Eur. J. Entomol. **105**: 73–80, 2008
http://www.eje.cz/scripts/viewabstract.php?abstract=1304
ISSN 1210-5759 (print), 1802-8829 (online)

Intronic sequences of the silkworm strains of *Bombyx mori* (Lepidoptera: Bombycidae): High variability and potential for strain identification

KEE YOUNG KIM[1], EUN MEE LEE[2], IN HEE LEE[2], MEE YEON HONG[3], PIL DON KANG[1], KWANG HO CHOI[1], ZHONG ZHENG GUI[4], BYUNG RAE JIN[5], JAE SAM HWANG[1], KANG SUN RYU[1], YEON SOO HAN[3] and IKSOO KIM[3]*

[1]Department of Agricultural Biology, The National Institute of Agricultural Science and Technology, Suwon 441-100, Republic of Korea

[2]Department of Life Science, Hoseo University, Asan-city, Chungchungnam-do 336-795, Republic of Korea

[3]College of Agriculture & Life Sciences, Chonnam National University, Gwangju 500-757, Republic of Korea

[4]Sericultural Research Institute, Chinese Academy of Agricultural Sciences, Zhenjiang 212018, China

[5]College of Natural Resources and Life Science, Dong-A University, Busan 604-714, Republic of Korea

Key words. Silkworm, *Bombyx mori*, intronic sequence, strain identification, genetic variation

Abstract. We sequenced nine introns of 25 silkworm (*Bombyx mori* L.) strains, assuming that the introns are particularly prone to mutation. Mean sequence divergence and maximum sequence divergence in each intronic sequence among 25 silkworm strains ranged from 0.81% (3.8 nucleotides) ~ 9.15% (85.2 nucleotides) and 1.2% (seven nucleotides) ~ 39.3% (366 nucleotides), respectively. The degree of sequence divergence in some introns is very variable, suggesting the potential of using intronic sequences for strain identification. In particular, some introns were highly promising and convenient strain markers due to the presence of a large indels (e.g., 403 bp and 329 bp) in only a limited number of strains. Phylogenetic analysis using the individual or the nine concatenated intronic sequences showed no clustering on the basis of known strain characteristics. This may further indicate the potential of the intronic sequences for the identification of silkworm strains.

INTRODUCTION

Due to their great economic value, more than 3000 genetically different silkworm (*Bombyx mori*) strains, some of which produce different qualities and yields of the silk, are maintained in Europe and Asia (Nagaraju, 2000). In the Republic of Korea, approximately 300 silkworm strains are maintained in The National Institute of Agricultural Science & Technology (NIAST), and some of these are also kept in other cocoon-producing countries, such as China, Japan and India. These strains are reared annually, and scores from indoor rearing are analyzed for consistent character maintenance.

Silkworm strains are described on the basis of several morphological and physiological characteristics such as origin, voltinism, number of moults or cocoon making. However, sorting one strain from another based on these characteristics is often difficult because of the high variability and environmental dependence of these characteristics. Furthermore, silkworm strains have been selected in order to maximize their commercial and regional suitability. Thus, compared to the diversity that exists within natural populations, the genetic diversity of silkworm strains is very much diminished. Additionally, the general genetic backgrounds of the strains are quite similar, even though some of the characteristics selected for commercial and regional purposes may differ. From a practical perspective, discriminating one strain from another is often necessary because silkworm larvae with similar external morphologies are often reared at the same place at the same time, and cross contamination between strains is possible. Once this occurs, the best procedure is to destroy the contaminated cultures, as it is impossible to guarantee the purity of the remaining larvae. This limitations has prompted some investigators to use molecular methods such as isozymes (Seong, 1997; Sohn et al., 2002), RAPD (random amplified polymorphic DNA; Hwang et al., 1995), RFLP (restriction fragment length polymorphism, Shi et al., 1995), and direct sequencing of mitochondrial DNA (mtDNA; Kim et al., 2000; Hwang et al., 1996, 1998) to identify strains. Most of these techniques resolved the origin-based evolutionary relationships among some silkworm strains, and the relationships between the domestic and wild silkworm, *B. mandarina*, presumed ancestor of the domestic silkworm, rather than discriminating between silkworm strains. Microsatellite DNA is an exception in this regard, in that some microsatellite DNAs reflected a certain character type in the silkworm strains (i.e. diapause vs. non-diapause) (Reddy et al., 1999).

It is suggested that introns are particularly prone to mutations (Serapion et al., 2004), possibly due to reduced selective pressure (Juszczuk-Kubiak et al., 2004; Ueda et al., 1984, 1985; Martinez et al., 2004). Thus, more variation may be revealed by intronic sequencing. There are several genomic sequences of *B. mori* in the GenBank. Thus, several intron regions from the GenBank were

* Corresponding author; e-mail: ikkim81@chonnam.ac.kr

TABLE 1. General information on the silkworm strains utilized in this study.

