Available online at www.sciencedirect.com
ScienceDirect
GENOMICS
Genomics 89 (2007) 630–637
www.elsevier.com/locate/ygeno

A BAC-based physical map of the apple genome

Yuepeng Han [a], Ksenija Gasic [a], Brandy Marron [b], Jonathan E. Beever [b], Schuyler S. Korban [a,*]

[a] *Department of Natural Resources and Environmental Sciences, University of Illinois, Urbana, IL 61801, USA*
[b] *Department of Animal Sciences, University of Illinois, Urbana, IL 61801, USA*

Received 31 October 2006; accepted 20 December 2006
Available online 31 January 2007

Abstract

Genome-wide physical mapping is an essential step toward investigating the genetic basis of complex traits as well as pursuing genomics research of virtually all plant and animal species. We have constructed a physical map of the apple genome from a total of 74,281 BAC clones representing ∼10.5× haploid genome equivalents. The physical map consists of 2702 contigs, and it is estimated to span ∼927 Mb in physical length. The reliability of contig assembly was evaluated by several methods, including assembling contigs using variable stringencies, assembling contigs using fingerprints from individual libraries, checking consensus maps of contigs, and using DNA markers. Altogether, the results demonstrated that the contigs were properly assembled. The apple genome-wide BAC-based physical map represents the first draft genome sequence not only for any member of the large Rosaceae family, but also for all tree species. This map will play a critical role in advanced genomics research for apple and other tree species, including marker development in targeted chromosome regions, fine-mapping and isolation of genes/QTL, conducting comparative genomics analyses of plant chromosomes, and large-scale genomics sequencing.
© 2007 Elsevier Inc. All rights reserved.

Keywords: BAC clone fingerprints; Contig map; Apple genome; Contig assembly

Apples are among the most popular and important fruit trees in the world. The domesticated apple, *Malus × domestica* Borkh., belongs to the Rosaceae family. The family consists of over 100 genera and 3000 species, most of which are perennial trees, shrubs, and herbs [1,2]. The apple is self-incompatible and highly heterozygous and displays a juvenile period of 6 to 10 years or more. These characteristics seriously hamper apple breeding efforts. To save time and space and reduce cost, it is imperative to identify young seedlings with desirable horticultural traits accurately using molecular marker-assisted selection. Hence, identifying molecular markers linked to major genes/quantitative trait loci (QTL) contributing to desirable economic traits has become an important goal in apple genetics studies. To meet this goal, genetic tools such as genetic linkage maps, bacterial artificial chromosome (BAC) libraries, and expressed sequence tags (ESTs) have been recently developed [3–8] (see also http://titan.biotec.uiuc.edu/apple/). To date, molecular markers that are either close to or within genes responsible for a few important traits have been developed [9,10].

It has been widely reported that genome-wide physical maps not only serve as platforms for large-scale genome sequencing efforts, but also are very helpful for various other purposes such as development of DNA markers for a genomic region of interest, QTL fine-mapping, effective positional cloning of genes, high-throughput EST mapping (functional genomics), and comparative genomics (synteny studies) [11,12]. To facilitate future advanced genomics research, such as gene and/or QTL fine-mapping as well as structural and functional analyses of the apple genome, it is necessary to develop a genome-wide physical map of the apple genome.

To date, physical maps have been constructed for human, various animals, and many other organisms [13–17]. In plants, physical maps have been established for *Arabidopsis thaliana* [18], sorghum [19], rice [20], and soybean [12]. Genome-wide physical maps have already proven to be powerful tools and infrastructures for advanced genomics research of human and several model species. To develop these various physical maps, several approaches have been developed, including BAC

* Corresponding author. Fax: +1 217 333 8298.
E-mail address: korban@uiuc.edu (S.S. Korban).

0888-7543/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2006.12.010

restriction-based fingerprinting [15,21], iterative hybridization [18], and sequence tag connectors (STCs) [22] involving use of BAC-end sequences for connecting BAC clones by sequence identity. The restriction-based fingerprinting method is less hindered by the presence of repeated sequences within a genome than the iterative hybridization method, and it is much faster and economical than the STC method. Therefore, restriction-based fingerprinting offers a reasonable and powerful means of rapid development of genome-wide physical maps. In fact, restriction-based fingerprinting has been successfully applied in the physical mapping of large complex genomes, including human [15], chicken [11,17], sorghum [19], rice [20,23], and soybean [12].

