Genes Genet. Syst. (2000) **75**, p. 105–113

Characterization of Early Follicular cDNA Library Suggests Evidence for Genetic Polymorphisms in the Inbred Strain C108 of *Bombyx mori*

David R. Mills[1] and Marian R. Goldsmith[2*]

[1]*Division of Biology and Medicine, Brown University, Box G, Providence, Rhode Island 02912-0001, USA*
[2]*Department of Biological Sciences, University of Rhode Island, Kingston, Rhode Island 02881-0816, USA*

(Received 25 September 1999, accepted 10 April 2000)

Recent work towards the completion of a saturated molecular genetic linkage map for the lepidopteran silkworm, *Bombyx mori* (*n*=28), has provided evidence for existing polymorphisms in the inbred strain C108. Two inbred parental strains, p50 and C108, were crossed to produce the F1 (P/C) hybrid offspring. The populations used in this project were comprised of a combination of 29 F2 (F1 *x* F1) and 31 reciprocal backcross (P/C *x* C/C, P/C *x* P/P) progeny. All restriction fragment length polymorphisms (RFLPs) for the initial analysis were hybridized with anonymous probes derived from a random early follicular cDNA (Rcf) library from *Bombyx*. A total of 19 Rcf probes were selected as showing scorable codominant polymorphic patterns when screened against F2 and backcross DNAs digested with the restriction enzymes *Eco*RI, *Hin*dIII, or *Pst*I, and Southern blotted to nylon membranes for hybridization. Of the newly reported Rcf probes, 7 (37%) were characterized as producing 'simple' polymorphic patterns, while 12 (63%) were characterized as producing 'complex' polymorphic patterns. Further characterization of the complex patterns subdivided this group into two general classes: polymorphisms that contained an additional allele, and multiple bands that contained an easily scored two banded polymorphism. Because the extra allele class was limited to the (P/C *x* C/C) backcross progeny, it is suggested that the inbred parental strain C108 harbors polymorphic loci that are inherited in a simple Mendelian fashion. A genetic analysis discussing plausible origins and maintenance of these polymorphisms is presented.

INTRODUCTION

Recent advances in recombinant DNA technologies have greatly enhanced progress in many disciplines including genomic mapping, linkage analysis and evolutionary biology. For example, following the advent of the polymerase chain reaction (PCR) (Saiki et al., 1985; Saiki et al., 1988), both microsatellites (Litt and Luty, 1989; Tautz, 1989; Weber and May, 1989) and the dominantly scored randomly amplified polymorphic DNAs (RAPDs) (Welsh and McClelland, 1990; Williams et al., 1991) provided the ability to generate a new class of DNA marker useful for observing inherited Mendelian segregation and for estimating levels of genetic variation that exist between and within populations. A second class of DNA marker that requires the hybridization of probes to complementary genomic sequences includes restriction fragment length polymorphisms (RFLPs) (Lander and Botstein, 1989), and fluorescent *in situ* hybridization (FISH) (Fan et al., 1990; Trask, 1991; Ellegren et al., 1994). Interestingly, the PCR contributed to RFLP analysis by increasing the ease by which hybridization probes may be constructed. For example, cloned cDNA fragments may be easily amplified by designing primers specific to the vector sequences that flank the insert's cloning site.

RFLPs, which arise through point mutations, inversions, insertions, and deletions, are inherited in a Mendelian fashion as codominant markers and appear to be randomly dispersed throughout most genomes. Although their analysis is technically involved and costly, they have a number of qualities that make them robust, flexible, and genetically informative for map construction. For example, RFLP probes may be composed of either known cloned sequences (i.e., expressed genes), or randomly-cloned DNAs of unknown function (e.g., anonymous

Edited by Masa-Toshi Yamamoto
* Corresponding author. E-mail: mki 101@uri.edu

genes). Furthermore, since the RFLP hybridization probes do not need to have homology to coding sequences, restriction polymorphisms may also be detected in noncoding regions, or regions of repetitive DNA (Botstein et al., 1980). Since traits of interest are often found to segregate with known molecular DNA markers in a Mendelian fashion, arranged crosses may be rapidly screened for specific polymorphisms.

The domesticated silkworm, *Bombyx mori (n* =28), has a rich history as an organism used in areas as diverse as commercial sericulture and basic genetic research. Historically, the silkworm has had a major economic impact as the primary organism involved in silk production, and, among insects, is second only to the well–studied fruit fly, *Drosophila melanogaster*, in detailed mechanistic genetic studies. Presently, the *Bombyx* genome is second only to *Drosophila* in its number of mapped morphological and molecular markers (Doira et al., 1992; Shi et al., 1995; Promboon et al., 1995). Our laboratory, in collaboration with others in the silkworm community, is engaged in the process of constructing a saturated molecular genetic linkage map for *B. mori*. Our lab has used a random early follicular cDNA library from *B. mori* as a source for generating genetic mapping probes that recognize polymorphic RFLPs between two of our parental strains, C108 and p50 (Shi et al., 1995). Briefly, randomly selected clones are screened against digested parental and F1 progeny DNAs. Markers that identify a size polymorphism between the two parental strains are selected for mapping against F2 progeny. During the course of our present *Bombyx* mapping project, which will be discussed elsewhere, we screened a set of novel cDNA clones against both F2 and backcross progeny and identified what we believe to be polymorphic loci segregating in a Mendelian fashion in our inbred parental C108 line. To our knowledge, this is the first such report describing this phenomenon in silkworms. The existence of these polymorphisms raises interesting questions regarding the evolutionary history of improved strains and it is suggested that laboratories working with this popular race and possibly other closely related Lepidoptera be advised of this condition.