Strain no.	Strain	Origin	Voltinism	Moultinism	Egg color	Blood color	Cocoon color/shape
24	N74	Japan	2	4	B	W	LYG /Spindle
33	Bibakjam	Japan	2	4	B	W	W/Peanut
34	BakanEBkwainggi	Japan	2	4	B	W	W/Oval
40	Usungrokeui	Japan	2	4	B	W	LYG /Long peanut
60	Lemon	Europe	1	4	B	Y	LYG/ Peanut
62	Baghdad	Europe	1	4	B	W	LG/Peanut
69	Youlkukjam	Europe	1	4	B	Y	LY/Spindle
79	Huka	Europe	1	4	B	Y	Y/Peanut
83	Kwasulpyung	China	2	4	B	W	W/Oval
122	Sammyunhong	China	2	3	B	Y	F/Spindle
128	Yonggakjam	China	2	4	B	W	W/Peanut
142	Hansunghukran	China	2	4	♀, B ♂, W	W	W/Oval
143	Hansungbanmun	China	2	4	B	W	W/Peanut
145	Hansammyun	Korea	2	3	B	W	LG/Peanut
148	Sun 3ho	Korea	2	3	B	Y	Y/Long peanut
157	HM	Tropics	M	4	B	W	LG/Spindle
158	Kyunsakjuk	Europe	1	4	B	Y	F/Oval
174	Il 111	Japan	2	4	B	Y	Y/Oval
181	Je 1bakran	unknown	unknown	4	Br	W	LG/Peanut
213	od Eujam	unknown	unknown	4	B	W	C/Peanut
266	NC Bakran	unknown	unknown	4	W	W	W/Oval
290	Eppanol	unknown	unknown	4	B	Y	Y/Spindle
296	Sandong sammyun	unknown	unknown	3	B	W	F/Spindle
319	Nd^{H}	unknown	unknown	4	B	W	W/—
324	Kore sammyun	Korea	2	3	B	W	W/Oval

M – multi-voltine strain; B – black; W – white; Br – brown; Y – yellow; LYG – light yellow green; F – Flesh; LG – light green; LY – light yellow; C – cream; "—" – no rigid cocoon shape.

selected and sequenced to determine the variability in the intronic sequences among the strains and to assess their potential for use in strain identification. Some intronic sequences showed substantial variation among the silkworm strains tested.

MATERIAL AND METHODS

Silkworm strains

Silkworm strains chosen for the present study represent a diverse range of genetic stocks: different geographic origin, voltinism, moultinism, cocoon colour, and cocoon shape (Table 1).

Genomic DNA extraction

Genomic DNA was isolated from eggs of *B. mori* strains that are maintained at NIAST, Republic of Korea. In the case of field-collected wild silkworm, *B. mandarina* (Suwon City, Korea), genomic DNA was extracted from a larval specimen. Approximately 100 eggs or individual larvae were crushed in a glass grinder and genomic DNA was extracted using the Wizard Genomic DNA Purification Kit, in accordance with the manufacturer's instructions (Promega, Madison, WI, USA).

Intron selection

Of the silkworm genes for which complete genomic structures are available, 13 intron regions, approximately 500–700 bp in length, were selected from the GenBank database for laboratory study. These intron regions are described in Table 2. Primers were designed based on the sequence information of the flanking exons (Table 2).

PCR amplification, cloning and sequencing

The polymerase chain reactions were performed using a PCR mix (Bioneer, Soeul, Republic of Korea) with primers, both at a concentration of 10 pmol, along with genomic DNA at a concentration of approximately 100 ng and H_2O up to a total volume of 20 µl. The following PCR protocol was used: 5 min at 94°C, followed by 40 cycles of 30 s at 94°C, 40 s at 50–60°C, and 45 s at 72°C, and a subsequent 7 min final extension at 72°C. The amplified PCR product was separated by electrophoresis in a 0.5% agarose gel (Sigma, St. Louis, MO, USA) with ethidium bromide. The amplicons were then cloned in pGEM-T Easy vector (Promega), and the resulting plasmid DNA was isolated using the Wizard Plus SV Minipreps DNA Purification System (Promega). Both strands of the PCR amplicons were cycle-sequenced using the ABI PRISM® BigDye® Terminator v1.1 Cycle Sequencing Kit and electrophoresed in each direction on an ABI PRISM® 310 Genetic Analyzer (PE Applied Biosystems, Foster City, CA, USA). When necessary, an additional internal primer was designed to complete sequences by primer walking.

Sequence analysis and phylogenetic analyses

Each intronic sequence was aligned with the original sequence registered in GenBank using the CLUSTAL X program (Thompson et al., 1997). Sequence divergence and phylogenetic analysis were performed using PAUP* (Phylogenetic Analysis Using Parsimony and Other Method*) ver. 4.0b10

TABLE 2. Summary of the silkworm intron regions utilized in this study.

Gene	Intron number (abbreviation)	Primer name (direction)	Location*	Sequence (5'-3')	Expected size (bp)	GenBank no.	Reference
Cytoplasmic actin (A4)	Intron 1 (A4 Intron 1)	A4-E1-F1 (F) A4-E2-R1 (R)	488 1134	CAGGGCGTTTAATTTTCAT TATTGCACAGCTTCGTTA	539	U49644	Mange et al. (1996)
Cytoplasmic actin (A4)	Intron 2 (A4 Intron 2)	A4-E2-F1 (F) A4-E3-R1 (R)	1117 1760	TAACGAAGCTGTGCAATA ATTGTCTACTACCAACG	533	U49644	Mange et al. (1996)
Prothoracicotropic hormone-Kinshu Showa	Intron 3 (PTTH Intron 3)	PTTH-E3-F1 (F) PTTH-E4-R1 (R)	1518 2241	AGCAAGCTATTCCGGAT TGAAGCGTGGAACAGAG	640	X75942	Adachi-Yamada et al. (1994)
Larval cuticle protein 30	Intron 3 (LCP30 Intron 3)	LCP30-E3-F1 (F) LCP30-E4-R1 (R)	1984 2639	ACAACTCTGGCCGTTACA ATCCAGTGTAGAAGCCA	566	X74321	Nakato et al. (1994)
Larval cuticle protein 30	Intron 4 (LCP30 Intron 4)	LCP30-E4-F1 (F) LCP30-E5-R1 (R)	2623 3319	TGGCTTCTACACTGGAT AGGGGAACAATGCTGGA	579	X74321	Nakato et al. (1994)
Sex-specific storage-protein 1	Intron 3 (SP1 Intron 3)	SP1-E3-F1 (F) SP1-E4-R1 (R)	3569 5009	ATGATCCGAGAGGGTAT CAATGTCCTCAGACTTC	459	X12978	Sakurai et al. (1988)
Vitellin-degrading protease precursor	Intron 4 (VDP Intron 4)	VDP-E4-F1 (F) VDP-E5-R1 (R)	6760 7609	CTCACCGATTTACGCAAT TATTCTGGTCTCGCACA	694	D16233	Ikeda et al. (1991)
Fibroin light-chain	Intron 3 (FLC Intron 3)	FL-E3-F1 (F) FL-E4-R1 (R)	10684 11462	ATCTATCTGCCGGTAT ACCGAGAGATTGTCTG	684	M76430	Kikuchi et al. (1992)
Xanthine dehydrogenase	Intron 3 (XDH Intron 3)	XD-E3-F1 (F) XD-E4-R1 (R)	12837 13377	ACGACGAGCTAGAAATA ACTCTCGCCACTATTCTA	455	AB005911	Komoto et al. (1999)