Recently, progress has been made in genomics research of various fruit trees and woody plants. For example, efforts for constructing a physical map and integrating the physical map with the linkage map in *Prunus*, almond, and peach, are well under way [24–28] (see also http://www.bioinfo.wsu.edu/gdr/). The draft genome sequence of the black cottonwood tree, *Populus trichocarpa*, has been completed using a shotgun-based sequencing strategy [29]. However, no genome-wide physical map has been reported so far for the apple, or any other member of the Rosaceae family, or for any tree species. The apple not only is a major economic fruit crop grown worldwide, but also serves as an important model species for functional genomics research of woody perennial angiosperms due to its relatively small genome size of 750 Mb/haploid and availability of various genomics resources including ESTs, genetic maps, and BAC libraries [3–10,30]. Therefore, developing a whole-genome map for the apple not only will play a critical role in our understanding of the apple genome structure and function, but also will be useful in pursuing plant comparative genomics studies, particularly between annual herbaceous and woody perennial plants. We report here the first genome-wide BAC-based physical map of the apple genome using an agarose gel-based restriction fingerprinting method [15,17].

Results

BAC fingerprinting

A total of 82,503 BAC clones, derived from two complementary BAC libraries, were fingerprinted using the agarose gel-based restriction fingerprinting method [13]. An example of a DNA fingerprinting agarose gel is shown in Fig. 1. Of these clones, 8222 (9.96%) were deleted during fingerprint editing due to either nonrecombinant clones or cross-contamination between clones. Thus, a total of 74,281 clones were successfully fingerprinted and entered into the FPC database for contig assembly. These clones represented ~10.5× haploid genome equivalents. Among those clones, 44.4%, equivalent to 4.4× haploid genomes, were from the *Bam*HI library, and 55.6%, equivalent to 6.1× haploid genomes, were from the *Hin*dIII library. Clones from the *Bam*HI library and *Hin*dIII library had an average of 23.6 and 26.0 bands per clone, respectively (Table 1). In addition, among useful fingerprints, 45 were derived from DNA of the wild crabapple species *Malus floribunda* 821, while all remaining others were derived from DNA of the apple cultivar “GoldRush.”

Fingerprint analysis and contig assembly

To assess the accuracy of the fingerprints, frequencies of each migration value for all clones from the *Hin*dIII library

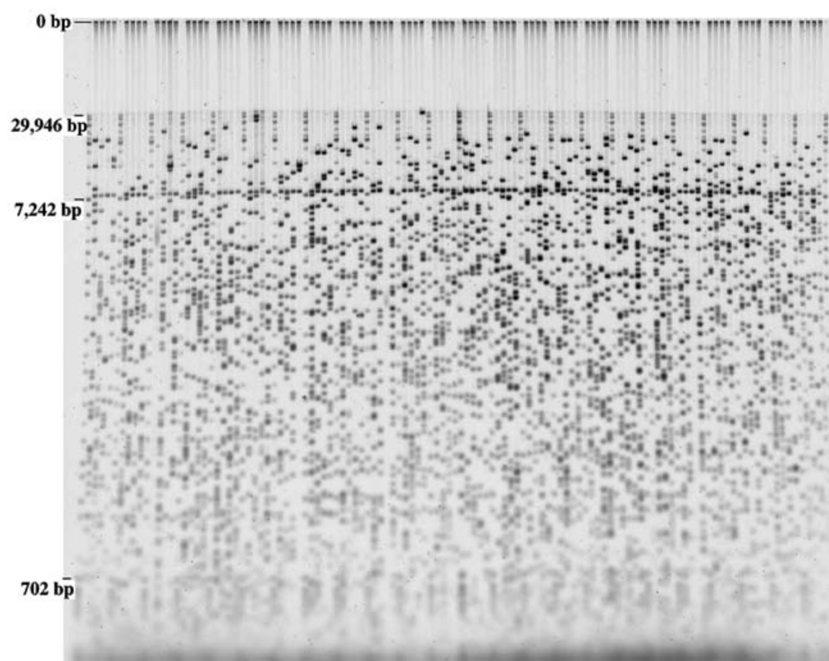

Fig. 1. Agarose gel exhibiting 96 BAC clones digested with *Hin*dIII. DNA size standards are present in every fifth lane. The gel was stained using SYBR green and visualized by fluorescence.

Table 1
Sources of BACs fingerprinted for the apple physical map

Cloning site	No. of clones fingerprinted	Mean insert size (kb)	No. of clones used in mapping	Valid bands per clone	Genome coverage
*Bam*HI	35,712	105	32,992	23.6	4.4×
*Hin*dIII	46,791	115	41,289	26.0	6.1×
Total	82,503	110	74,281	24.9	10.5×

were calculated. The results are presented in Fig. 2. Since BAC fingerprints were derived from *Hin*dIII complete digestion, all clones contained a vector fragment of ~7.5 kb in size. Vector migrations from different gels were very similar, and their values were close to ~1298 (Fig. 2). Moreover, vector fragments exhibited higher frequencies than any of the other bands (Fig. 2). These results clearly indicated that there was little variation among gels and that our fingerprint data were reliable.

