

DNA fingerprinting using AFLP markers to search for markers associated with yield attributes in the silkworm, *Bombyx mori*

Authors: Gaviria, Duverney A., Aguilar, Enrique, Serrano, Herman J., and Alegria, Alvaro H.

Source: Journal of Insect Science, 6(15) : 1-10

Published By: Entomological Society of America

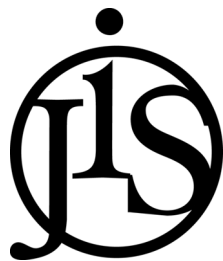
URL: https://doi.org/10.1673/2006_06_15.1

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.



DNA fingerprinting using AFLP markers to search for markers associated with yield attributes in the silkworm, *Bombyx mori*

Duverney A. Gaviria, Enrique Aguilar, Herman J. Serrano and Alvaro H. Alegria

Center for Molecular Biology and Biotechnology, Universidad Tecnológica de Pereira, Risaralda, Colombia

Abstract

This study was carried out on 11 Chinese and 12 Japanese silkworm strains maintained by the Center for the Technological Development of Sericulture (CDTS) germplasm bank, located in Pereira, Colombia. The goals were to determine the genetic population structure of the two groups and the association between molecular markers (AFLPs) and important productivity characters. Group analysis showed the separation of the strains according to their geographic origin. The molecular markers and the productivity characters were correlated by multiple variance analysis. The analysis permitted the identification of molecular markers associated with the cocoon weight or the shell weight separately. Some markers were associated with both characters.

Correspondence: duberney82@netscape.net, dugaar@utp.edu.co, serrano@utp.edu.co, AlvaroHAlegriaSA@netscape.net

Received: 15.2.2005 | Accepted: 24.1.2006 | Published: 1.9.2006

Copyright: Creative Commons Attribution 2.5 (<http://creativecommons.org/licenses/by/2.5>)

ISSN: 1536-2442 | Volume 2006, Number 15

Cite this paper as:

Gaviria DA, Aguilar E, Serrano HJ, Alegria AH. 2006. DNA fingerprinting using AFLP markers to search for markers associated with yield attributes in the silkworm, *Bombyx mori*. 10pp. *Journal of Insect Science* 6:15, available online: insectscience.org/6.15

Introduction

Molecular markers possess several advantages over morphological characteristics, since they are not modified by the environment and can be identified at any developmental stage. Several techniques have been used in an attempt to identify molecular polymorphisms in the silkworm, including enzyme markers such as hemolymph amylases in tropical and temperate zones (Abraham et al. 1992; Goldsmith 1995). Using RAPDs Nagaragu et al. (1995), separated 13 silkworm strains in two groups, six with diapause and seven without diapause. Several additional groups markers were also identified. Using a probe (BKm -2, GATA66 TA) the two groups were clearly separated, and sex-specific bands were also identified. Damodar et al. (1999) analyzed 28 microsatellite loci (SSR) and identified a marker, sat 211, with specific alleles for strains with and without diapause. A very polymorphic marker, sat 2763, was also identified with a total of 17 alleles and very high heterozygosity (90%). These results could be very useful in the study of geographical relationships and in the analysis of wild strains or those in danger of extinction. Similar results with these same strains were obtained by Reddy et al. (1999), using other techniques.

Cocoon weight and shell weight are the main traits evaluated for productivity in sericulture and have been used for more than half a century. Cocoon weight is an important commercial characteristic used to determine approximately the amount of raw silk that can be obtained. Shell weight gives a better measure, but cannot be determined in commercial cultures because it requires damaging the cocoon. The difference between the two measures is the weight of the pupa. Using AFLPs the correlation found between these two parameters in this study was 70%. It was concluded that at least four genes affect shell weight and eight affect cocoon weight. One of the genes involved has been tentatively located in the sex chromosomes, both for cocoon weight and shell weight. However no definitive location has been reported so far (He et al, 1998).

A large number of morphobiochemical markers have been linked to molecular markers (Shi et al. 1995; Yasakuchi 1998; Nagaragu 2000). However, little has been done to understand the genetics of productivity traits, except in the analysis of heritability and combinatorial ability (Tazima 1964; Shibukawa et al. 1986; Rao et al. 1991; Chaterjee et

al. 1993). DNA markers closely linked to a characteristic of interest could be used to select for that trait in different ways such as: 1) a segregating population of F₂ crosses, using an analysis of bulked segregants (Michelmore *et al.*1991); 2) recombinant endogamic lines (Reiter et al, 1992); 3) near isogenic lines produced by improvement programs (Martin et al. 1991). However, within a intraspecific context, the quantitative trait loci mapping of many important agronomical characteristics require using very informative markers that allow the discrimination of a large number of bands (Tan et al. 2001). Of course, the affectivity of such marker-assisted selection depends on the precision of the phenotypic classification of the trait of interest and the degree of linkage between the marker and the trait (Nagaragu et al. 2002). The advances in statistical modeling and mathematics applied to genetic processes are also helping to optimize the use of the molecular and phenotypic information in selection programs (Bovenhuis et al. 1997; Hoeschele et al. 1997). When the characteristics of a population prevent the use of the mentioned techniques, alternatives such as statistical models including multiple regression analysis, or discriminating factor analysis, may be used to identify a few markers strongly linked to one or several productivity components. These can be later examined by progeny analysis and productivity pattern segregation, leading to the identification of the genes associated with this character in particular (Chaterjee et al. 2003 (#b4-2006_06_15)).

