World Journal of Engineering Research and Technology
WJERT
www.wjert.org
SJIF Impact Factor: 5.218

YOLK PROTEIN GENOME ORGANIZATION OF SILKWORM *Bombyx mori* (L) AND THEIR ASSOCIATION WITH EST BASED RFLP MARKER FOR SILKWORM CONSERVATION

Vishnu Priya S.*, Somasundaram P. and Kamble C. K.

Central Sericultural Germplasm Resources Centre, Central Silk Board, Hosur-635109, Tamil Nadu, India.

Article Received on 11/05/2019 **Article Revised on 02/06/2019** **Article Accepted on 23/06/2019**

***Corresponding Author**
Vishnu Priya S.
Central Sericultural Germplasm Resources Centre, Central Silk Board, Hosur-635109, Tamil Nadu, India.

ABSTRACT

Germplasm characterization is an important for conservation and utilization of genetic resources. DNA markers allow researchers to identify accessions at the taxonomic level, assess the relative diversity within and among species, and locate diverse accessions for breeding purposes. An investigation was carried out in the present study to identify the yolk protein gene related restriction fragment length polymorphism using expression sequence tag (EST) were used to differentiate high and low yolk yielding silkworm strains. Three restriction enzymes viz., Hinf I, EcoR I, Tag I were used in this study for digesting the EST-PCR product of yolk protein DNA of ten different silkworm races possessing contrasting silk for yolk yield. Among the selected races Kollegal Jawan , MY-1, Pure Mysore and Hosa Mysore were shown a polymorphic and total of six cluster group were observed on EST –marker. The RFLP-PCR technique is flexible with regard to different environmental conditions as DNA is more stable. Using this study different silkworm strains will be characterized through restriction pattern and grouped together based on silk yield. This technique is rapid and inexpensive, enabling silkworm to be characterized in short periods.

KEYWORDS: DNA Markers, EST, RFLP, *Bombyx Mori.*

Highlights

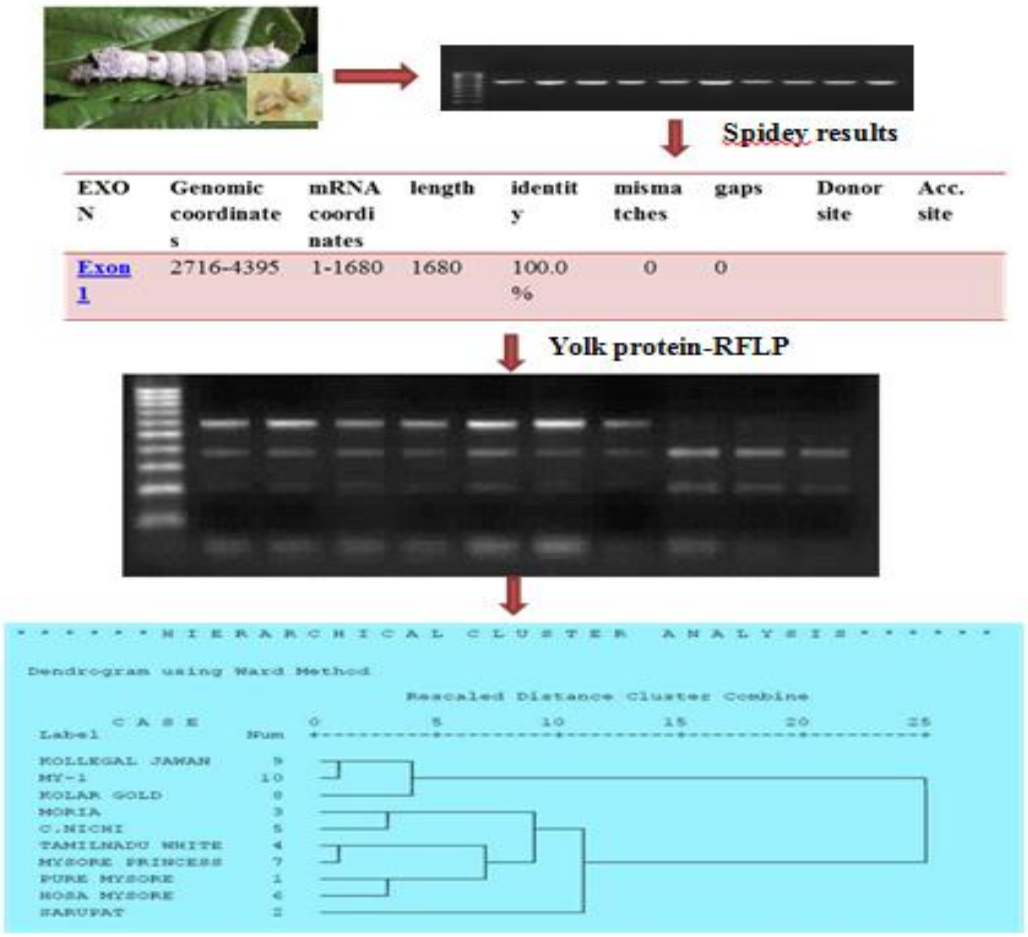
1. To know the gene expression in yolk protein in silkworm *Bombyx Mori*.

2. To identify the high yolk yielding strains among different germplasm stocks.

3. To select for better breeding for researcher and former.

4. Improve the economic status of the country.

Genome organization of yolk protein in silkworm *Bombyx mori* using EST_RFLP markers



Graphical abstract

1. INTRODUCTION

Germplasm characterization is important for conservation and utilization of genetic resources (Razi et al., 2019). DNA markers allow researchers to identify accessions at the taxonomic level, assess the relative diversity within and among species, and locate diverse accessions for breeding purposes. Since genetic differentiation is often correlated with geographic isolation, it may be appropriate to analyze accessions that represent a wide range of geographic regions, in order to estimate the genetic diversity within the breeding stock. Attempts have been made to understand the genetics of yield potential in *Bombyx mori* (Chatterjee et al., 1993b; Datta,

1984; Jingade et al., 2011). Similarly, linkage analysis of clownfish genome were reported by Casas et al., 2018.