MATERIALS AND METHODS

The *B. mori* stocks used for this project are maintained in this laboratory. Strains include the parentals C108 and p50, their F1 hybrids and a combination of 29 F2 (F1 *x* F1) and 31 reciprocal backcross (P/C *x* C/C, P/C *x* P/P) progeny. C108 was a kind gift of Dr. Y. Tazima, National Institute of Genetics, Mishima in 1972, and has been maintained in this laboratory by inbreeding small sib populations (average 30–100 larvae per rearing). p50 was a gift of Dr. H. Doira, Institute of Genetic Resources, Kyushu University, Fukuoka, in 1986, and was similarly maintained. During the rearing, the silkworms were fed fresh mulberry leaves [obtained from the University of Rhode Island's East Farm or campus greenhouse]. DNA for the genetic studies was extracted from the posterior silkglands of 3–4 day old fifth instar larvae as previously described (Gautreau et al., 1993; Shi et al., 1995). Briefly, pairs of posterior silkglands were dissected and ground in 1 ml of warmed extraction buffer (0.5% sarkosyl, 50 mM EDTA, pH 8.0, and 0.5 mg/ml proteinase K) and incubated for 2–3 hours at 50°C. The extract was then subjected to one round of chloroform (chloroform isoamyl alcohol 24:1) and 3 rounds of phenol (0.1% hydroxyquinoline, pH 8.0) extractions. Following a series of two more chloroform extractions, DNA was ethanol precipitated (1/2 volume 7.5 M ammonium acetate, 2X volume 100% ethanol), pelleted by centrifugation, and resuspended in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0). Once in solution the DNA was further purified by precipitation with 0.1 M spermidine, resuspended in TE buffer and quantified by fluorometry (TKO 100 dedicated mini fluorometer, Hoefer Scientific Instruments).

Ten µg of genomic DNA was digested with an excess (100 –140 units) of *Eco*RI, *Hin*dIII, or *Pst*I restriction enzyme (New England Bio Labs). The digest was allowed to proceed overnight at 37°C to ensure all DNA was cleaved to completion. Following the digest, the DNA was ethanol precipitated, pelleted by centrifugation, rinsed in 70% ethanol, dried under a gentle stream of N_2 gas, and resuspended in 15 µl of TE buffer. The restriction-digested DNA was resolved by overnight gel electrophoresis (0.7% agarose in 1X TAE buffer: 0.04 M Tris–acetate, 0.001 M EDTA, pH 8.0) at 25–30 volts. Prior to Southern blotting, the gel was soaked briefly in 0.25 N HCl followed by 0.5 N NaOH. The DNA fragments were transferred to a nylon membrane (Zeta-Probe GT, Bio-Rad) by capillary action with 10X SSC (1X SSC: 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0) essentially as described by Southern (1975). Following transfer, the blot was rinsed briefly in 2X SSC, air dried, and baked at 80°C for 1 hour.

Anonymous hybridization probes (designated Rcf's for Random cDNA from follicles) were obtained from a random early follicular cDNA library (donated by N. Spoerel, University of Connecticut Medical Center) which was constructed essentially as described by Sambrook et al., (1989). The cDNA inserts were cloned into the *Eco*RI site of the bacteriophage lambda *gt*11 and amplified for use as hybridizing probes using the polymerase chain reaction (PCR) (Saiki et al., 1985; Saiki et al., 1988). For amplification, 25 µl of the phage lysate was removed and denatured at 95°C for 5 min using a PTC–100 thermal cycler (MJ Research, Inc.). Upon completion this DNA was transferred to 25 µl of PCR reaction buffer containing 200 ng each of the lambda *gt*11 forward [5'd(TGGCGACGACTCCTGGAGCCCG)3'] and reverse [5'd(TGACACCAGACCAACTGGTAATGG)3'] primers (a gift of Dr.

Kostas Iatrou), 200 mM of each deoxynucleotide triphosphate (Promega), 10 mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl, and 1 unit of thermostable Taq DNA polymerase (Promega). Amplification was performed for 1 cycle of denaturation at 94°C for 2 min, annealing at 58°C for 90 sec, and extension at 74°C for 90 sec. This was followed by 30 cycles of denaturation at 92°C for 1 min, annealing at 58°C for 90 sec, and extension at 72°C for 2 min. A final extension was carried out at 72°C for 7 min. Amplified inserts were ethanol precipitated, quantified with a fluorometer (Hoeffer) and diluted to 10 ng/μl in 10:1 TE buffer.

Prior to use as RFLP hybridization probes, the amplified inserts were screened using the slot blot procedure (Erlich et al., 1986). The function of this technique was to identify and eliminate probes complementary to highly repeated sequences (i.e., > 1000 copies in the genome) and multicopy (10–100 copies) gene families, as well as known single copy genes whose expression is common to the developmental stage from which this cDNA library was constructed. For example, sequences related to the multicopy chorion (eggshell) genes (Nadel et al., 1980), as well as egg specific protein (ESP) (Sato and Yamashita, 1991), and ribosomal DNA, are known to be especially abundant in the cDNA library (Goldsmith and Shi, 1993) due to their high levels of expression in the follicular cells from which the library was derived or in the case of the rDNAs as contamination during construction of the library (M. Goldsmith, unpublished results). Unique copy inserts that failed to hybridize to these blots were selected to be screened against parental and F1 genomic Southern blots. Subsequently, only single or low copy number sequences that detected polymorphisms on these blots were chosen as hybridization probes for genetic segregation analysis.

Hybridized slot blot and RFLP marker bands were visualized using the ECL (Amersham) nonradioactive chemiluminescent detection system (Heslop-Harrison et al., 1990; Pollard-Knight et al., 1990) following instructions of the manufacturer with minor modifications (J. Shi, personal communication). The hybridization was carried out overnight at 42°C with gentle agitation and then treated with equal volumes of the chemiluminescent detection reagents (Amersham) and exposed to Hyperfilm-ECL luminescence detection film (Amersham) for 1 hour. Blots were successfully reused up to 20 times.