The present study was carried first to determine the degree of diversity and the existing relationships between the different accessions belonging to the silkworm germplasm bank of the Center for the Technological Development of Sericulture (CDTS), using amplified fragment length polymorphisms (AFLPs) (Vos et al. 1995). Then, an effort to identify molecular markers associated with productivity characteristics such as cocoon and shell weight was attempted, through a statistical approach based on multiple variance and simple correlation analysis.

Materials and Methods

Insects

The characterization by molecular methods was carried out on 23 silkworm lines, 12 from Japanese (KOI, KO2, KO5, K10, K20, K30, K40, K522, SG3, SG2, NG, KNA) and 11 from Chinese origin (CA, CBS, CC, CJ, CGS, CHS, CLS, CTS, SC1, SC2, SC3) that belong to the germplasm bank of the Center for

the Technological Development of Sericulture (CDTS). They are kept in the farm “El Píamo”, Pereira, Risaralda, Colombia.

For the linkage analysis between productivity traits and molecular markers, 160 cocoons were studied, distributed equally in Japanese lines (4 of high and 4 of low productivity), and Chinese lines (4 of high and 4 of low productivity). The productivity values were based on statistical data obtained at the Center for about ten years. Each line was separated into males and females and five replicas were carried out with each line to homogenize variance, for a total of ten individuals for each line (Table 1)

Quantitative data

Cocoon weight was determined in each case. The weight of the shell was also measured by eliminating the pupa in each cocoon. An electronic balance was used and the data obtained to two significant figures.

DNA extractions

DNA was obtained from two different tissues (silk gland and pupa) using a different protocol for each, Suzuki (1972) and Chatterjee (1993) respectively. In the extraction from the silk gland grinding of the tissue (0.04 g) was carried in 1 ml extraction buffer (NaCl 50mM, EDTA 50mM, SDS 1%), at 50°C with 10 µl of K-proteinase (100 mg/ml). For extraction of pupae, 1 ml of buffer was used (glucose 200 mM, EDTA 50 mM, Tris HCl 100 mM, SDS 0.5%, pH: 9.0) plus 2 µl K-proteinase (100 mg/ml) at 55°C. The tissue was incubated and ground periodically until there were no visible fragments; the samples were then centrifuged for 5 min at 14,000 rpm. 10 µl of potassium acetate (8M) were added to the extract before centrifugation. The aqueous phase was recovered and re-extracted with an equal volume of a chlorophorm- isoamyl alcohol mixture

(24:1) for 5 min. After centrifugation at 14,000 rpm (Eppendorf 5417 R centrifuge) for 5 min. two volumes of absolute ethanol and 1/10 sodium acetate 3M were added to the separated aqueous phase to precipitate nucleic acids. After 5 min of centrifugation the supernatant was discarded and the pellet was air-dried. The pellet was re-suspended in TE buffer (Tris HCl 10mM, EDTA 1mM, ph: 8.0) and incubated one hour at 37°C after the addition of 10 µl pancreatic RNase A (100 µg/ml). The DNA was re-extracted with phenol-chloroform-isoamyl alcohol and ethanol precipitated as before. The DNA was re-suspended in TE buffer, quantified by optical absorbance and examined by electrophoresis in an agarose gel (0.8%) stained with ethidium bromide. To further examine the DNA quality, samples were digested with the restriction endonuclease *MspI*. When the sample did not digest properly, the DNA was submitted to a new round of extraction as indicated above.

Primer selection for AFLPs

AFLPs reactions (Vos et al. 1995) were carried out using the “AFLP Analysis System I” kit from Invitrogen (Cat 10544-013), according to the manufacturer recommendations. Selective amplification by PCR was performed using a three nucleotide extension for both primers. To characterize the AFLPs four primer combinations were selected which gave a high number of non-ambiguous polymorphisms after 10 combinations were tried on four silkworm lines. For the identification of markers associated with productivity traits, five combinations were selected on a random basis.