Traditional method of characterization and estimation of genetic variability is not completely reliable due to inherent limitations. cDNA libraries prepared from various tissues and at different developmental stages that cover almost entire set of *Bombyx* genes, comprising 35,000 expressed sequence tags (EST) from 36 cDNA libraries. ESTs are popularly used in full genome sequencing and mapping program is underway for a number of organisms for identifying active genes thus helping in identification of diagnostic markers. EST is a very simple tool for genome analysis. The genetic evaluation and characterization based on EST-SSR markers in *Lagerstroemia* were studied by Ye et al., 2019. In addition, Ukoskit et al., 2019 analyzed the linkage mapping of sugar cane breeds used EST_SSR markers. Okuda et al., 2002 reported linkage analysis of maternal EST cDNA clones covering all twenty-eight chromosomes in the silkworm, *Bombyx mori*. ESTs provide researchers with a quick and inexpensive route for discovering new genes, for obtaining data on gene expression, regulation, and for constructing the genome maps. EST database can also be used to develop simple sequence repeat (SSR) markers for various species (Mita et al.,2003; Cheng et al., 2004). Linkage analysis of EST cDNA clones using a RFLP in the silkworm were reported by Hara et al., 2001. The EST-associated microsatellites were also reported in *Castanopsis sieboldii* Aoki et al.,2019. Recently, DNA based techniques are being extensively used to characterize crop plants and live stocks for understanding their genetic diversity and breed improvement Programme. Genome analysis in Indian silkworms, the application of PCR based RAPD and also DNA fingerprinting with minisatellite probes have been taken up to study the DNA profiling of silkworm genotypes (Nagaraja and Nagaraju,1995; Nagaraju et al., 1995, 2001; Mohandas et al.,2004).

Yolk proteins are mainly expressed during the stress conditions and are also involved in the normal physiological functions. The yolk protein which plays a role in molecular chaperones assembly, which bind and stabilize proteins that is in non-native conformation. Interaction with their unstable protein conformations prevent the formation of large protein aggregates and facilitate normal protein folding, membrane translocation and the degradation and removal of damaged proteins. Yolk protein gene families are highly conserved, but they can also be extremely sensitive and be rapid indicators of cellular stress when examined at the mRNA level. Furthermore, yolks are known stress related proteins, and are involved in

thermo tolerance mechanism of silkworm *Bombyx mori*. The present study is aimed to find the genetic polymorphism of ten selected silkworm races of *Bombyx mori* through gene specific marker EST molecular tools followed by RFLP.

2. MATERIALS AND METHODS

2.1 Selection of silkworm races

Ten silkworm races were selected based on geographical origin are shown in (Table-1).

2.2 DNA extraction and quantification

The DNA was extracted and quantified based on reference (Mohandas et al., 2004).

2.3 DNA sequences and database retrieval

The yolk protein gene sequences were retrieved and analyzed by BLAST search. EST database were retrieved from Tokyo University Silk-base (http://www.ab.a.u-tokyo.ac.jp/silkbase).

2.4 Identification of exons and introns

The exons and introns of yolk protein gene were identified by spidey programme (a tool for mRNA and genomic alignments) (Wheelan et al., 2001). The windows were examined carefully to get the percent identity per exon, the number of gaps per exon, the overall percent identity, the percent coverage of the mRNA, presence of an aligning or non-aligning poly-A tail, number of splice donor sites, the presence or absence of splice donor and acceptor sites for each exon, and the occurrence of an mRNA that has a 5' or 3' end (or both) that does not align to the genomic sequence. For the overall percent identity and percent length coverage that were above the user-defined cutoffs, a summary of report and a text alignment showing identities and mismatches were recorded.

Table. 1: Silkworm races and their geographical origin.

S. No	Acc. No	Race Name	Donor	Origin	Class	Parentage
1	BMI-0001	Pure Mysore	CSR,MYS	KAR	O(BCU)	O
2	BMI-0002	Sarupat	CSR,MYS	ASM	O(OB)	O
3	BMI-0003	Moria	CSR,MYS	ASM	O(OB)	O
4	BMI-0004	Tamilnadu White	CSR,MYS	TNU	E(OB)	PMxJ122
5	BME-005	C.Nichi	CSR,MYS	JAP	O(OB)	O
6	BMI-0006	Hosa Mysore	CSR,MYS	KAR	E(BCU)	PM.A4e
7	BMI-0007	Mysore Princess	CSR,MYS	KAR	E(OB)	(PM.NN6D) (Hosho. Shungetsu)
8	BMI-0008	Kolar Gold	CSR,MYS	KAR	E(OB)	(PM.NN6D) (Hosho. Shungetsu)
9	BMI-0009	Kollegal Jawan	CSR,MYS	KAR	E(OB)	(PM.NN6D) (Hosho. Shungetsu)
10	BMI-0010	My-1	CSR,MYS	KAR	E(BCU)	PM.Nistari

2.5 Primer designing

The gene specific primers were designed using on online software programmed (http://frodo.wi.mit.edu/cgi-bin/primer3).

2.6 EST-PCR analysis and amplification of product

The reaction was performed in an MJ research thermal cycler, PTC 200, by using 20µl of reaction mixture consist of 50-100ng of DNA, Primer (Forward & Reverse-100ng/µl),10X Taq buffer with MgCl2,MgCl2 (25mM),dNTPs (2mM),Taq polymerase (3000 unit/ml),Distilled water (nuclease free). The PCR was scheduled at 96°C for 3 min, 96°C for 30 sec and followed by annealing at 56°C for 30 sec and 72°C for 2 minutes for extension and followed by repeat the step for 45 cycles at final extension at 72°C for 7 min. The PCR amplified products were resoled on 1.5% of agarose gel with 1X TAE and electrophoresis was carried out 80V with standard molecular weight.

2.7 Polymorphism identification by restriction digestions

EST –PCR products were analyzed by RFLP using a three restriction enzyme viz, , ECOR I, Taq I, Hinf I and the enzymes were diluted with Tris –Hcl (p H-7.2-8), 1 mM DTT, 10 mM Mgcl2 and Nacl.

