2005

Simple sequence repeat-based consensus linkage map of *Bombyx mori*

Xue-Xia Miao

Shi-Je Xu

Ming-Hui Li

Mu-Wang Li

Jian-Hua Huang

See next page for additional authors

Citation/Publisher Attribution

Miao, X-X., Xu, S-J., Li, M-H., Li, M-H., Huang, J-H., Dai, F-Y., Marino, S. W.,...Huang, Y-P. (2005). Simple sequence repeat-based consensus linkage map of *Bombyx mori*. *PNAS, 102*(45), 16303-16308. doi: 10.1073/pnas.0507794102
Available at: https://doi.org/10.1073/pnas.0507794102

Simple sequence repeat-based consensus linkage map of *Bombyx mori*

Authors

Xue-Xia Miao, Shi-Je Xu, Ming-Hui Li, Mu-Wang Li, Jian-Hua Huang, Fang-Yin Dai, Susan W. Marino, David R. Mills, Peiyu Zeng, Kazuei Mita, Shi-Hai Jia, Yong Zhang, Wen-Bin Liu, Hui Xiang, Qiu-Hong Guo, An-Ying Xu, Xiang-Yin Kong, Hong-Xuan Lin, Yao-Zhou Shi, Gang Lu, Xianglin Zhang, Wei Huang, Yuji Yasukochi, Toshiyuki Sugasaki, Toru Shimada, Javaregowda Nagaraju, Zhong-Huai Xiang, Sheng-Yue Wang, Marian R. Goldsmith, Cheng Lu, Guo-Ping Zhao, and Yong-Ping Huang

Simple sequence repeat-based consensus linkage map of *Bombyx mori*

Xue-Xia Miao[a,b], Shi-Jie Xu[b,c], Ming-Hui Li[a,b], Mu-Wang Li[a,b,d], Jian-Hua Huang[a,b], Fang-Yin Dai[e], Susan W. Marino[f], David R. Mills[f], Peiyu Zeng[f], Kazuei Mita[g], Shi-Hai Jia[a], Yong Zhang[a], Wen-Bin Liu[a], Hui Xiang[a], Qiu-Hong Guo[a], An-Ying Xu[d], Xiang-Yin Kong[a], Hong-Xuan Lin[a], Yao-Zhou Shi[h], Gang Lu[c], Xianglin Zhang[c], Wei Huang[c], Yuji Yasukochi[g], Toshiyuki Sugasaki[i], Toru Shimada[i], Javaregowda Nagaraju[j], Zhong-Huai Xiang[e], Sheng-Yue Wang[c,k], Marian R. Goldsmith[f,k], Cheng Lu[e,k], Guo-Ping Zhao[a,c,k,l], and Yong-Ping Huang[a,k]

[a]Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 300 Fenglin Road, Shanghai 200032, China; [c]Chinese National Human Genome Centre, 250 Bi Bo Road, Zhang Jiang High Tech Park, Shanghai 201203, China; [d]Sericultural Research Institute, Chinese Academy of Agriculture Sciences, Zhengjiang 212018, China; [e]Southwest University, Baibei, Chongqing 400716, China; [f]Department of Biological Sciences, University of Rhode Island, Kingston, RI 02881; [g]Laboratory of Insect Genome, National Institute of Agrobiological Sciences, Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan; [h]National Engineering Center for Biochip, 151, Libing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, China; [i]Department of Agricultural and Environmental Biology, University of Tokyo, 113-0033 Tokyo, Japan; [j]Laboratory of Molecular Genetics, Centre for DNA Fingerprinting and Diagnostics, ECIL Road, Nacharam, Hyderabad 500-076, India; [l]State Key Laboratory of Genetic Engineering/Department of Microbiology and Microbial Engineering, School of Life Science, Fudan University, Shanghai 200433, China

Communicated by Jiazhen Tan, Fudan University, Shanghai, China, September 7, 2005 (received for review January 2, 2005)

We established a genetic linkage map employing 518 simple sequence repeat (SSR, or microsatellite) markers for *Bombyx mori* (silkworm), the economically and culturally important lepidopteran insect, as part of an international genomics program. A survey of six representative silkworm strains using 2,500 (CA)n- and (CT)n-based SSR markers revealed 17–24% polymorphism, indicating a high degree of homozygosity resulting from a long history of inbreeding. Twenty-nine SSR linkage groups were established in well characterized Dazao and C108 strains based on genotyping of 189 backcross progeny derived from an F_1 male mated with a C108 female. The clustering was further focused to 28 groups by genotyping 22 backcross progeny derived from an F_1 female mated with a C108 male. This set of SSR linkage groups was further assigned to the 28 chromosomes (established linkage groups) of silkworm aided by visible mutations and cleaved amplified polymorphic sequence markers developed from previously mapped genes, cDNA sequences, and cloned random amplified polymorphic DNAs. By integrating a visible mutation *p* (plain, larval marking) and 29 well conserved genes of insects onto this SSR-based linkage map, a second generation consensus silkworm genetic map with a range of 7–40 markers per linkage group and a total map length of ≈3431.9 cM was constructed and its high efficiency for genotyping and potential application for synteny studies of *Lepidoptera* and other insects was demonstrated.

silkworm | microsatellite

The mulberry silkworm, *Bombyx mori*, has been domesticated for silk production for ≈5,000 years (1). Currently, it is the major economic resource for >30 million families in countries such as China, India, Vietnam, and Thailand. With the development of biotechnology, *B. mori* has been used as an important bioreactor for the production of recombinant proteins (2, 3). In addition, *B. mori* is the model organism for Lepidoptera, the second most numerous order of insects, including many species important for agriculture and forestry. Advances in silkworm research not only have a great impact in improving sericulture, but also may facilitate the development of new strategies for pest control.

