THE UNIVERSITY OF RHODE ISLAND

University of Rhode Island
DigitalCommons@URI

Biological Sciences Faculty Publications | Biological Sciences

2008

A BAC-Based Integrated Linkage Map of the Silkworm *Bombyx mori*

Kimiko Yamamoto

Junko Nohata

See next page for additional authors

Creative Commons License

This work is licensed under a Creative Commons Attribution 2.0 License.

Follow this and additional works at: http://digitalcommons.uri.edu/bio_facpubs

Part of the Genomics Commons

Terms of Use
All rights reserved under copyright.

Citation/Publisher Attribution

Yamamoto, K., Nohata, J., Kadono-Okuda, K., Narukawa, J., Sasanuma, M., Sasanuma, S., Minami, H., Shimomura, M., Suetsugu, Y., Banno, Y., Osoegawa, K., de Jong, P. J., Goldsmith, M. R., & Mita, K. (2008). A BAC-based integrated linkage map of the silkworm *Bombyx mori*. *Genome Biology, 9*(1), R21-R21.14.
Available at: http://dx.doi.org/10.1186/gb-2008-9-1-r21

This Article is brought to you for free and open access by the Biological Sciences at DigitalCommons@URI. It has been accepted for inclusion in Biological Sciences Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.

Authors
Kimiko Yamamoto, Junko Nohata, Keiko Kadono-Okuda, Junko Narukawa, Motoe Sasanuma, Shun-ichi Sasanuma, Hiroshi Minami, Michihiko Shimomura, Yoshitaka Suetsugu, Yutaka Banno, Kazutoyo Osoegawa, Pieter J. de Jong, Marian R. Goldsmith, and Kazuei Mita

This article is available at DigitalCommons@URI: http://digitalcommons.uri.edu/bio_facpubs/4

Research

Open Access

A BAC-based integrated linkage map of the silkworm *Bombyx mori*

Kimiko Yamamoto[¤*], Junko Nohata[¤*], Keiko Kadono-Okuda[¤*], Junko Narukawa[*], Motoe Sasanuma[*], Shun-ichi Sasanuma[*], Hiroshi Minami[†], Michihiko Shimomura[†], Yoshitaka Suetsugu[*], Yutaka Banno[‡], Kazutoyo Osoegawa[§], Pieter J de Jong[§], Marian R Goldsmith[¶] and Kazuei Mita[*]

Addresses: [*]Insect Genome Research Unit, National Institute of Agrobiological Sciences, Owashi, Tsukuba, Ibaraki 305-8634, Japan. [†]Genome Project Department, Tsukuba Bank, Mitsubishi Space Software Co., Ltd, Takezono, Tsukuba, Ibaraki 305-8602, Japan. [‡]Laboratory of Insect Genetic Resources, Faculty of Agriculture, Kushu University, Fukuoka 812-8581, Japan. [§]Children's Hospital Oakland Research Institute, 52nd Street, Oakland, California 94609, USA. [¶]Biological Sciences Department, University of Rhode Island, Kingston, Rhode Island 02881-0816, USA.

¤ These authors contributed equally to this work.

Correspondence: Marian R Goldsmith. Email: mki101@uri.edu

Published: 28 January 2008

Genome ***Biology*** 2008, **9:**R21 (doi:10.1186/gb-2008-9-1-r21)

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2008/9/1/R21

Received: 9 August 2007
Revised: 17 December 2007
Accepted: 28 January 2008

© 2008 Yamamoto *et al.*; licensee BioMed Central Ltd.
This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: In 2004, draft sequences of the model lepidopteran *Bombyx mori* were reported using whole-genome shotgun sequencing. Because of relatively shallow genome coverage, the silkworm genome remains fragmented, hampering annotation and comparative genome studies. For a more complete genome analysis, we developed extended scaffolds combining physical maps with improved genetic maps.

Results: We mapped 1,755 single nucleotide polymorphism (SNP) markers from bacterial artificial chromosome (BAC) end sequences onto 28 linkage groups using a recombining male backcross population, yielding an average inter-SNP distance of 0.81 cM (about 270 kilobases). We constructed 6,221 contigs by fingerprinting clones from three BAC libraries digested with different restriction enzymes, and assigned a total of 724 single copy genes to them by BLAST (basic local alignment search tool) search of the BAC end sequences and high-density BAC filter hybridization using expressed sequence tags as probes. We assigned 964 additional expressed sequence tags to linkage groups by restriction fragment length polymorphism analysis of a nonrecombining female backcross population. Altogether, 361.1 megabases of BAC contigs and singletons were integrated with a map containing 1,688 independent genes. A test of synteny using Oxford grid analysis with more than 500 silkworm genes revealed six versus 20 silkworm linkage groups containing eight or more orthologs of *Apis* versus *Tribolium*, respectively.

Conclusion: The integrated map contains approximately 10% of predicted silkworm genes and has an estimated 76% genome coverage by BACs. This provides a new resource for improved assembly of whole-genome shotgun data, gene annotation and positional cloning, and will serve as a platform for comparative genomics and gene discovery in Lepidoptera and other insects.

Background

Genome analysis of insects has moved quickly in recent years, in part because insects are so widespread and diverse, and elucidating their characteristic biological phenomena will yield enormous resources for basic science, agriculture, and industry. Complete genome sequences have been published for 12 *Drosophila* spp. [1-3], *Anopheles gambiae* [4] and *Apis mellifera* [5], and that of *Tribolium castaneum* will appear shortly [6]. In 2004, the draft sequence of *Bombyx mori* was reported independently by groups in Japan [7] and China [8]. Because of relatively shallow genome coverage (3× and 6×, respectively) using the whole-genome shotgun (WGS) sequencing method, the silkworm genome is still somewhat fragmented, with an average contig length of under 6 kilobases (kb). This makes it difficult to identify and annotate genes effectively, or to obtain a global view of the silkworm's general and unique features as a model for Lepidoptera.

Well developed genetic resources for silkworm include more than 400 described morphologic and biochemical mutants [9], affecting such characters as chorion (eggshell) composition and structure; embryo development; larval cuticle transparency, pigmentation, segment identity, and body shape; hemolymph proteins; cocoon color, shape, and texture; and adult fertility, egg laying behavior, eye color, and wing pattern. These have been assigned to more than 200 loci on linkage maps. Additionally, molecular linkage maps composed of various markers including about 1,000 random amplified polymorphic DNAs [10,11], about 250 RFLPs (restriction fragment length polymorphism) [12-14], 545 amplified fragment length polymorphisms [15], more than 500 simple sequence repeats [16,17], more than 500 single nucleotide polymorphism (SNPs) [18], and more than 400 sequence tagged sites for cloned genes and expressed sequence tags (ESTs) [19] have been constructed. Three bacterial artificial chromosome (BAC) libraries [20,21] and more than 185,00 ESTs [22-24] are also available. For a more complete genome assembly and analysis, and to take full advantage of these extensive resources, it is of great importance to combine genetic maps with physical map information. This can be accomplished by connecting genetic mapping data to BAC clones; this is a well established approach that has not been employed in silkworm on a genome-wide basis.

Our aim in the present study was to conduct a complete genome analysis. We report here an integrated map between a high-density SNP genetic map and a physical map of BAC contigs using the following strategy: extension of the previous SNP linkage map using BAC end sequences to produce a second-generation map containing 1,755 SNP markers; construction of BAC contigs using two methods, namely restriction digest fingerprinting of BAC clones and hybridization of ESTs to a BAC library on high-density replica (HDR) filters; and assignment of 1,082 ESTs to existing linkage groups by Southern analysis using a nonrecombining female informative backcross. Finally, we searched for orthologs among 1,688 genes on the updated silkworm linkage maps and tested the level of synteny with honey bee and beetle chromosomes using the Oxford grid method.

Results

Linkage map construction

We previously constructed a linkage map by surveying the segregation patterns of 534 SNPs detected in 190 first-generation backcross (BC_1) individuals from a single pair mating between a p50T female and an F_1 male (p50T female × C108T male) [18]. Based on the analysis of additional data with Mapmaker/exp (version 3.0 [25]; LOD [log of the odds] score 3.0) using the same mapping panel, we successfully positioned a total of 1,755 SNPs on an expanded linkage map. The SNP markers segregated into 28 linkage groups, with a total recombination length of 1,413 cM. We assigned 26 of the SNP linkage groups to classical silkworm chromosomes 1 to 26, defined by morphologic markers (for example, cocoon color and larval markings) and protein polymorphisms (for instance, hemolymph proteins), as reported previously [18]. Unambiguous morphologic markers are not yet available for the remaining two classical linkage groups; therefore, we anchored linkage group 27 to a reference gene (vitellogenin), and arbitrarily defined linkage group 28.

