Copyright © 2006 by the Genetics Society of America
DOI: 10.1534/genetics.105.053801

Construction of a Single Nucleotide Polymorphism Linkage Map for the Silkworm, *Bombyx mori*, Based on Bacterial Artificial Chromosome End Sequences

Kimiko Yamamoto,*[,1] Junko Narukawa,* Keiko Kadono-Okuda,* Junko Nohata,* Motoe Sasanuma,* Yoshitaka Suetsugu,* Yutaka Banno,† Hiroshi Fujii,† Marian R. Goldsmith‡ and Kazuei Mita*

**Genome Research Department, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8634, Japan, †Laboratory of Insect Genetic Resources, Faculty of Agriculture, Kyushu University, Fukuoka 812-8581, Japan and ‡Biological Sciences Department, University of Rhode Island, Kingston, Rhode Island 02881-0816*

Manuscript received November 22, 2005
Accepted for publication March 6, 2006

ABSTRACT

We have developed a linkage map for the silkworm *Bombyx mori* based on single nucleotide polymorphisms (SNPs) between strains p50T and C108T initially found on regions corresponding to the end sequences of bacterial artificial chromosome (BAC) clones. Using 190 segregants from a backcross of a p50T female × an F_1 (p50T × C108T) male, we analyzed segregation patterns of 534 SNPs between p50T and C108T, detected among 3840 PCR amplicons, each associated with a p50T BAC end sequence. This enabled us to construct a linkage map composed of 534 SNP markers spanning 1305 cM in total length distributed over the expected 28 linkage groups. Of the 534 BACs whose ends harbored the SNPs used to construct the linkage map, 89 were associated with 107 different ESTs. Since each of the SNP markers is directly linked to a specific genomic BAC clone and to whole-genome sequence data, and some of them are also linked to EST data, the SNP linkage map will be a powerful tool for investigating silkworm genome properties, mutation mapping, and map-based cloning of genes of industrial and agricultural interest.

THE silkworm, *Bombyx mori*, is an agriculturally important insect that has been domesticated for an estimated 5000 years and used extensively for silk production. In addition, it is a key model of the Lepidoptera, the second most numerous group of holometabolous insects, which include many beneficials but also the most destructive agricultural pests. The strength of the silkworm as a model arises in part from its value as a genetic resource. Spontaneous mutations found while practicing sericulture were used to help establish principles of classical genetics, including sex linkage and the discovery of no crossing over in lepidopteran females (Tanaka 1913; Sturtevant 1915), maternal mutations (Toyama 1912), homeotic mutants and complex loci (Ichikawa 1943; Tsujita 1953), and the construction of early linkage maps (Fujii *et al.* 1998; Yasukochi *et al.* 2005). With the establishment of stable transformation (Yamao *et al.* 1999; Tamura *et al.* 2000), the silkworm has shown the potential to produce pharmaceutically important proteins in high yield (Tomita *et al.* 2003), opening up new applications for sericulture in medical, agricultural, and industrial fields.

Analysis of the silkworm genome began a few years ago because of its importance for breeding and genetic studies, for isolating valuable genes and promoters, and for comparative genomics (Goldsmith *et al.* 2005). Our group first initiated intensive sequencing of the silkworm genome using expressed sequence tags (ESTs; Mita *et al.* 2003). Recently, our group (Mita *et al.* 2004) and a second group (Xia *et al.* 2004) reported the results of whole-genome shotgun sequencing and provided public access to the assembled silkworm genome data (http://www.dna.affrc.go.jp/genome/; Wang *et al.* 2005; http://silkworm.genomics.org.cn/).

At present, several silkworm linkage maps based on molecular markers are available, including randomly amplified polymorphic DNA (RAPD) (Promboon *et al.* 1995; Yasukochi 1998), simple sequence repeat and RAPDs (Nagaraja *et al.* 2005), amplified fragment length polymorphisms (AFLP) (Tan *et al.* 2001), restriction fragment length polymorphisms (Shi *et al.* 1995; Kadono-Okuda *et al.* 2002; Nguu *et al.* 2005), and microsatellites (Prasad *et al.* 2005; Miao *et al.* 2005). Recently, karyotyping of the chromosomes using bacterial artificial chromosome (BAC)–FISH based on Yasukochi's RAPD map was also reported (Yoshido *et al.* 2005). However, no linkage map composed of single-nucleotide polymorphisms (SNPs) has been developed,

[1] *Corresponding author:* 1-2 Owashi, Tsukuba, Ibaraki 305-8634, Japan. E-mail: kiya@nias.affrc.go.jp

although a characterization of SNPs located in coding regions between different individuals and tissues has been reported (Cheng *et al.* 2004).

We report here construction of a linkage map of the silkworm genome comprising 534 SNPs. The SNPs were detected between two parental mapping strains, p50T and C108T, and their F_1 hybrid by PCR amplification and resequencing of the end-sequence regions of BAC clones that are part of an independent p50T BAC end-sequence project (K. Yamamoto, Y. Suetsugu, J. Narukawa, H. Minami, S. Sasanuma, M. Sasanuma, J. Nohata, K. Kadono-Okuda, M. Shimomura and K. Mita, unpublished results). We constructed the map by scoring discovered SNPs in 190 segregants from a male informative backcross. The SNP maps were assembled into 28 linkage groups, which were further assigned to the standard genetic linkage groups based on morphological markers. In addition, we colocalized 107 ESTs to the same BACs on which the SNPs were discovered. The future localization of SNPs and ESTs on a robust genetic and physical map based on BAC clones will provide a useful tool not only for investigation of genome properties of silkworm and other lepidopteran species, but also for mapping mutations of interest and cloning them by map-based procedures.

MATERIALS AND METHODS

Silkworm strains and crosses: The inbred silkworm strains p50T and C108T, maintained at the University of Tokyo, were used as parent strains for the mapping panel. For linkage map construction, 190 segregants of a single-pair backcross (BC_1) between a p50T female and an F_1 male (p50T female × C108T male) were used.

Genomic DNA extraction: Genomic DNA of parental strains and F_1 individuals was isolated from whole bodies of fifth instar larvae after removing midguts and hemolymph as described in a previous report (Nguu *et al.* 2005). Genomic DNA of individual BC_1 segregants was isolated from whole pupae using DNAzol (Invitrogen, San Diego) after freezing in liquid nitrogen and homogenization with stainless steel beads.

BAC libraries: Two BAC libraries constructed from strain p50T genomic DNA were used for end sequencing of individual clones. One, designated as RPCI-96, was constructed by Pieter de Jong's group (Children's Hospital Oakland Research Institute, Duarte, CA) with genomic DNA partially digested with *Eco*RI (Koike *et al.* 2003); the other was constructed with genomic DNA partially digested with *Bam*HI (C. Wu and H. Zhang, personal communication; GeneFinder Genomic Resources, Texas A&M University, College Station, TX).

Purification of BAC clones: *Escherichia coli* cells harboring single BAC clones were inoculated into individual wells of a 96-deep-well plate filled with 1.25 ml of 2× LB medium (2% tryptone peptone, 1% yeast extract, and 1% sodium chloride) containing 20 μg/ml of chloramphenicol and cultivated with shaking for 18–20 hr at 37°. The BAC DNA was prepared using a PI-1100 automatic DNA isolation system (KURABO) according to the manufacturer's instructions.

