor. Appl. Genet.* 121:449–463. doi:10.1007/s00122-010-1322-9
- Sargent, D.J., G. Cipriani, S. Vilanova, D. Gil-Ariza, P. Arús, D.W. Simpson, K.R. Tobutt, and A. Monfort. 2008. The development of a bin mapping population and the selective mapping of 103 markers in the diploid *Fragaria* reference map. *Genome* 51:120–127. doi:10.1139/G07-107
- Sargent, D.J., J. Clarke, D.W. Simpson, K.R. Tobutt, P. Arus, A. Monfort, S. Vilanova, B. Denoyes-Rothan, M. Rousseau, K.M. Foltá, N.V. Brassil, and N.H. Battey. 2006. An enhanced microsatellite map of diploid *Fragaria*. *Theor. Appl. Genet.* 112:1349–1359. doi:10.1007/s00122-006-0237-y
- Sargent, D.J., F. Fernández-Fernández, J.J. Ruiz-Roja, B.G. Sutherland, A. Passey, A.B. Whitehouse, and D.W. Simpson. 2009. A genetic linkage map of the cultivated strawberry (*Fragaria ×ananassa*) and its comparison to the diploid *Fragaria* reference map. *Mol. Breed.* 24:293–303. doi:10.1007/s11032-009-9292-9
- Sargent, D.J., A.M. Hadonou, and D.W. Simpson. 2003. Development and characterization of polymorphic microsatellite markers from *Fragaria viridis*, a wild diploid strawberry. *Mol. Ecol. Notes* 3:550–552. doi:10.1046/j.1471-8286.2003.00507.x
- Shulaev, V., D.J. Sargent, R.N. Crowhurst, R.N. Mockler, R.E. Veilleux, O. Folkerts, A.L. Delcher, P. Jaiswal, A. Liston, S.P. Mane, P. Burns, K. Mockaitis, T.M. Davis, J.P. Slovin, N. Bassil, R.P. Hellens, C. Evans, R.V. Jensen, A.C. Allan, T.P. Michael, J.C. Setubal, J.M. Celson, D.J.G. Rees, K. Williams, S.H. Holt, A. Dickerman, J.J. Ruiz Rojas, M. Chatterjee, B. Liu, H. Silva, L. Meisel, A.A.S. Filichkin, R. Velasco, M. Troggio, R. Viola, M. Borodovsky, T.L. Ashman, A. Aharoni, J. Bennetzen, P. Dharmawardhana, J. Elser, R. Raja, H.D. Priest, D.W. Bryant, S.E. Fox, S.A. Givan, S. Naithani, A. Christofels, D.Y. Salama, J. Carter, E. Lopez Girona, A. Zdepski, W. Wang, R.A. Kerstetter, S.L. Salzberg, W. Schwab, S.S. Korban, J. Davik, A. Monfort, B. Denoyes-Rothan, P. Arus, R. Mittler, B. Flinn, and K.M. Foltá. 2011. Whole genome sequencing and analysis of diploid strawberry, *Fragaria vesca*, a compact and versatile perennial Rosaceous fruit. *Nat. Genet.* 43:109–116. doi:10.1038/ng.740
- Spigler, R.B., K.S. Lewers, A.L. Johnson, and T.L. Ashman. 2010. Comparative mapping reveals autosomal origin of sex chromosome in octoploid *Fragaria virginiana*. *J. Hered.* 101:S107–S117. doi:10.1093/jhered/esq001