The extended SNP linkage map is illustrated in Figures 1 to 5[26] (Additional data file 1 for details of each marker, including BAC accession number). Basic map parameters are significantly improved in this version. The number of markers per linkage group varies from 20 (group 2) to 105 (group 4; Figure 1), and the recombination length for each linkage group ranges from 42.2 cM (group 15; Figure 3) to 68 cM (group 24; Figure 4). In the previously reported SNP map [18], the minimum and maximum number of markers were seven (groups 26 and A; Figure 5) and 32 (group 10; Figure 2), and the linkage map lengths ranged from 27 cM (group 20; Figure 4) to 64 cM (group 11; Figure 2). The number of markers for individual linkage groups in the revised map increased proportionally to the threefold rise in the total number of markers, but the extension of the linkage maps remained relatively small, because of a higher marker density. The average distance between the markers is 0.81 cM, which is much improved compared with that of the previous map (2.5 cM). The markers are not evenly distributed throughout the linkage map, and so different regions are more densely or sparsely populated. The number of gaps with lengths exceeding 10 cM decreased to five from 14 in the previous map, and the largest gap length decreased to 12.4 cM from 21.3 cM.

BAC contig construction by DNA fingerprinting

We fingerprinted a total of 81,024 BACs from three BAC libraries, each made with a different restriction enzyme (Table 1), using the large-scale agarose gel-based restriction fingerprinting method [27,28]. We used the computer program FPC V6.0 [29,30] to assemble BAC contigs from the

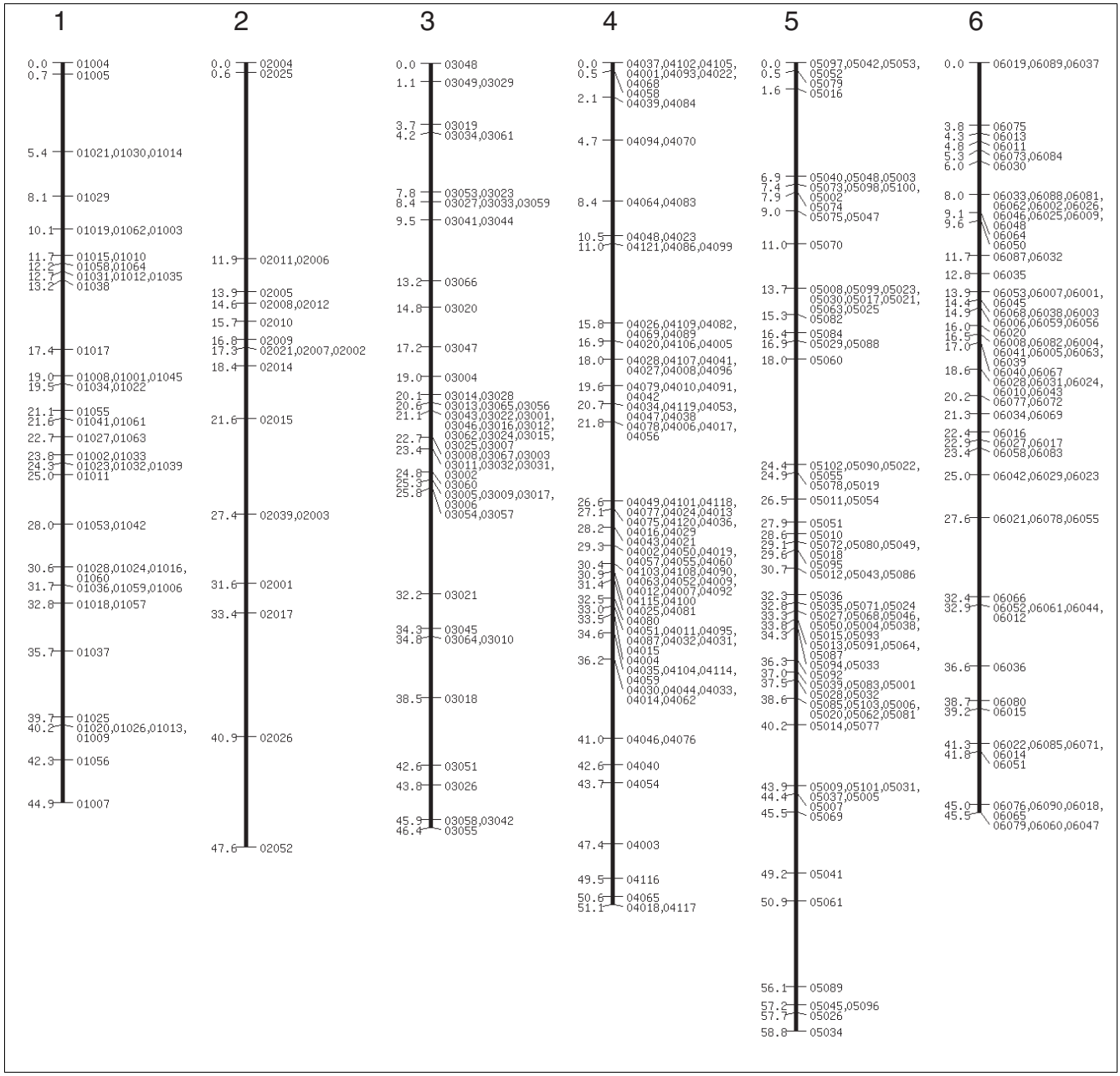

Figure 1

SNP linkage map comprising 1,755 markers: linkage groups 1 to 6. For additional details see Silkworm Genome Research Program [26]. SNP, single nucleotide polymorphism.

BAC fingerprints. We performed preliminary tests to determine appropriate tolerance and cut-off value parameters, and adopted values of 3 and $1 \times e^{-12}$, respectively. A lower stringency condition produced larger contigs, but it also increased the risk for false contigs because of a high density of repetitive sequences.

Of the clones fingerprinted, we deleted 2,246 BACs during fingerprint editing because of insert-empty clones, having too many bands because of possible contamination, or too few bands (fewer than five). In addition, we removed 1,152 BACs from contigs assembled by the FPC program because they formed an extensive false contig, in which several constituting BACs were assigned to different chromosomes by independent BAC-fluorescence *in situ* hybridization experiments (data not shown). The false BAC contig is likely to have formed from transposon-rich regions with similar restriction fingerprints, which were difficult to remove even by the use of high-stringency assembly conditions (namely, with a cut-off below $1 \times e^{-12}$).

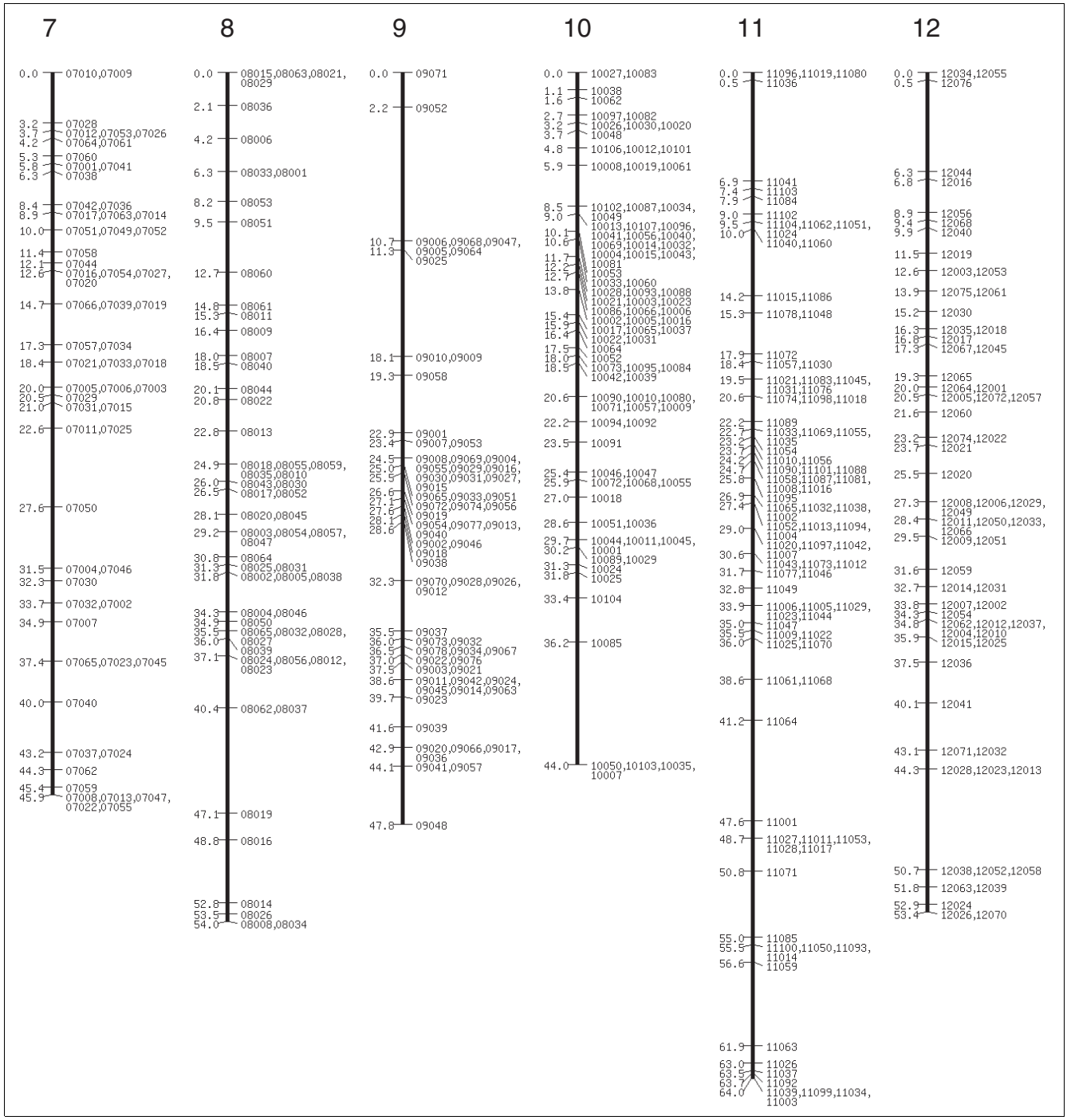

Figure 2
SNP linkage map continued: linkage groups 7 to 12. SNP, single nucleotide polymorphism.