BAC end sequencing: Sequencing reactions were performed using a reaction mixture composed of 3 μl Big Dye terminator (Applied Biosystems, Foster City, CA), 1.0 μl of 5× sequencing buffer, 0.5–1.0 μg of template DNA, 10 pmol of primer, and 4 mM $MgCl_2$. The thermal cycling reactions were conducted under the following conditions: 96° for 5 min; 99 cycles of 96° for 30 sec, 55° for 10 sec, and 60° for 4 min, followed by 4° on hold. Custom-made T7 and SP6 sequencing primers (Table 1) were used. The DNA was recovered by multiscreen 384SEQ (Millipore, Bedford, MA). Sequence trimming was conducted by processing the traces using the base-calling software PHRED (Ewing and Green 1998; Ewing *et al.* 1998). Altogether, the sequences of 73,728 ends from the *Eco*RI-digested library and 42,240 ends from the *Bam*HI-digested library were determined (K. Yamamoto, Y. Suetsugu, J. Narukawa, H. Minami, S. Sasanuma, M. Sasanuma, J. Nohata, K. Kadono-Okuda, M. Shimomura and K. Mita, unpublished results).

Survey of the SNPs between p50T and C108T: For the linkage map construction, SNPs, including small base insertions and deletions (indels), were identified in a large number of PCR amplicons designed from the sequence data obtained from the BAC end sequencing. We searched for nonredundant sequences using the BLAST algorithm (Altschul *et al.* 1990) with an *E*-value of 1e-50 as a threshold, and then we randomly selected 3840 nonredundant BAC end sequences for the SNP survey. For each end sequence, we designed a PCR primer pair using Primer3 (Rozen and Skaletsky 2000) and performed PCR amplification of the genomic DNA of the parental (p50T and C108T) and F_1 strain with ExTaq (TaKaRa) using the manufacturer's instructions. We detected the presence of SNPs in these amplicons by sequencing the 3840 amplicons for all three genotypes and analyzing the resulting traces using PolyPhred (Nickerson *et al.* 1997).

Linkage map construction: For the detection of polymorphisms, we used both the direct sequencing of the PCR amplicons from BAC end regions and a fluorescent polarization dye terminator SNP-detection assay (Nasu *et al.* 2002) in parallel. Details of the primers used for the amplification of each marker are given in supplemental Table S1 at http://www.genetics.org/supplemental/. For each of the SNPs detected between parent strains, the BAC end regions of 190 BC_1 segregants were amplified and sequenced, and the polymorphisms of the segregants were determined. The SNPs among the PCR amplicons were also detected by using the Acyclo-Prime FP SNP detection kit (PerkinElmer Life Science) with a Wallac 1420 $ARVO_{MX}$ instrument (PerkinElmer Life Science). In this system, following enzymatic removal of excess primers and nucleotides, the first PCR amplicons are used as templates for a second PCR reaction based on primers designed to terminate immediately upstream of the polymorphic site to incorporate one of two fluorescent terminators, representing the allelic SNP nucleotides, into products. The SNPs are detected by differential fluorescence polarization of the two terminator dyes.

Segregation patterns were analyzed using Mapmaker/exp (version 3.0; Lander *et al.* 1987) with the Kosambi mapping function (Kosambi 1944). The "GROUP" command (Linkage Groups at min LOD 3.00, max Distance 37.2) was used to cluster all informative markers into linkage groups. The "COMPARE" commands were used to construct a framework (draft) of each linkage group, and additional marker positions were assigned by using the "TRY" command. Sequence and typing errors were detected with the "ERROR DETECTION" option.

Mapping of the *p* locus: p50T larvae have dark body pigments ($+^p$ phenotype), but C108T larvae are unmarked (*p* plain phenotype). Thus, BC_1 segregants with the same intense dark pigment as p50T were scored as homozygous ($+^p/+^p$) phenotype, and those with light-colored pigment, compared to p50T, were scored as heterozygous ($+^p/p$). The segregation pattern obtained was analyzed together with SNP markers by Mapmaker/exp as described above.

Localization of ESTs on BAC clones: A p50T *Eco*RI–BAC library consisting of 36,864 clones with an average insert size of 168 kb was arrayed in duplicate in specific patterns onto two nylon membranes to make BAC high-density replica (HDR) filters (copies of BAC library and HDR filters are available through BACPAC Resources at the Children's Hospital Oakland Research Institute; http://www.chori.org/bacpac). HDR filters were hybridized with PCR-amplified inserts of individual cDNA clones representing random, nonredundant EST sequences from a large-scale sequencing project (Mita *et al.* 2003; K. Mita, K. Kadono-Okuda, K. Yamamoto, M. Shimomura, Y. Nagamura, J. Nohata, S. Sasanuma, M. Sasanuma, M. R. Goldsmith and T. Shimada, unpublished results). Labeling, hybridization, and detection were performed using the ECL direct nucleic acid labeling and detection system kit (Amersham-Pharmacia Biotec) exactly according to the manufacturer's instructions (Koike *et al.* 2003).

RESULTS

Preliminary analysis of SNP frequency and characteristics: Since minimal information was available on the frequency and characteristics of SNPs in the silkworm genome (Cheng *et al.* 2004), we carried out a survey using the end-sequence data of clones from two silkworm genomic BAC libraries constructed from strain p50T with partial *Eco*RI-digested or *Bam*HI-digested genomic DNA in conjunction with a whole-genome physical mapping project (K. Yamamoto, Y. Suetsugu, J. Narukawa, H. Minami, S. Sasanuma, M. Sasanuma, J. Nohata, K. Kadono-Okuda, M. Shimomura and K. Mita, unpublished results). We eliminated redundant sequences from the complete data set of 115,968 BAC ends using the BLAST algorithm and randomly selected 3840 nonredundant sequences (3072 from the *Eco*RI- and 768 from the *Bam*HI-digested library) for the SNP survey. In an initial characterization, we randomly choose 95 BAC end sequences from this subset, designed specific primer pairs for the respective end-sequence regions, and amplified the regions from genomic DNA of p50T and C108T by PCR. Then the amplicons from the two strains were resequenced. As a result, we found SNPs in 31 BAC ends and detected 133 SNPs within the total sequence length of 54,586 bp. This indicated that the frequency of SNPs for our data set was one/410 bp, sufficient for linkage mapping. Among these SNPs, transitions (A/G and C/T substitutions) accounted for 49%, transversions accounted for 46% [21, 17, and 8% for A/T, A/C (G/T), and G/C substitutions, respectively], and small base indels (Hayashi *et al.* 2004) accounted for 5%. A/T substitutions were the most frequent, whereas G/C substitutions were the least frequent. The ratio of transitions to transversions was 1.08.

SNP survey: We designed primers for each of the 3840 BAC end sequences and used them to amplify genomic DNA from p50T and C108T, two standard strains that have been used routinely for large-scale molecular linkage map construction, as well as their F_1 hybrid. We then resequenced the amplicons and examined the traces for evidence of SNPs. The results of this survey, including small nucleotide indels, are summarized in Table 1. Among the 3840 sequences analyzed, 3005 gave high-enough-quality sequence data for further analysis. Among them, 2109 did not have SNPs and 115 had SNPs between the parents but the F_1 did not show a heterozygous sequence peak at the expected position. Consequently, we initially identified SNPs in 781 sequences, which could be used for a genetic analysis of polymorphism. However, we did not obtain a high-enough-sequence quality for 247 of these SNPs in the backcross mapping panel, giving a final yield of 534 SNPs for the linkage analysis.