Contig assembly was performed using the program FPC version 7.2 [31] (see also http://www.agcol.arizona.edu/software/fpc/). To determine the appropriate cutoff value or "Sulston score," the tolerance and cutoff values were varied, and their effects on known overlapping clones were evaluated. A tolerance of 7 and a cutoff value of 3×10^{-9} were finally used for automatic contig assembly. Contigs with five or more questionable clones were rearranged at higher stringencies (lower cutoff values) using the DQer function of the FPC version 7.2. The DQer automatically reanalyzed those contigs with five or more questionable clones by reassembling clones up to three times, each time the cutoff value being lowered by a factor of 10. We ran the DQer function for several times at a final cutoff down to 3×10^{-16} (by setting at 3×10^{-13}). Following the automatic assembly, 68,058 BAC clones (92%) were assembled into 3943 contigs (Table 2). The physical length of the automated contigs was estimated to be 943.8 Mb, based on 242,001 unique bands, and each band was equivalent to 3.9 kb (Table 2).

Subsequent to automated map assembly, a manual review of the assembly was conducted as it is an essential step for refining the relative order of clones within contigs, identifying joints between contigs, and disassembling larger chimeric contigs. First, we manually checked every contig using the FPC functions of Calc CB map, the Contig window, and the Fingerprint window. Of 3943 contigs, only 351 (8.9%) contained 1–4 questionable clones (more than 50% of bands were unmatched), most of which had only 1 questionable clone. All questionable contigs were then either split at a higher stringency or rearranged after removing the questionable clone(s). Potential chimeric contigs that failed to overlap according to fingerprinting patterns of clones were disassembled. Second, to identify potential junctions, we used those clones at extreme ends of each contig to query the FPC database at a lower required fingerprint overlapping stringency (first cutoff at 3×10^{-8} and then at 3×10^{-7}) than was used during initial assembly. Contig pairs were merged if their terminal clones shared more than 10 bands and their overall fingerprint patterns supported the junction. Individual singleton clones were also added to contigs as needed to increase coverage of sparse regions. As a result, the total number of contigs of the physical map was reduced to 2702 (Table 2). The assembled 2702 contigs consisted of 237,763 unique bands collectively spanning 927.3 Mb in physical length. The longest contig comprised 287 clones, encompassing 702 unique bands

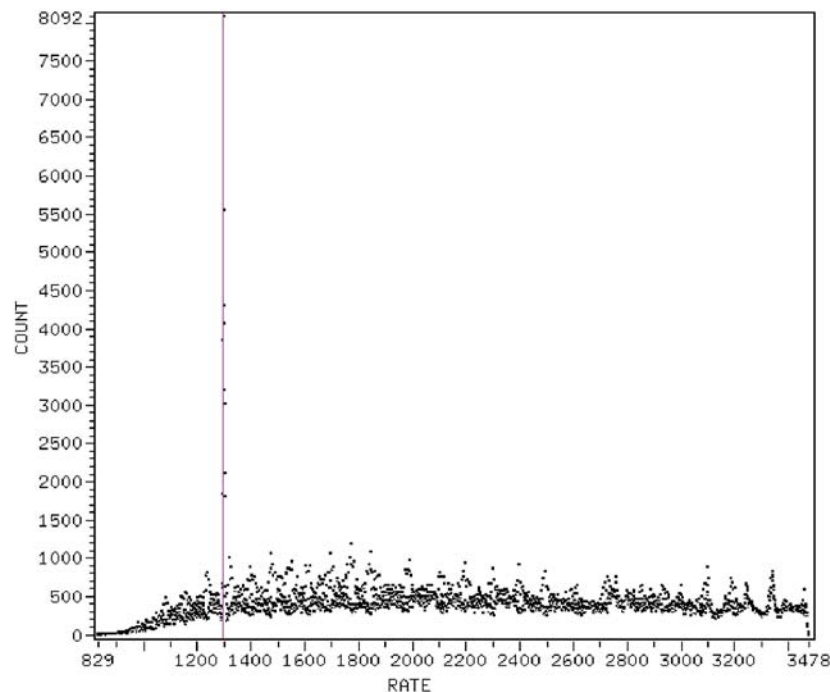

Fig. 2. Frequency distribution of band migrations from the *Hin*dIII library. Each band value in the FPC database of the *Hin*dIII library was graphed against the number of times it was found in the database. Vector fragments had a much higher frequency than any other bands, with an average migration of ~1298.