RESULTS

For several years our laboratory has screened an early follicular cDNA library from *B. mori* to identify random RFLP hybridization probes for use as genetic mapping markers. A recent screening of this library produced 19 novel cDNA–based probes that clearly demonstrated codominant RFLP patterns when hybridized to Southern blots of digested C108 and p50 parental and F1 (C *x* P) genomic DNAs. Each probe is referred to individually as an Rcf clone for Random cDNA from Follicles. The probes, their approximate size, and total detected polymorphisms are shown in Table 1. A total of 19 probes, which detected codominant polymorphisms with the restriction enzymes *Eco*RI (68.4%), *Hin*dIII (63.2%), and *Pst*I(52.6%), were selected for analysis against F2 and backcrossed mapping populations. The high percentage of polymorphism with these three restriction enzymes suggests that high levels of genetic diversity remain between the two C108 and p50 parental strains, even though both are derived from stocks of Chinese geographic origin.

Of the 19 selected Rcf probes, 7 (37%) were classified as producing a characteristic 'simple' polymorphic pattern (SPP). This classification was limited to single or low copy sequences giving unambiguous polymorphisms that were easily scored. Of the 7 SPPs detected, 4 of the Rcf probes (424 , 480, 508, and 523) were represented only as a single band polymorphism as shown by Rcf 424 (Fig. 1A), while the remaining 3 Rcf probes (471, 472, and 510) shared additional banding fragments such as a single monomorphic band or other low copy number sequences as shown by Rcf 471 (Fig. 1B). In contrast to the SPPs, 12 (63%) of the Rcf probes were designated as showing a 'complex' polymorphic pattern (CPP). This was characterized by a number of segregating banding patterns. Two classes of CPPs were observed: class II, showing multiple banding patterns that clearly contained an evident polymorphism suitable for scoring as either C, P, or F (the P/C F1 hybrid) (Fig. 1C) and class I, a polymorphic pat-

Table 1. Selected characteristics and classification of Rcf probes used for genetic analysis where (+) showed a polymorphism and (–) showed monomorphic or identical banding patterns.

Rcf Probe	Approximate Size (Kb)	*Eco*RI	*Hin*dIII	*Pst*I	Polymorphic Pattern
360	1.5	+	+	–	Complex I, II
372	1.5	+	–	+	Complex I
424	1.0	–	+	–	Simple
427	1.0	–	+	–	Complex I, II
432	2.0	+	+	–	Complex I, II
434	2.0	+	–	+	Complex I
438	1.5	–	–	+	Complex I
441	1.0	+	+	+	Complex II
471	0.8	+	–	–	Simple
472	0.9	+	–	–	Simple
477	1.0	+	–	+	Complex II
479	1.0	+	+	+	Complex II
480	1.2	–	+	–	Simple
483	1.7	+	–	+	Complex I
486	1.2	+	+	+	Complex II
508	1.2	–	+	+	Simple
510	1.8	–	+	–	Simple
516	1.0	+	+	–	Complex I
523	0.9	+	+	+	Simple

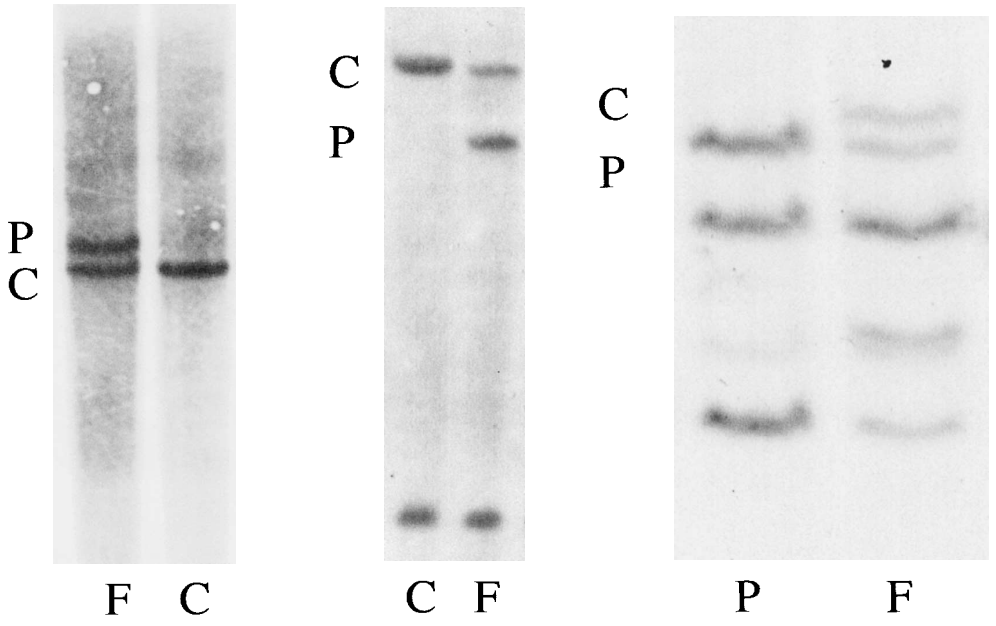

Fig. 1. Autoradiographs illustrating various classifications of polymorphic patterns for Rcf probes. C represents the parental C108 phenotype, P the parental p50 phenotype, and F the F1 hybrid phenotype (10 μg DNA/lane). Panel A, a single copy polymorphism represented by a blot of P/C *x* P/P backcross DNAs digested with *Hin*dIII and hybridized to Rcf 424. Panel B, a single copy polymorphism with an additional monomorphic band demonstrated with F2 DNAs digested with *Eco*RI and hybridized to Rcf 471. Panel C, complex polymorphic pattern producing multiple bands for Rcf 441 when screened against F2 DNAs digested with *Hin*dIII.

tern that contained an additional or third allele not immediately recognized as a parental C or P band when analyzed in the progeny (see below). The emphasis of this report will concern the class I complex polymorphic patterns.