- Tuskan, G.A., S. DiFazio, S. Jansson, J. Bohlmann, I. Grigoriev, U. Hellsten, N. Putnam, S. Ralph, S. Rombauts, A. Salamov, J. Schein, L. Sterck, A. Aerts, R.R. Bhallerio, R.P. Bhallerio, D. Blaudez, W. Boerjan, A. Brun, A. Brunner, V. Busov, M. Campbell, J. Carlson, M. Chalot, J. Chapman, G.L. Chen, D. Cooper, P.M. Coutinho, J. Couturier, S. Covert, Q. Cronk, R. Cunningham, J. Davis, S. Degroove, A. Dejardin, C. Depamphilis, J. Detter, B. Dirks, I. Dubchak, S. Duplessis, J. Ehrling, B. Ellis, K. Gendler, D. Goodstein, M. Gribskov, J. Grimwood, A. Groover, L. Gunter, B. Hamberger, B. Heinze, Y. Helariutta, B. Henrissat, D. Holligan, R. Holt, W. Huang, N. Islam-Faridi, S. Jones, M. Jones-Rhoades, R. Jorgensen, C. Joshi, J. Kangasjarvi, J. Karlsson, C. Kelleher, R. Kirkpatrick, M. Kirst, A. Kohler, U. Kalluri, F. Larimer, J. Leebens-Mack, J.C. Leple, P. Locascio, Y. Lou, S. Lucas, F. Martin, B. Montanini, C. Napoli, D.R. Nelson, C. Nelson, K. Nieminen, O. Nilsson, V. Pereda, G. Peter, R. Philippe, G. Pilate, A. Poliakov, J. Razumovskaya, P. Richardson, C. Rinaldi, K. Ritland, P. Rouze, D. Ryaboy, J. Schmutz, J. Schrader, B. Segerman, H. Shin, A. Siddiqui, F. Sterky, A. Terry, C.J. Tsai, E. Uberbacher, P. Unneberg, J. Vahala, K. Wall, S. Wessler, G. Yang, T. Yin, C. Douglas, M. Marra, G. Sandberg, Y. Van de Peer, and D. Rokhsar. 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313:1596–1604. doi:10.1126/science.1128691
- van Dijk, T., T. Dubos, Y. Noordijk, A. Pikunova, H. Yilmaz-Temel, and E. van de Weg. 2010. Increased effectiveness of mapping in octoploid strawberry through quantitative interpretation of SSR data. *In Plant and Animal Genomes XVIII Conf. Fruit and Nut Crops*, San Diego, CA. 9–13 Jan. 2010. Scherago International Inc., Jersey City, NJ.
- Van Ooijen, J.W. 2006. Joinmap 4. Software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen, The Netherlands.
- Velasco, R., A. Zharkikh, J. Affourtit, A. Dhingra, A. Cestaro, A. Kalyanaraman, P. Fontana, S.K. Bhatnagar, M. Troggio, D. Pruss, S. Salvi, M. Pindo, P. Baldi, S. Castelletti, M. Cavaiuolo, G. Coppola, F. Costa, V. Cova, A. Dal Ri, V. Goremykin, M. Komjanc, S. Longhi, P. Mag-nago, G. Malacarne, M. Malnoy, D. Micheletti, M. Moretto, M. Perazzolli, A. Si-Ammour, S. Vezzulli, E. Zini, G. Eldredge, L.M. Fitzgerald, N. Gutin, J. Lanchbury, T. Macalma, J.T. Mitchell, J. Reid, B. Wardell, C. Kodira, Z. Chen, B. Desany, F. Niaz, M. Palmer, T. Koepke, D. Jiwan, S. Schaeffer, V. Krishnan, C. Wu, V.T. Chu, S.T. King, J. Vick, Q. Tao, A. Mraz, A. Stormo, K. Stormo, R. Bogden, D. Ederle, A. Stella, A. Vecchiatti, M.M. Kater, S. Masiero, P. Lasserre, Y. Lespinasse, A.C. Allan, V. Bus, D. Chagne, R.N. Crowhurst, A.P. Gleave, E. Lavezzo, J.A. Fawcett, S. Proost, P. Rouze, L. Sterck, S. Toppo, B. Lazzari, R.P. Hellens, C.E. Durel, A. Gutin, R.E. Bumgarner, S.E. Gardiner, M. Skolnick, M. Egholm, Y. Van de Peer, F. Salamini, and R. Viola. 2010. The genome of the domesticated apple (*Malus × domestica* Borkh.). *Nat. Genet.* 42:833–839. doi:10.1038/ng.654