Table 2
Status of the apple physical map before and after manual editing

	Automatic contig assembly	After manual editing
Number of clones in FPC database	74,281	74,281
Number of singletons	6,223	5,953
Number of contigs	3,943	2,702
Contigs containing		
>200 clones	2	5
101–200 clones	21	67
51–100 clones	212	266
26–50 clones	522	537
10–25 clones	1,407	945
3–9 clones	1,534	754
2 clones	245	128
Unique bands of the contigs	242,001	237,763
Physical length of the contigs in megabase pairs	943.8	927.3

and spanning 2.7 Mb in physical length. An example of BAC contigs of the physical map and the distribution of the BACs from the apple libraries is presented in Fig. 3.

Contig reliability

Several different approaches were used to assess contig reliability. First, we determined the stability of contigs at different cutoff values. By increasing the stringency of contig assembly from 3×10^{-9} to 3×10^{-10}, the number of contigs increased from 3943 to 4718. Hundreds of contigs assembled at the higher stringency were randomly selected and compared with corresponding contigs assembled at the lower stringency. A major difference was observed in clone content due to contig split at the higher stringency, but this was not detected in clone order. Of the 3943 contigs, approximately 380 were split at the higher stringency, which was less than 10% of the initial total contig number. Second, we assembled contigs using separate fingerprint data from each of the *Bam*HI and *Hin*dIII BAC libraries. A total of 100 randomly selected contigs were assembled from each of the two libraries and compared with their corresponding contigs in the physical map. The results showed that 92 and 96% of the contigs from *Bam*HI and *Hin*dIII BAC libraries, respectively, were shown to be in complete agreement with their corresponding contigs in the physical map in both clone content and order.

For the third approach, we checked contig score and the number of extra bands for each contig using the consensus band maps (Fig. 4). The contig score is used as an indicator of group alignment of all clones within a contig [32]. The majority of the contigs had a contig score ranging from 0.88 to 1.0, whereas a small number of contigs had a score ranging between 0.81 and 0.88. Meanwhile, if a clone within a contig was not a questionable clone, but had more than 10 extra bands (only a few such cases were encountered in this study), the best match for the clone was determined. The clone was then either removed or rearranged within the same contig. Moreover, fingerprint patterns in the Fingerprint window were also used to evaluate contig reliability (Fig. 5). We checked fingerprint patterns of every contig to ensure that each clone within a contig was properly ordered with respect to its most closely related neighboring clones.

Finally, we checked three contigs using either DNA markers or PCR probes. The first was the contig spanning the region of the *Vf* gene, which is responsible for apple scab resistance. The *Vf*-linked SCAR marker ACS-3 was first used to screen the

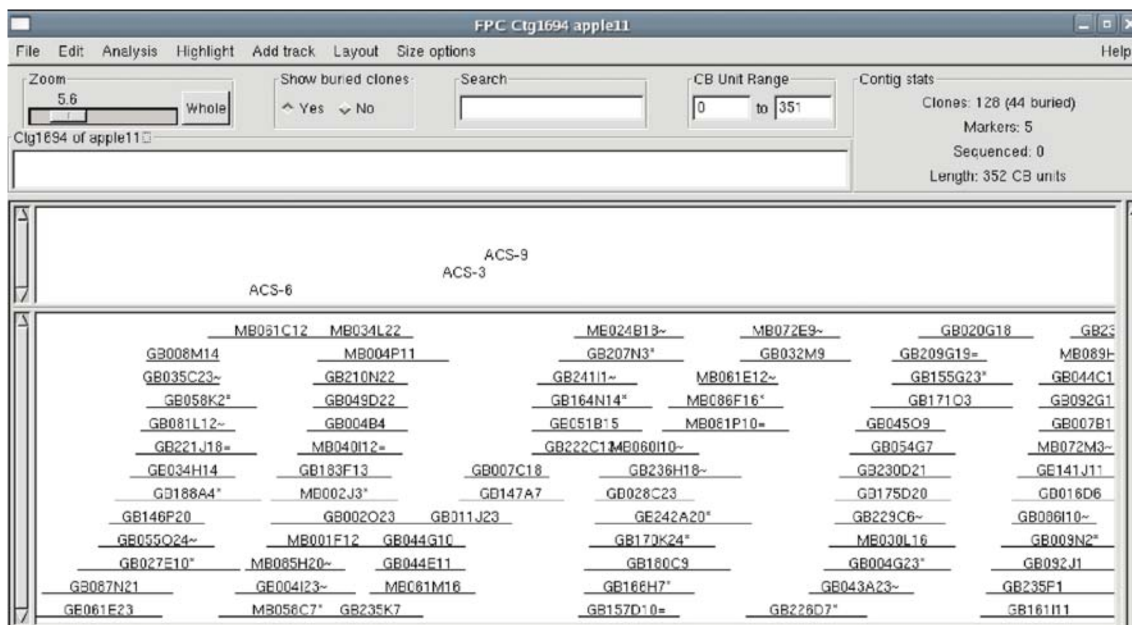

Fig. 3. Example of the BAC contigs of the apple physical map spanning the *Vf* gene region. Only partial clones of the contig are shown. The clones prefixed with "GB" were constructed from the cultivar GoldRush and those with "MB" from the wild crabapple *M. floribunda* 821. ACS-3, ACS-6, ACS-8, ACS-9, and ACS-10 were DNA-based markers tightly linked to the *Vf* locus.