* Locations are taken from the gene sequences in the original publications.

(Swofford, 2002). For tree construction, the maximum-parsimony (MP) method (Fitch, 1971) was performed with the heuristic search. Branches were collapsed if the maximum branch length was zero. Trees were evaluated using the bootstrap test (Felsenstein 1985) with 1,000 iterations. To root the tree, the homologous sequence of *B. mandarina*, which is assumed to be an ancestor of the domestic silkworm (Arunkumar et al., 2006), was utilized.

RESULTS AND DISCUSSION

Nucleotide composition and variability of intronic sequences

Among 13 intron regions, nine provided stable DNA amplicons that could be successfully sequenced. The GenBank accession numbers for 225 intronic sequences, composed of nine intron regions from 25 silkworm strains are DQ833532–DQ833750 and DQ852325–DQ852330. The nucleotide composition and genetic variability of each intronic sequence are presented in Table 3. The nucleotide composition of the intron regions was somewhat biased toward adenine and thymine, ranging from 57% (A4 Intron 1) – 70.9% (LCP30 Intron 4). At each intron, the mean sequence divergence among the 25 silkworm strains ranged from 0.81% (3.8 bp) to 9.15% (85.2 bp) and the maximum sequence divergence at each intron ranged from 1.2% (7 bp) to 39.3% (366 bp) (Table 3). The 5,897 bp of the nine concatenated intronic sequences resulted in 4.65% of the mean sequence divergence and 15.6% of the maximum sequence divergence (Table 4). In comparison with some previous sequence-based studies of silkworm strains from which comparable divergence estimates can be drawn, the divergence estimate obtained in this study is substantial. For example, the maximum sequence divergence among 11 silkworm strains of different origin is only 0.2% in the 410 bp section of the Cytochrome Oxidase Subunit I (COI) gene of mitochondrial DNA (mtDNA) (Kim et al., 2000). The sequence of a ~1000 nucleotide long single intron of the heavy-chain fibroin gene (H-fib) from five strains of *B. mori* and five geographic samples of *B. mandarina* revealed as much as 0.26% of the maximum sequence divergence (Martinez et

TABLE 3. Summary of sequence composition and genetic variability in each intron region among 25 silkworm strains.

Intron	Expected size (bp)	Range (bp)	*Mean nucleotide composition (%)				No. of polymorphic sites			Mean sequence divergence: bp (S. D.)	Maximum sequence divergence: bp (%)
			G	A	T	C	Sub.	indel	Sub. + indel		
A4 Intron 1	539	537–542	22.3	26.5	30.5	20.7	34	5	2	6.763 (3.298)	19 (3.5%)
A4 Intron 2	533	530–533	21.8	25.4	34.5	18.3	28	3	0	5.653 (2.795)	14 (2.6%)
PTTH Intron 3	640	640–1079	18.5	31.5	32.4	17.6	67	446	0	64.33 (28.72)	470 (43.5%)
LCP30 Intron 3	566	566–568	15.6	36.5	34.4	13.6	13	0	0	1.193 (0.790)	7 (1.2%)
LCP30 Intron 4	579	407–590	20.3	27.8	31.1	20.8	22	185	5	45.86 (20.57)	197 (33.2%)
SP1 Intron 3	459	459–460	17.9	30.6	32.1	19.4	30	2	0	3.250 (1.733)	31 (6.7%)
VDP Intron 4	694	350–694	18.5	32.9	29.6	19.0	37	354	31	98.27 (43.69)	366 (39.3%)
FLC Intron 3	684	652–684	15.8	36.2	31.6	16.4	83	47	2	43.34 (19.45)	87 (12.2%)
XDH Intron 4	455	455–481	17.4	35.9	28.8	17.9	74	28	2	43.57 (19.56)	83 (17.1%)

* Mean nucleotide composition (%) was obtained by averaging the nucleotide composition of each intronic sequence in 25 silkworm strains. Sub. – substitution. S. D. – standard deviation.

TABLE 4. Uncorrected pairwise distance of the nine concatenated intronic sequences among 25 silkworm strains.