TABLE 1

Sequences retained at each stage of the SNP survey

Stage	No. of sequences
Used for initial SNP survey	3840
Quality sufficient for mapping[a]	3005
Containing SNPs	781
Assigned to linkage groups	534

[a] Among the sequences, 2109 had no SNPs between parents, and 115 had SNPs but the F_1 did not have a heterozygote sequence peak.

Linkage map construction: To construct the linkage map, using the same PCR amplification and resequencing procedure in parallel with a fluorescent polarization dye terminator SNP-detection assay (Nasu *et al.* 2002) (AcycloPrime FP method), we surveyed the segregation patterns of the SNPs of 190 BC_1 individuals from a single pair mating between a p50T female and an F_1 male (p50T female × C108T male). On the basis of an analysis of the data using Mapmaker/exp (version 3.0; LOD score 3.0), we successfully positioned 534 SNPs on the linkage map. The results are summarized in Table 2. The SNP markers segregated into 28 linkage groups, with a total recombination length of 1305 cM.

Assignment of linkage groups: We assigned one of the SNP linkage groups to the Z chromosome using a sex-chromosome-specific property on the basis of the fact that the female is heterogametic (ZW) and the male is homogametic (ZZ). Consequently, a BC_1 female is predicted to have a sex chromosome pair of either Z-p50T/W-p50T or Z-C108T/W-p50T, whereas autosomes (A) will be A-p50T/A-p50T or A-C108T/A-p50T. Since Z and W chromosomes have virtually no correspondence, Z-linked markers are hemizygous in females. Therefore, we assigned the linkage group for which SNPs of the female population had either p50T or C108T-type SNP patterns to the Z chromosome, in contrast to the autosomes, which displayed homozygous p50T or heterozygous p50T/C108T phenotypes.

We assigned 25 of the autosomal SNP linkage groups to standard silkworm linkage groups 2–26 summarized

TABLE 2

Summary of linkage groups in SNP linkage map

Linkage group[a]	Morphological mutants[b]	No. of markers	Recombination length (cM)
1	—	23	45
2	*p*	11[c]	32
3	*Ze, lem*	13	35
4	*L*	26	51
5	*oc*	27	60
6	*E*	24	43
7	*q*	21	46
8	*st*	27	54
9	*Ia*	23	34
10	*w-2*	32	44
11	*K*	30	64
12	*C*	21	47
13	*ch*	26	47
14	*U*	8	52
15	*bl*	25	42
16	*cts*	14	51
17	*bts*	16	46
18	*mln*	21	48
19	*nb*	17	49
20	*oh*	8	27
21	*Lan*	19	45
22	*or*	21	59
23	*tub*	27	53
24	*Ym, sel*	10	48
25	*oy*	19	39
26	*so*	7	48
A	—	7	42
B	—	12	54
Total		535	1305

[a] A total of 534 SNP markers and one morphological marker segregated into 28 linkage groups. The numbering of the linkage groups corresponds to those of the standard silkworm linkage map (Fujii *et al.* 1998).

[b] The linkage group numbers except groups 1, A, and B were determined by using 27 morphological mutants specific to previously assigned groups (Fujii *et al.* 1998).

[c] The number of markers includes the *p* locus.

in Fujii *et al.* (1998) and Banno *et al.* (2005). For the assignment, we used one or two morphological mutations specific to each of linkage groups 2–26 (27 mutations total). The mutations used for each group are listed in Table 2.

Marker strains with morphological mutations specific to individual linkage groups of the standard maps were crossed with a normal (wild-type) strain, and the F_1 female was backcrossed to either the recessive homozygous mutant male or to a normal (wild-type) male as required to score segregants. The procedure relies on lack of crossing over in females to give complete linkage for markers on a given chromosome (Yasukochi *et al.* 2005). We selected a SNP marker from each of the linkage groups of our map, examined whether or not the markers cosegregated with the morphological mutations, and determined the SNPs corresponding to the mutations. Twenty-three backcrosses were carried out to marker stocks, and the cosegregations were scored in 6–22 individuals for each cross. By the analysis, we obtained 27 SNPs corresponding to 27 morphological mutations and assigned the linkage groups of our map to the standard silkworm linkage groups 2–26. We assigned the remaining two linkage groups provisionally to groups A and B, which do not have well-defined morphological markers.

The SNP linkage map is illustrated in Figure 1. The number of markers per linkage group varies from 7 (groups 26 and A) to 32 (group 10), and the recombination length for each linkage group ranges from 27 cM (group 20) to 64 cM (group 11). The average distance between the markers is 2.5 cM. The markers are not evenly distributed throughout the linkage map, and so they may be dense or sparse, depending on the region. For example, there are 14 gaps with lengths exceeding 10 cM, including large gaps on the proximal end of linkage group B (20.2 cM) and on the distal part of linkage group 14 (21.3 cM). Alternatively, several regions with relatively high marker density also exist (*e.g.*, the middle portion of linkage group 23, from T013C09 to T603H10, with an average distance of 0.78 cM). Details of each marker, including BAC accession number, are described in supplemental Table S1 at http://www.genetics.org/supplemental/.

Mapping of the *p* locus on the linkage map: Alleles for one morphological marker, the *p* locus, which affects larval body pigmentation, segregated in the initial SNP mapping panel. p50T carries a semidominant allele that is darkly pigmented compared to C108T, which is unpigmented or "plain." We were able to discriminate homozygotes and heterozygotes by scoring pigment intensity in fifth instar larvae. We mapped the *p* locus to a position 10.3 cM from the proximal end of the SNP linkage group; on the basis of the presence of this marker we assigned it to standard linkage group 2.

Mapping of EST markers onto the linkage map: In parallel, as part of an independent project, we have been proceeding with the construction of BAC contigs with the "overlapping" method, whereby BAC clones are arrayed on HDR filters and subjected to large-scale screening by hybridization with individual, nonredundant cDNAs representing ESTs (Mita *et al.* 2003). Contigs are constructed by the presence of one or more common ESTs on different BAC clones. So far, we have carried out ∼6000 hybridizations (data not shown). Upon inspection, we found that a number of BAC clones containing nonredundant ESTs corresponded to ones mapped by the SNP analysis reported here. We designed specific primer sets to amplify expected ESTs on selected BACs by PCR and, finally, confirmed the presence of 107 ESTs on 89 BACs of the mapped SNP-containing BACs (Table 3). Forty-nine of them (55%) were found to have deduced amino acid sequences with significant homology to known proteins.