The resulting physical map contained 6,221 contigs, among which 782 major contigs included eight or more BAC clones, as summarized in Table 2 (Additional data file 2). The total length of the 782 major contigs was 376 megabases (Mb), which corresponds to 79% of the silkworm genome (476 Mb measured by flow cytometry; Johnston JS, personal communication).

We evaluated the reliability of the predicted BAC contigs using two approaches. First, using BAC end sequences [31] we designed primer sets for the BAC clones belonging to a putative contig to determine the presence or absence of the sequence in each BAC clone independently by PCR. Using this method, we found mis-assemblies in six contigs out of the 782 major BAC contigs; we could not determine whether nine

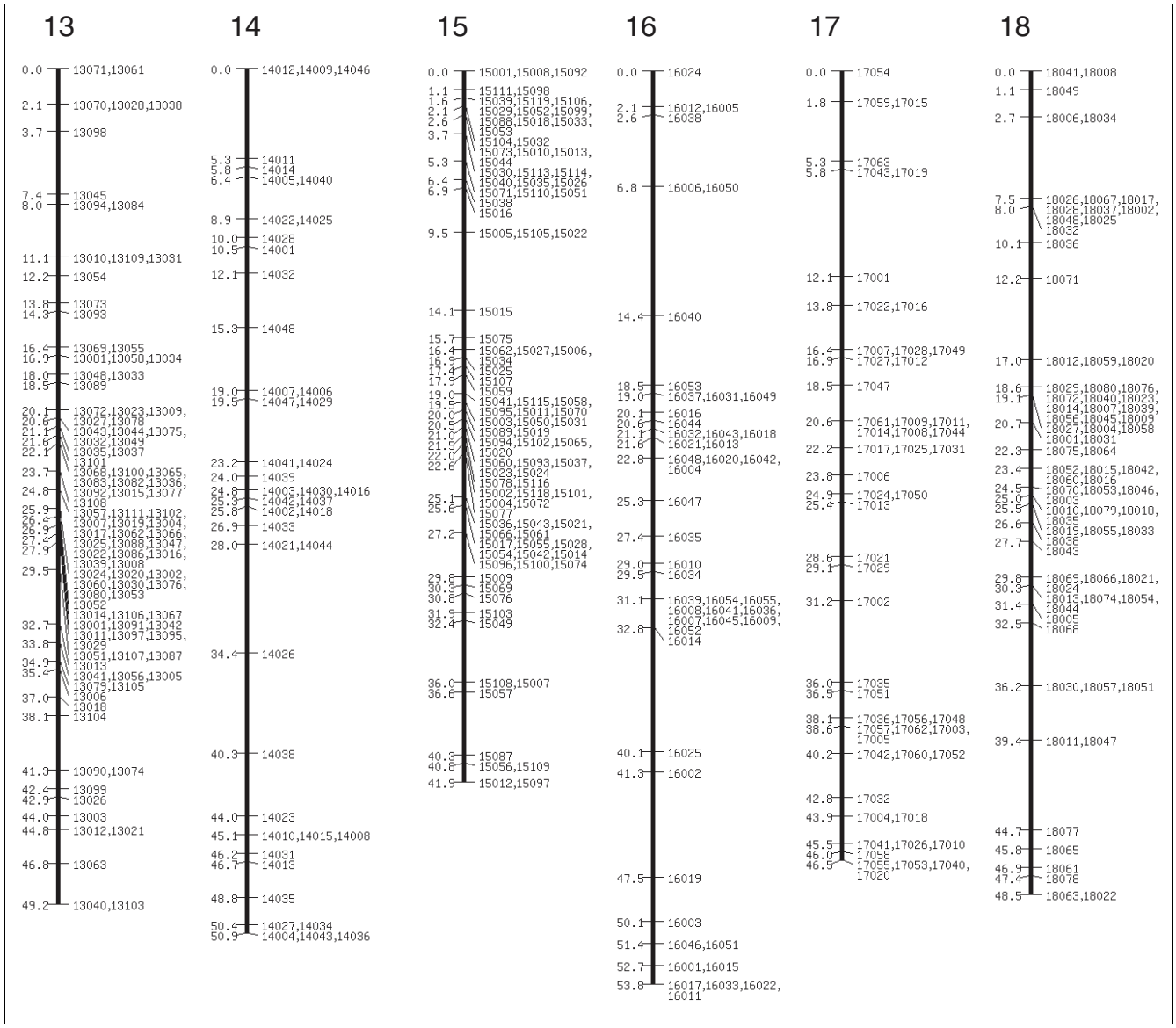

Figure 3
SNP linkage map continued: linkage groups 13 to 18. SNP, single nucleotide polymorphism.

additional contigs, which we denote as 'doubtful', were correctly formed. In most cases of doubtful contigs, only one BAC clone bridged two well defined clusters of overlapped BACs. Such a BAC clone is likely to be chimeric. Therefore, we removed such chimeric and doubtful BAC clones from the 15 BAC contigs involved, which no longer represented a contradiction in our evaluation.

The second approach was a comparison with SNP markers. If a BAC contig included more than two SNPs, then those markers should be positioned around the same locus. A total of 128 BAC contigs contained more than two SNP markers, among which we found contradictions in 11 BAC contigs. We checked the 11 contigs by PCR using primers designed from BAC end sequences, as described above. However, we were unable to resolve the contradictions because most of the BAC end sequences in question corresponded to repetitive transposon sequences, and PCR provided similar bands even if BACs in the false contig belonged to different chromosomes. This situation reinforced the supposition that BAC clones forming all or most of the false contigs were derived from transposon-rich domains that produced similar restriction band patterns.

Integration of the linkage map and BAC contigs

Altogether we mapped 581 BAC contigs containing 6,061 BACs onto 28 chromosomes through BAC clones containing SNP markers common to the linkage map. The length of mapped singletons and BAC contigs calculated for each chromosome is shown in Table 3. A total of 361.1 Mb are cov-