The class I CPPs proved to be an unexpected find and were the impetus behind initiating a characterization of our probe cDNA library. All 8 probes represented by this class (Table 1) were identified as belonging exclusively to the P/C *x* C/C backcross progeny, suggesting that the potential source of this allelic variation was specific to the parental C108 strain. The Rcf probes that produced patterns suggestive of an additional allele included 372, 434, 438, 483, 516, and also 360, 432, and 427, which also share qualities of the class II CPPs. These probes, when hybridized to both the P/C *x* P/P backcross and F2 blots, produced the expected banding pattern (i.e., simple or low copy polymorphisms) identified from their respective parental/F1 blots. Because the blots were hybridized a number of times against the various probes, all blots producing new allelic patterns were compared to their previous Rcf hybridization patterns to ensure that 'bleeding' of old bands was not producing a false interpretation of the current scoring.

The Rcf probes, 360 and 434, initially gave the simplest patterns for explaining the observed results in terms of heterozygous alleles derived from the C108 parent. When screened against the initial parental/F1 blotted DNAs, Rcf 360 detected 2 distinct but similarly scored polymorphic bands after digestion with *Eco*RI and *Hin*dIII; however, the lower molecular weight (MW) polymorphism with the *Hin*dIII digest was very faint. When Rcf 360 was later screened against the *Eco*RI digested P/C *x* C/C DNAs, the lower MW polymorphism revealed the expected backcross pattern (C or F); however, the upper MW polymorphism revealed a distinct doublet comprised of the originally scored C phenotype and a second slightly lower MW band. This doublet was not present when the P phenotype was scored, which allowed for unambiguous scoring of the hybrid F phenotype (Fig. 2A).

When digested with the restriction enzyme *Hin*dIII, the same P/C *x* C/C progeny again expressed the extra allele doublet in the samples scored as C (i.e., in the absence of the expressed P phenotype) (Fig. 2B). Also, the progeny scores (as C or F) were identical, as expected, for Rcf 360 with both the *Eco*RI and *Hin*dIII digested DNAs. By scoring the doublets as C and progeny segregating both P and C as F, linkage was detected between Rcf's 360 and 479. This scoring was later confirmed when this linkage group (Rcfs 360/479) was tested in both the P/C *x* P/P backcross and F2 progeny. The basis for assigning genetic linkage between a pair of Rcf probes relied on the observance of identical inheritance patterns between 2 or more probes within the described mapping populations and will be presented further elsewhere. A similar strategy to that of Rcf 360 was followed when characterizing Rcf 434. Because the *Pst*I digested P/C *x* C/C DNAs scored similarly to the original parental/F1 blot, but the same

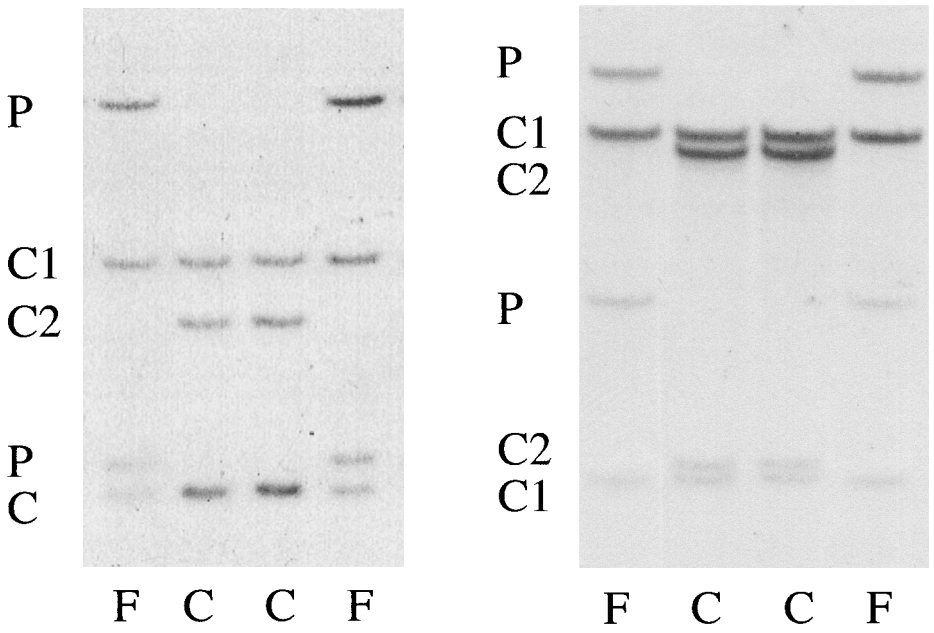

Fig. 2. Autoradiographs illustrating class I complex polymorphic patterns characterized as identifying an extra allele when hybridized to P/C x C/C backcross DNAs (10μg/lane). These segregation patterns would be expected from a PC^2 *x* C^1C^1 backcross mating (see Table 2). Panel A, Rcf 360 hybridized to *Eco*RI digested backcross DNAs. Panel B, Rcf 360 hybridized to backcross *Hin*dIII digested DNAs.

individuals digested with *Eco*RI produced the extra allele, it suggested that hybrid F offspring inherited the originally expected P and C bands whereas only those scored as the C phenotype expressed the doublet. Another useful strategy while scoring inheritance patterns was to compare the molecular weights of the observed bands to the size standard bacteriophage lambda *Hin*dIII. For instance, with Rcf 434, the 'extra allele' was evident because of its higher molecular weight than either the C or P bands scored earlier on the parental/F1 blot (not shown).

These results are interesting because they suggest that the parental strain C108 is not homozygous at all loci and not as genetically inbred as originally thought. A genetic model is presented in Table 2 diagramming crosses that predict observed patterns expressing an "extra" allele. Briefly, it shows the p50 strain as being homozygous at all loci thus far examined and containing 2 copies of the P allele while C108 may possess 2 different alleles at this locus designated as C^1 and C^2. Because C108 is known as an 'improved' strain, meaning that it has a history of matings between the most viable or productive progeny from a series of hybrid crosses selected for traits of economic interest, C108 has had more opportunity to acquire and maintain low levels of genetic variation than the p50 parental populations.