PCR products were mixed with 10 µl loading mixture for sequencing (formamide 98%, EDTA 10

Table 1. Productivity data

Name	Geographic origin	Productivity level	Mean cocoon weight Males ±SD	Mean cocoon weight Females ±SD	Mean shell weight Males ±SD	Mean shell weight Females ±SD
K 02	Japanese	High	1.62 ±0.11	1.88±0.12	0.37 ±0.05	0.37±0.03
K 20	Japanese	High	1.66±0.07	1.90±0.30	0.40± 0.03	0.38±0.04
K 30	Japanese	High	1.66 ±0.15	2.10±0.14	0.40 ±0.03	0.42±0.02
SG 2	Japanese	High	1.64±0.07	2.01±0.13	0.41 ±0.019	0.45±0.039
K 01	Japanese	Low	1.53±0.04	1.97±0.14	0.33 ±0.019	0.37±0.032
K 03	Japanese	Low	1.50± 0.08	1.95±0.06	0.32 ±0.05	0.36±0.03
K 05	Japanese	Low	1.48 ±0.14	1.84±0.08	0.32 ±0.04	0.33±0.03
K 10	Japanese	Low	1.53 ±0.09	2.01±0.14	0.34±0.02	0.38±0.03
SC 1	Chinese	High	1.77±0.05	2.14±0.17	0.42± 0.06	0.46±0.04
SC 2	Chinese	High	1.66±0.19	2.16±0.16	0.38 ±0.05	0.45±0.04
SC 3	Chinese	High	1.80 ±0.10	2.35±0.26	0.44± 0.03	0.48±0.07
CLS	Chinese	High	1.77±0.15	2.21±0.18	0.42 ±0.04	0.45±0.05
CA	Chinese	Low	1.61 ±0.03	2.11±0.004	0.37±0.0006	0.34±0.0009
CC	Chinese	Low	1.43±0.08	1.96±0.06	0.33±0.02	0.37±0.01
CBS	Chinese	Low	1.37±0.10	1.85±0.13	0.33 ±0.03	0.38±0.02
CGS	Chinese	Low	1.56±0.13	1.98±0.10	0.36±0.03	0.41±0.04

mM, xylencyanol 0.025%, bromophenol blue 0.025%), heated for 3 minutes at 90°C and then cooled on ice. 5 µl were then loaded on a 6% sequencing polyacrylamide gel and run at 90 W during 90 minutes. The gel was silver stained using the “Silver sequenceTM DNA Sequencing System” kit from Promega (www.promega.com), according to the manufacture recommendations. Permanent records were obtained by digital photography, transferring the images to TIFF format using a Nikon D’100 camera.

Statistical Analysis

DNA bands were taken as a genetic locus, assuming that allele markers corresponding to different loci do not co-migrate and that each locus can be taken as a two allele system (Lynch and Milligan 1994). Data matrices were built for the presence (1) and the absence (0) in each of the loci. Zero (0) and one (1) matrices were obtained with the program for image analysis for molecular data “Gene Profiler ver 4.05” from Scanalytics, (www.scanalytics.com).

Similarity and group analysis

The determination of the relationships between the 23 silkworm lines was based on matrices built with the similarity indices of Nei and Li (1979). Dissimilarity values were analyzed using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) and a dendrogram was built. A principal coordinate analysis (Gower 1966) was also carried out to show the distribution in three dimensions. Both graphic representations made use of the statistical package NTSYS-pc, ver 2.01 (Rohlf 1992).

Descriptive statistics and diversity

The homocedasticity assumption was revised using a Bartlett test (Sokal and Rohlf 1981) before each analysis was run. The programs POPGENE, ver 1.32 (Yeh et al. 1997) and TFPGA (Tools for Population Genetic Analysis), ver 1.3 (Miller 1999) were used to calculate, under the Hardy-Weinberg equilibrium assumption, the following parameters: 1) Recessive allele frequency and, accordingly, that of the dominant allele using the Taylor expansion method (Lynch and Milligan 1994); 2) Polymorphic loci percentage; 3) Non biased heterocigosity mean for all loci (Nei 1978), neutrality test, and exact test of population differentiation (Raymond and Rousset 1995). These analyses were carried out both with the raw data and data to which an analysis restriction had been applied to avoid the bias in bands with a frequency higher than $1-(3/N)$

(Lynch and Milligan 1991). Data discussed in Results and Discussion refer to the corrected data.

Population structure

The program AMOVA-PREP (Miller 1998) was used to transform the set of dominant data for the Molecular Analysis of Variance program (AMOVA, ver 1.55. Excofier et al. 1992). The data were analyzed as two populations and also as two subpopulations inside each population, in order to determine the part contributed by each one of the subdivisions to the total variance component. The analyses were carried out using both the raw data and data corrected with the factor $1-(3/N)$.

Genetic distance

Genetic distance was determined for the populations and subpopulations using the TFPGA program and the calculation described by Nei (1978).

Statistical analysis for the association of molecular markers and productivity traits

In order to determine which markers have a significant influence on productivity, multiple variance analyses were carried out for all samples. Shell weight first and then cocoon weights were considered as dependent variables and line, sex and marker as explanatory variables. Marker action on the character was determined by a lineal correlation analysis. All of these analyses were carried out using the Statgraphics plus, ver 5.1, software (www.statgraphics.com). Shell and cocoon weight were selected as study variables since they are the main parameters to measure productivity. Each variance analysis was repeated, eliminating each time the explanatory variable with the less significant contribution until all the remaining variables made a significant contribution to productivity. This model succession allows detecting those factors with global importance, that is, those that are significant not only when compared with the dependent variable but also in the presence of other determining factors. These analyses were made for the markers grouped according to the primer combinations. Significance was set at 95%. The proposed method allows not only for the identification of groups of important factors, but is also theoretically more satisfactory than multiple regression analysis. The variance analysis works for discrete variables without violating assumptions or assuming normality, which would be wrong in this case (Sokal and Rohlf 1981).