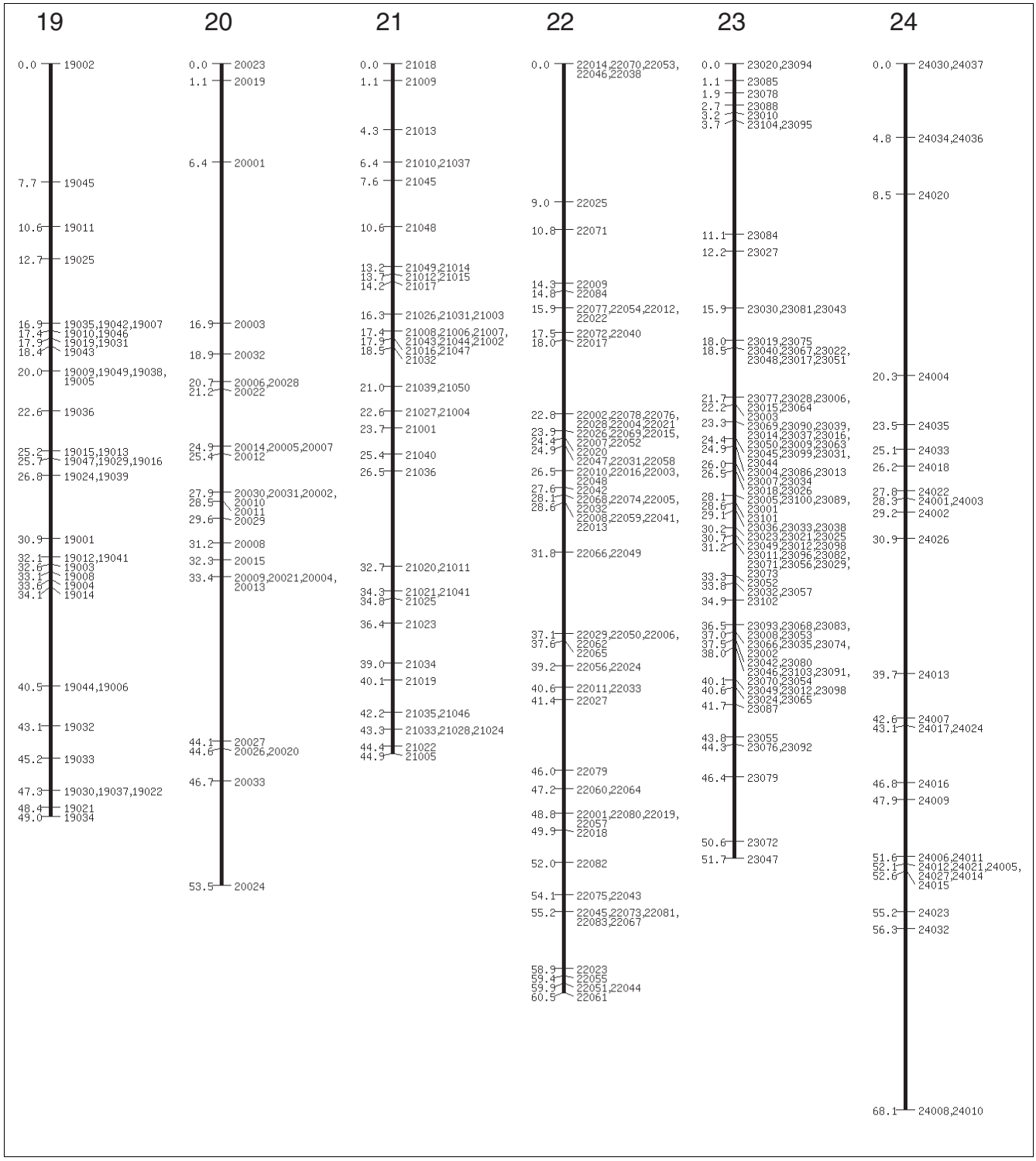

Figure 4
SNP linkage map continued: linkage groups 19 to 24. SNP, single nucleotide polymorphism.

ered by BACs, which corresponds to 76% coverage of the genome.

Mapping of EST markers onto SNP markers and BAC contigs

Screening of mapped BACs harboring functional genes by HDR filter hybridization using EST probes is a powerful tool

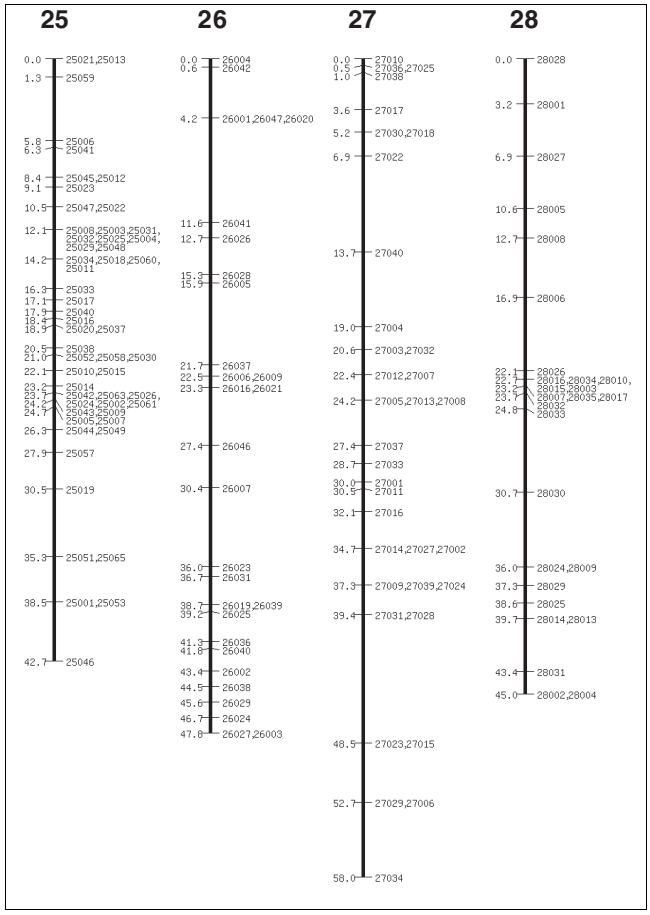

Figure 5
SNP linkage map continued: linkage groups 25 to 28. SNP, single nucleotide polymorphism.

for positional cloning. A large EST database is available in silkworm [22-24]. Therefore, in addition to using DNA fingerprinting for the construction of BAC contigs, we employed the 'overlapping' method with EST markers, whereby BAC clones arrayed on HDR filters are subjected to large-scale screening by hybridization with individual, nonredundant ESTs to identify clones carrying single copy sequences. This approach helped confirm BAC contigs and allowed us to identify functional genes on the combined physical-genetic map. For screening we used the RPCI-96 BAC library, consisting of 36,864 clones with an average insert size of 168 kb, which corresponds to 13× redundancy.

From the number of positive BAC clones screened by HDR filter hybridization, we could roughly estimate whether the probe cDNA was a single-copy or multiple-copy gene, or contained a repetitive sequence. Table 4 summarizes the results of the EST hybridization experiments. For EST mapping, we employed only putative single-copy genes, based on filter hybridization characteristics and the number of positive hits per filter. For a single-copy gene, approximately 13 BACs should give hybridization signals. Of 692 putative single-copy genes identified by this procedure, we were able to assign 523 EST markers to chromosomes by identifying BACs common to fingerprinted contigs that had been integrated with the genetic map via SNP markers (Additional data file 1). We identified an additional 353 ESTs on the mapped contigs by BLAST search of BAC end sequences. Of these 152 were duplicates, yielding a total of 724 mapped single-copy genes. For confirmation of the initial map assignments, we designed specific primer sets to amplify expected ESTs on HDR filter-screened BACs by PCR (Additional data file 1).

Linkage analysis of ESTs using RFLPs

In parallel with the HDR filter hybridization experiments, we carried out RFLP analysis of segregants from a backcross between an F_1 female (p50T × C108T) and a C108T male using a common set of ESTs as hybridization probes. The lack of meiotic crossing over in females results in complete linkage, enabling fast and efficient chromosome assignment of large numbers of markers using small segregant populations [13]. We assigned a total of 1,082 ESTs to linkage groups by RFLP analysis (Table 5 and Additional data file 3); of these, 118 ESTs were mapped in duplicate using HDR filter hybridization or BAC end SNPs. In addition to providing independent confirmation of linkage assignments on the integrated SNP-physical map, the linkage assignment of 964 new ESTs will enable future annotation and evaluation of mapped scaffolds in the WGS assembly now in progress [7,8] (Mita K, Xia Q, personal communication).

Synteny with other insects

Altogether we assigned 1,688 independent silkworm genes to 28 linkage groups (Table 5). Of these, there is positional

Table 1

Characteristics of BAC libraries used in this study

BAC library	Vector	Cloning site	Number of clones	Mean insert size (kb)	Clone coverage[a]
*Eco*RI-BAC	pBACe3.6	*Eco*RI	36,864	168	13×
*Bam*HI-BAC	pBeloBAC11	*Bam*HI	21,120	165	7.3×
*Hind*III-BAC	pBAC-lac	*Hind*III	23,040	125	6.1×

All libraries were constructed with strain p50T using mixed sexes. [a]To estimate the genome coverage of each library, we used a genome size of 476 megabases. BAC, bacterial artificial chromosome.

Table 2

Summary of fingerprinted BAC contigs

Number of fingerprinted BAC clones (FPC)	Number of singletons	Number of contigs	Number of major contigs with eight or more BACs	Total length major BAC contigs (Mb)
78,778	47,274	6,221	782	376

BAC, bacterial artificial chromosome; Mb, megabases.

information for 724 genes, whereas 964 ESTs are simply assigned to a chromosome. We tested these 1,688 genes for orthology between silkworm and other model holometabolous insects for which complete genome data were available, notably *A. mellifera, T. castaneum, A. gambiae*, and *Drosophila melanogaster*, in order to compare the syntenic relationships among them. We did not conduct an analysis of synteny relative to dipteran chromosomes because of their small number, which would produce many false connections, and their high rate of chromosome rearrangement [32-34], which would probably reduce the number of significant syntenic blocks within chromosome arms or segments. Among

Table 3

Summary of integrated SNP linkage maps and fingerprinted BAC contigs

Linkage group	Number of Markers	Recombination length (cM)	Number of mapped contigs	Sum of contig lengths (Mb)	Number of BACs in contigs	Total length BAC singletons and contigs (Mb)
1	53	44.9	8	2.2	32	9.5
2	20	47.4	8	2.1	60	4.1
3	57	46.4	24	8.3	304	12.8
4	105	50.6	32	11.2	431	20.8
5	94	58.6	31	9.2	258	18.1
6	84	45.5	20	6.5	246	15.4
7	62	45.9	20	5.9	177	12.7
8	60	53.8	15	4.6	131	11.6
9	68	47.8	22	6.6	211	14.0
10	91	43.8	32	10.0	302	18.4
11	98	63.9	27	8.5	259	19.9
12	68	53.3	33	11.6	405	16.7
13	104	49.0	31	10.5	325	21.9
14	44	50.8	19	6.2	224	9.7
15	99	42.2	32	10.3	359	20.1
16	49	53.9	19	5.8	182	10.4
17	54	46.5	22	6.8	208	11.4
18	77	48.0	20	6.7	273	15.5
19	40	49.2	15	4.6	144	8.4
20	29	53.5	13	4.4	192	6.6
21	46	44.9	17	5.2	163	9.7
22	77	60.2	24	8.6	309	16.5
23	99	51.7	33	9.7	290	20.0
24	32	68.0	10	2.8	81	6.4
25	54	42.8	20	5.8	141	11.1
26	29	47.6	10	2.9	104	5.7
27	35	58.2	14	4.8	158	8.3
28	27	45.0	10	2.9	92	5.4
Total	1,755	1,413.4	581	184.7	6,061	361.1

BAC, bacterial artificial chromosome; Mb, megabases; SNP, single nucleotide polymorphism.