Fig. 4. Example of a consensus band map of a BAC contig of the apple physical map.

Fig. 5. Example of the clone order fingerprints of a BAC contig of the apple physical map.

*Bam*HI BAC library, and three positive BAC clones were identified. The positive clones were then used to detect corresponding contigs by searching the FPC database, and a contig containing the three clones was found. The *Vf* contig derived from the automatic assembly consisted of 30 BAC clones and was 420 kb in size. Another two SCAR markers, ACS-6 and ACS-9, closely linked to the *Vf* gene were further used to verify contig reliability. All 30 clones within the *Vf* contig were screened using these two SCAR markers, and 8 and 3 positive BAC clones were detected for ACS-6 and ACS-9, respectively. These results were consistent with clone orders within the contig, thus indicating that the contig was correctly assembled (Fig. 3). The *Vf* contig was finally merged with the other two contigs at a lower stringency and extended to 1.6 Mb in size. The contig was further screened with two additional SCAR markers, ACS-1 and ACS-4, located on the opposite side of the *Vf* gene. A total of 94 clones from the two contigs, which were merged to the *Vf* contig, were tested. A total of 7 positive clones were identified, and they were correctly assembled within the contig, thus confirming the reliability of the *Vf* contig.

The second of the three contigs checked was a contig covering the apple starch branching enzyme *SbeI* gene, which is a single-copy gene in the apple genome as detected by Southern blot analysis [33]. The PCR probe, derived from the *SbeI* gene, was used to screen the *Bam*HI BAC library, and 3 positive clones were detected. These 3 clones were found to be located in the same contig after a search of the FPC database. The *SbeI* contig consisted of 14 clones and was 240 kb in size. The other 11 clones were screened with the PCR *SbeI* probe, and 5 clones from the *Hin*dIII BAC library were identified to be positive. Probes derived from end sequences of these BACs were further used to test the remaining 6 clones, and the results showed that they were in fact members of the *SbeI* contig.

The third of the three contigs checked was a contig covering the *Comt* gene encoding the caffeic acid *O*-methyltransferase. The contig consisted of 31 clones and was 380 kb in size. A PCR probe derived from the *Comt* gene was used to test for all 31 clones, and 13 positive clones were detected. The remaining 18 clones were further tested using probes derived from BAC-end sequences, and the results indicated that they were all overlapping, thus also suggesting that the contig was correctly assembled.

Discussion

The process of developing fingerprint clone maps is based upon complete restriction enzyme digests of clones representative of a target genome that ultimately consist of a near-contiguous path of clones across the genome [17]. These physical maps are useful in validating sequence assembly order, providing linking information of assembled sequences, anchoring sequences to the genetic map, and providing templates for closing intercontig gaps [17]. In this study, we have successfully fingerprinted a total of 74,281 BAC clones representing ~10.5× haploid genome equivalents of the apple. The cloned fingerprints have been used to construct a genome-wide physical map for the apple. The map consists of 2702 contigs collectively spanning 927.3 Mb in physical length. The total contig map length is ~177 Mb (23.6%) greater than the estimated 750 Mb size of the apple genome [30]. An observed discrepancy between the total physical length of contigs and the estimated genome size of the apple has been similarly reported in studies of genome-wide physical maps of the human [15], chicken [11], and soybean [12]. Reasons for these discrepancies may be primarily attributed to failures in detecting overlaps between adjacent contigs, although underestimation of the size of a genome may also contribute to this observed discrepancy and therefore cannot be ruled out. Detecting and merging overlapping contigs are critical steps in the development of genome-wide physical maps. For example, in the physical mapping effort of the chicken genome [17], a total of 6509 contigs were initially generated using the automated assembly, and following a manual review, these were joined to produce 320 contigs. Therefore, it is plausible that the contigs of the apple genome reported herein will be further merged and refined as additional information regarding DNA markers and probes derived from BAC-end sequence data becomes available.

Genomes of higher plants contain highly repeated DNA sequences [34]. The impact of repeated DNA sequences on the construction of the apple physical map has been estimated. If repeated sequences strongly influence contig assembly, many "deep" contigs (too many clones assembled within a small region) will be generated. In this study, of the 2702 assembled apple contigs, only 1 shows deep assembly. This contig has 455 clones, and it is 268 kb in size. As a single deep contig is observed in the apple physical map, it is reasonable to assume that this deep contig may be derived from chloroplast DNA rather than apple genomic DNA.