The problem of variance uniformity was studied among the obtained markers, since the model does not fit this assumption strictly. In this respect, markers with a presence or absence less than 5%, were eliminated to eliminate false positives, and for the rest, variances were given according to which the significance of the marker in the variance analysis was determined. Using a regression and identifying which markers show a positive slope, those with a positive influence were defined. This regression implies only that the slope is positive or negative by least squares, without additional statistical assumptions.

Results and Discussion

Similarity and group analysis

To carry out the analysis, groups were defined from the dendrogram clusters. Accordingly, two principal groups were set up: Chinese and Japanese lines. From these, subgroups were tentatively defined, based again on the dendrogram clusters. The significance in the selection of the groups was statistically proved using the population differentiated exact proof (Raymond and Rousset 1995). The structuring level corresponding to

subgroups was found non significant but the group level was significant.

Individual analysis of each of the used combinations gave similar results and presented correlation indices for the data of the different matrices higher than 90%, indicating that the number of combinations used was sufficient for exploring the diversity present in the silkworm lines studied. Since the combinations gave highly concordant results, it is assumed that the whole genome was being analyzed (Barker et al. 1999; Rouppe et al. 1997 (#b27-2006_06_15); Zhu et al. 1998). The dendrograms corresponding to each of the primer combinations gave similar groupings (Figures 1 and 2).

The level of polymorphism for each combination was in the range 65.6–75.8%, although a large number of restriction fragments (247 bands) were generated in individual samples by AFLPs (Table 2). For analysis only those fragments that were polymorphic and registered without ambiguity were used. This correction was made according to the criterion established by Lynch and Milligan 1994, and finally a total of 168 bands were analyzed

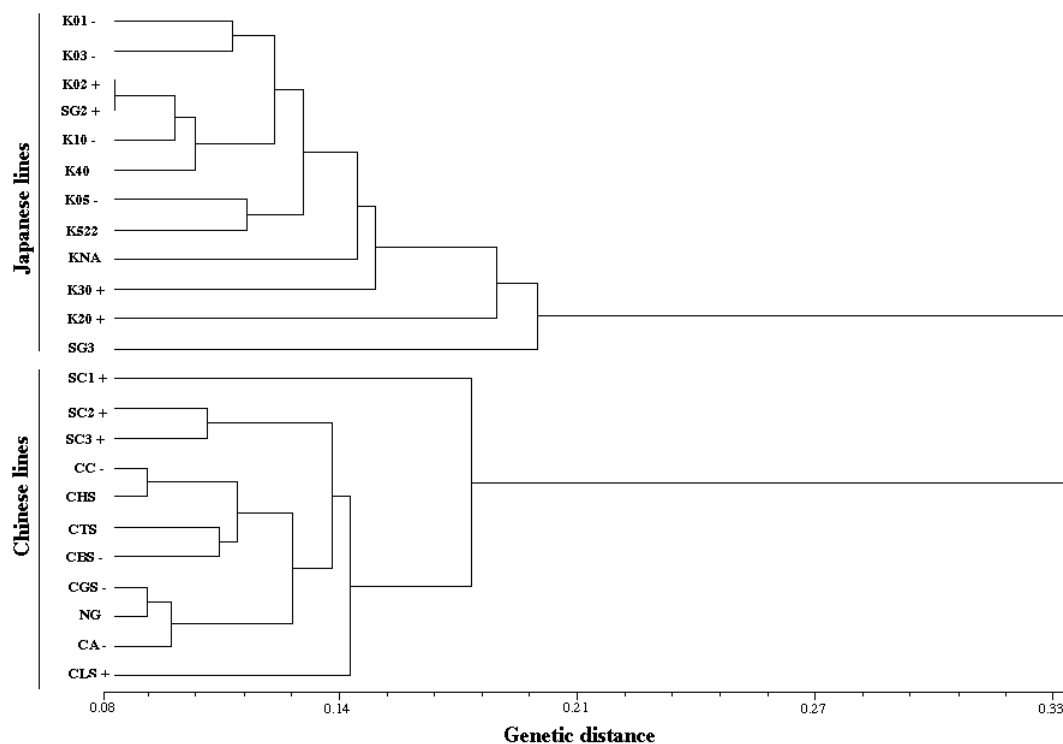


Figure 1. Dendrogram obtained by UPGMA indicating the relationships between the 23 CDTS silkworm lines. (168 AFLP bands). Signs (-) and (+) indicate lines with low or high productivity respectively.