Using this notation, if we return to Rcf 360 in Fig. 2, it is observed that the uppermost band is scored as a P allele and the lower bands show either the doublet comprised of the two C^1 and C^2 alleles or the homozygous C^1C^1 allele. Each individual for this cross may therefore be classified as having inherited either a PC^1 or C^1C^2 pair of alleles from this locus which would result directly from the PC^2 *x* C^1C^1 backcross mating (Table 2). A similar strategy was employed while scoring Rcf 434. In this case, however, the allele pairs C^1C^2 and PC^2 were segregated, which may be explained by the backcross mating PC^1 *x* C^2C^2 (Table 2).

The P/C *x* C/C patterns produced by the Rcf probes, 372, 432, 438, and 516, were initially more difficult to characterize and classify, but using the presented model and resulting strategies acquired from the interpretations of Rcf's 360 and 434, they were also shown to fit the 'extra

Table 2. Theorized P/C *x* C/C backcross mating arrangements which account for the presence of an extra C-specific allele.

Backcross Mating	Expected Progeny Classes	Scored As
PC^1 *x* C^1C^1	C^1C^1	C
	C^1P	F
PC^1 *x* C^1C^2	C^1C^1	C
	C^1C^2	C
	C^1P	F
	C^2P	F
PC^2 *x* C^1C^1	C^1C^2	C
	C^1P	F
PC^2 *x* C^1C^2	C^1C^2	C
	C^2C^2	C
	C^1P	F
	C^2P	F
PC^1 *x* C^2C^2	C^1C^2	C
	C^2P	F

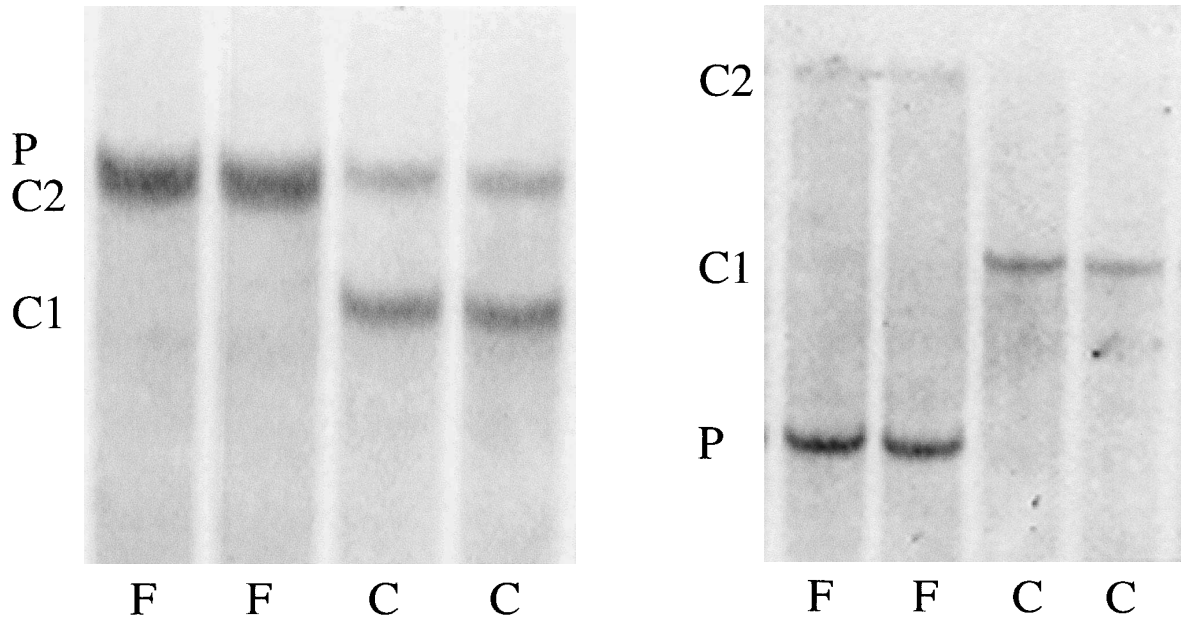

Fig. 3. Hybridization patterns demonstrating class I complex polymorphic patterns for Rcf 516. These segregation patterns would be expected for this locus by either the PC^2 *x* C^1C^2 or PC^1 *x* C^1C^2 backcross matings. Panel A, *Eco*RI digested P/C *x* C/C DNAs. Panel B, *Hin*dIII digested P/C *x* C/C DNAs.

allele' model. The initial problem presented by these 4 probes was the apparent presence of a homozygous P phenotype when hybridized to the backcrossed P/C *x* C/C DNAs. As an example, we will discuss the observed patterns and strategy used for interpreting and scoring Rcf 516. A similar approach for classification was applied to Rcf's 372, 432, and 483 for their final characterization.

When Rcf 516 was first hybridized to the parental/F1 blot, a single easily scored polymorphism was scored with both *Eco*RI and *Hin*dIII digested DNAs (data not shown). However, when Rcf 516 was probed against P/C *x* C/C DNAs digested with the restriction enzyme *Eco*RI, the initial scoring assigned an upper homozygous P (Fig. 3A, left 2 lanes) and heterozygous F phenotypes (Fig. 3A, right 2 lanes). Subsequent examination of this autoradiogram revealed the uppermost band to be of higher molecular weight than the uppermost P band on its parental/F1 blot. Also, the band initially scored as P was significantly more intense than the putative heterozygous F bands and more closely resembled a poorly separated doublet. The *Hin*dIII blot for Rcf 516 further complicated this analysis by producing homozygous C bands for individuals where putative F phenotypes were originally scored on the *Eco*RI digest (Fig. 3B, right 2 lanes). By applying the heterozygous C allele model for scoring these progeny, we see that in the *Eco*RI digest (Fig. 3A) the C^2 allele migrates just below P giving an F doublet, and C^1 migrates as the single lower band. Therefore, a C individual has inherited the C^1C^2 alleles while the F has inherited the PC^2 alleles. These segregation patterns would be expected by either PC^2 *x* C^1C^2 or PC^1 *x* C^1C^2 matings for this locus (Table 2). This scoring strategy was further supported following the observance of genetic linkage between Rcf's 516 and 483 in our F2 and P/C x C/C and P/C *x* P/P backcross mapping populations.