3. RESULTS

3.1 Retrieval of gene bank sequences

The genomic organization of yolk protein gene, the cDNA was searched from *Bombyx mori* EST database was retrieved from Tokyo University Silk –base (http://www.ab.a.u-tokyo.ac.jp/silkbase) (Mita et al., 2003). The complete cDNA sequence of this was searched from the database and it was downloaded. This available cDNA of yolk protein gene was retrieved from www.ncbi.nih.nlm.gov. (Genbank sequences are shown in Table 2). Alignment of these sequences allowed accurate determining the nucleotide sequences and reconstituting the full-length cDNA. In parallel EST sequences were subjected to BLAST search to find out similar sequences present in other eukaryotic organisms (Reddy et al., 1999a).

Table 2. *Bombyx mori* gene for egg-specific protein, complete cds - GenBank: D12521.1

1 atgaagacta tatacgcatt gctgtgtctg acgctggtgc agagcatctc ctgcagcatc

61 ttcatgacga aacaacacag tcaggatgac atcattcaac accctctgga ctatgtcgaa

121 cagcagatcc atcagcagaa acaaaaacta cagaagcaaa ccttgaacaa gaggagccac

181 cagcactctg attcggattc ggattcggcg tctcgtgcgg cggcgtcaca ttcagcctcc

241 caatcttcga gttcacaaag ctccagttca caagaggatg aagctaagca cgtgcaagat
301 aagatgaacg tgaaacacca ctcgccggtg tattctgtca ttatgaaact caagaaagaa
361 gttgatatca atcacggcga ttccgtcgtt tggaagaata tagaaatggc ctccggccct
421 aactcgccgg ttcagacaga gcaagatatt gaggatattt tcggtgactc cctcaagacg
481 tgggatcatt tcactgacga tgcaaagaaa aataccttcc acgacgctat cagtgaaact
541 caaagggaaa acaatgagga cttccaccta aacgctaccg agctgctcaa gaaacaccaa
601 taccctgtag aagaacacac ggtcgccacc gacgacggtt accatttgac tgtcctgcgc
661 attccaccca cgcaccaaac cagggacgat aagaagaagc cggtcgcgct tctaatgcac
721 ggcttactcg gaagtgctga cgactggtta ctgatgggtc ccagtaagtc actcgcttac
781 atgctctctg acgccggcta cgacgtatgg ctgggtaatg ttcgtggaaa taaatattcc
841 cgctctcacg tcagcaagca cccagcactc aatgacttct ggaagtttag caatgacgag
901 atcgctcttc acgacttacc cgctataatt gaccacgttt tggatattag cggccaagag
961 agacttcatt acataggcca ttctcaaggc gcgaccacct tcttcgccct gatgtctgaa
1021 cagccttcgt acaacgaaaa gatcgtttcg atgcacgcgt tgtctcctat tgtttacatg
1081 aattatgtac gctcgcccct cttccgtatg atcgcgccca cgagcaagtt ctaccagtat
1141 atacacgacc aagtcggtca cggagccttc gaacctggca aacacttgat cgaaaccttc
1201 ggcggcgccg cgtgcaggga aaaactcggt tgcaggcacg tctgtaacaa cttgaactac
1261 gtcatatcgg gtatcaacgt ttacaaccag gatgcggata tagttcccgt tgtgatggcc
1321 cacctgccag ccggcacatc cgcccgggtc atgaaacaat acggtcagaa tgtggcgtcg
1381 cacgatttta gaaaatacaa ctacggagcc gaaaccaaca tgaaagtgta cggcgcttcg
1441 gaaccaccta gttacgactt gagcaaagtc agcgcgcctg tcaatcttta ccacagccac
1501 gatgcctggt tggcccatcc caaggacgtg gagaaactcc aagaaaacct acctaatgtg
1561 aagcagtctt tcgaagttcc agagcaacaa cacttcacgg acctggactt ccaattctcg
1621 aagaaagccc ccgataccgt ataccagaaa ctgatggaaa acatgcagaa taactcataa

3.2 Genomic DNA isolation and Quantification

The ten multivoltine (MV) silk races of origin of different class and parentage were chosen because of their divorce cocoon morphology and quantitative variation. The genomic DNA was isolated from ten multivoltine races viz., Pure Mysore, Sarupat, Moria, Tamilnadu White, C.nichi, Hosa Mysore, Mysore Princess, Kolar gold, Kollegal Jawan and MY-1(Fig1).

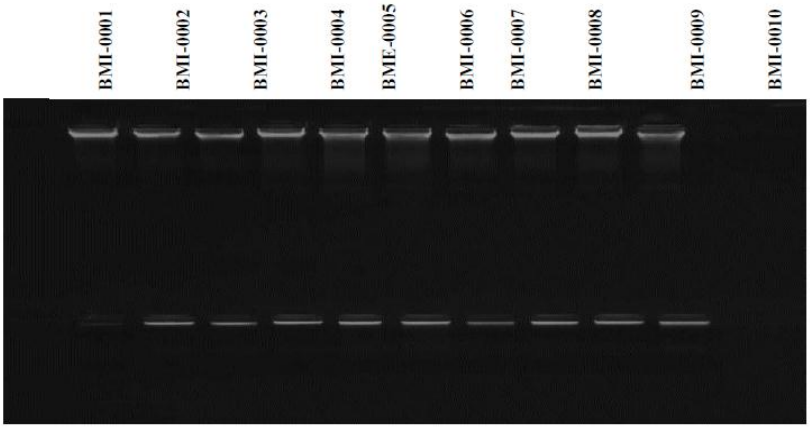


Figure. 1. Isolated DNA banding profiles of ten silk worm races resolved in 0.8% of agarose and its quantification.

3.3 EST and RFLP: Gene specific primers of yolk gene were designed and shown in Table 4. The DNA fragments representing yolk protein gene was amplified through PCR. The amplified products were electrophoresed on 1.5 % Agarose gel, it shown a single fragment yolk protein gene. The results indicated that all ten races possess the yolk protein gene (Fig 2). The total cDNA of each gene was compared with *Bombyx mori* genomic DNA contigs, to find out the coding region (exons) and noncoding region (introns) using spidey programme-mRNA and genomic alignment tool (Table 2). The results obtained through this programme were used to identify the number of exons and introns, splicing sites and poly (A) positions of yolk protein gene (SPIDEY results were shown in Table 3). The PCR product of yolk protein gene forward and reverse sequences were shown in Table 4. The product were further digested with three restriction enzymes and the digested products were analyzed in 1.5% Agarose gel (Fig 3, 4 & 5).Similar studies were reported, molecular genetic approach for identifying markers associated with yield traits in the silkworm, *Bombyx mori* using RFLP-STS primers (Reddy et al.,1999b).