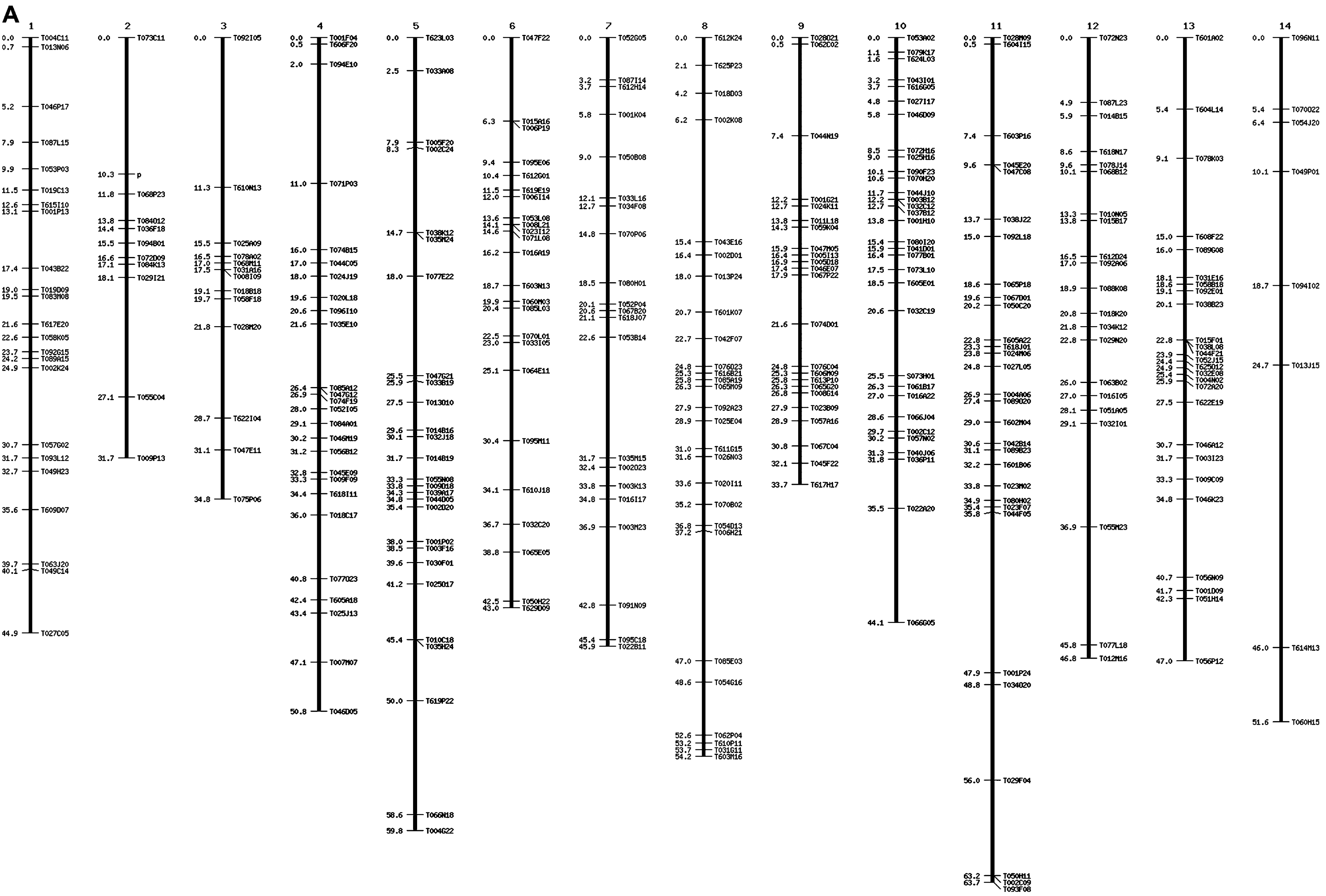

Figure 1.—*B. mori* SNP linkage map based on 534 SNP markers segregated into 28 linkage groups, represented by vertical lines. The BAC clones corresponding to the mapped SNP markers are shown at the right of the vertical lines; the recombination distances between the markers are indicated at the left. The *p* locus was mapped on linkage group 2 at the position of 10.3 cM from the proximal end.