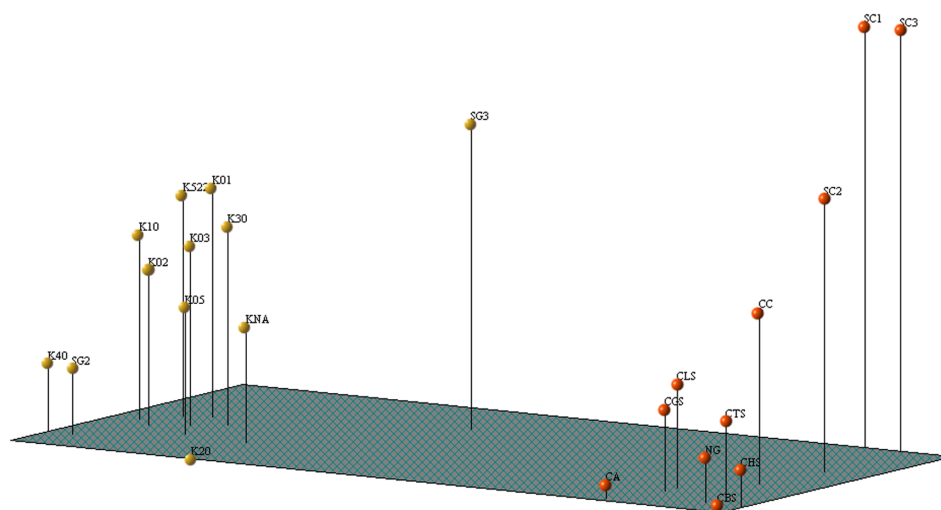


Figure 2. Principal coordinate analysis using a simple pairing matrix, for the data (168 AFLP bands). Yellow color are Japanese lines, red color are Chinese lines.

with an average polymorphism of 90.4% (Table 2). The rest of the analyses were carried out with the corrected data only.

Table 2. Used AFLPs combinations

Combination	Loci number	% Polymorphic loci
E-ACG/M-CAC	72	72.2
E-ACG/M-CAT	46	69.5
E-ACG/M-CAA	67	65.6
E-AGG/M-CAC	62	75.8
Total loci	247	54.6
Total loci (Lynch y Milligan Correction)	168	90.4

The dendrogram generated by UPGMA analysis of the 168 bands (four primer combinations) resolved the 23 silkworm lines into two very well defined groups corresponding to the Chinese and Japanese lines (Figure 1), with a distance of 0.22 for the Japanese lines and 0.18 for the Chinese lines. The analysis identified all the individuals.

Descriptive statistics and diversity

For analysis, the following parameters were calculated for each combination: polymorphic loci percentage, heterozygosity, genetic flow, genetic distance, and G_{st} values (Fisher 1954). The data obtained using the four primer combinations show both in the dendrogram and in the principal coordinate analysis that the Japanese branch (distance 0.22) is more heterogeneous than the Chinese branch (distance 0.18) (Figure 1 and 2). This observation was confirmed by the heterozygosity data, which were always higher for the Japanese lines. The estimate of the degree of population differentiation, given by the G_{st} value for each loci shows a high degree of differentiation (Table 3)

Population structure

The variance components were obtained by AMOVA (AMOVA, ver 1.55, Excofier *et al.*1992)

Table 3. Descriptive statistical analysis for each primer combination and their sum.

Combination	Analysis level	Heterozygosity	Genetic flow (Nm)	Genetic distance	G_{st}
E-ACG/M-CAC	Population 1	0.2199	0.6550	0.3599	0.4329
	Population 2	0.1333			
	Population total	0.3044			
E-ACG/M-CAT	Population 1	0.2435	0.5957	0.3476	0.4563
	Population 2	0.0738			
	Population total	0.2794			
E-ACG/M-CAA	Population 1	0.1838	0.9350	0.2378	0.3484
	Population 2	0.1684			
	Population total	0.2659			
E-AGG/M-CAC	Population 1	0.2632	0.3486	0.3114	0.3486
	Population 2	0.1667			
	Population total	0.3108			
Total	Population 1	0.2446	0.7712	0.3117	0.3933
	Population 2	0.1474			
	Population total	0.2983			
Total (Lynch and Milligan)	Population 1	0.2585	0.7068	0.5052	0.4143
	Population 2	0.2253			
	Population total	0.3785			

using the corrected samples. Differences were analyzed between the evaluated races and inside each one. This analysis revealed that the variance components between populations were somewhat higher than inside each population, indicating some differentiation between the two races (Table 4). This result was further confirmed by a comparison of paired populations using Fst (Fisher 1954). Variance analysis was also carried out for the subpopulations defined by conglomerate analysis and group behavior was not observed, indicating that in these populations no subdivision levels with a lesser order exist (Table 5).

Table 4. Corrected AMOVA: Variance component analysis between populations and inside each silkworm population.

Variation source	df	MS	VC	%Total	P	B
Variance between populations	1	353	28.97	58.57%	<0.001	<0.001
Variance inside populations	21	20.5	20.5	41.43%		

Table 5. Corrected AMOVA: Variance component analysis between populations and subpopulations and inside populations and subpopulations.