3.4 Races and their classification

Among the ten races chosen four are original land races viz., Pure Mysore, Sarupat, Moria, C.Nichi. Tamilnadu white and Hosa Mysore shared their same parentage PM X J122 and PM.A4e and originated from Pure Mysore. Remaining races Mysore Princess, Kolor Gold, Kollegal Jawan, MY-1 also sharing the land race Pure Mysore as one of the parentage. The details of accession number, race, donor, origin, class and parentage are provided in (Table1).

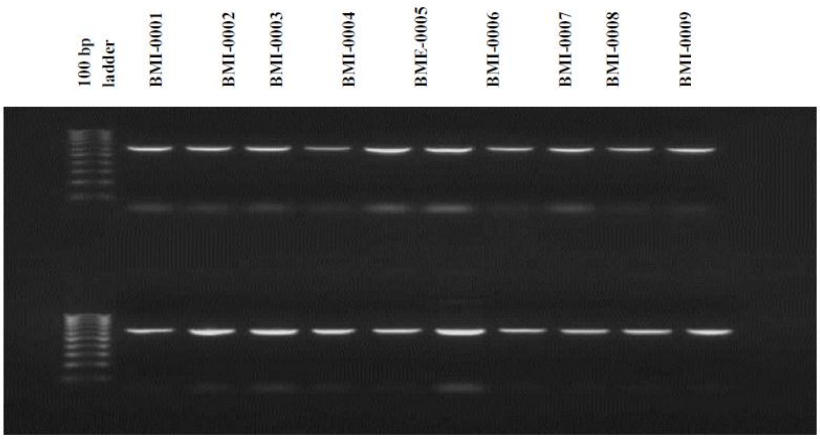


Figure. 2. Amplified Yolk protein gene EST- PCR banding profiles of genomic DNA of ten silk worm races as resolved on 1.5% of agarose gel stained with EtBr).

Table. 3. Spidey programme results.

Alignment is on plus strand of genomic sequence and on plus strand of mRNA sequence
mRNA coverage: 100%
Overall percent identity: 100.0%

2716							4395	
EXON	**Genomic coordinates**	**mRNA coordinates**	**length**	**identity**	**mismatches**	**gaps**	**Donor site**	**Acc. site**
Exon 1	2716-4395	1-1680	1680	100.0%	0	0		

Exon 1: 2716-4395 (genomic); 1-1680 (mRNA)

2716
TGTTGGCAACATGAAGACTATATACGCATTGCTGTGTCTGACGCTGGTGC

1 ATGAAGACTATATACGCATTGCTGTGTCTGACGCTGGTGC
M K T I Y A L L C L T L V

2756
AGAGCATCTCCTGCAGCATCTTCATGACGAAACAACACAGTCAGGATGAC

41 AGAGCATCTCCTGCAGCATCTTCATGACGAAACAACACAGTCAGGATGAC
Q S I S C S I F M T K Q H S Q D D

2806
ATCATTCAACACCCTCTGGACTATGTCGAACAGCAGATCCATCAGCAGAA

91 ATCATTCAACACCCTCTGGACTATGTCGAACAGCAGATCCATCAGCAGAA
I I Q H P L D Y V E Q Q I H Q Q K

2856
ACAAAAACTACAGAAGCAAACCTTGAACAAGAGGAGCCACCAGCACTCTG

141 ACAAAAACTACAGAAGCAAACCTTGAACAAGAGGAGCCACCAGCACTCTG
Q K L Q K Q T L N K R S H Q H S

2906
ATTCGGATTCGGATTCGGCGTCTCGTGCGGCGGCGTCACATTCAGCCTCC

191 ATTCGGATTCGGATTCGGCGTCTCGTGCGGCGGCGTCACATTCAGCCTCC
D S D S D S A S R A A A S H S A S

2956
CAATCTTCGAGTTCACAAAGCTCCAGTTCACAAGAGGATGAAGCTAAGCA

241 CAATCTTCGAGTTCACAAAGCTCCAGTTCACAAGAGGATGAAGCTAAGCA
Q S S S S Q S S S S Q E D E A K H

3006
CGTGCAAGATAAGATGAACGTGAAACACCACTCGCCGGTGTATTCTGTCA

291 CGTGCAAGATAAGATGAACGTGAAACACCACTCGCCGGTGTATTCTGTCA
V Q D K M N V K H H S P V Y S V

3056
TTATGAAACTCAAGAAAGAAGTTGATATCAATCACGGCGATTCCGTCGTT

341 TTATGAAACTCAAGAAAGAAGTTGATATCAATCACGGCGATTCCGTCGTT
I M K L K K E V D I N H G D S V V

3106
TGGAAGAATATAGAAATGGCCTCCGGCCCTAACTCGCCGGTTCAGACAGA

391 TGGAAGAATATAGAAATGGCCTCCGGCCCTAACTCGCCGGTTCAGACAGA
W K N I E M A S G P N S P V Q T E

3156
GCAAGATATTGAGGATATTTTCGGTGACTCCCTCAAGACGTGGGATCATT

441 GCAAGATATTGAGGATATTTTCGGTGACTCCCTCAAGACGTGGGATCATT
Q D I E D I F G D S L K T W D H

3206
TCACTGACGATGCAAAGAAAAATACCTTCCACGACGCTATCAGTGAAACT
491 TCACTGACGATGCAAAGAAAAATACCTTCCACGACGCTATCAGTGAAACT
F T D D A K K N T F H D A I S E T