	*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	24	—	0.100	0.024	0.025	0.036	0.022	0.069	0.034	0.035	0.023	0.034	0.029	0.023	0.045	0.080	0.018	0.082	0.014	0.097	0.037	0.037	0.037	0.020	0.041	0.050
2	33	592	—	0.108	0.088	0.093	0.087	0.136	0.102	0.098	0.109	0.097	0.105	0.092	0.130	0.156	0.094	0.150	0.100	0.153	0.102	0.101	0.108	0.094	0.104	0.136
3	34	139	640	—	0.036	0.023	0.033	0.078	0.020	0.019	0.010	0.021	0.008	0.034	0.048	0.090	0.028	0.094	0.024	0.080	0.018	0.017	0.041	0.033	0.049	0.043
4	40	150	521	210	—	0.025	0.008	0.059	0.029	0.021	0.037	0.025	0.031	0.012	0.056	0.080	0.022	0.078	0.023	0.080	0.027	0.023	0.031	0.019	0.027	0.061
5	60	213	552	137	148	—	0.024	0.076	0.013	0.012	0.019	0.007	0.019	0.029	0.069	0.089	0.028	0.085	0.038	0.066	0.015	0.013	0.045	0.029	0.042	0.064
6	62	130	515	196	48	141	—	0.058	0.028	0.018	0.034	0.022	0.029	0.011	0.055	0.080	0.019	0.076	0.021	0.078	0.024	0.024	0.031	0.018	0.023	0.061
7	69	409	804	460	346	447	342	—	0.068	0.067	0.077	0.071	0.073	0.055	0.034	0.109	0.064	0.101	0.060	0.113	0.076	0.075	0.054	0.062	0.056	0.039
8	79	201	604	116	174	76	167	400	—	0.015	0.016	0.010	0.016	0.026	0.065	0.086	0.024	0.080	0.035	0.075	0.019	0.017	0.039	0.024	0.044	0.060
9	83	208	577	115	125	69	104	396	89	—	0.018	0.006	0.014	0.023	0.064	0.091	0.029	0.088	0.031	0.066	0.012	0.012	0.041	0.030	0.034	0.059
10	122	136	645	59	218	113	201	452	97	104	—	0.014	0.015	0.034	0.055	0.086	0.024	0.089	0.025	0.082	0.025	0.025	0.048	0.034	0.050	0.049
11	128	199	573	122	147	44	128	422	62	37	84	—	0.016	0.027	0.066	0.086	0.024	0.083	0.036	0.070	0.012	0.013	0.042	0.028	0.038	0.060
12	142	173	619	46	183	110	170	430	92	85	89	94	—	0.029	0.053	0.085	0.024	0.089	0.030	0.076	0.013	0.013	0.036	0.029	0.045	0.048
13	143	138	543	201	68	171	64	324	151	138	203	162	170	—	0.054	0.081	0.019	0.073	0.020	0.084	0.028	0.029	0.027	0.015	0.028	0.060
14	145	263	767	286	332	409	326	203	385	381	325	389	315	319	—	0.112	0.048	0.113	0.043	0.127	0.059	0.063	0.062	0.056	0.071	0.012
15	148	472	923	530	474	523	472	644	509	535	510	510	503	477	660	—	0.070	0.022	0.082	0.037	0.094	0.094	0.078	0.078	0.081	0.115
16	157	108	554	163	127	165	110	376	144	169	143	144	139	113	284	416	—	0.074	0.020	0.091	0.030	0.033	0.031	0.017	0.033	0.054
17	158	483	889	553	460	505	450	597	474	517	524	488	526	434	668	128	439	—	0.082	0.034	0.091	0.090	0.073	0.073	0.078	0.119
18	174	83	590	143	134	224	125	354	208	185	147	210	176	118	256	487	121	487	—	0.093	0.037	0.038	0.038	0.026	0.038	0.049
19	181	571	901	475	475	390	460	669	446	390	484	414	451	495	750	218	536	202	549	—	0.073	0.071	0.088	0.090	0.077	0.122
20	213	218	603	104	158	87	143	450	111	69	146	73	75	167	350	558	179	539	219	432	—	0.008	0.038	0.031	0.040	0.058
21	266	218	594	101	134	77	143	446	103	73	148	77	75	171	375	554	196	533	223	422	48	—	0.037	0.029	0.042	0.058
22	290	219	637	242	186	268	184	320	232	243	282	247	215	158	365	461	181	429	227	519	223	219	—	0.027	0.022	0.065
23	296	118	556	196	111	172	108	365	143	178	201	166	171	89	331	463	99	430	151	529	181	171	159	—	0.034	0.060
24	319	243	613	289	158	246	137	330	261	199	296	223	265	166	417	477	193	461	226	457	237	246	128	200	—	0.074
25	324	293	806	255	363	376	358	233	355	348	290	356	284	357	69	680	318	701	292	719	342	344	382	353	440	—

Numbers above the diagonal are mean distance values; numbers below the diagonal are absolute distance values. * Strain number.

TABLE 5. Insertions and deletions (> 6 bp) in the intronic sequences among 25 silkworm strains.

Intron	Indel	Strain number*
PTTH Intron 3	430 UI	33
LCP30 Intron 4	178 SD	69, 145, 324
VDP Intron 4	329 SD	148, 158, 181
LCP30 Intron 4	6 UI	296

UI – unique insertion; SD – shared deletion; * silkworm strains sharing the same deletion.

al., 2004). Further, an initial sequence analysis of five silkworm strains by ~500 bp of a hypervariable A+T-rich region of mtDNA revealed an identical sequence (data not shown). Taking these results into consideration, the degree of sequence divergence of the intron regions in this study is substantial. Specifically, PTTH Intron 3, LCP30 Intron 4, and VDP Intron 4 are highly polymorphic with a maximum of 43.5%, 33.2%, and 39.3% sequence divergence, respectively, including insertion and deletion (indel) (Table 3), suggesting that these intron regions may provide a means of strain discrimination.