The economic and scientific significance of silkworms have made them the subject of intensive genetic studies since the last century, and thus, the most important insect genetic model after *Drosophila melanogaster*. More than 400 mutations have been identified, and >1,000 silkworm strains are maintained as genetic resources (4–6). These mutations affect many fundamental aspects of the insect life cycle, including egg and egg shell formation, early embryonic pattern formation, development and diapause, larval feeding behavior, and molting (7). Genetic linkage analysis of silkworm has evolved since the first identification of visible mutations, and now employs molecular markers. The earliest report of linkage can be traced back to the beginning of the 20th century (8, 9). Up to now, >200 visible and biochemical mutations have been placed onto a chromosomal linkage map with 28 established linkage groups (ELGs, equivalent to chromosomes; refs. 5 and 6) covering 900.2 cM. In 1995, molecular linkage maps were constructed employing random amplified polymorphic DNA (RAPD) and restriction fragment length polymorphic (RFLP) markers (10, 11). After a high-density RAPD map with a marker interval of ≈2 cM (≈500 kb) was reported (12), an amplified fragment length polymorphism (AFLP) genetic linkage map was constructed with a logarithm of odds score of 3.0 (13), and more recently a RFLP map with nearly 200 ESTs (14). Although RAPDs, AFLPs, and RFLPs are convenient tools for generating polymorphic markers, in general, the polymorphic level of RAPD markers is much lower than that of AFLP markers, and both RAPD and AFLP markers capture only dominant loci (15), which may not be present in different mapping populations. In addition, the dense RAPD linkage map was constructed employing an F_2 population (12), which complicates the integration of maps derived from each parent in silkworm, which has achiasmatic meiosis in one sex. RFLPs are usually codominant, but require relatively large amounts of progeny DNA, and are not suitable for high throughput screening. Because of the disadvantageous characteristics of these markers, the molecular genetic maps contained a large number of linkage groups with only a few assignments integrated into the ELG map (11, 16).

There has been considerable interest in simple sequence repeat (SSR, or microsatellite) markers because they are PCR-based, highly reproducible and polymorphic, generally codominant, and abundant in animal and plant genomes. SSRs have been reported in the silkworm genome (17), and, in many cases, DNA sequences

Conflict of interest statement: No conflicts declared.

Freely available online through the PNAS open access option.

Abbreviations: ELG, established linkage group; RAPD, random amplified polymorphic DNA; RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; SSR, simple sequence repeat; CAPS, cleaved amplified polymorphic sequence.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. DQ249917, AB023091, AB023097, AB062684, AB064496, AB090243, AF024618, AF332550, AF529422, AY227000, AY272037, AY387408, AY426343, AY429304, AY769269, AY769275, AY769283, AY769291, AY769331, AY769341, D12521, D13338, D66906, D76418, DQ242653–DQ243686, U07847, U94328, AB195971, X07552, and X75942).

[b]X.-X.M., S.-J.X., M.-H.L., M.-W.L., and J.-H.H. contributed equally to this work.

[k]To whom correspondence may be addressed. E-mail: yphuang@sibs.ac.cn, gpzhao@sibs.ac.cn, lucheng@swau.cq.cn, mki101@uri.edu, or sywang@chgc.sc.cn.

Table 1. Polymorphism rates (%) of SSR loci in six silkworm strains

Strains	Characters		Dazao	C108	JS	L10	F50B	54A
	Voltinism	Origin						
Dazao	2–3	China	0					
C108	2	China	22.2	0				
JS	2	China	20.8	16.9	0			
L10	2	China	22.3	22.4	18.5	0		
F50B	1	Europe	24.0	23.7	20.7	22.4	0	
54A	2	Japan	22.0	22.1	19.7	22.0	19.3	0

Silkworm strains were selected based on voltinism, origin, and economic importance.

flanking them are found in different strains and even conserved between taxa (18). Recently, a preliminary microsatellite based linkage map and a Z-chromosome linkage map comprised of SSRs, inter-SSRs, and RAPDs have been reported (19).

As part of the international effort in silkworm genomics, we constructed a comprehensive SSR-based linkage map using backcross populations. Here, we describe the construction of linkage maps using 518 previously unidentified SSRs and the assignment of the SSR linkage groups to the ELGs using visible mutations. With the assignment of these markers, we demonstrate the advantages of employing this second-generation linkage map of silkworm for rapid genotyping and studies of synteny.

Materials and Methods

Silkworm Strains. Silkworm strains are listed in Table 1. We used strain Dazao (equivalent to p50) for plasmid library construction and C108 and Dazao as parental strains for genotyping based on an international consensus to use them for genetic and genomic studies (refs. 10–13; Lepidoptera Consortium, http://papilio.ab.a.u-tokyo.ac.jp/lep-genome/new_lepgenome.htm). We selected a panel of four additional silkworm strains (JS, L10, 54A, and F50B) to represent different economic traits, geographic origin and voltinism (generations per year) for polymorphism analysis of the SSR markers.

Backcross Design. We made backcrosses using single-pair matings between (*i*) an F_1 male (from a cross between a Dazao female and a C108 male, i.e., Dazao × C108) with a C108 female (dcBC1M) and (*ii*) an F_1 female (Dazao × C108) with a C108 male (dcBC1F). We scored the progeny for the *p* locus (4), and used 250 larvae (125 males and 125 females) from the dcBC1M backcross and 22 larvae from the dcBC1F backcross (11 males and 11 females) for DNA isolation and scoring.