3256
CAAAGGGAAAACAATGAGGACTTCCACCTAAACGCTACCGAGCTGCTCAA

541 CAAAGGGAAAACAATGAGGACTTCCACCTAAACGCTACCGAGCTGCTCAA
Q R E N N E D F H L N A T E L L K

3306
GAAACACCAATACCCTGTAGAAGAACACACGGTCGCCACCGACGACGGTT

591 GAAACACCAATACCCTGTAGAAGAACACACGGTCGCCACCGACGACGGTT
K H Q Y P V E E H T V A T D D G

3356
ACCATTTGACTGTCCTGCGCATTCCACCCACGCACCAAACCAGGGACGAT

641 ACCATTTGACTGTCCTGCGCATTCCACCCACGCACCAAACCAGGGACGAT
Y H L T V L R I P P T H Q T R D D

3406
AAGAAGAAGCCGGTCGCGCTTCTAATGCACGGCTTACTCGGAAGTGCTGA

691 AAGAAGAAGCCGGTCGCGCTTCTAATGCACGGCTTACTCGGAAGTGCTGA
K K K P V A L L M H G L L G S A D

3456
CGACTGGTTACTGATGGGTCCCAGTAAGTCACTCGCTTACATGCTCTCTG

741 CGACTGGTTACTGATGGGTCCCAGTAAGTCACTCGCTTACATGCTCTCTG
D W L L M G P S K S L A Y M L S

3506
ACGCCGGCTACGACGTATGGCTGGGTAATGTTCGTGGAAATAAATATTCC

791 ACGCCGGCTACGACGTATGGCTGGGTAATGTTCGTGGAAATAAATATTCC
D A G Y D V W L G N V R G N K Y S

3556
CGCTCTCACGTCAGCAAGCACCCAGCACTCAATGACTTCTGGAAGTTTAG

841 CGCTCTCACGTCAGCAAGCACCCAGCACTCAATGACTTCTGGAAGTTTAG
R S H V S K H P A L N D F W K F S

3606
CAATGACGAGATCGCTCTTCACGACTTACCCGCTATAATTGACCACGTTT

891 CAATGACGAGATCGCTCTTCACGACTTACCCGCTATAATTGACCACGTTT
N D E I A L H D L P A I I D H V
3656
TGGATATTAGCGGCCAAGAGAGACTTCATTACATAGGCCATTCTCAAGGC

941 TGGATATTAGCGGCCAAGAGAGACTTCATTACATAGGCCATTCTCAAGGC
L D I S G Q E R L H Y I G H S Q G

3706
GCGACCACCTTCTTCGCCCTGATGTCTGAACAGCCTTCGTACAACGAAAA

991 GCGACCACCTTCTTCGCCCTGATGTCTGAACAGCCTTCGTACAACGAAAA
A T T F F A L M S E Q P S Y N E K

3756
GATCGTTTCGATGCACGCGTTGTCTCCTATTGTTTACATGAATTATGTAC

1041 GATCGTTTCGATGCACGCGTTGTCTCCTATTGTTTACATGAATTATGTAC
I V S M H A L S P I V Y M N Y V

3806
GCTCGCCCCTCTTCCGTATGATCGCGCCCACGAGCAAGTTCTACCAGTAT

1091 GCTCGCCCCTCTTCCGTATGATCGCGCCCACGAGCAAGTTCTACCAGTAT
R S P L F R M I A P T S K F Y Q Y

3856
ATACACGACCAAGTCGGTCACGGAGCCTTCGAACCTGGCAAACACTTGAT

1141 ATACACGACCAAGTCGGTCACGGAGCCTTCGAACCTGGCAAACACTTGAT
I H D Q V G H G A F E P G K H L I

3906
CGAAACCTTCGGCGGCGCCGCGTGCAGGGAAAAACTCGGTTGCAGGCACG

1191 CGAAACCTTCGGCGGCGCCGCGTGCAGGGAAAAACTCGGTTGCAGGCACG
E T F G G A A C R E K L G C R H

3956
TCTGTAACAACTTGAACTACGTCATATCGGGTATCAACGTTTACAACCAG

1241 TCTGTAACAACTTGAACTACGTCATATCGGGTATCAACGTTTACAACCAG
V C N N L N Y V I S G I N V Y N Q

4006
GATGCGGATATAGTTCCCGTTGTGATGGCCCACCTGCCAGCCGGCACATC

1291 GATGCGGATATAGTTCCCGTTGTGATGGCCCACCTGCCAGCCGGCACATC
D A D I V P V V M A H L P A G T S

4056
CGCCCGGGTCATGAAACAATACGGTCAGAATGTGGCGTCGCACGATTTTA
1341 CGCCCGGGTCATGAAACAATACGGTCAGAATGTGGCGTCGCACGATTTTA
A R V M K Q Y G Q N V A S H D F