As mentioned above, the apple is highly self-incompatible, thus contributing to its high level of heterozygosity [1,2]. In this study, more than 92% of the apple BAC clones have been successfully assembled into contigs, with an average estimated size of ~340 kb in physical length. This suggests that heterozygosity has little influence on contig assembly of the apple genome. This is consistent with early reports that polymorphisms among individuals or cultivars do not contribute to problems in contig assembly, e.g., in the human [15] and soybean [12] genomes.

Polyploidy is a significant evolutionary process in higher organisms. It has been reported that genomes of flowering plants are likely to have incurred one or more polyploidization events during evolution [35]. The cultivated apple (*M.×domestica*) is a diploidized allopolyploid species with 17 haploid chromosomes [36,37]. In this study, it is demonstrated that the diploidized genome of the apple is amenable to physical mapping. This also suggests that divergence of the apple must have occurred prior to and/or subsequent to formation of the original allopolyploid genome.

The fingerprints reported herein are derived from the completed digestion of apple BAC DNAs using the *Hin*dIII enzyme, which commonly recognizes 6-bp sequences. On average, *Hin*dIII cuts DNA every 4^6 or 4096 bp; however, in this study the average band size in our FPC database is only

~3.9 kb. A total of 81 BAC clones have been randomly selected from the *Bam*HI BAC library of GoldRush, and their DNAs have been separately digested using two different restriction enzymes, *Hin*dIII and *Bam*HI, having 6-bp recognition restriction sites. For each of the BAC clones, an average of either 24.7 or 18.5 bands for *Hin*dIII or *Bam*HI digestion, respectively, is obtained. Since the recognition sequences for *Hin*dIII and *Bam*HI are AAGCTT and GGATCC, respectively, altogether these results suggest that the GC and AT contents in the apple genome are not in equilibrium. The analysis of apple ESTs has already shown that the average ratio of the GC content is 44% [8]. Based on results obtained in this study, an average GC content ratio of less than 50% for the whole apple genome can be further inferred.

In this study, we found that the average insert size of apple BAC clones estimated by fingerprinting using *Hin*dIII digestion is slightly smaller than that previously estimated by CHEF electrophoresis using *Not*I digestion [4]. This discrepancy may be attributed to the presence of repetitive DNA sequences as well as approaches used for gel editing following fingerprinting. We have previously reported that the presence of repetitive DNA sequences in the apple genome contributes to observed differences in apple BAC inserts; e.g., insert size of an apple BAC DNA is 200 kb in fiber-FISH analysis compared to 115 kb as estimated by CHEF electrophoresis [4]. Moreover, to obtain accurate BAC fingerprints, several strategies are utilized during the process of band calling using the Image program. First, we found that small marker fragments, i.e., 587 and 540 bp, often penetrated to neighboring lanes, thus contributing to difficulties in band calling of small DNA fragments of BAC clones. Moreover, bands of DNA fragments of less than 702 bp in size are relatively faint and are difficult to call. Likewise, variability in the mobility of these small DNA fragments is considerable and contributes to low estimates of overlaps between clones during contig assembly. Hence, only bands ranging between 702 and 29,950 bp in size are collected for contig assembly. In addition, it is also difficult to call multiple fragments located at nearly the same position on the gel correctly. To increase the reliability of contig assembly, these very close fragments are collapsed into single bands. Finally, BAC clones with more than 50 bands are deleted to reduce the likelihood of cross-contamination between clones. In summary, deletion of small fragments and coalition of very close bands may be responsible for the observed low estimates of BAC insert sizes obtained by the fingerprinting method.

The apple physical map described here represents the first report of a genome-wide physical map for any member of the Rosaceae family or for any tree species. This physical map of the diploidized apple genome, of an allopolyploid origin, will play a critical role in: (a) genome assembly; (b) marker-assisted analysis of germ plasm and breeding materials; (c) dissecting and characterizing the genetic basis of complex traits; (d) fundamental understanding of the apple genome structure and function; (e) studies on polyploidization, postpolyploidization, and evolution of polyploid plants; and (f) accelerating development of plant comparative genomics in higher plants. Moreover, the apple physical map, as well as its ever-expanding genomics resources, together with genomic advances in other tree species such as *Prunus* and *Populus* [28,29], will also provide new opportunities for advanced comparative and functional genomics research of woody perennial angiosperms.

Materials and methods

Source BAC libraries

Two *Bam*HI BAC libraries were previously developed in our laboratory [4,5], one from the cultivated apple, *M.*×*domestica* cv. GoldRush, and the other from the wild crabapple species, *M. floribunda* 821. Each library represented about 5× haploid genome equivalents. To increase the probability of coverage of the whole genome, a complementary library was recently constructed for GoldRush using *Hin*dIII. This library consisted of 46,791 clones and was equivalent to ~7× haploid genome. Nearly all BAC clones used for physical map construction were from the two GoldRush libraries, except for a few BACs from the *M. floribunda* 821 library.