Additional matings were used to identify the correspondence between the SSR-based linkage groups and the ELGs. We generated 28 vmBC1F*x* backcrosses [*x* refers to one of the 28 visible mutants (vm, Table 2)] by backcrossing an F_1 female, derived from a single pair mating of a strain *x* female (with the target *x* visible mutation) and a Dazao male, to a homozygous tester male. A wild-type male was used as the tester for a dominant visible mutation, and a homozygous mutant male was used as the tester for a recessive mutation. To confirm the initial ELG assignments, 11 additional representative visible mutations (*ch-2*, *cf*, *Nd-s*, *oh*, *Ze*, *bp*, *w-3*, *I*, *re*, *bts*, and *sch*) (6) were used with the same mating scheme (Table 2). Our assignments were further confirmed by sequence-tagged-site (STS) primers developed from genetically mapped *B. mori* BACs based on BAC-FISH chromosomal identification (20).

Genomic Library Construction and Screening for SSR Loci. Genomic DNA was prepared from individual final instar larvae following Yasukochi (12). Two genomic plasmid libraries were constructed for the Dazao strain. The genomic DNA was partially digested with restriction enzymes *Sau*3AI or *Tsp*509, and size-selected fragments (>7 kb) were cloned into plasmid pUC18. Ligation products were transformed into DH10B electro-competent cells and library colonies were arrayed in 384-well microtiter dishes.

Radioactively labeled probes of the repeat sequences, $(CA)_{15}$ and $(CT)_{15}$, were hybridized with the genomic libraries under screening conditions that favored positive clones bearing four or more repeats. These positive clones were sequenced with a combination of several anchored nonspecific primers, $(GT)_7X$ (X = A,T,C) or $(GA)_7X$ (X = A,T,C). Based on the first sequencing results, we designed forward primers for each confirmed nonredundant SSR locus, which were used for a second sequencing of the same clones. We subsequently designed reverse primers based on the second sequencing information and tested for amplification by using Dazao strain DNA as template. At this stage, we only retained 2,500 primer pairs that could amplify single bands in Dazao and C108. All primers were designed with PRIMER 5.00 (Premier Biosoft, Palo Alto, CA).

The genomic densities of the $(CA)_n$ and $(CT)_n$ repeats were calculated based on the number of unique repeat loci confirmed by sequencing both ends of the repeats and the genome size of silkworm obtained from the whole genome draft sequence (428.7 Mb, ref. 21). Because the biases in sequencing and clone redundancy needed to be adjusted, both events were analyzed. We noticed that the redundancy reduction rates were intrinsic but similar for either repeat, whereas the sequencing success rates, including the percentage of clones picked for sequencing and the

Table 2. Integration of SSR linkage map with the established chromosome (Chr.) linkage map

Silkworm Chr.	SSR linkage groups	Visible mutations*	Genetic distance	Percentage†	Marker total (SSR, others)‡
1	29	*od* and *sch*	67.8	2.0	7 (6, 1)
2	1	*p*	87.9	2.6	16 (15, 1)
3	13	*lem* and *Ze*	150.4	4.4	20 (18, 2)
4	12	*L*	106.7	3.1	21 (21, 0)
5	26	*oc* and *re*	142.3	4.1	22 (18, 4)
6	24	E^{KP}	102.0	3.0	14 (13, 1)
7	20	*q*	78.8	2.3	13 (12, 1)
8	17	*st*	106.2	3.1	21 (20, 1)
9	19	*Ia* and *I*	87.7	2.6	14 (13, 1)
10	15	*w-2* and *w-3*	134.8	3.9	24 (22, 2)
11	10 and 14	*K* and *bp*	217.9	6.3	40 (40, 0)
12	3	*ms*	167.4	4.9	20 (20, 0)
13	6	*ch* and *cf*	86.1	2.5	14 (14, 0)
14	7	*U* and *Nd-s*	106.2	3.1	26 (26, 0)
15	4	*bl*	94.5	2.8	14 (12, 2)
16	23	*cts*	60.8	1.8	15 (13, 2)
17	28	*Bm* and *bts*	62.8	1.8	10 (9, 1)
18	2	*mln* and *ch-2*	142.3	4.1	20 (20, 0)
19	21	*nb*	81.4	2.4	15 (14, 1)
20	11	*ci* and *oh*	131.0	3.8	21 (20, 1)
21	8	*rb*	175.6	5.1	19 (18, 1)
22	5	*or*	133.6	3.9	18 (17, 1)
23	9	*tub*	88.8	2.6	22 (22, 0)
24	18	*Sel* and *Xan*	189.8	5.5	34 (32, 2)
25	25	*Nd*	99.4	2.9	23 (23, 0)
26	22	*so*	260.8	7.6	27 (25, 2)
27	27	*III*§	118.9	3.5	19 (19, 0)
28	16	*E-tr*	150.0	4.4	19 (16, 3)
Total	29	39	3431.9	100	548 (518, 30)

*The visible mutations refer to Doira (4) and Lu *et al.* (6).
†The ratio (%) of the genetic distance of each linkage group to the total genetic distance of the genome.
‡The total 548 markers include 518 SSR markers, 27 CAPS markers, 2 RAPD markers and 1 visible mutations; see Fig. 1 and Table 3.
§No visible mutation available.