4106
GAAAATACAACTACGGAGCCGAAACCAACATGAAAGTGTACGGCGCTTCG

1391 GAAAATACAACTACGGAGCCGAAACCAACATGAAAGTGTACGGCGCTTCG
R K Y N Y G A E T N M K V Y G A S

4156
GAACCACCTAGTTACGACTTGAGCAAAGTCAGCGCGCCTGTCAATCTTTA

1441 GAACCACCTAGTTACGACTTGAGCAAAGTCAGCGCGCCTGTCAATCTTTA
E P P S Y D L S K V S A P V N L Y

4206
CCACAGCCACGATGCCTGGTTGGCCCATCCCAAGGACGTGGAGAAACTCC

1491 CCACAGCCACGATGCCTGGTTGGCCCATCCCAAGGACGTGGAGAAACTCC
H S H D A W L A H P K D V E K L

4256
AAGAAAACCTACCTAATGTGAAGCAGTCTTTCGAAGTTCCAGAGCAACAA

1541 AAGAAAACCTACCTAATGTGAAGCAGTCTTTCGAAGTTCCAGAGCAACAA
Q E N L P N V K Q S F E V P E Q Q

4306
CACTTCACGGACCTGGACTTCCAATTCTCGAAGAAAGCCCCCGATACCGT

1591 CACTTCACGGACCTGGACTTCCAATTCTCGAAGAAAGCCCCCGATACCGT
H F T D L D F Q F S K K A P D T V

4356
ATACCAGAAACTGATGGAAAACATGCAGAATAACTCATAAATAATTACGT

1641 ATACCAGAAACTGATGGAAAACATGCAGAATAACTCATAA
Y Q K L M E N M Q N N S *

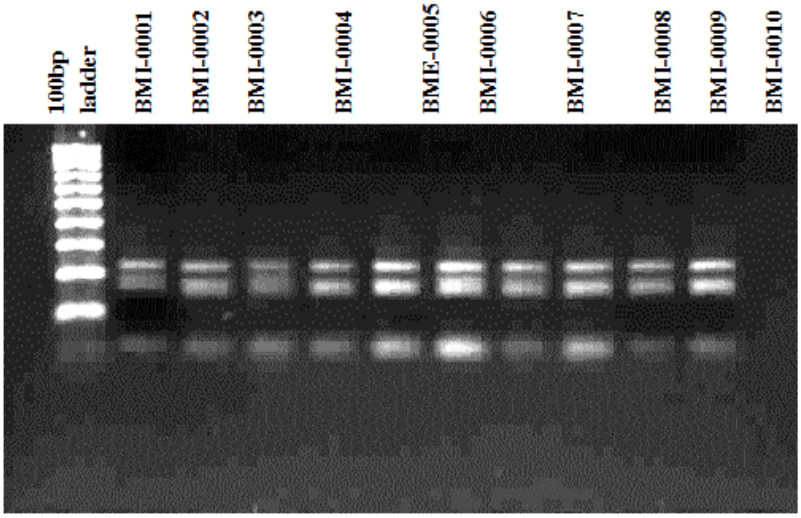


Figure. 3. Amplified Yolk protein gene EST- PCR banding profiles of genomic DNA of ten silk worm races were digested with restriction enzyme Hinf I.

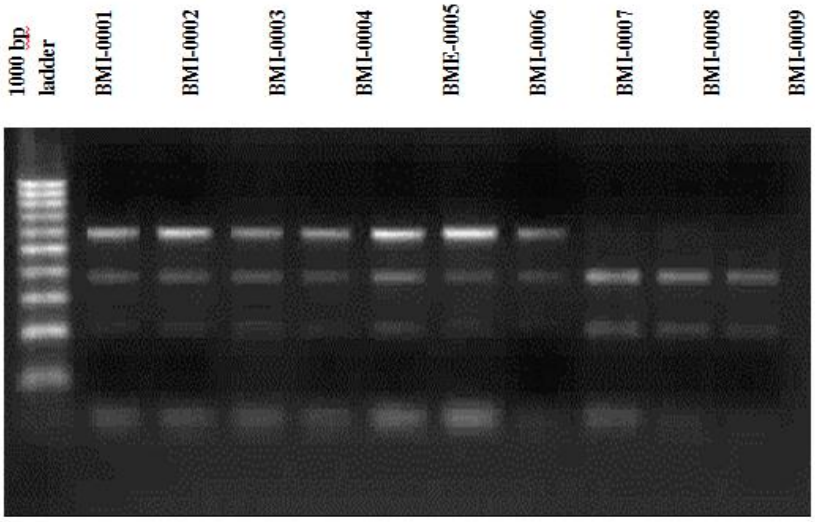


Figure. 4. Amplified Yolk protein gene EST- PCR banding profiles of genomic DNA of ten silk worm races were digested with restriction enzyme EcoRI .

Table. 4. Yolk protein genes forward reverse sequence.

Primer	Gene Coding	Primer Length
F ATGCTATTGTTTCGCTTTTC	Yolk Protein	20
R CTCTATTAGTGTCTGTTTCG	Yolk Protein	20

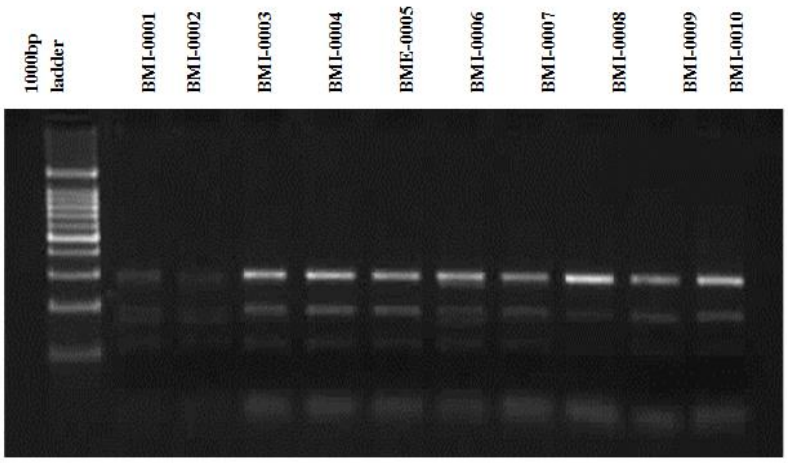


Figure. 5. Amplified Yolk protein gene EST- PCR banding profiles of genomic DNA of ten silk worm races were digested with restriction enzyme Taq I.

3.5 Finding of Allelic variation

Based on these base pair variation the migrating bands in the gel were designated a, b, c from the most anode and cathode region and this assign of a to c letters formed bases for genotype identification as ab, ac, bb, bc and cc there are about five genotype identify from yolk protein gene among ten multivoltine sequence studied. The number of genotypes identified as five in

numbers were found distributed in different frequency (Table 5) out of thirty alleles identified among ten multivoltine races 11 number of ab genotype with 36.66% frequency, 3 bb genotype with 10% frequency, 11bc genotype 36.66% frequency, 2ac genotype with 6.66% frequency and cc genotype with 10% genotype were observed among 10 multivoltine races studied each genotype pattern against each race was given in (Table 6).

3.6 Cluster grouping

A total of 6 cluster groups were found distributing a ten MV silk races into their similar band bp alignments among these 6 cluster group Kollegal Jawan, MY-1 formed in first cluster groups , Kolar Gold represents a distinct second cluster group, Moria and C.Nichi are grouped into 1,3 cluster group. Tamil Nadu White, Mysore Princess following in 4th cluster group.5th cluster group comprise Pure Mysore and Hosa Mysore form 5th cluster group. Sarupat had shown a distinct separate 6th cluster group (Fig.6 & Table 7).

Table 5: Genotyping and significance characters of selected races.