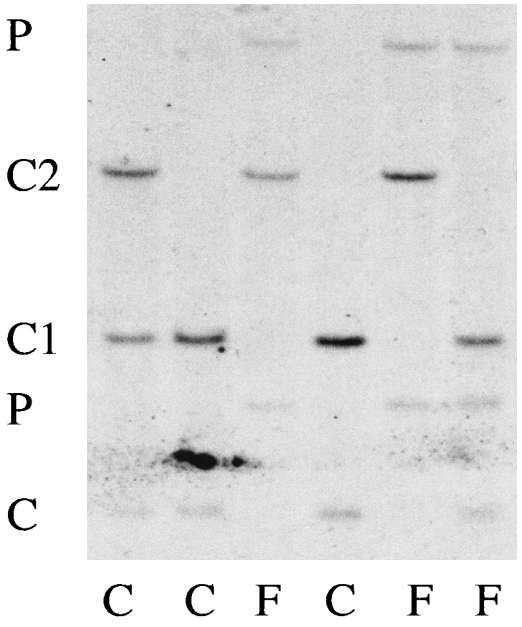

Fig. 4. Rcf 427 class I complex polymorphic hybridization pattern of P/C *x* C/C backcross DNAs digested with the *Hin*dIII restriction enzyme. These segregation patterns would be expected from the PC^1 *x* C^1C^2 backcross mating.

Another interesting pattern was produced by Rcf 427. This probe, which detects two similarly scored polymorphic bands against *Hin*dIII digested backcross and F2

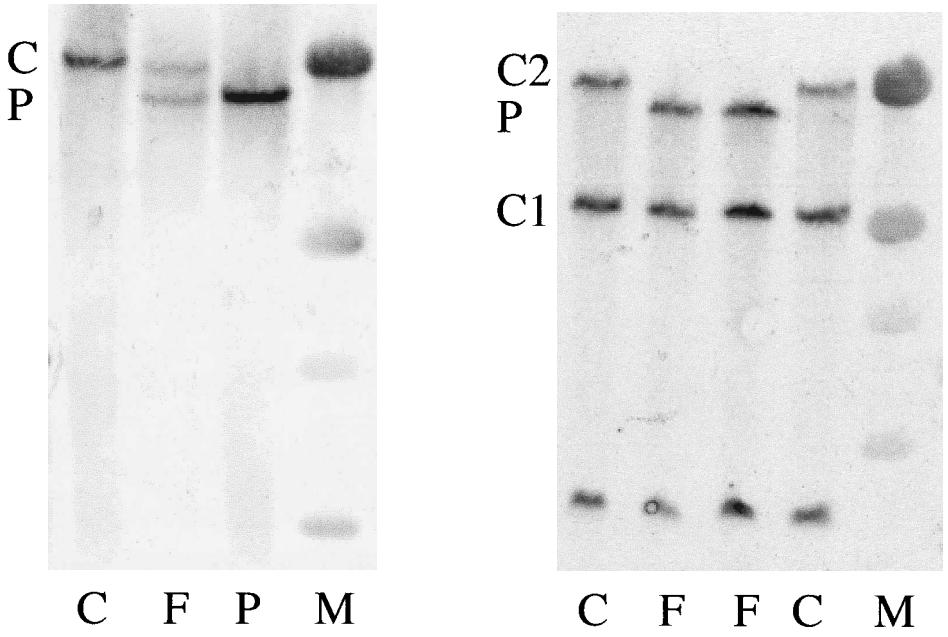

Fig. 5. Rcf 438 class I complex polymorphic hybridization pattern of P/C *x* C/C backcross DNAs digested with the *Pst*I restriction enzyme (10 μg DNA/lane). Lane M is the size standard lambda *Hin*dIII. Panel A, genomic restriction pattern for Rcf 438 when probed against parental/F1 DNAs in the order C108, F1, p50. Panel B, Rcf 438 polymorphic pattern when probed against P/C *x* C/C backcross DNAs. The observed C^1P and C^1C^2 segregants are predicted by the PC^2 *x* C^1C^1 backcross mating.

progeny, reveals an extra allele only on the upper polymorphic band when screened against the P/C *x* C/C progeny (Fig. 4). This additional allele was initially thought to segregate independently, but it is now more plausible that this band is a second C specific allele. With the P/C *x* C/C progeny, the upper set of polymorphic bands provides evidence for 4 classes of segregating allelic pairs: C^1C^1, C^1C^2, C^1P, and C^2P. Table 2 shows that these classes are expected from the PC^1 *x* C^1C^2 backcross mating. Because of the faintness of the lower polymorphic bands, they could not be reliably scored.

One remaining probe, Rcf 438, was the most difficult of the class I CPPs to interpret. When first screened against the parental/F1 blot, it identified polymorphic bands migrating with the top fragment of the lambda marker as shown in Fig. 5A. When later screened against backcrossed *Pst*I digested P/C *x* C/C DNAs, the C and P bands expected from the parental/F1 blot were present; however, hybrid F phenotypes were not initially observed. Instead, all progeny showed an extra smaller molecular weight band. If scored as an extra C band, two classes of progeny comprising the F hybrid and C doublet could be distinguished. We believe that the uppermost fragment is a C^2 allele as shown on the parental/F1 blot, the band migrating below this is the P band, and the smallest or fastest migrating band is a C^1 allele (Fig. 5B). As shown in Table 2, the backcross PC^2 *x* C^1C^1 predicts the two classes C^1P and C^1C^2 that result from this suggested scoring.