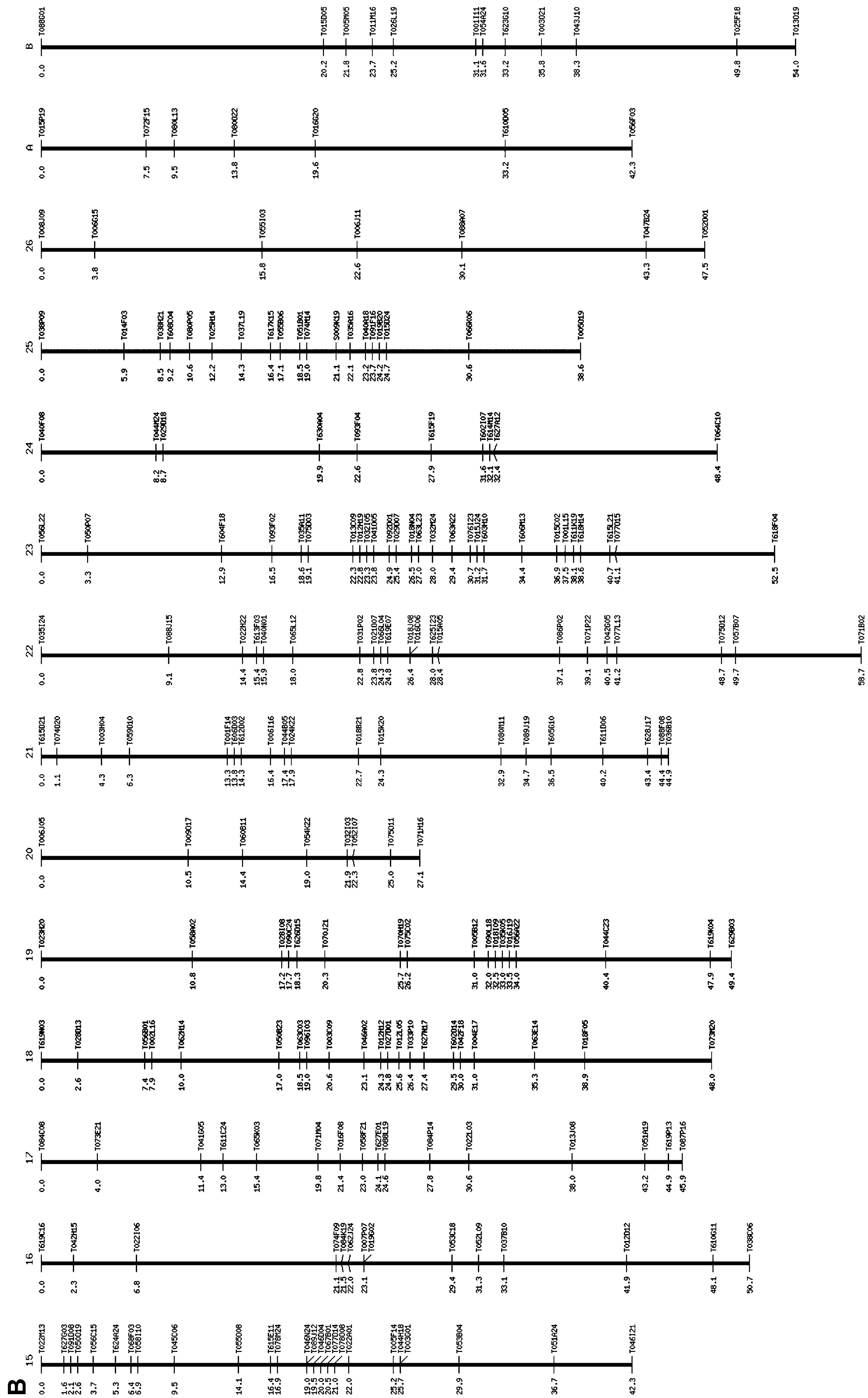

FIGURE 1.—*Continued.*

TABLE 3

Mapped ESTs

Linkage group	Position (cM)	BAC ID	EST ID[a]	EST accession no.	Description[b]	Score (bits)	*E*-value
1	19.5	083M08	wv40151	AU005825	ref|XP_396476.1| similar to ENSANGP00000009569 (*Apis mellifera*)	190	2e-47
	23.7	092G15	wdS30172	AU005039	ref|XP_315683.1| ENSANGP00000021837 (*Anopheles gambiae*)	139	6e-32
2	17.1	084K13	heS30665	AV402881	ref|XP_308839.1| ENSANGP00000021738 (*A. gambiae*)	96	1e-26
3	17.5	031A16	heS30235	AV402629			
	17.5	031A16	prgv0160	AV404557			
	17.5	008I09	wdS00806	AU003984			
4	18.0	024J19	wdS30342	AU005170			
	26.4	085A12	wdS30999	AU005669			
	26.4	047G12	heS30087	AV402533			
	26.4	047G12	n0006	AU002482	ref|XP_320836.1| ENSANGP00000017965 (*A. gambiae*)	188	4e-47
	26.4	047G12	heS00144	AV401774	ref|XP_319975.1| ENSANGP00000016783 (*A. gambiae*)	106	5e-22
	26.9	074F19	heS30428	AV402746	gb|AAM50732.1| GM29503p (*Drosophila melanogaster*)	164	6e-40
	47.1	007M07	NV021016	AV398042			
5	14.7	038K12	e96h0575	AV401340	ref|XP_307965.1| ENSANGP00000013477 (*A. gambiae*)	231	1e-59
	14.7	035M24	heS30184	AV402596			
	27.5	013O10	n0671	AU002914	ref|XP_309541.1| ENSANGP00000012655 (*A. gambiae*)	171	6e-42
	29.6	014B16	NV060186	AV398700	ref|XP_397415.1| similar to CG3204-PA (*A. mellifera*)	73	3e-12
	29.6	014B16	fbpv0469	BP125169			
	33.8	009D18	wdS00230	AU003516			
	38.0	001P02	n0242	AU002614			
	39.6	030F01	e96h0213	AV401089	ref|NP_001002720.1| zgc:86649 (*Danio rerio*)	65	6e-16
	45.4	010C18	heS30303	AV402666			
	45.4	035H24	e96h0691	AV401418	ref|XP_321293.1| ENSANGP00000018447 (*A. gambiae*)	84	2e-15
	45.4	035H24	wdS00334	AU003610	ref|XP_394153.1| similar to RIKEN cDNA 4921516M08 (*A. mellifera*)	189	4e-47
6	6.3	015A16	msgV0729	AV403591			
	6.3	006P19	heS00306	AV401906	gb|AAL76013.1| putative carboxylesterase (*Aedes aegypti*)	207	1e-52
	14.1	008L21	n0666	AU002911	ref|XP_311717.1| ENSANGP00000014281 (*A. gambiae*)	94	4e-18
	14.1	023I12	wdS00704	AU003886	gb|AAH73002.1| MGC82583 protein (*Xenopus laevis*)	117	3e-25
	14.1	023I12	NV060169	AV398683			
7	20.6	067B20	heS30988	AV403051			
8	31.6	026N03	NV060214	AV398725			
	33.6	020I11	wdS00070	AU003368	gb|AAD38624.1| BcDNA.GH08388 (*D. melanogaster*)	178	1e-43
	48.6	054G16	heS30160	AV402579	ref|XP_394434.1| similar to CG1782-PA (*A. mellifera*)	169	4e-41
9	13.8	011L18	wdS00098	AU003395			
10	8.5	072H16	heS00167	AV401792			
	9.0	025H16	wdS20879	AU004814			
	9.0	025H16	wdS00081	AU003379	ref|XP_393498.1| similar to CG2182-PA (*A. mellifera*)	201	1e-50
	12.2	003B12	wv40101	AU005784			
	12.2	032C12	wdS30748	AU005485	ref|XP_315790.1| ENSANGP00000018745 (*A. gambiae*)	149	7e-35
	29.7	002C12	heS30901	AV403016			
	30.2	057N02	wv41052	AU006456			

(*continued*)

TABLE 3

(Continued)