S. No	Acc. No	Race Name	Hinf I	EcoR I	Taq I	Fec (No)	VLD(h)	TLD(h)	Cocoon wt(g)	Shell wt (g)
1	BMI-0001	Pure Mysore	bb	bc	cc	457	186	626	1	0.14
2	BMI-0002	Sarupat	ab	bc	bb	405	149	569	1.16	0.17
3	BMI-0003	Moria	bb	ac	bc	388	145	564	1.13	0.16
4	BMI-0004	Tamilnadu White	ab	bc	bc	429	147	566	1.21	0.19
5	BME-0005	C.Nichi	ab	ac	bc	397	112	539	1	0.12
6	BMI-0006	Hosa Mysore	ab	bc	cc	498	144	559	1.36	0.21
7	BMI-0007	Mysore Princess	ab	bc	bc	488	144	558	1.16	0.18
8	BMI-0008	Kolar Gold	ab	ab	cc	477	134	553	1.26	0.19
9	BMI-0009	Kollegal Jawan	ab	ab	bc	480	132	549	1.28	0.19
10	BMI-0010	My-1	ab	ab	bc	452	146	564	1.36	0.19

Fec=Fecundity, Hat=Hatching, VLD=Vth instar larval duration, TLD=Total larval duration, SCW=Single Cocoon weight, SSW= Single Shell weight.

Table. 6: Genotyping of MV races based on percentage of frequency.

S. No	Genotype	No of genotype observed	% frequency of genotype
1	Ab	11/30	36.66
2	Bb	3/30	10
3	Bc	11/30	36.66
4	Ac	2/30	6.66
5	Cc	3/30	10

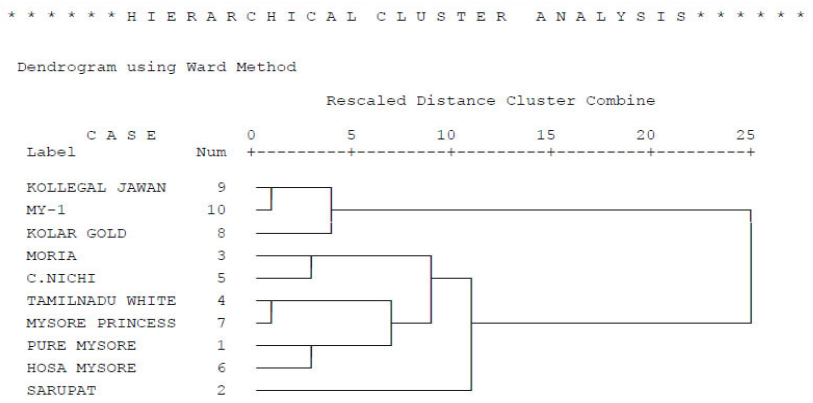


Figure. 6. Dendrogram of ten silkworm races profile using Ward method.

Table. 7: Cluster group of yolk protein gene primer amplified with restriction enzyme.

S. No	Cluster W	Race Name	Parameters
1	I	Kollegal Jawan,MY-1	1.36
2	II	Kolar Gold	1.26
3	III	Moria,C.Nichi	1.13
4	IV	Tamilnadu White,Mysore Princess	1.21
5	V	Pure Mysore,Hosa Mysore	1.36
6	VI	Sarupat	1.16

4. DISCUSSION

The grouping results shown Pure Mysore, Sarupat, Moria, TamilNadu White and C.Nichi possessed the same parentage. Mysore Princess, Kolar Gold and Kollegal Jawan these races were having separate groups of parentage. Other races like Hosa Mysore and TamilNadu White and MY-1 which are having the distinct parentage. Similar studies were reported in silkworm *Bombyx mori* using their growth-rate,pupation rate and analyzing their quantitative trait characters (Gamo and Hirabayashi, 1983). Among the three restriction enzymes pattern of yolk protein gene shown genetic variation in the form of genotype frequency bc and ab were providing a highest variation.

The allelic pattern of RFLP-STS was analyzed with Cluster W. These results were showed that the races Kollegal Jawan, MY-1 Pure, Mysore and Hosa Mysore having the top most level of parameters. The total of six cluster group were observed by the EST –marker. Among the six groups Kollegal Jawan, MY-1, Pure Mysore and Hosa Mysore showed their highest parameters. The amplified fragment length polymorphism techniques were shown distinct polymorphic patterns in silkworm varieties were reported by Tan et al., 2001. Similarly, human genome was characterized by using single nucleotide polymorphisms in coding regions of the hCNT3 gene (Cargill et al., 1999; Damaraju et al., 2000). The above results

were concluded that the genomic variations of ten races based on their yolk gene polymorphism shown their thermal resistance behaviour and this study helps to identify the hardy silkworm breeds which can be utilized for crop improvement programs. The genomic variations in silkworm races were also reported in several studies (Nagaraju and Goldsmith, 2002). Awasthi et al., 2008 studied the bivoltine, multivoltine and mutant silkworm races by RAPD, ISSR and RFLP_STS markers for analyzing their molecular evolutionary relationship. This study is very useful in bringing out the inherent genetic characteristics of silkworm stocks that are being maintained in the Central Sericultural Germplasm Genetic Resources Centre at Hosur. Furthermore, the breeder's involved in breeding programmers across the country can choose the identified breeds of the germplasm as parents to improve the silk production in different parts of the country.

5. CONCLUSION

Expressed sequence tag based genomic composition identification is an easiest way to identify the silkworm races. Furthermore, based on silkworm races productivity, origin, genetic makeup as well as disease resistant varieties will be analyzed by EST. In view of the above,the current study identified the yolk protein gene for silk productivity yield using an EST –PCR based RFLP technique. The obtained results of Kollegal Jawan, MY-1, Pure Mysore and Hosa Mysore were separated unique group has a high yolk yielding silkworm races. Hence, the EST based RFLP tool will be a greater tool to identify the silkworm races corresponding to their better silk yield.

CONFLICT OF INTEREST

The authors declare no financial or commercial conflict of interest.

ACKNOWLEDGEMENTS

The corresponding author are thankful to the Director, Central Sericultural Germplasm Resources Centre (CSGRC), Hosur, for providing silk worm races and lab facilities, and also acknowledge her thanks to scientists for their help and guidance.