Although in theory there is the possibility of heterozygosity, intra-strain variation in the intronic sequences is expected to be minimal in such long maintained silkworm strains, due to purifying selection. In fact, sequencing of five clones each from A4 Intron 1, A4 Intron 2, PTTH Intron 3, LCP30 Intron 4, and VDP Intron 4 at the begin-

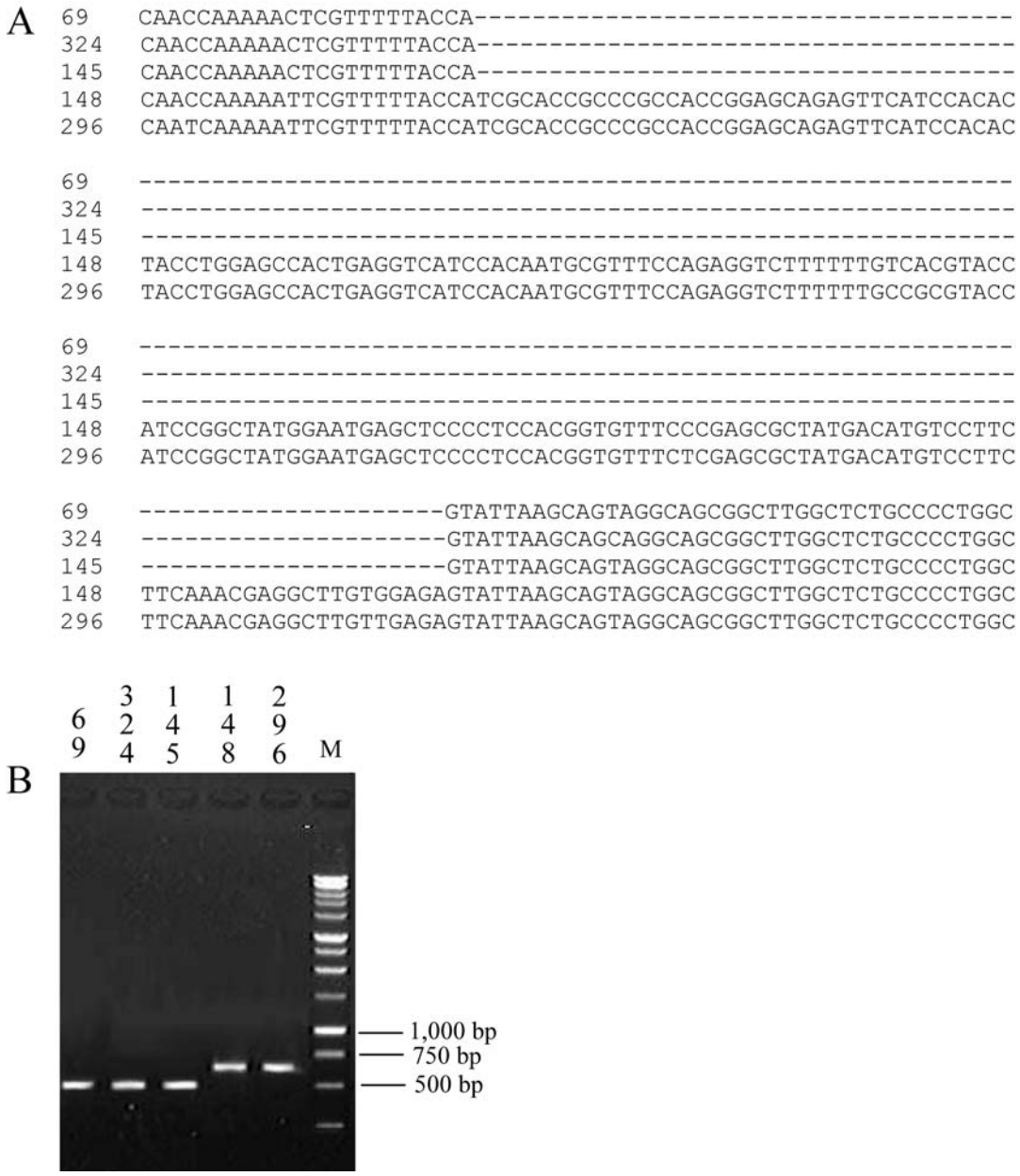

Fig. 1. An example of a large deletion found in the intronic sequence of LCP30 Intron 4: (A) a partial sequence alignment of the strains 69, 324 and 145, showing a 187 bp deletion and of strains 148 and 296 that do not have such a deletion; and (B) PCR products of strains 69, 324 and 145, showing a 187 bp deletion and strains 148 and 296 that do not have such deletion. M, molecular size marker.