DISCUSSION

Successful genetic linkage analysis requires the availability of polymorphic mapping markers. A recent screening of our *Bombyx* early follicular cDNA library identified 19 codominant restriction fragment length polymorphisms (RFLPs) in three sets of mapping populations comprised of reciprocal backcross and F2 progeny. In addition, because different banding patterns were observed between the F2 and backcross progeny for several individual Rcf hybridization probes, autoradiographic patterns were classified into two major classes: those producing 'simple' polymorphic patterns (SPPs), and those producing 'complex' polymorphic patterns (CPPs). The complex polymorphic patterns were further subdivided into two subclasses: (i) polymorphic bands containing an extra fragment not immediately recognizable as the expected segregated C or P phenotype but likely to have originated as an additional allele in the C108 parent, and (ii) multiple banding patterns that were scored reliably as a C, P, or F genotype.

The two parental strains, C108 and p50, were originally chosen for this project because of known differences in morphological characteristics and developmental traits, a long history of being easily reared under laboratory conditions, and for having been inbred over many generations to substantial genetic homozygosity. Nevertheless, the backcross populations, which were expected to score only 2 predicted phenotypes (the recurrent parental or the F1 hybrid), revealed that the parental stock C108 harbored

a previously undetected level of genetic variation at a number of loci. In this study, approximately 30% of our isolated Rcf probes detected a second C–specific allele limited to the P/C *x* C/C backcross DNAs. Characterization of the resulting hybridization patterns for these probes identified two allelic C phenotypes, C^1 and C^2, and the single P phenotype. Theorized P/C *x* C/C allelic crosses are presented in Table 2 that predict the progeny classes observed in offspring segregating the extra C–specific allele. Although results suggestive of genetic variation in the parental C108 strain were found because the allelic variation was limited to one strain, C108, and not shared by homologous loci in p50, we have demonstrated that these probes remain viable genetic mapping markers and that their usefulness for linkage mapping has not been reduced between these two laboratory populations. If scoring difficulties do arise, it is possible to further confirm the analysis by comparing the data to the reciprocal backcross progeny or to the F2 data.

The presence of the extra allele initially confounded the scoring of the P/C *x* C/C backcross DNAs; however, several strategies were applied to resolve their final characterization. For example, several probes detected the extra allele when screened against a specific set of digested DNAs (e.g., *Eco*RI) but lacked the extra allele when scored against differently digested DNAs (e.g., *Pst*I) of the same individuals. Because scoring for a given probe is expected to be identical between differently digested DNAs for a single individual, it was possible to confirm the extra allele doublet as a C phenotype and the progeny segregating both a C and a P phenotype as the F phenotype. In many cases this scoring was further confirmed when linkage to other markers was detected with both backcross and F2 populations. A second approach for identifying the extra allele was by fragment size. The approximate size or molecular weight of the polymorphic bands was scored from the parental/F1 blot and then used as a comparison to identify and characterize the new allele from the expected P and C bands.

The allelic variation found in the C108 genome was an unexpected and interesting find. One explanation may be a result of its status as a 'genetically improved' strain, which has a history of artificial selection and hybrid matings between progeny possessing economic and morphological traits of interest. Although these artificial matings alone may suffice to explain introduced variation, it is also plausible that the mating process increased fitness for a number of desired traits or phenotypes favoring their selection for future matings. Because many of the economic traits of interest, such as silk yield and quality, growth rate, number of eggs per laying, hatching rate, and survival, are most likely under quantitative or polygenic control, a large number of new alleles may have been introduced to the stocks. From population genetic theory, inbreeding, coupled with migration of new alleles, small population size, and viability selection, will maintain allelic variation at a given locus regardless of linkage (Hartl and Clark, 1989).

It is also possible that spontaneously occurring genetic events may be responsible for producing or maintaining the levels of genetic variation observed in the C108 strain. For example, the *Bombyx* genome contains a large number of retroposon elements such as Bm1 and Bm2 (Adams et al., 1986; Eichbush, 1995) and retrotransposable elements such as mag (Michaille et al., 1990; Garel et al., 1994), K1.4 (Ueda et al., 1986; Tamura et al., 1993), and BmC1 (Ogura et al., 1994). Not only is it very likely that these transposable elements contribute to the levels of polymorphism detected between the two parental strains, it is also plausible that they are in part a mechanism for introducing the detected levels of allelic variation found specific to this C108 mapping population.

It has been demonstrated by Regier et al. (1978) that differences in electrophoretic mobilities of silkworm chorion proteins result from differences in their primary structure. Previously, while examining the electrophoretic mobilities of numerous chorion proteins, Goldsmith and Basehoar (1978) reported evidence suggestive of low levels of heterozygosity specific to C108, a laboratory race largely believed at the time to be homozygously inbred. At that time, however, the authors were unable to rule out the possibility that the responsible genes had been either deleted or were not expressed. To the best of our knowledge, we do not believe that the topic of maintained heterozygosity in inbred silkworm races has been revisited since this information was last presented. We believe that we now provide genetic evidence for maintained polymorphic loci in the inbred race C108. We also believe that this may explain the protein level variants first described by Goldsmith and Basehoar (1978) although we do not have formal evidence linking our presently described Rcf probes to the chorion protein variants described previously. Work addressing this goal is presently in progress.

REFERENCES

Adams, D. S., Eickbush, T. H., Herrera, R. J., and Lizardi, P. J. (1986) A highly reiterated family of transcribed oligo(A)-terminated, interspersed DNA elements in the genome of *Bombyx mori*. Journal of Molecular Biology **187**, 465–478.

Botstein, D., White, R. L., Scholnick, M. H., and Davis, R. W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics **32**, 314–331.

Doira, H., Fujii, H., Kawaguchi, Y., Kihara, H., and Banno, Y. (1992) Genetical stocks and mutations of *Bombyx mori*: Important genetic resources. Isseido Press, Fukuoka, Japan.