Table 4

Summary of BAC HDR filter hybridizations with EST probes

Number of probes detected	Number of single copy genes	Number of 2-copy genes	> 3 copy genes	Repetitive sequences
1,960	692 (35.3%)	585 (29.8%)	211 (10.8%)	469 (23.9%)

We used two high-density replica (HDR) filters containing a total of 36,864 independent bacterial artificial chromosome (BAC) clones for hybridization experiments (13× genome coverage). We did not include the expressed sequence tag (EST) if one of the two filters gave poor results because of smearing or high background, or if the two filters gave imbalanced numbers of positive signals (for instance, the ratio of positive signals between two filters exceeded 5). Cut-off values used to define gene copy number were as follows: single copy gene, 2 to 15 positive signals; two copy gene, 16 to 29 positive signals; > 3 copy gene, > 30 positive signals; and repetitive sequence, > 50 copies. We mapped 523 single copy genes represented by these hybridized ESTs via single nucleotide polymorphism (SNP) markers.

1,688 silkworm genes, we found 769 orthologs for *A. mellifera* and 790 for *T. castaneum* that could be used for a test of synteny. We then checked the distribution of silkworm gene orthologs in honey bee and beetle chromosomes (Additional data file 3).

Table 5

Summary of ESTs assigned to linkage groups

Linkage group	Number of ESTs mapped by filter hybridization	Number of ESTs mapped by BAC end sequences	Number of ESTs mapped by RFLPs	Total number of independent ESTs[a]
1	17	7	21	36
2	15	8	25	37
3	16	10	26	43
4	25	18	37	73
5	37	18	50	87
6	16	8	46	65
7	12	20	34	57
8	16	16	50	71
9	19	13	45	65
10	29	15	47	77
11	38	26	54	95
12	22	9	42	64
13	18	16	42	67
14	6	6	20	30
15	46	32	65	113
16	24	11	30	54
17	17	10	56	75
18	15	14	39	62
19	12	8	40	59
20	13	6	38	50
21	12	12	36	52
22	21	14	59	86
23	27	19	48	78
24	10	7	31	39
25	19	13	34	55
26	8	9	15	28
27	11	7	32	44
28	2	5	20	26
Total	523	357	1,082	1,688

[a]Duplicates were removed from the total number of independent expressed sequence tags (ESTs). BAC, bacterial artificial chromosome; RFLP, restriction fragment length polymorphism.

(a) Apis

Silkworm

LG	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Tot
1	2	2	2	0	0	3	4	1	0	0	0	2	0	0	1	2	19
2	3	1	0	0	1	0	0	0	0	0	2	0	1	6	0	0	14
3	2	1	1	0	0	0	2	1	0	1	3	2	1	2	0	0	16
4	1	1	3	1	0	0	1	0	3	2	4	4	7	1	0	3	31
5	2	3	0	0	2	1	0	1	2	6	3	1	0	1	2	0	24
6	3	1	0	6	1	0	1	1	3	2	1	2	0	0	2	0	23
7	3	1	2	1	1	3	0	2	0	1	3	1	2	0	0	1	21
8	7	2	5	0	0	2	0	3	0	1	3	2	4	0	0	0	29
9	0	0	0	0	1	0	4	2	1	1	3	2	0	2	2	1	19
10	0	1	9	0	1	0	1	2	1	2	3	1	3	0	0	0	24
11	4	0	1	2	0	0	2	2	4	2	3	0	0	3	1	3	27
12	3	1	2	10	1	1	0	2	1	1	1	1	0	0	2	0	26
13	0	4	0	1	0	1	1	4	1	2	0	0	0	1	1	1	17
14	1	1	0	1	0	1	0	0	0	0	0	0	0	0	1	0	5
15	2	0	5	3	3	1	0	1	1	9	2	1	2	0	1	0	31
16	7	1	0	1	0	2	1	4	0	1	0	1	2	1	1	0	22
17	5	1	1	2	3	1	0	1	1	1	3	0	0	9	4	0	32
18	0	2	0	0	0	0	1	0	4	6	0	0	1	2	0	0	16
19	1	3	0	0	4	1	1	0	4	0	0	1	0	2	1	0	18
20	2	0	0	3	0	0	0	1	0	2	1	1	1	0	0	4	15
21	3	0	1	2	1	3	0	2	0	0	2	1	0	2	2	0	19
22	3	4	0	1	0	1	0	9	0	3	1	2	0	0	0	4	28
23	2	4	0	1	0	1	0	2	0	0	0	1	0	0	5	1	17
24	4	0	0	0	2	1	1	1	0	0	0	1	1	1	1	0	13
25	2	0	0	1	6	0	0	1	2	0	1	1	1	0	0	0	15
26	2	0	1	0	0	2	0	0	0	0	1	0	0	1	1	0	8
27	2	4	0	0	1	0	2	2	1	1	1	0	0	1	0	0	15
28	1	0	0	0	8	0	1	0	0	1	0	0	0	0	0	0	11
Tot	67	38	33	36	36	25	23	45	29	45	41	28	26	35	28	20	555

(b) Tribolium

Silkworm

LG	1	2	3	4	5	6	7	8	9	10	Tot
1	0	0	8	3	4	2	1	0	1	0	19
2	0	0	2	1	0	0	11	0	0	1	15
3	1	0	1	1	1	1	1	18	0	0	24
4	0	1	0	0	27	0	1	2	1	1	33
5	5	0	4	0	5	2	3	0	3	1	23
6	1	10	3	2	2	1	2	0	0	4	25
7	2	0	1	1	7	0	3	0	0	0	14
8	1	2	4	1	1	1	4	7	8	0	29
9	5	8	0	0	2	1	1	2	0	6	25
10	1	0	2	0	1	0	0	0	16	0	20
11	0	2	12	1	3	4	1	4	4	2	33
12	2	5	0	9	3	1	2	1	5	2	30
13	0	10	0	10	0	2	1	0	3	1	27
14	0	1	2	3	0	0	0	1	0	0	7
15	2	5	7	1	14	1	2	3	0	5	40
16	0	1	4	1	1	14	1	1	0	1	24
17	0	2	6	1	1	1	4	18	1	1	35
18	0	1	0	12	2	2	0	1	1	0	19
19	0	0	1	1	1	1	15	2	0	0	21
20	0	0	0	3	3	0	12	0	2	0	20
21	2	5	2	3	2	2	0	0	3	1	20
22	2	1	6	0	15	0	1	1	13	1	40
23	1	9	1	0	1	4	2	2	2	3	25
24	1	3	3	0	2	1	3	0	0	0	13
25	1	0	2	1	2	0	11	0	0	0	17
26	0	0	1	0	0	0	1	0	3	0	5
27	2	0	2	1	2	2	1	3	1	0	14
28	0	0	1	0	2	1	6	1	0	0	11
Tot	29	66	75	56	104	44	90	67	67	30	628

Figure 6

Oxford grids. Shown are Oxford grids displaying a matrix of cells comparing the number of orthologous genes on chromosomes of two species. **(a)** Silkworm-honey bee comparison. **(b)** Silkworm-*Tribolium* comparison. Shadowed cells show high synteny conservation (eight or more orthologs).

Figure 6 presents Oxford grids showing the number of shared orthologs mapped in honeybee (555 total) and beetle (628 total) on silkworm chromosomes. Based on the simplifying assumption that chromosomes are of equal length and shared orthologs are uniformly distributed throughout the genome, the ratios of observed versus expected silkworm orthologs per chromosome should be 1.0. The ratios of shared orthologs per chromosome fell in the range of 0.75 to 1.3 in the 16 honey bee chromosomes [35] and 0.8 to 1.2 in the ten beetle chromosomes [36], which is consistent with the possibility that the mapped silkworm genes were randomly sampled for both genomes.

A striking feature of this analysis is that a very poor syntenic relationship exists between silkworm and honey bee (Figure 6a) compared with that of silkworm and beetle (Figure 6b). Thus, only six chromosomes of the silkworm genome, namely linkage groups (LGs) 10, 12, 15, 17, 22 and 28 (shaded in Figure 6a), share high syntenic conservation with honey bee chromosomes (eight or more shared orthologs). In contrast, 20 silkworm chromosomes, namely LGs 1, 2, 3, 4, 6, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 22, 23 and 25 (shaded in Figure 6b), exhibit high correspondence of synteny conservation relative to beetle chromosomes. Consistent with these findings, honey bee chromosomes had fewer shared silkworm orthologs than did beetle (average 9 versus 12.7 orthologs per chromosome), and the fraction of shared orthologs in syntenic groups was much smaller (0.07 in honey bee versus 0.35 in beetle). For most of the orthologous genes, homology scores between silkworm and honey bee were similar to those between silkworm and beetle, indicating that the rate of evolutionary change was comparable, whereas those for Diptera were frequently lower (data not shown).