- Velasco, R., A. Zharkikh, M. Troggio, D.A. Cartwright, A. Cestaro, D. Pruss, M. Pindo, L.M. FitzGerald, S. Vezzulli, J. Reid, G. Malacarne, D. Iliev, G. Coppola, B. Wardell, D. Micheletti, T. Macalma, M. Facci, J.T. Mitchell, M. Perazzolli, G. Eldredge, P. Gatto, R. Oyzerski, M. Moretto, N. Gutin, M. Stefanini, Y. Chen, C. Segala, C. Davenport, L. Dematte, A. Mraz, J. Battilana, K. Stormo, F. Costa, Q.Z. Tao, A. Si-Ammour, T. Harkins, A. Lackey, C. Perbost, B. Taillon, A. Stella, V. Solovyev, J.A. Fawcett, L. Sterck, K. Vandepoele, S.M. Grando, S. Toppo, C. Moser, J. Lanchbury, R. Bogden, M. Skolnick, V. Sgaramella, S.K. Bhatnagar, P. Fontana, A. Gutin, Y. Van de Peer, F. Salamini, and R. Viola. 2007. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PLoS ONE* 2:18. doi:10.1371/journal.pone.0001326
- Vilanova, S., D.J. Sargent, P. Arus, and A. Monfort. 2008. Synteny conservation between two distantly-related Rosaceae genomes: *Prunus* (the stone fruits) and *Fragaria* (the strawberry). *BMC Plant Biol.* 8:12. doi:10.1186/1471-2229-8-67
- Vogel, J.P., D.F. Garvin, T.C. Mockler, J. Schmutz, D. Rokhsar, M.W. Bevan, K. Barry, S. Lucas, M. Harmon-Smith, K. Lail, H. Tice, J. Grimwood, N. McKenzie, N.X. Huo, Y.Q. Gu, G.R. Lazo, O.D. Anderson, F.M. You, M.C. Luo, J. Dvorak, J. Wright, M. Febrer, D. Idziak, R. Hasterok, E. Lindquist, M. Wang, S.E. Fox, H.D. Priest, S.A. Filichkin, S.A. Givan, D.W. Bryant, J.H. Chang, H.Y. Wu, W. Wu, and A.P. Hsia. S. SP, A. Kalyanaraman, B. Barbazuk, T.P. Michael, S.P. Hazen, J.N. Bragg, D. Laudencia-Chingcuanco, Y.Q. Weng, G. Haberer, M. Spannagl, K. Mayer, T. Rattei, T. Mitros, S.J. Lee, J.K.C. Rose, L.A. Mueller, T.L. York, T. Wicker, J.P. Buchmann, J. Tanskanen, A.H. Schulman, H. Gundlach, A.C. de Oliveira, L.D. Maia, W. Belknap, N. Jiang, J.S. Lai, L.C. Zhu, J.X. Ma, C. Sun, E. Pritham, J. Salse, F. Murat, M. Abrouk, R. Bruggmann, J. Messing, N. Fahlgren, C.M. Sullivan, J.C. Carrington, E.J. Chapman, G.D. May, J.X. Zhai, M. Ganssmann, S.G.R. Gurazada, M. German, B.C. Meyers, P.J. Green, L. Tyler, J.J. Wu, J. Thomson, S. Chen, H.V. Scheller, J. Harholt, P. Ulvskov, J.A. Kimbrel, L.E. Bartley, P.J. Cao, K.H. Jung, M.K. Sharma, M. Vega-Sanchez, P. Ronald, C.D. Dardick, S. De Bodt, W. Verelst, D. Inze, M. Heese, A. Schnittger, X.H. Yang, U.C. Kalluri, G.A. Tuskan, Z.H. Hua, R.D. Vierstra, Y. Cui, S.H. Ouyang, Q.X. Sun, Z.Y. Liu, A. Yilmaz, E. Grotewold, R. Sibout, K. Hematy, G. Mouille, H. Hofte, J. Pelloux, D. O'Connor, J. Schnable, S. Rowe, F. Harmon, C.L. Cass, J.C. Sedbrook, M.E. Byrne, S. Walsh, J. Higgins, P.H. Li, T. Brutnell, T. Unver, H. Budak, H. Belcram, M. Charles, B. Chalhoub, and I. Baxter. 2010. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 463:763–768.