Table 3. Genes mapped on SSR linkage map of *B. mori* and synteny analysis to butterfly and fruit fly

CAPS	Gene description	GenBank accession number	Chromosome location of silkworm	Polymorphic characters	Mapped to silkworm chromosome	Synteny to fruit fly chromosome
C0101*	Rcf96	DQ249917	1	PCR	1	No match
C0301	*B. mori* ribosomal protein L13A	AY769283	Unknown	TaqI	3	X
C0302	*B. mori* ribosomal protein S26	AY769341	Unknown	BclI	3	2L
C0501	*B. mori* glucosidase gene	AY272037	Unknown	XhoI	5	No match
C0502	*B. mori* wing disc-specific protein	AB062684	Unknown	BclI	5	No match
C0503*	Rcf71	U94328	5	DdeI	5	No match
C0504	*B. mori* elongation factor 1 alpha	D13338	Unknown	EcoRI	5	No match
C0601	*B. mori* attacin gene	D76418	Unknown	MluI	6	2R
C0701*	Rcf47	AF024618	7	AluI	7	3L
C0801	*B. mori* truncated α-amylase gene	U07847	Unknown	SspI/BclI	8	2R
C0901	*B. mori* molybdenum cofactor sulfurase gene	AB090243	9	PCR	9	No match
C1001	*B. mori* CP8 precursor gene	AY387408	Unknown	HhaI	10	No match
C1002*	*B. mori* orphan nuclear receptor E75A (E75)	AF332550	10	Cfr13I	10	3L
C1501	*B. mori* Boceropsin gene	AB064496	Unknown	PCR	15	3R
C1502	*B. mori* ribosomal protein L7A	Ay769275	Unknown	SacI	15	X
C1601*	*B. mori* DNA, Q8.86b-r RAPD sequence	AB023091	16	PCR	16	No match
C1602	*B. mori* ribosomal protein P2	AY769269	Unknown	SspI	16	2L
C1701	*B. mori* ribosomal protein L22	AY769291	Unknown	MboI	17	X
C1901†	*B. mori* gene egg-specific protein	D12521	19	DdeI	19	No match
C2001†	*B. mori* low molecular lipoprotein 30K	X07552	20	MspI	20	No match
C2101*	*B. mori* sorbitol dehydrogenase gene	D66906	21	PCR	21	3R
C2201*	*B. mori* PTTH-KS gene for prepro PTTH	X75942	22	PCR	22	No match
C2401	*B. mori* Suil gene	AY426343	Unknown	MboI	24	3L
C2402	*B. mori* ADP/ATP translocase (ANT) gene	AY227000	Unknown	SacII	24	X
C2601*	*B. mori* DNA, U16.105b-r RAPD sequence	AB023097	26	PCR	26	No match
C2602	*B. mori* ribosomal protein S15A	AY769331	Unknown	PCR	26	X
C2801	*H. melpomene* rosina STRI-B-548A cubitus interruptus gene	AY429304	Unknown	NcoI	28	No match
C2802	*B. mori* cubitus interruptus gene	AF529422	Unknown	HaeIII/MluI	28	4
C2803	Fruit fly embryonal lethal gene	wdS00443	Unknown	BalI	28	2R

*CAPS markers initially mapped using RFLPs and RAPDs (unpublished data).
†Mapped gene sequences in the NCBI database.

ratio of clones successfully sequenced, were more or less artificial and varied significantly for these two different kinds of repeats. Therefore, the nonredundant repeats were reestimated based on the total number of positive clones screened with the statistical total redundancy reduction rates, assuming 100% sequencing success rates (Table 4, which is published as supporting information on the PNAS web site).

PCR and Analysis of Amplified SSR Loci. Standard PCRs (15 μl) for SSR locus amplification contained 1.5 μl 10× buffer, 0.2 mM dNTP, 0.4 pM each primer, 0.38 unit of Taq polymerase (TaKaRa, 5 units/μl), and 15 ng template DNA. The reactions were initiated at 95°C for 2 min followed by 16 cycles of 94°C (30 s) denaturation, 63–56°C (1 min) annealing (see Table 5, which is published as supporting information on the PNAS web site, for the list of primer-pairs) and 72°C (1 min) extension. Twenty-four additional cycles were performed under the same conditions with a fixed annealing temperature of 56°C. The final elongation step was 10 min at 72°C. The amplified products (3 μl) from individual PCRs were size fractionated by electrophoresis (8% acrylamide gels in 1× TBE buffer at 110 volts for 8 h) in parallel with pUC19 DAN/MspI molecular weight markers.

We carried out polymorphism analysis using all of the selected SSR loci that could be amplified in the Dazao strain on a panel of six silkworm strains, including Dazao, C108, Jinsong, Lan10, 54A, and F50B (Table 1).

SSR Marker Scoring. The SSR markers that showed polymorphism between Dazao and C108 and had lengths from 100 to 350 bp were used to score mapping populations. Forward primers were 5′-labeled with TAMRA, HEX, or FAM. We genotyped 189 dcBC1M and 22 dcBC1F offspring using the same SSR markers. Genotyping was carried out on an Applied Biosystems 377 DNA sequencer (Applied Biosystems) for 2 h with 96 lanes run per gel.

We scored individual offspring as either homozygous (designated 1) or heterozygous (designated 2) for each SSR marker. We analyzed the segregation pattern of each marker in the dcBC1M progeny; the ratio of homozygotes versus heterozygotes for autosomal markers was 1:1, as expected (χ^2 test; $P > 0.05$), whereas the ratio for six markers did not differ from 3:1 ($P > 0.05$), as expected for *Z* chromosome markers. *Z* chromosome markers would appear as pseudo homozygous in female progeny due to their hemizygous status, indicating sex linkage. Linkage groups were generated by using MAPMAKER (version 3.0, ref. 22) with a LOD score of 5.0 and the Kosambi mapping function was used to calculate the distances between marker loci in cM.