Discussion

The 3.3-fold increase in SNP markers from 534 in our previous study to 1,755 in the present one represents a significant improvement in the quality of the linkage map. This is reflected by a roughly proportional decrease in average marker spacing from 2.5 cM to 0.81 cM, or approximately 270 kb. Despite the increase in markers, the total recombination length only increased from 1,305 cM [18] to 1,413.4 cM, or 1.08-fold, suggesting that map lengths determined by our method are reaching asymptotes. This may reflect factors that would reduce detection of crossing over in a relatively small mapping population (190 individuals), such as a high frequency of double crossovers and the presence of gene-dense regions along chromosomes. The recombination map length obtained in this study is approximately half that obtained in

projects using random amplified polymorphic DNA (3,229 cM [19]) and simple sequence repeat (3,432 cM [16]) markers, which is not easily explained and awaits integration of markers from independent studies into a single linkage map. The total length of the genome covered by mapped BACs amounted to 361.1 Mb, corresponding to 76% genome coverage in the present integrated map, in which BACs and BAC contigs produced by fingerprinting were aligned on the SNP linkage map through common BAC markers (Table 3). This suggests that the present SNP markers cover most of the silkworm genome.

Two linkage groups, 11 (64 cM) and 24 (68 cM), are significantly longer than the others. One explanation for this is the high frequency of double crossovers in these two chromosomes. The double-crossover frequency of group 24 is especially high, at 7.9% (15/190) of all detected recombinants, with group 11 next, at 6.3% (12/190). By contrast, double crossover frequencies of the other linkage groups are in the range of 0% to 2%. Consistent with these observations is that these two chromosomes are observed cytologically to be longer than the others [37]. In addition, group 11 contains the nucleolus organizer region, and group 24 has strongly stained DAPI (4,6 diamidino-2-phenylindole, dihydrochloride) regions, which are speculated to be heterochromatin [37]. Such special chromosome structures might cause the observed high frequency of crossing over, yielding longer recombination lengths for the two linkage groups. Further investigation is required to clarify the role and mechanism of double crossovers in the two longest linkage groups, as well as to determine the strain and population distribution of recombination frequency variation. In potentially related observations, polygenic factors have been proposed to be responsible for differences in recombination frequency between morphological markers on linkage group 2 in artificially selected high and low recombination silkworm lines [38,39].

The difference in relative divergence times between honey bee, beetle, and silkworm based on phylogenomic analyses [40,41] provides a partial explanation for the difference in synteny that we observed, with longer evolutionary time enabling more chromosome rearrangements. Another potential contributing factor is the average genome-wide recombination rate, which is 2.97 cM/Mb (1,413 cM/476 Mb) for silkworm in our study. This is within the same order of magnitude reported for *T. castaneum* (2.87 cM/Mb), *D. melanogaster* (1.59 cM/Mb), and *A. gambiae* (0.84 cM/Mb), whereas a much higher rate, 19 cM/Mb, was reported for *A. mellifera* [42]. This, together with changes in chromosome number, could account for more extensive gene reorganization in honey bee than beetle relative to silkworm.

Of interest is that we found a larger number of silkworm orthologs on conserved linkage groups in both honey bee and beetle than an earlier study, which was based on the 6× WGS assembly (six and four orthologs listed in synteny in the *Tribolium* and *Apis* genomes, respectively [41]). A likely explanation is that the genome-wide scaffold reported here is much more extensive than the initial WGS genome assembly, presenting a larger chromosome-anchored dataset for comparison. Recent studies reveal extensive synteny among lepidopteran insects [19,43,44]. The extent of microsynteny, or conservation of close linkage among shared orthologs within chromosomes, remains an important question for understanding patterns of genome evolution among Lepidoptera and other holometaboous insects, which will await a more complete assembly and annotation of the silkworm genome.

Conclusion

The integrated genetic-physical map with 76% genome coverage by BACs provides a powerful basis for construction of minimum BAC tiling paths, which is an essential resource for positional cloning. The average distance between SNP markers of 270 kb, in combination with new sequencing data generated for a forthcoming re-assembly of the initial WGS data [7,8], will allow us to obtain super-scaffolds of megabase order in the near future. Our assignment of nearly 10% of predicted silkworm genes [7,8] to 28 chromosomes will not only facilitate construction of accurate scaffolds and annotation of the silkworm genome, but also provide a valuable resource for testing microsynteny and gene discovery in Lepidoptera and other insects.

Materials and methods

Silkworm strains and crosses

The inbred silkworm strains p50T and C108T, maintained at the University of Tokyo, were used as parental strains for the mapping panels. For linkage map construction, the same population of 190 segregants of a single-pair backcross (BC_1) between a p50T female and an F_1 male (p50T female × C108T male) was used as the first-generation SNP map [18].

Genomic DNA extraction

Genomic DNA of parental strains, F_1 individuals and segregants from the female-informative backcross was isolated from whole bodies of fifth instar larvae after removing midguts and hemolymph, as described in a previous report [14]. Genomic DNA of individual BC_1 segregants was isolated from whole pupae using DNAzol (Invitrogen Japan K.K., Tokyo, Jpn) after freezing in liquid nitrogen and homogenization with stainless steel beads.

Survey of the SNPs between p50T and C108T

For the linkage map construction, SNPs, including small base insertions and deletions, were used. The SNPs were identified in a large number of PCR amplicons that were synthesized using primers designed from the data obtained by BAC end sequencing, as reported previously [18]. Briefly, for each end sequence, we designed a PCR primer pair using Primer3 [45],

and performed PCR amplification of the genomic DNA of the parental (p50T and C108T) and F_1 strains with ExTaq (TaKaRa Bio Inc., Otsu, Shiga, Jpn), using the manufacturer's instructions. We detected the presence of SNPs in these amplicons by sequencing all three genotypes and analyzing the resulting traces using PolyPhred [46].

Linkage map construction

Linkage map construction was carried out as previously reported [18]. Genomic DNA of the same 190 BC_1 segregants was amplified with primer sets corresponding to the SNPs detected between parent strains, and polymorphism in the segregants was determined (Additional data file 1). Segregation patterns were analyzed using Mapmaker/exp (version 3.0 [25]) with the Kosambi mapping function [47]. For genotyping polymorphisms, we directly sequenced PCR amplicons from BAC end regions and used a fluorescent polarization dye terminator SNP detection assay [18,48] in parallel.

BAC libraries

Three BAC libraries prepared using different restriction enzymes were emplyed in the present study; their characteristics are summarized in Table 1. A BAC library was constructed from genomic DNA of strain p50T fifth instar day 3 posterior silk glands partially digested with *Eco*RI [21], designated RPCI-96 (RP96), and is available from BACPAC Resources of the Children's Hospital Oakland Research Institute [49]. The *Hind*III-BAC library was prepared using *Hind*III partial digestion of the same p50T strain DNA [20]. The *Bam*HI library, similarly prepared with *Bam*HI using DNA from a strain derived from the same line as p50T (Wu C, Goldsmith M, personal communication) was purchased from the Laboratory for Plant Genomics and GENEfinder Genomic Resource of Texas A&M University [50]. All three libraries were fingerprinted for the construction of genome-wide BAC contigs. The RPCI-96 and *Bam*HI-BAC libraries were used for SNP analysis; the *Eco*RI-BAC library was used for EST mapping by HDR filter hybridization.

BAC HDR filter hybridization with EST probes

RPCI-96-BAC clones arrayed in duplicate in a specific pattern onto nylon membranes (HDR filters) were obtained from BACPAC Resources of the Children's Hospital Oakland Research Institute [49]. Labeling, hybridization, and detection were performed using the ECL Direct Nucleic Acid Labeling and Detection System (GE Heathcare UK Ltd., Little Chalfont, Buckinghamshire, UK), in exact accordance with the manufacturer's instructions [21]. Probes derived from ESTs were obtained by PCR amplification of plasmid cDNA inserts using universal sequencing primers corresponding to the plasmid vectors (see below). EST library construction and analysis have been reported [22].

BLAST search

The BLASTN [51] search of 1,755 BAC end sequences was carried out against an in-house collection of *B. mori* cDNA/ESTs containing 185,765 sequences. In this search, the e value (a probability cut-off value) was set to $1 \times e^{-50}$ and no complexity filter was used. The cDNA/ESTs of *B. mori* were retrieved from NCBI-GenBank Flat Files (release 159.0) with a custom Perl script (Additional data file 1).