<table>
<tr><th>Linkage group</th><th>Position (cM)</th><th>BAC ID</th><th>EST ID[a]</th><th>EST accession no.</th><th>Description[b]</th><th>Score (bits)</th><th>E-value</th></tr>
<tr><td>11</td><td>0.0</td><td>028M09</td><td>heS30831</td><td>AV402984</td><td>ref|XP_392005.1| similar to ENSANGP00000019468 (A. mellifera)</td><td>207</td><td>2e-52</td></tr>
<tr><td></td><td>9.6</td><td>047C08</td><td>msgV0004</td><td>AV403097</td><td>dbj|BAD00700.1| sericin 1B (B. mori)</td><td>75</td><td>4e-13</td></tr>
<tr><td></td><td>13.7</td><td>038J22</td><td>wdS20766</td><td>AU004713</td><td>ref|XP_308558.1| ENSANGP00000016021 (A. gambiae)</td><td>367</td><td>e-100</td></tr>
<tr><td></td><td>13.7</td><td>038J22</td><td>NV021864X</td><td>AV398486</td><td>gb|AAH47966.1| Ruvbl2-prov protein (X. laevis)</td><td>226</td><td>2e-58</td></tr>
<tr><td></td><td>19.6</td><td>067D01</td><td>heS30263</td><td>AV402645</td><td>ref|XP_310759.1| ENSANGP00000023871 (A. gambiae)</td><td>131</td><td>8e-31</td></tr>
<tr><td></td><td>24.8</td><td>027L05</td><td>wdS00999</td><td>AU004165</td><td>gb|AAL90438.1| SD10213p (D. melanogaster)</td><td>265</td><td>8e-70</td></tr>
<tr><td></td><td>24.8</td><td>027L05</td><td>msgV0666</td><td>AV403540</td><td></td><td></td><td></td></tr>
<tr><td></td><td>26.9</td><td>004A06</td><td>wdS30801</td><td>AU005526</td><td></td><td></td><td></td></tr>
<tr><td></td><td>26.9</td><td>004A06</td><td>prgv0990</td><td>AV405259</td><td>ref|XP_308810.1| ENSANGP00000009906 (A. gambiae)</td><td>150</td><td>2e-35</td></tr>
<tr><td></td><td>30.6</td><td>042B14</td><td>wv41029</td><td>AU006435</td><td></td><td></td><td></td></tr>
<tr><td></td><td>35.8</td><td>044F05</td><td>msgV0402</td><td>AV403384</td><td></td><td></td><td></td></tr>
<tr><td></td><td>48.8</td><td>034O20</td><td>heS00552</td><td>AV402088</td><td>ref|NP_608909.1| CG8680-PA (D. melanogaster)</td><td>139</td><td>3e-32</td></tr>
<tr><td></td><td>63.2</td><td>002C09</td><td>e40h269</td><td>AU000228</td><td>ref|XP_393536.1| similar to WD repeat protein Bub3 (A. mellifera)</td><td>187</td><td>1e-46</td></tr>
<tr><td>12</td><td>27.0</td><td>016I05</td><td>wdS30491</td><td>AU005281</td><td></td><td></td><td></td></tr>
<tr><td></td><td>36.9</td><td>055M23</td><td>wdS30367</td><td>AU005188</td><td>gb|AAH02240.1| 2810410M20Rik protein (Mus musculus)</td><td>100</td><td>3e-20</td></tr>
<tr><td>13</td><td>18.1</td><td>031E16</td><td>e96h0224</td><td>AV401096</td><td></td><td></td><td></td></tr>
<tr><td></td><td>18.1</td><td>031E16</td><td>fbpv0854</td><td>BP125451</td><td></td><td></td><td></td></tr>
<tr><td></td><td>22.8</td><td>015F01</td><td>wv40022</td><td>AU005726</td><td></td><td></td><td></td></tr>
<tr><td></td><td>22.8</td><td>038L08</td><td>prgv0635</td><td>AV404953</td><td></td><td></td><td></td></tr>
<tr><td></td><td>22.8</td><td>044F21</td><td>fbpv0831</td><td>BP125432</td><td>ref|NP_648779.1| CG16979-PA (D. melanogaster)</td><td>117</td><td>2e-25</td></tr>
<tr><td></td><td>41.7</td><td>001D09</td><td>e40h431</td><td>AU000345</td><td></td><td></td><td></td></tr>
<tr><td>14</td><td>6.4</td><td>054J20</td><td>wv40057</td><td>AU005752</td><td></td><td></td><td></td></tr>
<tr><td>15</td><td>0.0</td><td>022M13</td><td>wdS00110</td><td>AU003407</td><td></td><td></td><td></td></tr>
<tr><td></td><td>0.0</td><td>022M13</td><td>wdS30866</td><td>AU005569</td><td>ref|XP_397214.1| similar to hypothetical protein FLJ11171 (A. mellifera)</td><td>241</td><td>1e-62</td></tr>
<tr><td></td><td>3.7</td><td>056C15</td><td>wdS30669</td><td>AU005430</td><td>ref|XP_309182.1| ENSANGP00000018770 (A. gambiae)</td><td>169</td><td>2e-41</td></tr>
<tr><td></td><td>9.5</td><td>045C06</td><td>wdS20853</td><td>AU004791</td><td></td><td></td><td></td></tr>
<tr><td></td><td>19.0</td><td>046N24</td><td>wdS30871</td><td>AU005573</td><td></td><td></td><td></td></tr>
<tr><td></td><td>19.0</td><td>046N24</td><td>fbpv0538</td><td>BP125209</td><td></td><td></td><td></td></tr>
<tr><td></td><td>19.0</td><td>089J12</td><td>wdS30521</td><td>AU005309</td><td></td><td></td><td></td></tr>
<tr><td></td><td>19.5</td><td>046D04</td><td>wdS00064</td><td>AU003362</td><td></td><td></td><td></td></tr>
<tr><td></td><td>22.0</td><td>022A01</td><td>prgv0415</td><td>AV404765</td><td></td><td></td><td></td></tr>
<tr><td></td><td>25.7</td><td>044H18</td><td>n1053</td><td>AU003230</td><td>emb|CAD35493.1| acidic ribosomal protein P1 (B. mori)</td><td>140</td><td>9e-33</td></tr>
<tr><td></td><td>25.7</td><td>003G01</td><td>n0090</td><td>AU002534</td><td>gb|AAO38522.1| pumilio RBD (Schistocerca americana)</td><td>233</td><td>8e-61</td></tr>
<tr><td>16</td><td>2.3</td><td>042H15</td><td>e40h872</td><td>AU000684</td><td>gb|AAL39863.1| LP02196p (D. melanogaster)</td><td>182</td><td>5e-45</td></tr>
<tr><td></td><td>23.1</td><td>007P07</td><td>wdS30025</td><td>AU004944</td><td></td><td></td><td></td></tr>
<tr><td></td><td>23.1</td><td>019G02</td><td>wdS00972</td><td>AU004141</td><td></td><td></td><td></td></tr>
<tr><td></td><td>23.1</td><td>019G02</td><td>wv40171</td><td>AU005842</td><td></td><td></td><td></td></tr>
<tr><td></td><td>33.1</td><td>037B10</td><td>msgV0270</td><td>AV403279</td><td>ref|XP_214554.2| similar to asparaginyl-tRNA synthetase, cytoplasmic (Rattus norvegicus)</td><td>171</td><td>8e-42</td></tr>
</table>

(*continued*)

TABLE 3

(Continued)

<table>
<tr><th>Linkage group</th><th>Position (cM)</th><th>BAC ID</th><th>EST ID[a]</th><th>EST accession no.</th><th>Description[b]</th><th>Score (bits)</th><th>E-value</th></tr>
<tr><td>17</td><td>11.4</td><td>041G05</td><td>e40h320</td><td>AU000267</td><td>ref|XP_393352.1| similar to ENSANGP00000010223 (A. mellifera)</td><td>177</td><td>9e-44</td></tr>
<tr><td></td><td>15.4</td><td>065K03</td><td>wv40116</td><td>AU005796</td><td>gb|AAL26577.1| ribosomal protein L29 (Spodoptera frugiperda)</td><td>72</td><td>2e-11</td></tr>
<tr><td></td><td>30.6</td><td>022L03</td><td>fbpv0625</td><td>BP125270</td><td></td><td></td><td></td></tr>
<tr><td></td><td>38.0</td><td>013J08</td><td>wdS30444</td><td>AU005242</td><td></td><td></td><td></td></tr>
<tr><td></td><td>38.0</td><td>013J08</td><td>wdS30864</td><td>AU005567</td><td></td><td></td><td></td></tr>
<tr><td>18</td><td>23.1</td><td>046A02</td><td>wv41050</td><td>AU006454</td><td></td><td></td><td></td></tr>
<tr><td></td><td>23.1</td><td>046A02</td><td>e96h0001</td><td>AV400935</td><td></td><td></td><td></td></tr>
<tr><td></td><td>31.0</td><td>004E17</td><td>wdS30171</td><td>AU005038</td><td>ref|XP_318676.1| ENSANGP00000022240 (A. gambiae)</td><td>114</td><td>2e-24</td></tr>
<tr><td></td><td>38.9</td><td>018F05</td><td>wdS00027</td><td>AU003326</td><td></td><td></td><td></td></tr>
<tr><td></td><td>38.9</td><td>018F05</td><td>e40h958</td><td>AU000761</td><td>emb|CAG09542.1| unnamed protein product (Tetraodon nigroviridis)</td><td>114</td><td>2e-24</td></tr>
<tr><td>19</td><td>32.5</td><td>018I09</td><td>NV060108</td><td>AV398630</td><td>ref|NP_724756.1| CG8068-PE (D. melanogaster)</td><td>286</td><td>3e-76</td></tr>
<tr><td></td><td>33.0</td><td>035K05</td><td>e40h854</td><td>AU000668</td><td>ref|XP_315645.1| ENSANGP00000021747 (A. gambiae)</td><td>278</td><td>3e-74</td></tr>
<tr><td>20</td><td>10.5</td><td>009O17</td><td>prgv0351</td><td>AV404708</td><td></td><td></td><td></td></tr>
<tr><td></td><td>10.5</td><td>009O17</td><td>wdS00362</td><td>AU003633</td><td>ref|XP_317592.1| ENSANGP00000010195 (A. gambiae)</td><td>228</td><td>8e-59</td></tr>
<tr><td></td><td>21.9</td><td>032I03</td><td>e40h739</td><td>AU000586</td><td></td><td></td><td></td></tr>
<tr><td>21</td><td>6.3</td><td>059O10</td><td>heS30277</td><td>AV402652</td><td></td><td></td><td></td></tr>
<tr><td>22</td><td>26.4</td><td>018J08</td><td>wdS00826</td><td>AU004002</td><td>ref|XP_316177.1| ENSANGP00000019739 (A. gambiae)</td><td>104</td><td>5e-22</td></tr>
<tr><td></td><td>26.4</td><td>016C06</td><td>wdS00218</td><td>AU003507</td><td></td><td></td><td></td></tr>
<tr><td></td><td>40.5</td><td>042G05</td><td>NV021119X</td><td>AV398091</td><td>ref|XP_320303.1| ENSANGP00000016464 (A. gambiae)</td><td>99</td><td>7e-20</td></tr>
<tr><td>23</td><td>26.5</td><td>018N04</td><td>prgv0356</td><td>AV404712</td><td>ref|XP_311443.1| ENSANGP00000018481 (A. gambiae)</td><td>139</td><td>6e-32</td></tr>
<tr><td></td><td>28.0</td><td>032M24</td><td>prgv0316</td><td>AV404674</td><td></td><td></td><td></td></tr>
<tr><td></td><td>31.2</td><td>015J24</td><td>wdS00876</td><td>AU004049</td><td>ref|XP_394735.1| similar to ENSANGP00000013413 (A. mellifera)</td><td>115</td><td>3e-26</td></tr>
<tr><td></td><td>36.9</td><td>015C02</td><td>wdS30898</td><td>AU005593</td><td></td><td></td><td></td></tr>
<tr><td>25</td><td>8.5</td><td>038H21</td><td>mg0587</td><td>AU002179</td><td></td><td></td><td></td></tr>
<tr><td></td><td>22.1</td><td>035A16</td><td>e96h0349</td><td>AV401191</td><td></td><td></td><td></td></tr>
<tr><td></td><td>23.7</td><td>091F16</td><td>wdS30954</td><td>AU005638</td><td></td><td></td><td></td></tr>
<tr><td>A</td><td>19.6</td><td>016G20</td><td>NV021794X</td><td>AV398424</td><td>ref|XP_319486.1| ENSANGP00000020215 (A. gambiae)</td><td>190</td><td>1e-47</td></tr>
<tr><td>B</td><td>23.7</td><td>011M16</td><td>wdS00408</td><td>AU003662</td><td>ref|XP_393101.1| similar to ENSANGP00000001418 (A. mellifera)</td><td>86</td><td>6e-16</td></tr>
</table>