REFERENCES

1. Aoki K, Tamaki I and Nakao K, et al., "Approximate Bayesian computation analysis of EST-associated microsatellites indicates that the broadleaved evergreen tree Castanopsis sieboldii survived the last glacial maximum in multiple refugia in Japan", Heredity, 2019; 122: 326-340.

2. Awasthi A K, Kar P K and Srivastava P P, et al., "Molecular evaluation of bivoltine, polyvoltine and mutant silkworm (*Bombyx mori* L.) with RAPD, ISSR and RFLP-STS markers", Indian Journal of Biotechnology, 2008; 7: 188-194.
3. Cargill M, Altshuler D and Ireland J, et al., "Characterization of single-nucleotide polymorphisms in coding regions of human genes", Nature Genetics, 1999; 23L 373-373.
4. Casas L, Saenz-Agudelo P and Irigoien X, "High-throughput sequencing and linkage mapping of a clownfish genome provide insights on the distribution of molecular players involved in sex change", Sci Rep. 2018; 8.
5. Chatterjee S N, Rao C G P, Aswath S K and Patnaik A K, "Correlation between yield and biochemical parameters in the mulberry silkworm, *Bombyx mori* L", Theor Appl Genet. 1993b; 87: 385–391.
6. Cheng T C, Xia Q Y and Qian JF, et al., "Mining single nucleotide polymorphisms from EST data of silkworm, *Bombyx mori*, inbred strain Dazao", Insect Biochemistry and Molecular Biology, 2004; 34: 523-530.
7. Damaraju S, Zhang J and Visser F, et al., "Identification and functional characterization of variants in human concentrative nucleoside transporter 3, hCNT3 (SLC28A3), arising from single nucleotide polymorphisms in coding regions of the hCNT3 gene", Pharmacogenetics and Genomics, 2005; 15: 173-182.
8. Datta R K, Improvement of silkworm races, *Bombyx mori* (L). in India, Sericologia, 1984; 24: 93–415.
9. Gamo T and Hirabayashi T, "Genetic-analysis of growth-rate, pupation rate and some quantitative characters by Diallel cross in the silkworm, Bombyx-mori L." Japanese Journal of Breeding, 1983; 33: 178-190.
10. Guo Y S, Liu C J and Zhao YH, et al., "Construction of a molecular genetic linkage map for longan based on RAPD, ISSR, SRAP and AFLP markers", Iii International Symposium on Longan, Lychee, and Other Fruit Trees in Sapindaceae Family, 2010; 863: 141-147.
11. Jingade A H, Vijayan K, Somasundaram P, Srinivasababu G K. and Kamble C K, "A review of the implications of heterozygosity and inbreeding on germplasm biodiversity and its conservation in the silkworm, *Bombyx mori*." Journal of Insect Science, 2011; 11.
12. Kadono-Okuda K, Kosegawa E, Mase K and Hara W, "Linkage analysis of maternal EST cDNA clones covering all twenty-eight chromosomes in the silkworm, *Bombyx mori*", Insect Molecular Biology, 2002; 11: 443-451.

13. Mita K, Morimyo M, Okano K, Koike Y and Nohata J, et al., "The construction of an EST database for *Bombyx mori* and its application", Proc Nati Acad Sci USA, 2003; 100: 14121-14126.
14. Mohandas T P, Sethuraman B N, Saratchandra B and Chatterjee SN, "Molecular genetics approach for identifying markers associated with yield traits in the silkworm, *Bombyx mori* using RFLP-STS primers", Genetica, 2004; 122: 185-197.
15. Nagaraja G M and Nagaraju J, "Genome fingerprinting of the silkworm *Bombyx mori* using random arbitray primers", Electrophoresis, 1995; 16: 1633-1638.
16. Nagaraju J and Goldsmith M R, "Silkworm genomics - progress and prospects". Current Science, 2002; 83: 415-425.
17. Nagaraju J, Reddy K D and Nagaraja G M, et al., "Comparison of multilocus RFLPs and based marker systems for genetic analysis of the Silkworm, *Bombyx mori*", Heredity, 2001; 86: 588-597.
18. Nagaraju J, Sharma A and Sethuraman B N, et al., "DNA fingerprinting in Silkworm *Bombyx mori* using banded krait minor satellite DNA derived probe", Electrophoresis, 1995; 16: 1639-1642.
19. Razi M, Darvishzadeh R, Amiri M E, Doulati-Banehd H and Martinez-Gomez P, "Molecular characterization of a diverse Iranian table grapevine germplasm using REMAP markers: population structure, linkage disequilibrium and association mapping of berry yield and quality traits", Biologia, 2019; 74: 173-185.
20. Reddy K D, Abraham E G and Nagaraju J, "Microsatellites in the silkworm, *Bombyx mori*: Abundance, polymorphism and strain characterization", Genome, 1999a; 42: 1057-1065.
21. Reddy K D, Nagaraju J and Abraham E G, "Genetic characterization of the silkworm *Bombyx mori* by simple sequence repeat (SSR)-anchored PCR", Heredity, 1999b; 83: 681-687.
22. Tan Y D, Wan C L, Zhu Y F, Lu C, Xiang Z H and Deng H W, "An amplified fragment length polymorphism map of the silkworm", Genetics, 2001; 157: 1277-1284.
23. Tsukuba, Ibaraki, Kosegawa E, Mase K and Nagaoka S, et al., "Linkage analysis of EST cDNA clones by using RFLP in the silkworm, *Bombyx mori*", Journal of Sericultural Science of Japan, 2001; 70(3): 135-143.
24. Ukoskit K, Posudsavang G and Pongsiripat N, et al., "Detection and validation of EST-SSR markers associated with sugar-related traits in sugarcane using linkage and association mapping", Genomics, 2019; 111: 1-9.

25. Wheelan S I, Church D M and Ostell J M, "Spidey a tool for mRNA to genomic alignments", Genome research, 2001; 11: 1952-1955.
26. Ye Y J, Feng L and Liang X H, et al., "Characterization, validation, and cross-species transferability of newly developed EST-SSR markers and their application for genetic evaluation in crape myrtle (*Lagerstroemia spp*)", Molecular Breeding, 2019; 39(2): 26.