BLAST Search. We used BLASTX (23) with default parameters at NCBI and FlyBase (release 4.1) to obtain homology for sequences represented by cleaved amplified polymorphic sequence (CAPS) markers, known genes, and cDNAs. Criteria for relatedness were a maximum *E* value of −20 and bitscore of 80 (range from 41–96% amino acid identity for minimum 55 amino acids sequenced) (Table 3).

Results and Discussion

Genomic Distribution of SSR Loci. From 5,400 positive clones containing CA-repeats and 8,200 clones containing CT-repeats derived from two 7-kb insertion genomic libraries (230,000 clones, ≈3× silkworm genome coverage), we identified 1,272 CA-repeats and 1,418 CT-repeats confirmed by sequencing from both ends. Because of significant bias in the process of picking clones for

sequencing and slight bias in the ratio of clones successfully sequenced, the total numbers of nonredundant SSR loci for either repeat were reestimated by adjusting these biases (*Materials and Methods*), resulting in 2,657 loci for $(CA)_n$ and 4,174 loci for $(CT)_n$. Thus, the average densities of these repeats were calculated as ≈6.20 per Mb for $(CA)_n$ (161 kb per CA repeat) and ≈9.74 per Mb for $(CT)_n$ (103 kb per CT repeat). Because we favored colonies with relatively strong signals in the hybridization process to identify loci suitable for genotyping, loci with low-number ($n < 4$) repeats were underrepresented (*Materials and Methods*). In contrast, although previous estimates included all available repeats with various size distributions, they were obtained from much smaller (7, 17, 18) or biased (24) samples. Therefore, this estimation of genomic abundance for high repeat $(CA)_n$ and $(CT)_n$ loci ($n > 4$) is still a useful reference.

Polymorphism of SSR Markers in Representative Silkworm Strains. We subsequently identified ≈2,500 markers from confirmed SSR loci that produced robust PCR products of expected size employing chromosomal DNA from a test strain, Dazao. Analysis of a panel of six silkworm strains indicated that the highest polymorphism (24.0%) appeared between F50B, a strain of European origin, and Dazao, a Chinese strain, whereas the lowest polymorphism (16.9%) appeared between two Chinese strains, JS and C108 (Table 1). This low polymorphism (compared to that of rat, ranging from 35.0% to 55.0%, ref. 25) is in accordance with other analyses in silkworm using SSR markers (18). This is probably due to the fact that all cultivated silkworm strains originated from a common ancestor (1) with a long history of inbreeding under consistently strong selective conditions for strain maintenance and improvement (18).

As well characterized strains widely used for silkworm molecular genetic studies (10–13), Dazao and C108 were selected for map construction even though only 555 loci displayed polymorphism (22.2%) between them (Table 1). BLAST analysis of these 555 loci to the recently announced 3× shotgun sequence (26) and 6× shotgun sequence (21) yielded 68.6% and 85% perfect matches, respectively. In addition, only single matches were identified between SSR markers and the assembled DNA sequence contigs, indicating that no contig was long enough to cover the physical distance between any two of the 555 SSR loci. These results are consistent with the observation that the 3× and 6× shotgun sequences are estimated to cover 75% and 90% of the genome, respectively (21, 26).

Linkage Map Construction and Map Characteristics. The 555 SSR polymorphic loci between Dazao and C108 strains were used to genotype dcBC1M progeny. The segregation data for the inheritance patterns of 518 SSR loci were integrated into 29 linkage groups by using MAPMAKER at a threshold LOD score of 5.0. The remaining 37 markers could not be used for clustering because they were heterozygous in C108. Because there was one additional linkage group than the number of chromosomes ($n = 28$), we used a second backcross dcBC1F, which enabled a direct test for linkage of the SSR markers. This analysis not only confirmed the sorting results of the dcBC1M-based genotyping, but also found that the small linkage group 10 had the same inheritance pattern as group 14. Integration of these two groups resulted in identification of 28 independent linkage groups, which coincided with the haploid chromosome number of silkworm, covering 27 autosomes and the *Z* chromosome (Table 2 and Fig. 1).

Our SSR linkage maps are a second-generation linkage map for silkworm derived solely from SSR markers. The number of markers on individual linkage groups ranged from 6 to 40 (Table 2). Compared to the other high-density linkage maps of silkworm (12, 13), the SSR markers were more evenly distributed over the genome, enabling us to establish compact linkage groups (29 groups) with a limited number of markers (518) (Table 5) employing the dcBC1M mapping strategy. The recombination map length for individual linkage groups ranged from 60.8 cM (linkage group 16) to 260.8 cM (linkage group 26), with an average distance between markers of 6.27 cM in most of the linkage groups. The total length of the SSR linkage maps was 3431.9 cM, which fell between the two previous estimates, i.e., 1800 cM for the dense RAPD map (12) and 6000 cM for the AFLP map (13). Although many scenarios for the discrepancies in map length are possible, including differences in mating strategy and strains used, the distribution of markers is a likely contributing factor. The high reliability of the SSR scoring system and the even distribution of the SSR markers suggest that our estimate may better represent the overall recombination events between C108 and Dazao. Increased marker density should converge on a more realistic map length value.