Fingerprinting analysis and BAC contig construction

DNA fingerprinting analysis was used to construct BAC contigs. The detailed protocol is reported in Marra and coworkers [27], and we followed the steps described by Osoegawa and coworkers [28] exactly. BAC DNA was isolated from 1.2 ml cultures of LB medium containing chloramphenicol using an automated method (PI-1100; Kurabo Industries Ltd, Osaka, Japan) and dissolved in 50 μl of TRIS-EDTA buffer. Approximately 6 ng (3 μl) of DNA was digested in a 20 μl reaction containing five units of *Eco*RI at 37°C overnight. When the digestion was completed, 4 μl of 6× loading buffer (15% Ficoll [Sigma-Aldrich Japan K. K., Tokyo, Jpn], 0.25% bromophenol blue, and 0.25% xylene cyanol) was added and 3 μl of *Eco*RI-digested DNA was separated on a 1% agarose gel in 1ξ Tris-Acetate-EDTA buffer at 90 V for 15 minutes, then at 45 V for 12 hours. Electrophoresis was carried out in Horizon 20-25 electrophoresis tanks (Life Technologies, Gaithersburg, MD, USA) at 16°C. Size markers containing 12.5 ng/μl Analytical Marker DNA, Wide Range Ladder (Promega K. K., Tokyo, Jpn), and 2 ng/μl Marker V (Roche Diagnostics K.K., Tokyo, Jpn) were loaded in every fifth lane. After electrophoresis, the gel was stained in 500 ml of 1:1,000 dilution of SYBR Green 1 (FMC BioProducts, Rockland, ME, USA) in 1× TAE for 30 minutes and scanned using a Molecular Imager FX (Bio-Rad Laboratories Inc., Tokyo, Jpn). The gel image was analyzed using Image 3.9d software [52,53], followed by FPC v.6.0 software [29,30].

RFLP linkage analysis

Southern blotting was used to assign EST clones to linkage groups. After sequencing and homology search [22], unique clones were amplified by PCR with plasmid-specific primers to encompass the inserts such as M13 reverse or T3/M13M3, M13M4, or T7 primer sets. Amplified EST DNA was labeled using the ECL Direct Nucleic Acid Labeling and Detection System (Amersham-Pharmacia Biotech). BC_1 segregants from a single female informative pair mating ([p50T × C108T] female × C108T male) were used for analysis of linkage, using 16 samples per probe. When the DNA of a sample was used up a new sample was taken from the same family and assayed with a reference set of anchor loci to determine the inheritance pattern for each linkage group [13]. To test polymorphism, digestion was carried out on 2.4 μg parental DNA with one of six restriction enzymes, *Bam*HI, *Bgl*II, *Eco*RI, *Hin*dIII, *Kpn*I, or *Sac*I, and subjected to 0.8 % agarose gel electrophoresis for 16 hours at 14°C. Gel treatment for denaturation, depurination, neutralization, blotting onto nylon membranes (nylon membranes, positively charged; Roche Diagnostics K. K., Tokyo, Jpn), and probe labeling and hybridization were performed in exact accordance with the

manufacturer's instructions. Selection of suitable enzymes to detect the RFLPs between the two parents and the analysis of linkage in the BC_1 segregants was carried out using the method of Kadono-Okuda and coworkers [13].

Abbreviations

BAC = bacterial artificial chromosome; BC_1 = first-generation backcross; BLAST = basic local alignment search tool; EST = expressed sequence tag; HDR = high-density replica; kb = kilobases; LG = linkage group; Mb = megabases; PCR = polymerase chain reaction; RFLP = restriction fragment length polymorphism; SNP = single nucleotide polymorphism; WGS = whole-genome shotgun.

Authors' contributions

KY designed and performed SNP analysis using BAC end sequences and construction of the SNP linkage map. JNo performed fingerprinting of whole BAC clones and construction of BAC contigs by FPC. KK designed RFLP linkage analysis and performed synteny analysis by Oxford grid and EST hybridization experiments. JNa participated in SNP analysis. MS participated in RFLP analysis and EST hybridization experiments. SS performed all DNA sequencing. HM contributed in construction of the integrated map. MS participated in integrated map construction and Oxford grid analysis. YB participated in the assignment of the SNP linkage groups to the classical linkage groups with visual markers. KO contributed in BAC library construction, characterization of BACs, and BAC contig construction with FPC. PJ also contributed in BAC library construction and provided the knowledge on BAC contig construction with FPC. MG provided the overall knowledge of linkage maps and participated in manuscript preparation. KM conducted the whole project, and participated in EST hybridization experiments and manuscript preparation. All authors read and approved the final manuscript.

Additional data files

The following additional data are available with the online version of this paper. Additional data file 1 is a table listing 1,755 SNP markers. Additional data file 2 is a table listing 6,221 BAC contigs constructed by fingerprinting 81,024 BACs from three BAC libraries made with different restriction enzymes. Additional data file 3 is a table listing 1,688 genes mapped onto 28 silkworm linkage groups by EST hybridization, RFLP, and BAC end sequence analyses, and their orthologs in *Apis* and *Tribolium*.

Acknowledgements

We thank Dr Toru Shimada for providing silkworm strains. We also thank Reiko Komatsuzaki, Yoko Fukusaki, Satsuki Tokoro, and Keiko Shiiba for technical assistance, and Alexie Papanicolaou for adding hyperlinks to ButterflyBase and NCBI in additional data files. This work was supported by grants from the Ministry of Agriculture, Forestry and Fisheries of Japan (Integrated research project for plant, insect and animal using genome technology), and from the Program for Promotion of Basic Research Activities for Innovative Biosciences (PROBRAIN).

References

1. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen LX, Brandon RC, Rogers YH, Blazej RG, Champe M, Pfeiffer BD, Wan KH, Doyle C, Baxter EG, Helt G, Nelson CR, *et al.*: **The genome sequence of *Drosophila melanogaster*.** *Science* 2000, **287:**2185-2195.
2. Richards S, Liu Y, Bettencourt BR, Hradecky P, Letovsky S, Nielsen R, Thornton K, Hubisz MJ, Chen R, Meisel RP, Couronne O, Hua S, Smith MA, Zhang P, Liu J, Bussemaker HJ, van Batenburg MF, Howells SL, Scherer SE, Sodergren E, Matthews BB, Crosby MA, Schroeder AJ, Ortiz-Barrientos D, Rives CM, Metzker ML, Muzny DM, Scott G, Steffen D, Wheeler DA, *et al.*: **Comparative genome sequencing of *Drosophila pseudoobscura* : Chromosomal, gene, and cis-element evolution.** *Genome Res* 2005, **15:**1-18.
3. Drosophila 12 Genomes Consortium: **Evolution of genes and genomes on the *Drosophila* phylogeny.** *Nature* 2007, **450:**203-218.
4. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, Wincker P, Clark AG, Ribeiro JM, Wides R, Salzberg SL, Loftus B, Yandell M, Majoros WH, Rusch DB, Lai Z, Kraft CL, Abril JF, Anthouard V, Arensburger P, Atkinson PW, Baden H, de Berardinis V, Baldwin D, Benes V, Biedler J, Blass C, Bolanos R, Boscus D, Barnstead M, *et al.*: **The genome sequence of the Malaria mosquito *Anopheles gambiae* .** *Science* 2002, **298:**129-149.
5. The Honeybee Genome Sequencing Consortium: **Insights into social insects from the genome of the honeybee *Apis mellifera*.** *Nature* 2006, **443:**931-949.
6. Tribolium Genome Sequencing Consortium: **The genome of the developmental model beetle and pest *Tribolium castaneum*.** *Nature* 2008 in press.
7. Mita K, Kasahara M, Sasaki S, Nagayasu Y, Yamada T, Kanamori H, Namiki N, Kitagawa M, Yamashita H, Yasukochi Y, Kadono-Okuda K, Yamamoto K, Ajimura M, Ravikumar G, Shimomura M, Nagamura Y, Shin-I T, Abe H, Shimada T, Morishita S, Sasaki T: **The genome sequence of silkworm, *Bombyx mori* .** *DNA Res* 3004, **11:**27-35.
8. Xia Q, Zhou Z, Lu C, Cheng D, Dai F, Li B, Zhao P, Zha X, Cheng T, Chai C, Pan G, Xu J, Liu C, Lin Y, Qian J, Hou Y, Wu Z, Li G, Pan M, Li C, Shen Y, Lan X, Yuan L, Li T, Xu H, Yang G, Wan Y, Zhu Y, Yu M, Shen W, *et al.*: **A Draft sequence for the genome of the domesticated silkworm (*Bombyx mori*).** *Science* 2004, **306:**1937-1940.
9. Fujii H, Banno Y, Doira H, Kihara H, Kawaguchi Y: **Genetical stocks and mutations of *Bombyx mori* : Important genetic resources.** Fukuoka, Jpn: Kyusyu Univ; 1998.
10. Promboon A, Shimada T, Fujisawa H, Kobayashi M: **Linkage map of random amplified polymorphic DNAs (RAPDs) in the silkworm.** *Bombyx mori* 1995, **66:**1-7.
11. Yasukochi Y: **A dense genetic map of the silkworm, *Bombyx mori*, covering all chromosomes based on 1018 molecular markers.** *Genetics* 1998, **150:**1513-1525.
12. Shi J, Heckel DG, Goldsmith MR: **A genetic linkage map for the domesticated silkworm, *Bombyx mori*, based on restriction fragment length polymorphisms.** *Genet Res* 1995, **66:**109-126.
13. Kadono-Okuda K, Kosegawa E, Mase K, Hara W: **Linkage analysis of maternal EST cDNA clones covering all twenty-eight chromosomes in the silkworm, *Bombyx mori* .** *Insect Mol Biol* 2002, **11:**443-451.
14. Nguu EK, Kadono-Okuda K, Mase K, Kosegawa E, Hara W: **Molecular linkage map for the silkworm, *Bombyx mori*, based on restriction fragment length polymorphism of cDNA clones.** *J Insect Biotechnol Sericol* 2005, **74:**5-13.
15. Tan YD, Wan C, Zhu Y, Lu C, Xiang Z, Deng HW: **An amplified fragment length polymorphism map of the silkworm.** *Genetics* 2001, **157:**1277-1284.
16. Miao XX, Xub SJ, Li MH, Li MW, Huang JH, Dai FY, Marino SW, Mills DR, Zeng P, Mita K, Jia SH, Zhang Y, Liu WB, Xiang H, Guo QH, Xu AY, Kong XY, Lin HX, Shi YZ, Lu G, Zhang X, Huang W, Yasukochi Y, Sugasaki T, Shimada T, Nagaraju J, Xiang ZH, Wang SY, Goldsmith