BAC fingerprinting

BAC clones maintained in a 384-well microplate were inoculated in four 96-deep-well plates containing 1.5 ml of 2× LB medium plus 12.5 μl/ml chloramphenicol using a 96-well replicator pin tool. Plates were covered with AirPore gas-permeable plate sealant and incubated at 37 °C with continuous shaking at 325 rpm for 20–24 h. BAC DNA was then isolated using a modified alkaline lysis method. Purified BAC DNA was digested with *Hin*dIII at 37 °C for 3 h. Restriction fragments were separated by agarose gel electrophoresis on custom Latitude HT 121-well precast gels (Cambrex, Rockland, ME, USA) at 60 V for 15.5 h with buffer recirculation at 4 °C. DNA fragments were stained using SYBR Green I (Invitrogen, Carlsbad, CA, USA) and visualized by fluorescence using the Typhoon 8600 Imager (Amersham Biosciences, Piscataway, NJ, USA).

Fingerprint data collection and BAC contig assembly

Captured gel images were analyzed using Image version 3.10b (http://www.sanger.ac.uk/Software/Image/). Bands ranging from 702 to 29,950 bp in size were collected for contig assembly. Meanwhile, those fragments located at nearly the same position on the gel were collapsed into single bands. Clones failing fingerprinting, due to contamination, lack of marker data, or lacking inserts, were manually deleted. All band data were standardized using DNA markers II and V (Fermentas, Hanover, MD, USA) and converted into migration rates ranging from 891 to 3463. Contig assembly was carried out using the FPC version 7.2 program [31] (see also http://www.agcol.arizona.edu/software/fpc). A series of tests was conducted in which fingerprints of a set of overlapping clones were compared using different tolerances (from 4 to 7) and cutoffs (from 3×10^{-8} to 3×10^{-15}). Based on these tests, a tolerance of 7 and a primary cutoff of 3×10^{-9} were finally selected for contig assembly.

BAC library screening

The *Bam*HI GoldRush BAC library was used for DNA marker analysis. DNA markers and probes were derived from the *Vf* locus (scab resistance gene) [9] and BAC-end sequences, respectively. To facilitate screening of the BAC library by PCR, super and single pools were constructed. The pools consisted of 35,328 BAC clones representing about 5× apple haploid genome equivalents. Positive BAC clones were identified according to the three-step PCR-based screening procedure as described [5].

Acknowledgments

This project was supported by the USDA Cooperative State Research, Education and Extension Service–National Research Initiative, Plant Genome Program Grant 2005-35300-15538.

References

[1] S.S. Korban, H. Chen, Biotechnology of Apples, in: F. Hammerschlag, R. Litz (Eds.), Biotechnology of Fruit Tree Crops, CAB Int., Oxford, 1992, pp. 203–227.

[2] J. Janick, et al., Apples, in: J. Janick, J.N. Moore (Eds.), Fruit Breeding, Tree and Tropical Fruits, vol. 1, Wiley, New York, 1996, pp. 1–77.

[3] B.A. Vinatzer, H.B. Zhang, S. Sansavini, Construction and characterization of a bacterial artificial chromosome library of apple, Theor. Appl. Genet. 97 (1998) 1183–1190.

[4] M. Xu, J. Song, Z. Cheng, J. Jiang, S.S. Korban, A bacterial artificial chromosome (BAC) library of *Malus floribunda* 821 and contig construction for positional cloning of the apple scab resistance gene *Vf*, Genome 44 (2001) 1104–1113.

[5] M. Xu, S.S. Korban, J. Song, J. Jiang, Constructing a bacterial artificial chromosome library of the apple cultivar GoldRush, Acta Horticult. 595 (2002) 103–112.

[6] R. Liebhard, B. Koller, L. Gianfranceschi, C. Gessler, Creating a saturated reference map for the apple (*Malus*×*domestica* Borkh.) genome, Theor. Appl. Genet. 106 (2003) 1497–1508.

[7] S. Naik, C. Hampson, K. Gasic, G. Bakkeren, S.S. Korban, Development and linkage mapping of E-STS and RGAs for functional gene homologues in apple, Genome 49 (2006) 959–968.

[8] R.E. Newcomb, et al., Analysis of expressed sequence tags from apple, Plant Physiol. 141 (2006) 147–166.

[9] M.L. Xu, S.S. Korban, Saturation mapping of the scab resistance gene *Vf* using AFLP markers, Theor. Appl. Genet. 101 (2000) 844–851.

[10] R. Liebhard, et al., Development and characterization of 140 new microsatellites in apple (*Malus*×*domestica* Borkh.), Mol. Breed. 10 (2002) 217–241.