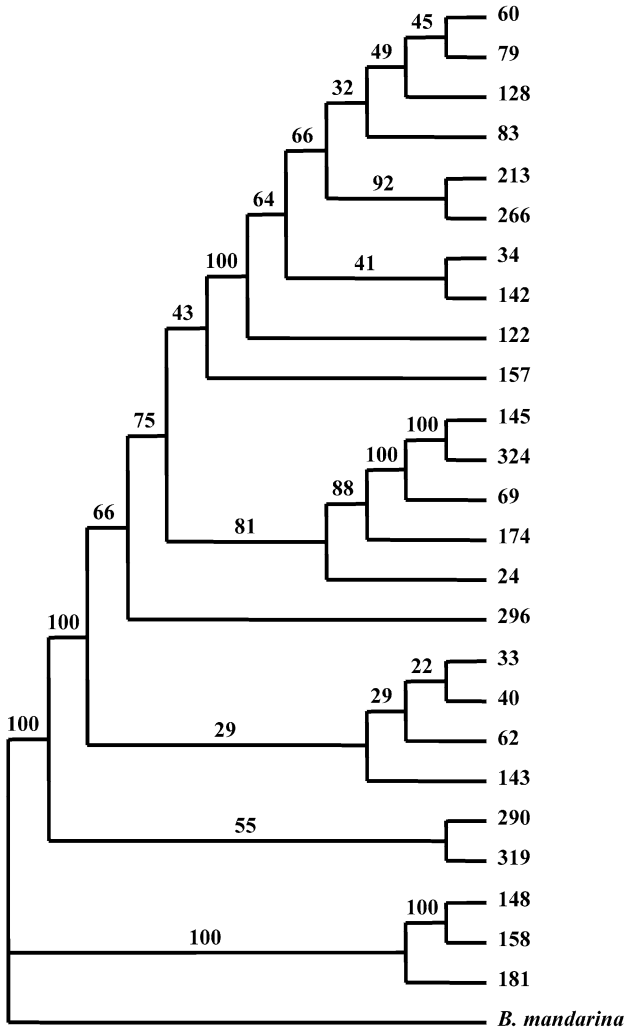

Fig. 2. The result of maximum parsimony analysis of 25 silkworm strains using the nine concatenated intronic sequences that comprise 5,897 bp including indels. The tree was obtained using heuristic search incorporated in PAUP (Swofford, 2002). This single-most parsimonious tree (tree length = 2930, consistent index = 0.658, retention index = 0.703, homoplasy index = 0.342) was obtained from an unweighted parsimony analysis. Numbers at each node indicate the number of times the nodes were supported in an analysis of a 1000 bootstrap replicate dataset. The outgroup selected was the wild silkworm, *B. mandarina*, which is presumed to be an ancestral species of the domestic silkworm, *B. mori*.

ning of this study did not reveal a single mutation. This probably implies that the intra-strain variation might be minimized during purifying selection. Nevertheless, an extra analysis of the remaining four intron regions is needed to confirm possible heterozygosity.

Large indels

Some intron regions in some silkworm strains have long indels (Table 5). There is a 178 bp deletion in LCP30 Intron 4 in three strains, numbers 69, 145, and 324 (Fig. 1). Additionally, there is a unique 430 bp insertion in PTTH Intron 3 in strain 33 and a 329 bp deletion in VDP Intron 4 in three strains, 148, 158, and 181 (data not shown). Except for a few runs of T nucleotides from the 178 bp region of LCP30 Intron 4, no repeating sequence was found either in the 430 bp region of PTTH Intron 3 or 329 bp region of VDP Intron 4. These indels seem to be highly promising and convenient strain markers, which do not require direct sequencing for strain identification, but more clones and strains need to be tested. There are also several 3~5-bp indels shared by only a small number of strains.

Analysis of relationships among strains

To test whether or not the intronic sequences reflect any known morphological characteristic or strain origin, a phylogenetic analysis of the 25 silkworm strains was performed using individual or nine concatenated intronic sequences. In this analysis, indel mutations were included in the construction of a phylogenetic tree, following the method suggested by Kawakita et al. (2003), in which long indels or short gaps are highly reliable sources of phylogenetic information, at least at lower taxonomic levels (i.e. among *Bombus* bumble bee species). In the trees obtained using individual intronic sequences, no group was formed based on any known strain characteristics, such as voltinism, moultinism, egg colour, blood colour, cocoon colour, or cocoon shape (data not shown). Furthermore, the tree obtained using the nine concatenated intronic sequences comprising 5,897 bp including indels resulted in a similar conclusion (Fig. 2). For example, although strains 148 and 158 have different origins (Korea vs. Europe), they formed a strong sister group with the highest bootstrap support (Fig. 2) and a relatively low sequence divergence, 2.2% (128 bp), in the range 0.6% (33 bp) ~ 15.6% (923 bp) (Table 4). Furthermore, these two strains share an identical sequence in LCP30 Intron 3 (Table 6). Nevertheless, the two strains only share brown egg colour and yellow blood colour (Fig. 3), but differ in voltinism and moultinism, cocoon colour, cocoon shape, larval morphology and origin (Table 1). Similar examples can be found at many other nodes of the tree. Thus, it is suggested that these intronic sequences do not reflect any known strain characteristics. Instead, the intronic sequences appear to be the product of neutral evolution with respect to gene expression. In fact, previous investigations of the 5' end region of the H-fib intron of *B. mori* and *B. mandarina* show no diagnostic difference between the two species. This was explained by the neutrality of the intronic variations with respect to the H-fib expression (Ueda et al., 1985; Kusuda et al., 1986). These results suggest that the intronic sequences obtained in this study may have better resolving power for the identification of silkworm strains than of the long-term evolutionary relationships among strains.

Identical sequences among strains

Sequence analysis of the nine intron regions provided varying numbers of identical sequences among the 25 silkworm strains, except for FLC Intron 3 (Table 6), for which substantially high mean and maximum sequence divergences of 5.99% (42.6 bp) and 12.2% (87 bp), respectively, were obtained (Table 3). Some intronic sequences, including LCP30 Intron 3 (identical sequences