Variation source	df	MS	VC	%Total	P	B
V. between populations	1	353.039	28.033	56.67%		
V. between populations inside subpopulations	2	30.464	1.986	4.02%	<0.001	<0.001
V. inside subpopulations	19	19.448	19.448	39.31%		

Variation source	df	MS	VC	%Total	P	B
V. between subpopulations	3	137.989	20.866	51.76%		
V. inside subpopulations	19	19.448	19.448	48.24%	<0.001	<0.001

Variation source	df	MS	VC	%Total	P	B
V. between populations	1	353.039	28.971	58.57%	<0.001	<0.001
V. inside populations	21	20.497	20.497	41.43%		

Association between AFLP markers and productivity characteristics

The number of markers generated from the analysis of 16 silkworm lines with five AFLP primer combinations was between 45 and 71, with an average of 60 AFLP markers distributed between 30–500 bp (Table 2).

To identify relationships (association, not linkage) between the DNA markers and the productivity characteristics, variance analysis for two characteristics against 302 molecular markers was made. Examples of the results are presented in

Tables 6 and 7. The analysis revealed the high degree of significance of sex and silkworm line in the explanation of productivity. The markers selected in each of the primer combinations for each of the productivity parameters are listed in Table 8.

Table 6. Variance analysis for cocoon weight/ square summation type III for the combination E-AAC M-CTA

Marker	Square error	df	MSE	F-Value	P-Value
Band 21	0.177945	1	0.177945	9.63	0.0023
Band 26	0.092459	1	0.092459	5	0.0269
Band 27	0.090538	1	0.090538	4.9	0.0285
Band 29	0.083203	1	0.083203	4.5	0.0357
Band 37	0.152334	1	0.152334	8.24	0.0047
LINE	2.19096	15	0.146064	7.9	0
SEX	7.02013	1	7.02013	379.8	0

Table 7. Variance analysis for cocoon shell weight/ square summation type III for the combination E-AAC M-CTA

Marker	Square error	df	MSE	F-Value	P-Value
Band 20	0.008208	1	0.008208	6.38	0.0127
Band 21	0.011105	1	0.011105	8.64	0.0039
Band 37	0.006182	1	0.006182	4.81	0.03
Band 40	0.007785	1	0.007785	6.05	0.0151
Band 41	0.008076	1	0.008076	6.28	0.0134
LINE	0.243386	15	0.016226	12.62	0
SEX	0.042757	1	0.042757	33.25	0

Marker identification was obtained through a process of successive elimination, one by one, of the markers giving non-significant results in the variance analysis. The number of markers used to carry out the selection varied from 2 to 8 for the combinations used for the evaluated characteristics. As seen in Table 8, common bands can be identified. In all 7 common bands were identified for the analyzed characteristics, one band for each primer combination on the average (eg. band 48 and 49 in the combination E-ACG/M-CAA) while others are specific for one of the characteristics only (eg. band 6 and 38 in the combination E-ACG/M-CAA). The type of contribution of each one of the markers for the evaluated traits was determined through a correlation analysis between each band and the particular trait. The number of selected markers for each individual primer combination is shown in Table 8. The number is variable and does not seem to have a direct relationship with the total number

Table 8. Identified markers for each of the evaluated characteristics

Combination	Band number	Band for cocoon weight	Band for cocoon shell weight	Number of comon bands	Total of selected bands
E-ACG/M-CAA	68	6, 16, 18, 48, 49	2, 38, 48, 49, 64	2	10
E-ACC/M-CAC	69	2, 9, 41, 63, 65, 66	2, 5, 10, 17, 18, 19, 20, 55	1	14
E-ACA/M-CAG	71	33, 36, 38, 65	4, 65	1	6
E-AAG/M-CAT	49	19, 26, 29, 35	29, 48	1	6
E-AAC/M-CTA	45	21, 26, 27, 29, 37	20, 21, 37, 40, 41	2	10
Total	302	24	22		46

of bands for each combination. The total number of bands identified for each evaluated character was rather similar.

Line genetic relationships

The present study shows that AFLPs can be used successfully in the silkworm to reveal DNA polymorphisms useful as genetic markers. The results demonstrate that this is a valuable technique for the discovery of genetic variation. In general, the information obtained by each primer combination seems to be the same. It can be concluded, therefore, that the analysis is covering the whole of the genome. The study allowed the identification of specific bands for each of the examined races, which can be very useful for identification of intellectual property rights and determining the purity of germoplasm both for pure lines and commercial hybrids.

The conglomerates formed by this technique identified perfectly each of the lines according to their geographic origin, clearly indicating their Chinese or Japanese origin; this relationship is probably due to the process of endogamy by which the lines are maintained. In spite of the fact that the observed diversity was low (<10%), the results suggest that the lines that originated the Japanese race were originally less homogeneous than those for the Chinese race. In the CDTS collection, two Japanese lines were shown to be almost identical. Although more of the Chinese lines are similar, none showed this low degree of diversity. The geographic varieties of silkworm share a large genomic component and this degree of similarity makes it necessary to study and evaluate a higher number of characteristics in each line to determine if they should be maintained and replicated as a stock for improvement, in spite of their high degree of similarity. It would also be important to further evaluate lines such as SG3, K20 and K30 in the Japanese group and, above all, the SC1 Chinese line, which showed AFLP patterns that locate them distantly in their cluster.