Assignment of SSR-Based Linkage Groups to ELGs. We selected 39 visible mutations, representing one or two mutations corresponding to each ELG, and carried out vmBC1F*x* backcrosses using Dazao as the reference strain. As shown in Table 2, we were able to assign each of the 27 SSR linkage groups to a single silkworm ELG. In earlier studies, two visible mutations, *Sel* and *Xan,* were assigned to ELGs 24 and 27, respectively (5, 6). However, our analysis found that these mutations were both linked to SSR linkage group 18. The same linkage pattern was obtained in additional backcrosses and a three-point linkage analysis (data not shown). We concluded that these two visible mutations were located on the same silkworm chromosome, as reported recently (16). The difficulties in judging the phenotype of these visible mutations may explain the previous misassignment. Another visible mutation, *Gc*, previously assigned to ELG 15 (5), was found to link with SSR markers assigned to ELG 28. There are two scenarios for this contradictory result. First, the previous assignment using visible mutations for *Gc* was not correct, possibly because of confusing the phenotype. Second, there is probably a fourth gene (*Gd*) controling cocoon color in addition to *Ga*, *Gb* (the two complementary genes controlling green cocoon, ref. 5), and *Gc* (the independent gene controlling green cocoon, ref. 5). The two examples described here indicated that the SSR linkage map not only offered a convenient platform for gene mapping but also identified past mapping errors.

To confirm the chromosome assignments, we designed CAPS markers for 10 cloned genes or cDNA sequences (unpublished data) and developed single-copy CAPS markers from two RAPDs (10), all of which had been assigned to ELGs using backcrosses to visible markers. Although all 12 markers were successfully mapped to the SSR-based molecular linkage groups and the ELG assignments were confirmed using visible mutations (Table 3), none was found on SSR linkage group 27 (Table 2). To make the assignment, we used the STS primer developed from BAC clone assigned to the 27th linkage group by chromosomal FISH (20).

Applications of the SSR Map. This SSR linkage map should be an efficient tool for future map-based positional cloning. We mapped the *p* locus in our backcross population to the proximal end of SSR linkage group 1, corresponding to ELG 2 (1), with an estimated map distance of 10.2 cM to its nearest SSR marker, S0215 (Fig. 1 and Table 2). Many *p*-alleles are known, affecting larval epidermal markings and pigmentation. The complex phenotypic variation suggests that it might be under strong selection pressure in wild species and therefore an important positional cloning target.

We developed 15 CAPS markers from silkworm single-copy nuclear gene sequences in GenBank but as yet unmapped onto chromosomes. We determined their linkage assignments and calculated their map positions based on the dcBC1M mapping panel (Table 3). The rapid integration of these markers indicated that the SSR linkage map offers a convenient platform for locating unmapped genes.

Lepidopteran chromosomes are typically holocentric, with multiple microtubule attachment sites. Whether lepidopteran genomes remain syntenic in the presence of this distinctive chromosomal