[a] For EST IDs, refer to Mita *et al.* (2003).

[b] Obtained by deduced amino acid sequence homology search against public protein database NCBI–NR using BLASTX 2.2.10. *E*-value of 1e-10 was used as a threshold.

DISCUSSION

Recently, the importance of insects, especially the silkworm, as genetic and bio-material resources has increased. In Japan, the genetic data on the silkworm have accumulated throughout the long history of sericultural study. The silkworm is also important as a model insect for Lepidoptera, which include the most highly destructive agricultural pests. Due to these industrial and agricultural interests, the genome analysis of the silkworm is urgently required. In this study, we have carried out the first linkage analysis of this insect based on SNP information within the BAC end-sequence regions. We also assigned the SNP linkage groups to those of previous linkage maps using morphological mutants.

The silkworm genome is expected to contain a high frequency of repetitive sequences, such as transposable elements (Mita *et al.* 2004; Xia *et al.* 2004). We made an effort to eliminate as many SNPs as possible that originated from repetitive sequences by selecting nonredundant sequences from the initial BAC end-sequence population using the BLAST algorithm. Of these, we then randomly selected 3840 sequences for the SNP survey. Finally, we used the SNPs detected in this subset for detailed characterization and map construction.

The SNP density of silkworm BAC end sequences was nearly twofold higher than that reported previously by Cheng *et al.* (2004) (1/775 bp *vs.* 1/410 bp). This seems reasonable, since we characterized SNPs between two different strains, p50T and C108T, whereas Cheng *et al.* (2004) detected those between individuals or tissues of the same strain. Furthermore, the previous analysis was performed on SNPs located in coding regions (EST sequences). In contrast, we used BAC end sequences, which contained both coding and noncoding regions. Although we sequenced in total >45 Mbp of BAC ends, the annotation of the genome is still incomplete at present and so it was difficult to estimate the extent of coding regions in our data set. Noncoding, intronic, and intergenic regions are expected to have a higher SNP frequency than regions coding for amino acids, and thus the noncoding portion of the end sequences is most likely to have contributed to the observed increase in SNP density.

We also mapped the *p* locus onto standard linkage group 2 at a position with good agreement with previous linkage maps (Banno *et al.* 2005). The nearest marker to the *p* locus in our SNP map is T068P23, which was 1.6 cM away. If SNP markers (*i.e.*, BAC clone markers) are mapped closer to the *p* locus, it will be a great help in isolating a gene that has many allelic variants affecting larval pigment patterns and thus of considerable interest for understanding the molecular basis of a key lepidopteran trait.

The recombination lengths for our linkage groups varied from 27 to 64 cM, which were shorter than those of previously reported linkage maps. For example, the Z chromosomes of the RAPD and AFLP maps constructed by Yasukochi (1998) and Tan *et al.* (2001) have recombination lengths of ~94 and 417.8 cM, respectively. On the other hand, the Z chromosome of our SNP map was 45 cM, which was around one-ninth to one-half the size of the others. The maps were composed of similar numbers of markers, indicating that marker density was an unlikely explanation for the length differences. It is possible that our map did not reach the ends of the chromosome. To test for this, we detected the position of the markers mapped at the proximal and distal ends of the Z chromosome map, T004C11 and T027C05, by fluorescent *in situ* hybridization using the corresponding BACs. We found these markers did in fact map to positions at the chromosome ends (S. Kuwazaki, J. Narukawa, K. Mita and K. Yamamoto, unpublished data). In addition, whereas markers were unevenly distributed over the chromosomes in all three maps, the marker distribution patterns were dissimilar, suggesting that they did not reflect true variations in physical distance. Although the reasons for these observations are not clear at present, an increase of marker density may provide information needed to understand these phenomena.

The SNP markers of this map were based on BAC end sequences and thus are directly linked to BAC clones. Consequently, integrated analyses among the genetic linkage maps, physical maps based on BAC contigs, EST database, and whole-genome sequence contigs will become possible through the BAC end-sequence data. As a first step of this synthesis, we performed the mapping of ESTs onto our linkage map using BAC HDR filters. As a result, we associated 107 EST markers with 89 of our 534 mapped BACs containing SNPs. When a SNP linkage map with a higher density of BAC clone markers is realized in the near future, much more effective integrated analyses will become possible for the investigation of silkworm genome properties. Furthermore, the SNP map will provide a reference to enable integration of previously reported silkworm linkage maps with the physical map, since many of the linkage maps also used the same pair of silkworm strains, p50T and C108T. Finally, SNP linkage analysis will be a powerful tool for gene isolation by using map-based cloning methods.

We thank Toru Shimada for providing silkworm strains; Reiko Komatsuzaki, Yoko Fukusaki, Satsuki Tokoro, and Keiko Shiiba for technical assistance; and Shun-ichi Sasanuma for sequencing. We are also grateful to Hiroshi Minami and Michihiko Shimomura for data analysis and to two anonymous reviewers and the editor for helpful comments. This work was supported by funds from the Ministry of Agriculture, Forestry and Fisheries of Japan, and the Bio-oriented Technology Research Advancement Institution.

LITERATURE CITED

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman, 1990 Basic local alignment search tool. J. Mol. Biol. **215:** 403–410.

Banno, Y., H. Fujii, Y. Kawaguchi, K. Yamamoto, K. Nishikawa *et al.*, 2005 A guide to the silkworm mutants: 2005 gene name and gene symbol. Kyusyu University, Fukuoka, Japan, p. 29.