The variance components suggest that it is necessary to broaden the genetic base, introducing new lines from the same or different geographic origin, if improvement programs are to be implemented to obtain higher productivity hybrids.

Association between molecular markers and productivity components

The silkworm genome was found to be an excellent subject for analysis in fingerprinting programs for

the identification of markers associated with productivity characteristics. Cocoon and shell weight are two very important economic traits that researchers have tried to physically identify in the genome for more than half a century. The present study tried to define markers present in a group of individuals with contrasting productivity characteristics without having an extreme distribution.

Different statistical approaches were applied to get an approximation to the association between molecular markers and productivity attributes. The fact that the information for each identifier (marker) is its presence or absence, prevent the use of regression inference because of the normality assumption. On the other hand, a multiple variance analysis does not make this requirement to its variables and allows conclusions about the significance of the contribution of each variable. The only point to discuss is the assumption of the uniformity of the variances. It can be argued, in this respect, that extreme cases in which the presence proportion was lower or higher than 5% were avoided and the method was used without evaluation of the individual variances. Multiple variance analysis allowed the identification of 46 marker bands, using a small number of AFLP primer combinations showing a significant association with the two productivity components studied.

Such markers might help in programs of marker-assisted selection. The success of such selection programs depends exclusively on the degree of linkage between the markers and relevant loci such as quantitative trait loci (QTLs). These require additional studies for specific markers selected from this analysis.

The analysis showed the importance of sex in this type of approximation. This characteristic becomes one that must be always evaluated, to avoid the identification of characteristics associated with sex and that can be falsely associated with productivity. Females are bigger and eat more than males. However these qualities do not result in higher productivity, since the additional energy is used to make eggs, not silk.

Correlation analysis between bands and productivity characteristics showed both positive and negative associations. In some cases markers for both cocoon weight and shell weight were observed. The contribution of these common

markers was positive for both in some cases and positive for one and negative for the other in other cases. This result shows that a higher cocoon weight does not necessarily reflect a higher shell percentage. These characteristics seem to be controlled by different genes, confirming the evidence reported by He et al. (1998).

Acknowledgements

The authors thank the support of COLCIENCIAS (project 1110-12-11603) and the Universidad Tecnológica de Pereira, Colombia.

References

- Abraham EG, Nagaraju J, Datta RK. 1992. Biochemical studies of amylases in the silkworm, *Bombyx mori* L.: Comparative analysis in the diapausing and nondiapausing strains. *Insect Biochemistry and Molecular Biology* 22: 687-873.
- Barker JHA, Matthes M, Arnold GM, Edwards KJ, Åhman I, Larsson S, Karp A. 1999. Characterisation of genetic diversity in potential biomass willows (*Salix* spp.) by RAPD and AFLP analyses. *Genome* 42: 173-183.
- Bovenhuis H, van Arendonk JAM, Davis G, Elsen JM, Haley CS, Hill WG, Baret PV, Hetzel DJS, Nicholas FW. 1997. Detection and mapping of quantitative trait loci in farm animals. *Livestock Production Science* 52: 135-144.
- Chatterjee SN, Mohandas TP. 2003. Identification of ISSR markers associated with productivity traits in silkworm, *Bombyx mori* L. *Genome* 46: 439-447.
- Chatterjee SN, Rao CGP, Chatterjee GK, Aswath SK, Patnaik AK. 1993. Correlation between yield and biochemical parameters in the mulberry silkworm, *Bombyx mori* L. *Theoretical and Applied Genetics* 87: 385-391.
- Damodar K, Abraham EG, Nagaraju J. 1999. Microsatellite in the silkworm, *Bombyx mori*: Abundance, polymorphism, and strain characterization. *Genome* 42: 1057-1065.
- Excofier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.
- Fisher RA. 1954. *Statistical methods for research workers*. 12th ed. Oliverand Boyd, Edinburgh.
- Geneprofiler 4.05 for Windows. Genotyping and DNA analysis software. Scanalytics, Inc.
- Goldsmith MR. 1995. Genetics of the silkworm: revisiting an ancient model system. In: Molecular Model Systems in the Lepidoptera (Goldsmith MR and Wilkins AS, editors), Cambridge University Press, New York. pp 21-76.
- Gower JC. 1966. Some distance properties of latent roots and vector methods used in multivariate analysis. *Biometrika* 53: 325-338.
- Hoeschele I, Ulimari P, Grignola FE, Zangh Q, Cage KM. 1997. Advances in statistical methods to map quantitative trait loci in outbred populations. *Genetics* 147: 1445-1457.
- He KR, Huang JH, Zhu XR. 1998. The loci of genes located on sex-chromosome of silkworm (*Bombyx mori*) responsible for the cocoon weight and cocoon shell weight and their effect analysis. *Sericologia* 38: 43-49.
- Lynch M, Milligan BG. 1994. Analysis of population genetic structure with RAPD markers. *Mol Ecol* 3: 91-99.
- Martin GB, Williams JGK, Tankley SD. 1991. Rapid identification of markers linked to a *Pseudomonas* resistance gene in tomato by using random primers and near isogenic lines. *Proceedings of the National Academy of Science USA* 88: 2336-2340.
- Michelmore RW, Paran I, Kesseli RV. 1991. Identification of markers linked to disease resistance gene bulked segregant analysis: a rapid method to detect markers in specific genomic region using segregation population. *Proceedings of the National Academy of Science USA* 88: 9828-9832.
- Miller MP. 1998. AMOVA-PREP 1.01. A program for the preparation of ANOVA input files from dominant marker raw data. Computer software distributed by the author.
- Miller MP. 1999. Tools for population genetic analysis (TFPGA) 1.3: A Windows program for the analysis of allozymes and molecular population genetic data. Computer software distributed by the author.
- Nagaraju J. 2000. Recent advances in molecular genetics of the silk moth, *Bombyx mori*. *Current Science* 78: 151-161.
- Nagaraju J, Goldsmith MR. 2002. Silkworm genomics-progress and prospects. *Current Science* 83: 415-425.
- Nagaraju J, Sharma A, Sethuraman BN, Rao GV, Singh L. 1995. DNA fingerprinting in the silkworm *Bombyx mori* using banded krait minor satellite DNA-derived probe. *Electrophoresis* 16: 1639-1642.
- Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.
- Nei M, Li WH. 1979. Mathematical-model for studying genetic-variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States Of America* 76: 5269-5273.
- Rao GS, Das SK, Das NK. 1991. Genetic divergence among fifteen multivoltine genetic stocks of silkworm (*Bombyx mori* L.). *Indian Journal of Sericulture* 30: 72-74.
- Reddy KD, Nagaraju J, Abraham EG. 1999. Genetic characterization of the silkworm *Bombyx mori* by simple sequence repeat (SSR)-anchored PCR. *Heredity* 83: 681-687.