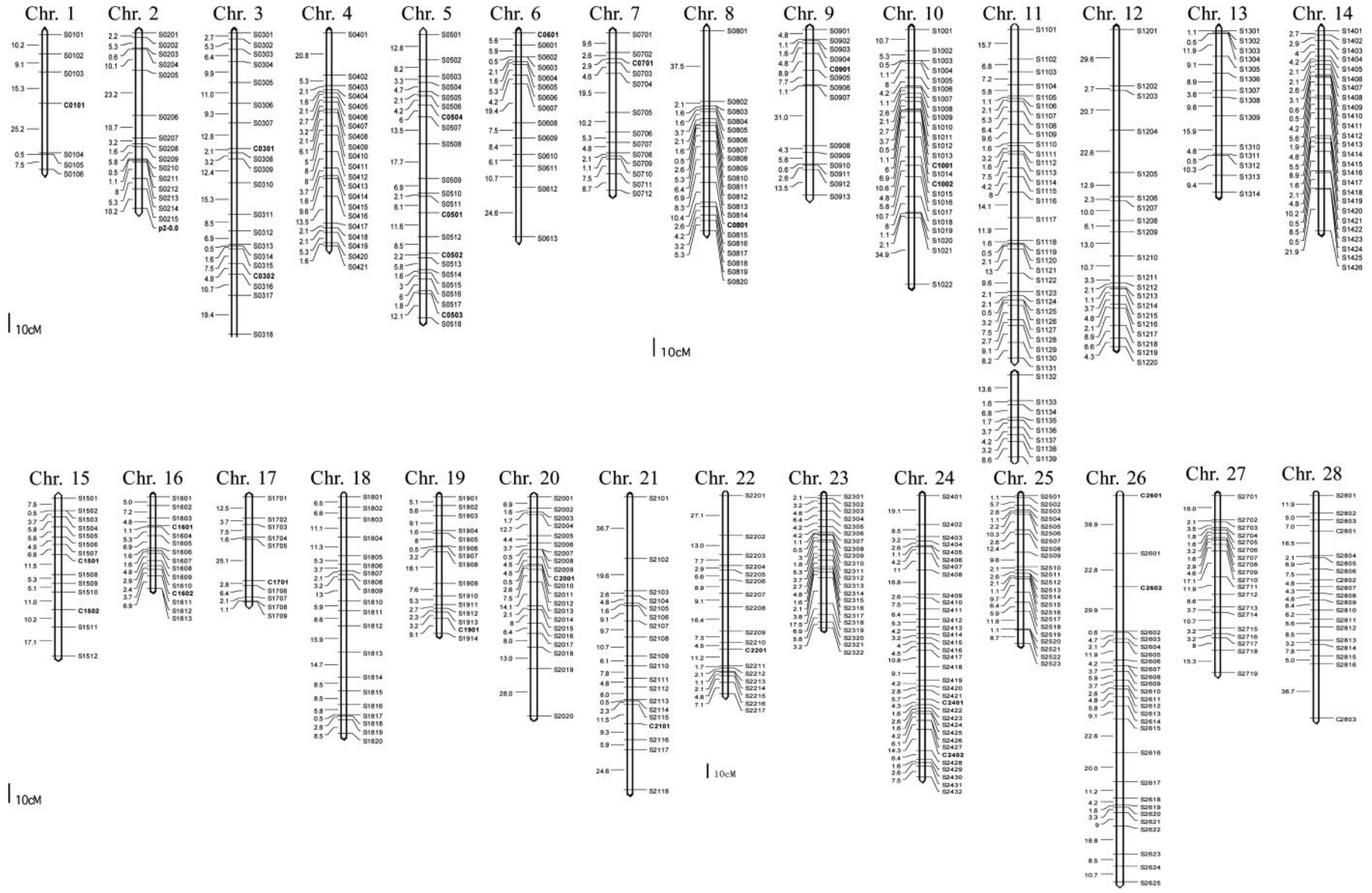

Fig. 1. SSR genetic map for the silkworm. The map is composed of 548 markers and covers 28 linkage groups. Markers are labelled S (SSR markers) or C (CAPS markers) with four digital numbers. Chromosome 11 contained two linkage groups based on the pcBC1M and was integrated into one chromosome through pcBC1F clustering. The numbers on the left of each linkage group are the recombination distances in each interval measured in cM.

structure is unknown, but would present important advantages for comparative genomic studies. Even though *Drosophila* species and dipterans in general show a high degree of intrachromosomal rearrangement, chromosome arms have remained relatively syntenic over long evolutionary periods (27), suggesting that the fruit fly (*D. melanogaster*) might be a suitable reference for testing synteny with *B. mori* as a model lepidopteran. Among the 27 nuclear genes mapped, 18 had the same function or high homology to genes in the fruit fly based on BLAST analysis (23). With only four chromosomes, it was not surprising to find that the same chromosome arm of fruit fly carried genes located on separate silkworm linkage groups (Table 3). For example, genes on silkworm ELGs 5 (EF1-alpha), 6 (attacin), 8 (truncated α-amylase), 15 (Hsp20.1), and 16 (RpP2) were located on fruit fly chromosome 2R. Furthermore, four ribosomal protein genes (L13A, L7A, L22, and S15A) located on chromosome X of the fruit fly were scattered on different silkworm ELGs (3, 15, 17, and 26, respectively). On the other hand, some genes located on the same silkworm ELG were present on different fruit fly chromosomes. For instance, one ribosomal protein gene (S26) located on silkworm ELG 3 was found on the fruit fly chromosome 2L, same as the P2 ribosomal protein gene located on the silkworm ELG 16. Although, preliminary, these results suggest that considerable translocation and/or internal chromosome rearrangements have occurred since the divergence of Lepidoptera and Diptera, estimated at 240 million years ago (28). Evidence for regions of microsynteny may still be found, with the availability of a denser map and more detailed comparisons.

Recently, linkage maps were established for two species of mimetic butterflies, *Heliconius melpomene* (29) and *Heliconius erato* (30), using a combination of AFLPs, SSRs, allozymes, and CAPS markers; additional linkage assignments were made with well conserved CAPS markers from an EST project for the two species (31). Two of the CAPS markers mapped here have also been placed on the butterfly maps: locus EF1-alpha (*H. melpomene* linkage group 10 mapped onto *B. mori* ELG 5; ref. 29), and cubitus interruptus (*H. erato* linkage group LG 3 mapped onto *B. mori* ELG 15; ref. 30 and Table 3). The SSR map will enable mapping many additional well conserved genes identified as ESTs (32) and in the silkworm genome (21, 26) to serve as anchors for a broad comparative analysis of synteny among Lepidoptera and other insects (30, 31).

Conclusions

The SSR-based linkage map described here is the densest map reported for a lepidopteran with a large number of chromosomes ($n = 28$). With its one-to-one assignment of the molecular linkage groups to ELGs, this map is a unique landmark for silkworm genetics and an enormous technological resource. Considering the incompleteness of the silkworm genomic sequence drafts (75–90%) (21, 26) and the limited number of mapped SSR and related molecular markers, it is probably premature to establish a quantitative relationship between the physical and genetic maps; this conclusion is reinforced by the fact that only single matches were identified between SSR markers and sequenced contigs. With the gradual increase of molecular markers on the genetic map and the establishment of longer contigs or supercontigs, these genetic linkage groups will facilitate the estimation of marker density needed for map-based cloning, the ordering of genomic contigs, and the establishment of a complete physical map.

Although the marker density of this molecular genetic map is relatively low because of the high level of homozygosity between the strains used in this study, the even distribution of markers and the overall 6.27 cM average marker distance offers sufficient resolution for preliminary gene mapping and genetic dissection of quantitative trait loci. Applying the latter strategy for economically important characters will enable marker-assisted selection, which is not yet used in sericulture. The map can also be used for synteny studies of Lepidoptera and other insects. Furthermore, we are genotyping the existing 2,500 markers for different pairs of silkworm strains, which may display higher levels of polymorphism than Dazao and C108. By integrating these data into a common linkage map, marker density may reach as high as ≈2–3 cM. This effort will make this map useful for both geneticists and practical breeders for a range of applications.

We thank Profs. R. J. Cogdell, S. J. Benkovic, and X. X. Zhu for their critical reading of the manuscript. The research was funded by the Knowledge Innovation Program (KSCX2-SW-318) from the Chinese Academy of Sciences (CAS), High Tech Research and Development (863) Program 2004AA2Z1020, Fundamental Research Project 2005CB121000, and three programs funded by the Shaghai Commission of Science and Technology (Projects 03JC14077 and O45607032 and Innovation Programs to The Chinese Human Genome Center at Shanghai). J.N. acknowledges the Department of Biotechnology, Government of India for financial assistance.