Cheng, T. C., Q. Y. Xia, J. F. Qian, C. Liu, Y. Lin *et al.*, 2004 Mining single nucleotide polymorphisms from EST data of silkworm, *Bombyx mori*, inbred strain Dazao. Insect Biochem. Mol. Biol. **34:** 523–530.

Ewing, B., and P. Green, 1998 Basecalling of automated sequencer trace using phred. II. Error probabilities. Genome Res. **8:** 186–194.

Ewing, B., L. Hillier, M. Wendi and P. Green, 1998 Basecalling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. **8:** 175–185.

Fujii, H., Y. Banno, H. Doira, H. Kihara and Y. Kawaguchi, 1998 Genetical stocks and mutations of *Bombyx mori*: important genetic resouces. Kyusyu University, Fukuoka, Japan, p. 54.

Goldsmith, M. R., T. Shimada and H. Abe, 2005 The genetics and genomics of the silkworm, *Bombyx mori*. Annu. Rev. Entomol. **50:** 71–100.

Hayashi, K., N. Hashimoto, M. Daigen and I. Ashikawa, 2004 Development of PCR-based SNP markers for rice blast resistance genes at the Piz locus. Theor. Appl. Genet. **108:** 1212–1220.

Ichikawa, N., 1943 Genetical and embryological studies of a dominant mutant, "new additional crescent," of the silkworm, *Bombyx mori L.* Jpn. J. Genet. **19:** 182–188.

Jakubowski, J., and K. Kornfeld, 1999 A local, high density, single-nucleotide polymorphism map, used to clone *Caenorhabditis elegans* cdf-1. Genetics **153:** 743–752.

Kadono-Okuda, K., E. Kosegawa, K. Mase and W. Hara, 2002 Linkage analysis of maternal EST cDNA clones covering all twenty-eight chromosomes in the silkworm, *Bombyx mori*. Insect Mol. Biol. **11:** 443–451.

Koike, Y., K. Mita, M. G. Suzuki, S. Maeda, H. Abe *et al.*, 2003 Genomic sequence of a 320-kb segment of the Z chromosome of *Bombyx mori* containing a kettin ortholog. Mol. Genet. Genomics **269:** 137–149.

Kosambi, D. D., 1944 The estimation of map distances from recombination values. Ann. Eugen. **12:** 172–175.

Lander, E. S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly *et al.*, 1987 MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:** 174–181.

Miao, X.-X., S.-J. Xu, M.-H. Li, M.-W. Li, J.-H. Huang *et al.*, 2005 Simple sequence repeat-based consensus linkage map of *Bombyx mori*. Proc. Natl. Acad. Soc. USA **102:** 16303–16308.

Mita, K., M. Morimyo, K. Okano, Y. Koike, J. Nohata *et al.*, 2003 The construction of an EST database for *Bombyx mori* and its application. Proc. Natl. Acad. Soc. USA **100:** 14121–14126.

Mita, K., M. Kasahara, S. Sasaki, Y. Nagayasu, T. Yamada *et al.*, 2004 The genome sequence of silkworm, *Bombyx mori*. DNA Res. **11:** 27–35.

Nagaraja, G. M., G. Mahesh, V. Satish, M. Madhu, M. Muthulakshmi *et al.*, 2005 Genetic mapping of Z chromosome and identification of W chromosome-specific markers in the silkworm, *Bombyx mori*. Heredity **95:** 148–157.

Nasu, S., J. Suzuki, R. Ohta, K. Hasegawa, R. Yui *et al.*, 2002 Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (*Oryza sativa*, *Oriza rufipogon*) and establishment of SNP markers. DNA Res. **9:** 163–171.

Nguu, E. K., K. Kadono-Okuda, K. Mase, E. Kosegawa and W. Hara, 2005 Molecular linkage map for the silkworm, *Bombyx mori*, based on restriction fragment length polymorphism of cDNA clones. J. Insect Biotechnol. Sericology **74:** 5–13.

Nickerson, D. A., V. O. Tobe and S. L. Taylor, 1997 PolyPhred: automating the detection and genotyping of single nucleotide substitution using fluorescence-based resequencing. Nucleic Acids Res. **25:** 2745–2751.

Prasad, M. D., M. Muthulakshmi, M. Madhu, S. Archak, K. Mita *et al.*, 2005 Survey and analysis of microsatellites in the silkworm, *Bombyx mori*: frequency, distribution, mutations, marker potential and their conservation in heterologous species. Genetics **169:** 197–214.

Promboon, A., T. Shimada, H. Fujisawa and M. Kobayashi, 1995 Linkage map of random amplified polymorphic DNAs (RAPDs) in the silkworm. *Bombyx mori*. Genet. Res. **66:** 1–7.

Rozen, S., and H. Skaletsky, 2000 Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. **132:** 365–386.

Shi, J., D. G. Heckel and M. R. Goldsmith, 1995 A genetic linkage map for the domesticated silkworm, *Bombyx mori*, based on restriction fragment length polymorphisms. Genet. Res **66:** 109–126.

Sturtevant, A. H., 1915 No crossing over in the female of the silkworm moth. Am. Nat. **49:** 42–44.

Tamura, T., C. Thibert, C. Royer, T. Kanda, E. Abraham *et al.*, 2000 Germline transformation of the silkworm, *Bombyx mori* L. using a *piggyBac* transposon-derived vector. Nat. Biotechnol. **18:** 81–84.

Tan, Y. D., C. Wan, Y. Zhu, C. Lu, Z. Xiang *et al.*, 2001 An amplified fragment length polymorphism map of the silkworm. Genetics **157:** 1277–1284.

Tanaka, Y., 1913 A study of Mendelian factors in the silkworm, *Bombyx mori*. J. Coll. Agric. Tohoku Imp. Univ. **5**(4): 91–113.

Tomita, M., H. Munetsuna, T. Sato, T. Adachi, R. Hino *et al.*, 2003 Transgenic silkworms produce recombinant human type III procollagen in cocoons. Nat. Biotechnol. **21:** 52–56.

Toyama, K., 1912 Maternal inheritance and Mendelism. J. Genet. **2:** 352–405.

Tsujita, M., 1953 Studies on the semi-allelic E-series in the silkworm. Annu. Rep. Natl. Inst. Genet. Misima **3:** 20–24.

Wang, J., Q. Xia, X. He, M. Dai, J. Ruan *et al.*, 2005 SilkDB: a knowledge base for silkworm biology and genomics. Nucleic Acids Res. **33:** D399–D402.

Xia, Q., Z. Zhou, C. Lu, D. Cheng, F. Dai *et al.*, 2004 A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). Science **306:** 937–940.

Yamao, M., N. Katayama, H. Nakazawa, M. Yamakawa, Y. Hayashi *et al.*, 1999 Gene targeting in the silkworm by use of a baculovirus. Genes Dev. **13:** 511–516.

Yasukochi, Y., 1998 A dense genetic map of the silkworm, *Bombyx mori*, covering all chromosomes based on 1018 molecular markers. Genetics **150:** 1513–1525.

Yasukochi, Y., Y. Banno, K. Yamamoto, M. R. Goldsmith and H. Fujii, 2005 Integration of molecular and classical linkage groups of the silkworm, *Bombyx mori* ($n = 28$). Genome **48:** 626–629.

Yoshido, A., H. Bando, Y. Yasukochi and K. Sahara, 2005 The *Bombyx mori* karyotype and the assignment of linkage groups. Genetics **170:** 675–685.

Communicating editor: A. D. Long