- Wenzl, P., H.B. Li, J. Carling, M.X. Zhou, H. Raman, E. Paul, P. Hearnden, C. Maier, L. Xia, V. Caig, J. Ovesna, M. Cakir, D. Poulsen, J.P. Wang, R. Raman, K.P. Smith, G.J. Muehlbauer, K.J. Chalmers, A. Kleinhofs, E. Huttner, and A. Kilian. 2006. A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. *BMC Genomics* 7:206. doi:10.1186/1471-2164-7-206
- Wu, S.B., T.K. Franks, P. Hunt, M.G. Wirthensohn, J.P. Gibson, and M. Sedgley. 2010. Discrimination of SNP genotypes associated with complex haplotypes by high resolution melting analysis in almond: Implications for improved marker efficiencies. *Mol. Breed.* 25:351–357. doi:10.1007/s11032-009-9324-5
- Yamamoto, T., T. Kimura, S. Terakami, C. Nishitani, Y. Sawamura, T. Saito, K. Kotobuki, and T. Hayashi. 2007. Integrated reference genetic linkage maps of pear based on SSR and AFLP markers. *Breed. Sci.* 57:321–329. doi:10.1270/jsbbs.57.321
- Yu, J., S.N. Hu, J. Wang, G.K.S. Wong, S.G. Li, B. Liu, Y.J. Deng, L. Dai, Y. Zhou, X.Q. Zhang, M.L. Cao, J. Liu, J.D. Sun, J.B. Tang, Y.J. Chen, X.B. Huang, W. Lin, C. Ye, W. Tong, L.J. Cong, J.N. Geng, Y.J. Han, L. Li, W. Li, G.Q. Hu, X.G. Huang, W.J. Li, J. Li, Z.W. Liu, L. Li, J.P. Liu, Q.H. Qi, J.S. Liu, L. Li, T. Li, X.G. Wang, H. Lu, T.T. Wu, M. Zhu, P.X. Ni, H. Han, W. Dong, X.Y. Ren, X.L. Feng, P. Cui, X.R. Li, H. Wang, X. Xu, W.X. Zhai, Z. Xu, J.S. Zhang, S.J. He, J.G. Zhang, J.C. Xu, K.L. Zhang, X.W. Zheng, J.H. Dong, W.Y. Zeng, L. Tao, J. Ye, J. Tan, X.D. Ren, X.W. Chen, J. He, D.F. Liu, W. Tian, C.G. Tian, H.G. Xia, Q.Y. Bao, G. Li, H. Gao, T. Cao, J. Wang, W.M. Zhao, P. Li, W. Chen, X.D. Wang, Y. Zhang, J.F. Hu, J. Wang, S. Liu, J. Yang, G.Y. Zhang, Y.Q. Xiong, Z.J. Li, L. Mao, C.S. Zhou, Z. Zhu, R.S. Chen, B.L. Hao, W.M. Zheng, S.Y. Chen, W. Guo, G.J. Li, S.Q. Liu, M. Tao, J. Wang, L.H. Zhu, L.P. Yuan, and H.M. Yang. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp *indica*). *Science* 296:79–92. doi:10.1126/science.1068037
- Zhang, J., and H. Deng. 2010. Analysis on microsatellites derived from strawberry ESTs. *Agric. Sci. Technol. Hunan* 11:164–166.