Eickbush, T. H. (1995) Mobile elements of lepidopteran genomes. In: Molecular Model Systems in the Lepidoptera (eds.: M. R. Goldsmith and A. S. Wilkins), pp. 77–105. New York: Cambridge University Press.

Ellegren, H., Chowdhary, B. P., Johansson, M., Marklund, L.,

Fredholm, M., Gustavsson, I., and Andersson, L. (1994) A primary linkage map of the porcine genome reveals a low rate of genetic recombination. Genetics **137**, 1089–1100.

Erlich, H. A., Horn, G. T., Saiki, R. T., Scharf, S. J., Mullis, K. B., and Bugawan, T. (1986) Genetic analysis using enzymatic amplification of specific genomic sequences. In: DNA Probes, Current Communications in Molecular Biology (ed.: L. S. Lerman), pp. 107–112. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Fan, Y. S., Davis, L. M., and Shows, T. B. (1990) Mapping small DNA sequences by fluorescence *in situ* hybridization directly on banded metaphase chromosomes. Proceedings of the National Academy of Sciences, USA **87**, 6223–6227.

Garel, A., Nony, P., and Prudhomme, J. C. (1994) Structural features of mag, a gypsy–like retrotransposon of *Bombyx mori*, with unusual short terminal repeats. Genetica **93**, 125–137.

Gautreau, D., Zetlan, S. R., Mazur, G. D., and Goldsmith, M. R. (1993) A subtle defect underlies altered lamellar orientation in the Gr^{16} chorion mutant of *Bombyx mori*. Developmental Biology **157**, 60–72.

Goldsmith, M. R., and Basehoar, G. (1978) Organization of the chorion genes of *Bombyx mori*, a multigene family. I. Evidence for linkage to chromosome 2. Genetics **90**, 291–310.

Goldsmith, M. R., and Shi, J. (1993) A molecular map for the silkworm: Constructing new links between basic and applied research. In: Silk Polymers: Materials Science and Biotechnology (eds.: David Kaplan, W. Wade Adams, Barry Fermer, and Christopher Viney), pp.45–58. American Chemical Society, symposium series 544.

Hartl, D., and Clark, A. G. (1989) Principles of Population Genetics, Second Edition. Sunderland, Massachusetts: Sinauer Associates, Inc.

Heslop–Harrison J. S., Leitch, A. R., Schwarzacher, T., and Anamthawat–Jonsson, K. (1990) Detection and characterization of 1B/1R translocations in hexaploid wheat. Heredity **65**, 385–392.

Lander, E. S., and Botstein, D. (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121**, 185–199.

Litt, M., and Luty, J. A. (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. American Journal of Human Genetics **44**, 397–401.

Michaille, J. J., Mathavan, S., Gaillard, J., and Garel, A. (1990) The complete sequence of Mag, a new retrotransposon in *Bombyx mori*. Nucleic Acids Research **18**, 674.

Nadel, M. R., Goldsmith, M. R., Goplerud, J., and Kafatos, F. C. (1980) Specific protein synthesis in cellular differentiation. V. A secretory defect of chorion formation in the Gr^{col} mutant of *Bombyx mori*. Developmental Biology **75**, 41–58.

Ogura, T., Okano, K., Tsuchida, K., Miyajima, N., Tanaka, H., Takada, N., Izumi, S., Tomino, S., and Maekawa, H. (1994) A defective non-LTR retrotransposon is dispersed throughout the genome of the silkworm, *Bombyx mori*. Chromosoma **103**, 311–323.

Pollard–Knight, D., Read, C. A., Downes, M. J., Howard, L. A., Leadbetter, M. R., Pheby, S. A., McNaughton, E., Syms, A., and Brady, M. A. W. (1990) Nonradioactive nucleic acid detection by enhanced chemiluminescence using probes directly labeled with horseradish peroxidase. Analytical Biochemistry **185**, 84–89.

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

Regier, J. C., Kafatos, F. C., Goodfliesh, R., and Hood, L. (1978) Silkworm chorion proteins: sequence analysis of the products of a multigene family. Proceedings of the National Academy of Sciences, USA **75**, 390–394.

Saiki, R. K., Scharf, S., Falona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. (1985) Enzymatic amplification of beta–globin genomic sequences and restriction site analysis for diagnosis of sickle–cell anemia. Science **230**, 1350–1354.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) Primer–directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**, 487–491.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning, A Laboratory Manual. Second edition. Cold Spring Harbor Laboratory Press, NY. pp 2.60–63.

Sato, Y., and Yamashita, O. (1991) Structure and expression of a gene coding for egg–specific protein in the silkworm, *Bombyx mori*. Insect Biochemistry **21**, 495–503.

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

Southern, E. M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology **98**, 503–517.

Tamura, T., Kanda, T., Yukuhiro, K., Yasukouchi, N., Hinomoto, N., Shimizu, K., Kosegawa, E., and Okajima, T. (1993) RFLP patterns detected by the transposons, mag and k1.4, in silkworm races. Gamma Field Symposia no. 32, Institute of Radiation Breeding, National Institutes Annual Reports, Ministry of Agriculture, Forestry, and Fisheries, Japan.

Tautz, D. (1989) Hypervariability of simple sequences as general source for polymorphic DNA markers. Nucleic Acids Research **17**, 6463–6471.

Trask, B. J. (1991) Fluorescence *in situ* hybridization: Applications in cytogenetics and gene mapping. Trends in Genetics **7**, 149–155.

Ueda, H., Mizuno, S., and Shimura, K. (1986) Transposable genetic element found in the 5'-flanking region of the fibroin H-chain gene in a genomic clone from the silkworm *Bombyx mori*. Journal of Molecular Biology **190**, 319–327.

Weber, J. L., and May, P. E. (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. American Journal of Human Genetics **44**, 388–396.

Welsh, J., and McClelland, M. (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research **18**, 7213–7218.

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A., and Tingey, S. V. (1991) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research **18**, 6531–6535.