MR, Lu C, *et al.*: **Simple sequence repeat-based consensus linkage map of *Bombyx mori*.** *Proc Natl Acad Sci USA* 2006, **102:**16303-16308.
17. Prasad MD, Muthulakshmi M, Madhu M, Archak S, Mita K, Nagaraju J: **Survey and analysis of microsatellites in the silkworm, *Bombyx mori* : frequency, distribution, mutations, marker potential and their conservation in heterologous species.** *Genetics* 2005, **169:**197-214.
18. Yamamoto K, Narukawa J, Kadono-Okuda K, Nohata J, Sasanuma M, Suetsugu Y, Banno Y, Fujii H, Goldsmith MR, Mita K: **Construction of a single nucleotide polymorphism linkage map for the silkworm, *Bombyx mori*, based on bacterial artificial chromosome end sequences.** *Genetics* 2006, **173:**151-161.
19. Yasukochi Y, Ashakumary LA, Baba K, Yoshido A, Sahara K: **A second-generation integrated map of the silkworm reveals synteny and conserved gene order between lepidopteran insects.** *Genetics* 2006, **173:**1319-1328.
20. Wu C, Asakawa S, Shimizu N, Kawasaki S, Yasukochi Y: **Construction and characterization of bacterial artificial chromosome libraries from the silkworm, *Bombyx mori* .** *Mol Gen Genet* 1999, **261:**698-706.
21. Koike Y, Mita K, Suzuki MG, Maeda S, Abe H, Osoegawa K, deJong PJ, Shimada T: **Genomic sequence of a 320-kb segment of the Z chromosome of *Bombyx mori* containing a kettin ortholog.** *Mol Genet Genomics* 2003, **269:**137-149.
22. Mita K, Morimyo M, Okano K, Koike Y, Nohata J, Kawasaki H, Kadono-Okuda K, Yamamoto K, Suzuki MG, Shimada T, Goldsmith MR, Maeda S: **The construction of an EST database for *Bombyx mori* and its application.** *Proc Natl Acad Sci USA* 2003, **100:**14121-14126.
23. Cheng TC, Xia QY, Qian JF, Liu C, Lin Y, Zha XF, Xiang ZH: **Mining single nucleotide polymorphisms from EST data of silkworm, *Bombyx mori*, inbred strain Dazao.** *Insect Biochem Mol Biol* 2004, **34:**523-530.
24. **ButterflyBase v. 2.92** [http://butterflybase.ice.mpg.de/]
25. Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L: **MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations.** *Genomics* 1987, **1:**174-181.
26. **Silkworm Genome Research Program** [http://sgp.dna.affrc.go.jp/index.html]
27. Marra MA, Kucaba TA, Dietrich NL, Green ED, Brownstein B, Wilson RK, McDonald KM, Hillier LW, McPherson JD, Waterston RH: **High throughput fingerprint analysis of large-insert clones.** *Genome Res* 1997, **7:**1072-1084.
28. Osoegawa K, Tateno M, Woon PY, Frengen E, Mammoser AG, Catanese JJ, Hayashizaki Y, de Jong PJ: **Bacterial artificial chromosome libraries for mouse sequencing and functional analysis.** *Genome Res* 2000, **10:**116-128.
29. Gregory S, Soderlund C, Coulson A: **Contig assembly by fingerprinting.** In *Genome Mapping: A Practical Approach* Edited by: Dear PH. Oxfordm UK: Oxford University Press; 1997:227-254.
30. Soderlund C, Longdem I, Mott R: **FPC: a system for building contigs from restriction fingerprinted clones.** *Comput Appl Biosci* 1997, **13:**523-535.
31. Suetsugu Y, Minami H, Shimomura M, Sasanuma S, Narukawa J, Mita K, Yamamoto K: **End-sequencing and characterization of silkworm (*Bombyx mori*) bacterial artificial chromosome libraries.** *BMC Genomics* 2007, **8:**314.
32. Bolshakov VN, Topalis P, Blass C, Kokoza E, della Torre A, Kafatos FC, Louis C: **A comparative genomic analysis of two distant diptera, the fruit fly, *Drosophila melanogaster*, and the malaria mosquito, *Anopheles gambiae*.** *Genome Res* 2002, **12:**57-66.
33. Gonzalez J, Ranz JM, Ruiz A: **Chromosomal elements evolve at different rates in the *Drosophila* genome.** *Genetics* 2002, **161:**1137-1154.
34. Severson DW, DeBruyn B, Lovin DD, Brown SE, Knudson DL, Morlais I: **Comparative genome analysis of the yellow fever mosquito *Aedes aegypti* with *Drosophila melanogaster* and the malaria vector mosquito *Anopheles gambiae*.** *J Hered* 2004, **95:**103-113.
35. ***Apis mellifera* (honey bee) genome view** [http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=7460]
36. ***Tribolium castaneum* (red flour beetle) genome view** [http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=7070]
37. Yoshido A, Bando H, Yasukochi Y, Sahara K: **The *Bombyx mori* karyotype and the assignment of linkage groups.** *Genetics* 2005, **170:**675-685.
38. Turner JRG: **Genetic control of recombination in the silkworm. I. Multigenic control of chromosome 2.** *Heredity* 1979, **43:**273-293.
39. Ebinuma H, Yoshitake N: **The genetic system controlling recombination in the silkworm.** *Genetics* 1981, **99:**231-245.
40. Savard J, Tautz D, Richards S, Weinstock GM, Gibbs RA, Werren JH, Tettelin H, Lercher MJ: **Phylogenomic analysis reveals bees and wasps (Hymenoptera) at the base of the radiation of Holometabolous insects.** *Genome Res* 2006, **16:**1334-1338.
41. Zdobnov E, Bork P: **Quantification of insect genome divergence.** *Trends Genet* 2007, **23:**16-20.
42. Beye M, Gattermeier I, Hasselmann M, Gempe T, Schioett M, Baines JF, Schlipalius D, Mougel F, Emore C, Rueppell O, Sirviö A, Guzmán-Novoa E, Hunt G, Solignac M, Page RE Jr: **Exceptionally high levels of recombination across the honey bee genome.** *Genome Res* 2006, **16:**1339-1344.
43. Pringle EG, Baxter SW, Webster CL, Papanicolaou A, Lee SF, Jiggins CD: **Synteny and chromosome evolution in the lepidoptera: evidence from mapping in *Heliconius melpomene* .** *Genetics* 2007, **177:**417-426.
44. Sahara K, Yoshido A, Marec F, Fuková I, Zhang HB, Wu CC, Goldsmith MR, Yasukochi Y: **Conserved synteny of genes between chromosome 15 of *Bombyx mori* and a chromosome of *Manduca sexta* shown by five-color BAC-FISH.** *Genome* 2007, **50:**1061-1065.
45. Rozen S, Skaletsky H: **Primer3 on the WWW for general users and for biologist programmers.** *Methods Mol Biol* 2000, **132:**365-386.
46. Nickerson DA, Tobe VO, Taylor SL: **PolyPhred: automating the detection and genotyping of single nucleotide substitution using fluorescence-based resequencing.** *Nucleic Acids Res* 1997, **25:**2745-2751.
47. Kosambi DD: **The estimation of map distances from recombination values.** *Ann Eugen* 1944, **12:**172-175.
48. Nasu S, Suzuki J, Ohta R, Hasegawa K, Yui R, Kitazawa N, Monna L, Minobe Y: **Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (*Oryza sativa*, *Oriza rufipogon*) and establishment of SNP markers.** *DNA Res* 2002, **9:**163-171.
49. **BACPAC Resources Center** [http://bacpac.chori.org/]
50. **GENEfinder Genomic Resource** [http://hbz7.tamu.edu/]
51. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucl Acids Res* 1997, **25:**3389-3402.
52. Sulston J, Mallett F, Staden R, Durbin R, Horsnell T, Coulson A: **Software for genome mapping by fingerprinting techniques.** *Comput Appl Biosci* 1988, **4:**125-132.
53. Sulston J, Mallett F, Durbin R, Horsnell T: **Image analysis of restriction enzyme fingerprint autoradiograms.** *Comput Appl Biosci* 1989, **5:**101-106.