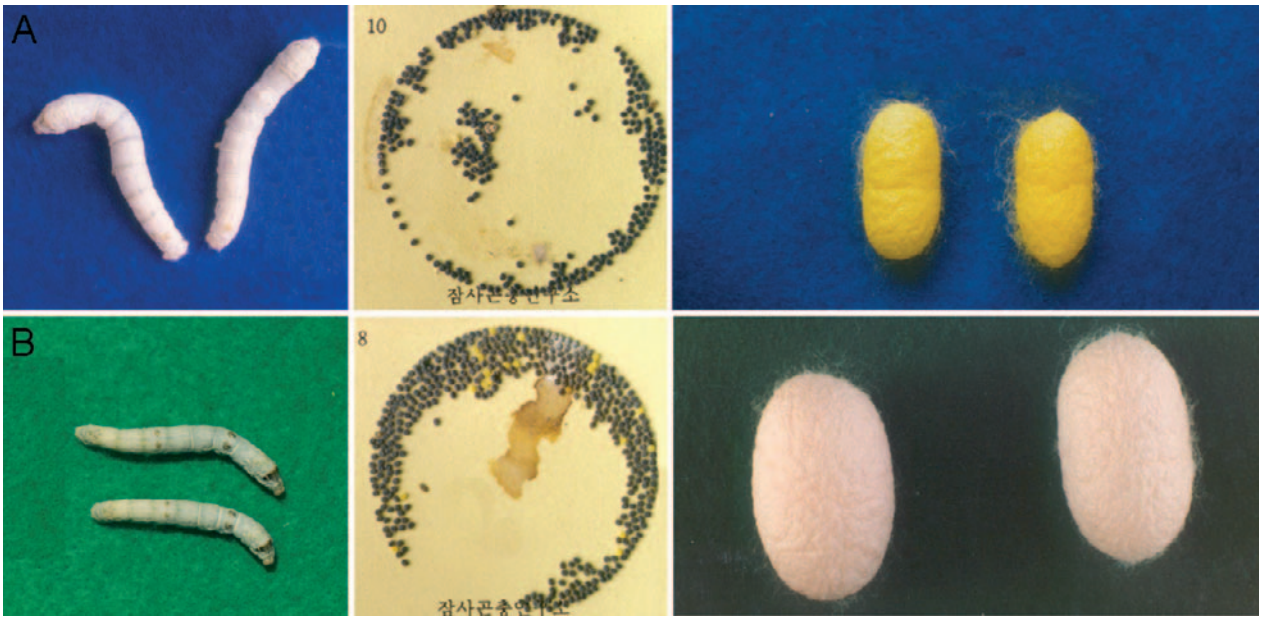

Fig. 3. Examples of silkworm larvae, eggs, and cocoons of strains (A) 148 and (B) 158, respectively.

in 13 strains) and SP1 Intron 3 (identical sequences in 12 strains), resulted in identical sequences among several strains, but other intronic sequences such as PTTH Intron 3 (identical sequences in three strains), LCP30 Intron 4 (identical sequences in two strains in each of three sequence types), and VDP Intron 4 (identical sequences in three strains) showed identical sequences in only a few strains. Thus, either using one intronic sequence, such as FLC Intron 3 alone or concatenation of some intronic sequences, or by removing highly conserved sequences (i.e. LCP30 Intron 3 and SP1 Intron 3), it is currently possible to discriminate the 25 strains tested in this study. Further, the strains sharing identical sequences, in PTTH Intron 3, LCP30 Intron 4 and VDP Intron 4, are very different. Collectively, concatenation of these intronic sequences can also distinguish all 25 strains tested in this study. Thus, these intronic sequences may be suitable for strain identification after more clones and strains are tested.

More than 3000 silkworm strains are currently kept around the world, including those in Korea. These are always exposed to accidental contamination, but there is

TABLE 6. List of identical intronic sequences among silkworm strains.

Intron	Silkworm strains with identical sequence
A4 Intron 1	○ 83 = 128
	○ 79 = 34 = 62 = 33 = 181 = 24
A4 Intron 2	○ 83 = 128
	○ 40 = 148 = 142 = 34 = 181 = 60
	○ 157 = 62
	○ 145 = 213 = 319
	○ 143 = 290
LCP30 Intron 3	○ 181 = 128 = 62 = 148 = 143 = 174 = 319 = 157 = 145 = 266 = 40 = 158 = 324
LCP30 Intron 4	○ 60 = 33
	○ 83 = 122
	○ 266 = 181
	○ 24 = 158
PTTH Intron 3	○ 145 = 290 = 324
SP1 Intron 3	○ 213 = 122 = 143 = 324 = 319 = 40 = 142 = 34 = 128 = 62 = 181 = 60
VDP Intron 4	○ 145 = 142 = 157
FLC Intron 3	None
XDH Intron 3	○ 79 = 83 = 142 = 60 = 34 = 213
	○ 157 = 290

○ – silkworm strains with identical sequence. None, no silkworm strain shares identical sequence.

no sequence information that discriminates between similar-looking strains. By utilizing the accumulated sequence information on the silkworm genome found in GenBank, the possibility of using intronic sequences for discriminating silkworm strains were tested, assuming that introns are more prone to mutations than other portions of the genome. We found that some intronic sequences provided more variability than other previous sequence-based analysis and may prove useful for strain discrimination.

REFERENCES

ADACHI-YAMADA T., IWAMI M., KATAOKA H., SUZUKI A. & ISHIZAKI H. 1994: Structure and expression of the gene for the prothoracicotropic hormone of the silkmoth Bombyx mori. *Eur. J. Biochem.* **220**: 633–643.

ARUNKUMAR K.P., METTA M. & NAGARAJU J. 2006: Molecular phylogeny of silkmoths reveals the origin of domesticated silkmoth, Bombyx mori from Chinese Bombyx mandarina and paternal inheritance of Antheraea proylei mitochondrial DNA. *Mol. Phyl. Evol.* **40**: 419–427.

FELSENSTEIN J. 1985: Confidence limits on phylogenics: an approach using the bootstrap. *Evolution* **29**: 783–791.

FITCH W.M. 1971: Toward defining the course of evolution: minimal change for a specific tree topology. *Syst. Zool.* **20**: 406–416.

HWANG J.S., LEE J.S., KANG H.A., LEE S.M. & SUH D.S. 1995: Analysis of genetic relationships among the silkworm, Bombyx mori, strains using RAPD-PCR. *Kor. J. Genet.* **17**: 291–300.