1. Xiang, Z. H. (1995) in *Genetics and Breeding of the Silkworm* (Chinese Agriculture Press, Beijing), pp. 273–289.
2. Tamura, T., Thibert, C., Royer, C., Kanda, T., Abraham, E., Kamba, M., Komoto, N., Thomas, J. L., Mauchamp, B., Chavancy, G., *et al.* (2000) *Nat. Biotechnol.* **18,** 81–84.
3. Tomita, M., Munetsuna, H., Sato, T., Adachi, T., Hino, R., Hayashi, M., Shimizu, K., Nakamura, N., Tamura, T. & Yoshizato, K. (2003) *Nat. Biotechnol.* **21,** 52–56.
4. Doira, H. (1992) *Linkage Maps and List of Genetical Stocks Maintained in Kyushu University* (Institute of Genetic Resources, Kyushu University, Fukuoka, Japan).
5. Fujii, H., Banno, Y., Doira, H., Kihara, H. & Kawaguchi, Y. (1998) *Important Genetic Resources* (Institute of Genetic Resources, Kyushu University, Fukuoka, Japan), 2nd Ed.
6. Lu, C., Dai, F. & Xiang, Z. (2003) *Scientia Agricultura Sinica* **36,** 968–975.
7. Nagaraju, J. & Goldsmith, M. R. (2002) *Curr. Sci.* **83,** 415–425.
8. Sturtevant, A. H. (1915) *Am. Nat.* **49,** 42–44.
9. Tazima, Y. (1951) *The Recent Advances in Gene Analysis of the Silkworm* (Technology Report Press, Tokyo).
10. Promboon, A., Shimada, T., Fujiwara, H. & Kobayashi, M. (1995) *Genet. Res.* **66,** 1–7.
11. Shi, J., Heckel, D. G. & Goldsmith, M. R. (1995) *Genet. Res.* **66,** 109–126.
12. Yasukochi, Y. (1998) *Genetics* **150,** 1513–1525.
13. Tan, Y.D., Wan, C. L., Zhu, Y. F., Lu, C., Xiang, Z. H. & Deng, H. W. (2001) *Genetics* **157,** 1277–1284.
14. Nguu, E. K., Kadono-Okuda, K., Mase, K., Kosegawa, E. & Hara, W. (2005) *J. Insect. Biotechnol. Sericol.* **74,** 5–13.
15. Sharma, S. K., Knox, M. R. & Ellies, T. H. N. (1996) *Theor. Appl. Genet.* **93,** 751–758.
16. Yasukochi, Y., Banno, Y., Yamamoto, K., Goldsmith, M. R. & Fujii, H. (2005). *Genome*, **48,** 626–629.
17. Reddy, K. D., Abraham, E. G. & Nagaraju, J. (1999) *Genome* **42,** 1057–1065.
18. Prasad, M. D., Muthulakshmi, M., Madhu, M., Archak, S., Mita, K., Nagaraju, J. (2005) *Genetics.* **169,** 197–214.
19. Nagaraja, G. M., Mahesh, G., Satish, V., Madhu, M., Muthulakshmi, M. & Nagaraju, J. (2005) *Heredity* **95,** 148–157.
20. Yoshido, A., Bando, H., Yasukochi, Y. & Sahara, K. (2005) *Genetics*, **17,** 674–685.
21. Xia, Q. Y., Zhou, Z. Y., Lu, C., Cheng, D. J., Dai, F. Y., Li, B., Zhao, P., Zha, X. F., Cheng, T. C., Chai, C. L., *et al.* (2004) *Science* **306,** 1937–1940.
22. Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M. J., Lincoln, S. E. & Newburg, L. (1987) *Genome* **1,** 174–181.
23. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402.
24. Li, B., Xia, Q. Y., Lu, C., Zhou, Z. Y. & Xiang, Z. H. (2004) *Genome Prot. Bioinfo.* **2,** 24–31.
25. Watanabe, T. K., Ono, T., Okuno, S., Miyakita, A. M., Yamasaki, Y., Kanemoto, N., Hishigaki, H., Oga, K., Takahashi, E., Irie, Y., *et al.* (2000) *Mamm. Genome* **11,** 300–305.
26. Mita, K., Kasahara, M., Sasaki, S., Nagayasu, Y., Yamada, T., Kanamori, H., Namiki, N., Kitagawa, M., Yamashita, H., Yasukochi, Y., *et al.* (2004) *DNA Res.* **11,** 27–35.
27. Zdobnov, E. M., von Mering, C., Letunic, I., Torrents, D., Suyama, M., Copley, R. R., Christophides, G. K., Thomasova, D., Holt, R. A., Subramanian, G. M., *et al.* (2002) *Science* **298,** 149–159.
28. Mitsialis, S. A. & Kafatos, F. C. (1985) *Nature* **317,** 453–456.
29. Jiggins, C. D., Mavarez, J., Beltrán, M., McMillan, W. O., Johnston, J. S. & Bermingham, E. (2005) *Genetics* **171,** 557–570.
30. Tobler, A., Kapan, D., Flanagan, N. S., Gonzalez, C., Peterson, E., Jiggins, C. D., Johnston, J. S., Heckel, D. & McMillan, W. O. (2005) *Heredity* **94,** 408–417.
31. Papanicolaou, A., Joron, M., McMillan, W. O., Blaxter, M. L. & Jiggins, C. D. (2005) *Mol. Evol.* **14,** 2883–2897.
32. Mita, K., Morimyo, M., Okano, K., Koikem Y., Nohata, J., Kawasaki, H., Kadono-Okuda, K., Yamamoto, K., Suzuki, M. G., Shimada, T., *et al.* (2003) *Proc. Natl. Acad. Sci. USA* **100,** 14121–14126.