[11] C. Ren, et al., A BAC-based physical map of the chicken genome, Genome Res. 13 (2003) 2754–2758.

[12] C. Wu, et al., A BAC- and BIBAC-based physical map of the soybean genome, Genome Res. 14 (2004) 319–326.

[13] J. Hodgkin, R.H.A. Plasterk, R.H. Waterston, The nematode *Caenorhabditis elegans* and its genome, Science 270 (1995) 410–414.

[14] R.A. Hoskins, et al., A BAC-based physical map of the major autosomes of *Drosophila melanogaster*, Science 287 (2000) 2271–2274.

[15] The International Human Genome Mapping Consortium, A physical map of the human genome, Nature 409 (2001) 934–941.

[16] S.G. Gregory, et al., A physical map of the mouse genome, Nature 418 (2002) 743–750.

[17] J.W. Wallis, et al., A physical map of the chicken genome, Nature 432 (2004) 761–764.

[18] T. Mozo, et al., A complete BAC-based physical map of the Arabidopsis thaliana genome, Nat. Genet. 22 (1999) 271–275.

[19] P.E. Klein, et al., A high-throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map, Genome Res. 10 (2000) 789–807.

[20] M. Chen, et al., An integrated physical and genetic map of the rice genome, Plant Cell 14 (2002) 537–545.

[21] A. Coulson, J. Sulston, S. Brenner, J. Karn, Toward a physical map of the genome of the nematode *C. elegans*, Proc. Natl. Acad. Sci. USA 83 (1986) 7821–7825.

[22] G.G. Mahairas, et al., Sequence-tagged connectors: a sequence approach to mapping and scanning the human genome, Proc. Natl. Acad. Sci. USA 96 (1999) 9739–9744.

[23] Q. Tao, et al., Bacterial artificial chromosome-based physical map of the rice genome constructed by restriction fingerprint analysis, Genetics 158 (2001) 711–724.

[24] R. Horn, et al., Candidate gene database and transcript map for peach, a model species for fruit trees, Theor. Appl. Genet. 110 (2005) 1419–1428.

[25] Y. Wang, et al., High-throughput targeted SSR marker development in peach (*Prunus persica*), Genome 45 (2002) 319–328.

[26] T. Joobeur, et al., Construction of a saturated linkage map for *Prunus* using an almond × peach F_2 progeny, Theor. Appl. Genet. 97 (1998) 1034–1041.

[27] L.L. Georgi, et al., Construction of a BAC library and its application to the identification of simple sequence repeats in peach [*Prunus persica* (L.) Batsch], Theor. Appl. Genet. 105 (2002) 1151–1158.

[28] S. Jung, et al., GDR (Genome Database for Rosaceae): integrated Web resources for Rosaceae genomics and genetics research, BMC Bioinform. 5 (2004) 130.

[29] G.A. Tuskan, et al., The genome of black cottonwood, *Populus trichocarpa* (Torr. and Gray), Science 313 (2006) 1596–1604.

[30] T. Tatum, S. Stepanovic, D.P. Biradar, A.L. Rayburn, S.S. Korban, Variation in nuclear DNA content in Malus species and cultivated apples, Genome 48 (2005) 924–930.

[31] W.M. Nelson, et al., Whole-genome validation of high-information-content fingerprinting, Plant Physiol. 139 (2005) 27–38.

[32] C. Soderlund, S. Humphray, A. Dunham, L. French, Contigs built with fingerprints, markers, and FPC V4.7, Genome Res. 10 (2000) 1772–1787.

[33] Y. Han, K. Gasic, F. Sun, M.L. Xu, A gene encoding starch branching enzyme I (SBEI) in apple (*Malus*×*domestica*, Rosaceae) and its phylogenetic relationship to *Sbe* genes from other angiosperms. Mol. Phylogenet. Evol. doi:10.1016/j.ympev.2006.09.001 (in press).

[34] S. Uozu, H. Ikehashi, N. Ohmido, H. Ohtsubo, E. Ohtsubo, K. Fukui, Repetitive sequences: cause for variation in genome size and chromosome morphology in the genus Oryza, Plant Mol. Biol. 35 (1997) 791–799.

[35] J. Masterson, Stomatal size in fossil plants: evidence for polyploid in majority of angiosperms, Science 264 (1994) 421–424.

[36] J.S. Challice, Rosaceae chemotaxonomy and the origin of the Pomoidae, Bot. J. Linn. Soc. 69 (1974) 239–259.

[37] D. Lespinasse, L. Grivet, V. Troispoux, M. Rodier-Goud, F. Pinard, M. Seguin, Identification of QTLs involved in the resistance to South American leaf blight (*Microcyclus ulei*) in the rubber tree, Theor. Appl. Genet. 100 (2000) 975–984.