- Reiter RS, Williams JGK, Feldeman KA, Rafalski JA, Tingey SV. 1992. Global and Local Genome Mapping in *Arabidopsis thaliana* by Using Recombinant Inbred Lines and Random Amplified Polymorphic DNAs. *Proceedings of the National Academy of Science USA* 89: 1477-1481.
- Rohlf FJ. 1992. *NTSYS-PC numerical taxonomy and multivariate analysis system*, Version 2.01. Owner's manual. Exeter publication, Setauket, NY, USA.
- Roupe van der Voort JNA, van Zandvoort P, van Eck HJ, Folkertsma RT, Hutten RCB, Draaistra J, Gommers FJ, Jacobsen E, Helder J, Bakker J. 1997. Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. *Molecular and General Genetics* 255: 438-447.
- Raymond M, Rousset F. 1995. An exact test for population differentiation. *Evolution* 49: 1280-1283.
- Shi J, Heckel DG, Goldsmith MR. 1995. A genetic linkage map for the domesticated silkworm, *Bombyx mori*, based on restriction fragment length polymorphisms. *Genetical Research* 66: 109-126.
- Shibukawa A, Eguchi R, Shimajaki A, Ichida M. 1986. Studies on the cocoon quality of the silkworm stock culture of the sericulture experiment station. *Technical Bulletin of Sericultural Experiment Station* 129: 1-34.
- Sokal R R, Rohlf F J. (1981). *Biometry*. 2nd ed. Freeman, NY.
- STATGRAPHICS 7.0. *Statistical graphics systems*. Magnugistics Inc., Rockville, MD, 1993.
- Suzuki Y, Gage LP, Brown DD. 1972. The genes for fibroin in *Bombyx mori*. *Journal of Molecular Biology* 70: 637-649.
- Tan YD, Wan C, Zhu Y, Lu C, Xiang Z, Deng HW. 2001. An amplified fragment length polymorphism map of the silkworm. *Genetics* 157: 1277-1284.
- Tazima Y. 1964. *The genetics of silkworm*. Logos Press, New York, NY
- Vos P, Hogers R, Bleeker M, Reijans M, van deer lee T, Hornes M, Frijters A, Pot J, Pelemen J, Kuiper M, Zabeau M. 1995. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407-4414.
- Yasakuchi Y. 1998. A dense genetic map of silkworm, *Bombyx mori*, covering all chromosomes based on 1018 nuclear markers. *Genetics* 150: 1513-1525.
- Yeh FC, Yang RC, Boyle TB, Ye ZH, Mao JX. 1997. *POPGENE, the user-friendly shareware for population genetic analysis*. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton, Alta.
- Zhu J, Gale MD, Quarrie S, Jackson MT, Bryan GJ. 1998. AFLP markers for the study of rice biodiversity. *Theoretical and Applied Genetics* 96: 602-611.