HWANG J.S., LEE J.S., LEE S.M., KANG H.A., HWANG S.J. & SUH D.S. 1996: Fundamental study for RAPD-PCR analysis in the silkworm, Bombyx mori. *Kor. J. Seric. Sci.* **38**: 7–12.

HWANG J.S., LEE J.S., GOO T.W., KANG H.A., SOHN H.R. & KIM H.R. 1998: Analysis of molecular relationships between Bombyx mandarina and Bombyx mori using RAPD-markers. *Kor. J. Life Sci.* **8**: 426–430.

IKEDA M., YAGINUMA T., KOBAYASHI M. & YAMASHITA O. 1991: cDNA cloning, sequencing and temporal expression of the protease responsible for vitellin degradation in the silkworm, Bombyx mori. *Comp. Biochem. Physiol. (B)* **99**: 405–411.

JUSZCZUK-KUBIAK E., SAKOWSKI T., FLISIKOWSKI K., WICINSKA K., OPRZADEK J. & ROSOCHACKI S.J. 2004: Bovine μcalpain (*CAPN*1)-gene: new SNP within intron 14. *J. Appl. Genet.* **45**: 457–460.

KAWAKITA A., SOTA T., ASCHER J.S., ITO M., TANAKA H. & KATO M. 2003: Evolution and phylogenetic utility of alignment faps within intron sequences of three nuclear genes in bumble bees (Bombus). *Mol. Biol. Evol.* **20**: 87–92.

KIKUCHI Y., MOR K., SUZUKI S., YAMAGUCHI K. & MIZUNO S. 1992: Structure of the Bombyx mori fibroin light-chain-encoding gene: upstream sequence elements common to the light and heavy chain. *Gene* **110**: 151–158.

KIM I., BAE J.S., SOHN H.D., KANG P.D., RYU K.S., SOHN B.H., JEONG W.B. & JIN B.R. 2000: Genetic homogeneity in the domestic silkworm, Bombyx mori, and phylogenetic relationship between B. mori and the wild silkworm moth, B. mandarina, using mitochondrial COI gene sequences. *Int. J. Indust. Entomol.* **1**: 9–17.

KOMOTO N., YUKUHIRO K. & TAMURA T. 1999: Structure and expression of tandemly duplicated xanthine dehydrogenase genes of the silkworm (Bombyx mori). *Insect Mol. Biol.* **8**: 73–83.

KUSUDA J., TAZIMA Y., ONIMARU K., NINAKI O. & SUZUKI Y. 1986: The sequence around the 5' end of the fibroin gene from the wild silkworm, Bombyx mandarina, and comparison with that of the domesticated species, B. mori. *Mol. Gen. Genet.* **203**: 359–364.

MANGE A., COUBLE P. & PRUDHOMME J.C. 1996: Two alternative promoters drive the expression of the cytoplasmic actin A4 gene of Bombyx mori. *Gene* **183**: 191–199.

MARTINEZ L., ALMAGRO J.C., COLL J.L. & HERRERA R.J. 2004: Sequence variability in the fibroin-H intron of domesticated and wild silk moths. *Insect Biochem. Mol. Biol.* **34**: 343–352.

NAGARAJU J. 2000: Recent advances in molecular genetics of the silkmoth, Bombyx mori. *Current Sci.* **2**: 151–161.

NAKATO H., SHOFUDA K., IZUMI S. & TOMINO S. 1994: Structure and developmental expression of a larval cuticle protein gene of the silkworm, Bombyx mori. *Biochim. Biophys. Acta* **121**: 64–74.

REDDY K.D., ABRAHAM E.G. & NAGARAJU J. 1999: Mocrosatellites in the silkworm, Bombyx mori: Abundance, polymorphism, and strain characterization. *Genome* **42**: 1057–1065.

SAKURAI H., FUJII T., IZUMI S. & TOMINO S. 1988: Complete nucleotide sequence of gene for sex-specific storage protein of Bombyx mori. *Nucleic Acids Res.* **16**: 7717–7718.

SEONG S.I. 1997: Genetic relationships of silkworm stocks in Korea inferred from isozyme analyses. *Kor. J. Seric. Sci.* **39**: 119–133.

SERAPION J., WALDBIESER G.C., WOLTERS W. & LIU Z.J. 2004: Development of type I markers in channel catfish through intron sequencing. *Anim. Genet.* **35**: 463–466.

SHI J., HECKEL D.G. & GOLDSMITH M.R. 1995: A genetic linkage map for the domesticated silkworm, Bombyx mori, based on restriction fragment length polymorphisms. *Genet. Res. (Cambridge)* **66**: 109–126.

SOHN B.H., KIM H.S., KANG P.D., LEE S.U. & SEONG S.I. 2002: Examination of genetic relationships of silkworm stocks in Korea by additive isozyme analysis. *Int. J. Indust. Entomol.* **5**: 205–211.

SWOFFORD D.L. 2002: PAUP*: *Phylogenetic Analysis Using Parsimony (*and other methods). ver 4.10.* Sinauer Associates, Sunderland, MA.

THOMPSON, J.D., GIBSON T.J., PLEWNIAK F., JEANMOUGIN F. & HIGGINS D.G. 1997: The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**: 173–216.

UEDA H., HYODO A., TAKEI F., SASAKI H., OHSHIMA Y. & SHIMURA K. 1984: Sequence polymorphism in the 5'-upstream region of the fibroin-H-chain gene in the silkworm, Bombyx mori. *Gene* **28**: 241–248.

UEDA H., MIZUNO S. & SHIMURA K. 1985: Sequence polymorphism around the 5'-end of the silkworm fibroin-H-chain gene suggesting the occurrence of crossing-over between heteromorphic alleles. *Gene* **34**: 351–355.

Received June 20, 2007; revised and accepted July 26, 2007