Recent advances in molecular genetics of the silk moth, *Bombyx mori*

J. Nagaraju

Centre for DNA Fingerprinting and Diagnostics, CDFD Interim Building Complex, Nacharam, ECIL Road, Hyderabad 500 076, India e-mail: jnagaraju@mailcity.com; jnagaraju@cdfd.org.in

The domesticated silk moth, *Bombyx mori*, the lepidopteran molecular model and an important economic insect is emerging as an ideal molecular genetic resource for solving a broad range of biological problems. The well-developed genetics of this species includes more than 400 mapped mutations, hundreds of geographical races and genetically improved strains that represent an array of differences for various qualitative and quantitative traits, and a number of well studied cloned genes encoding proteins with diverse functions. The recent progress in the construction of molecular genetic maps, BAC libraries and a variety of molecular marker assays will further widen the scope of genetic analysis of this organism.

By making use of a large body of genetic information in *Drosophila melanogaster* it is possible to study the dynamics of genome structure and organization in an evolutionary context besides identifying new genes and gene systems for transgenesis, develop improved strains through DNA marker utilization and molecular dissection of quantitative trait loci.

This review invites attention to *Bombyx* as a genetic resource and discusses the status of silkworm molecular genetics research.

MOST silk moth species belong to either the family of Bombycidae or Saturniidae, in the superfamily of Bombycoidea. Phylogenetic studies based on interspecific hybridization[1], chromosome pairing behaviour[2] and gene structure comparison[3–5] suggest that the present day domesticated silk moth, *B. mori*, is closely related to *B. mandarina*, the wild silk moth. Cytological investigations of inter-specific hybrids of Japanese *B. mandarina* (n = 27) and *B. mori* (n = 28) have revealed the formation of a trivalent with one chromosome of the former and two chromosomes of the latter suggesting that one mandarina chromosome might have split into two during the history of domestication[2].

As in most Lepidoptera, chromosomes of the silkworm are holocentric, i.e. they possess centromeres throughout the chromosome body. This was shown by the persistence of chromosome fragments after irradiation during cell division[2], the dispersion of microtubule attachments[6], the absence of recombination nodules in females[7], the presence of supernumerary chromosomes in some close relatives of the silkworm[8], and the chromosome pairing behaviour and fertility of interspecific hybrids of two Saturniid species with diverse chromosome number[9]. *B. mori* chromosomes are highly condensed and appear dot-shaped at most meiotic and mitotic metaphase stages. Diffused centromeres and lack of special features make them difficult to identify individually. This has resulted in limited application of modern cytogenetic tools for genetic and molecular studies.

The silk moth, *B. mori* has been used as a model system for formal genetic studies since the discovery of Mendelian inheritance at the turn of the century because of its large size, ease of rearing in the laboratory, and economic importance in silk production. The well-developed genetics of this species includes more than 400 mutations which have been mapped to 28 linkage groups or chromosomes (see Box 1 for details)[10]. In addition, the existence of

hundreds of geographic races and genetically improved strains used for silk production which differ not only in Mendelian traits but also in the complex of quantitative traits such as body size, feeding duration, thermal tolerance and disease resistance, makes them amenable to systematic analysis using modern genetic tools.

Genetic studies in the silkworm

Research on silkworm was initiated in Japan as early as 1900 under the patronage of the sericulture industry which rapidly promoted it as a valuable model system for fundamental studies. In 1906, Toyama reported the utilization of the principles of hybrid vigour to enhance viability and yield by crossing different silkworm genotypes[11]. The phenomenon of sex-linkage in *B. mori* was first reported by Tanaka[12], shortly after its initial report in *D. melanogaster*. Many genetic phenomena such as maternal inheritance, parthenogenesis, polyploidogenesis, radiation-induced translocations and complex loci were reported in the early 1930s (ref. 13).

Together with the many inherent merits of the silkworm, the growing collection of world genetic stocks, encompassing 3000 genotypes, places *B. mori* second

Box 1. Silkworm 'primer'

Biology	
Generation time	50 days (non-diapause), 6 months (diapause)* *Made to hatch anytime by following refrigeration and acid treatment schedules to terminate diapause
Number of offspring per brood	~400
Larval growth from hatching to silk secretion	10000 times
Larval size	~8 cm
Genome and genetics	
Haploid genome size	550 million basepairs (0.5 pg)
Number of chromosomes (*n*)	28 (Holocentric)
Linkage map (Classical)	900 cM
Number of characterized mutations	~400
Number of silkworm strains	~3000
Current number of mapped markers	61 RFLPs, 168 RAPDs, 140 RAPD primer pairs
Available molecular marker resources	RFLPs, RAPDs, SSRs, ISSRs, STS, SCARs, ESTs, BACs
Current number of cloned genes	>20
Homeotic genes	Similar to Bithorax (Bx-C) and Antennapedia (ANT-C) complexes of *Drosophila*.
Sex determination mechanism	ZW females, ZZ males
Transgenesis	Transient expression and piggy bac transposon-mediated germline integration achieved
Heterosis	Very high in cross between diapause and non-diapause varieties
Parthenogenesis	
Gynogenesis	Simple heat treatment at meiotic I divisional stage of the eggs.
Androgenesis	Cold treatment of the prezygotic eggs

only to *Drosophila* as an insect model for genetic studies. The formulation of artificial diets as early as in 1960 (ref. 14), freed the silkworm rearing from the use of fresh mulberry leaves and helped to expand the study of *Bombyx* to academic laboratories. The accumulation of information on the nutritional requirements of the silkworm led to the formulation of artificial diet. The composition of artificial diet has been gradually improved and there are only small differences in the rate of growth and development of larvae reared on mulberry leaves or artificial diets[14].

Many of the research findings have already provided important insights into the understanding of basic biological processes[15]. In silkworm, females are heterogametic (ZW) and males are homogametic (ZZ) in sex chromosome constitution. The change of ratio of sex chromosomes to autosomes has no effect on sex determination and gynandromorphs, that is, sexual mosaics, rather than morphologically blended intersexes were produced in such studies, suggesting that sex determinants are active on both Z and W-chromosomes. The triploid and tetraploid silkworms which carry ZZW and ZZZW chromosome constitution invariably develop into females, suggesting that female determinants are possibly localized on W-chromosome[13]. In light

of detailed investigations on sex-determination mechanism in *Drosophila*, it may be worth investigating the primary signals in *Bombyx* sex determination and the degree to which the underlying genetic and molecular mechanisms are conserved.

In silkworm, there appears to be no dosage compensation. This has been demonstrated using an anonymous Z-specific transcript[16] and *Bm kettin*, homologous of the X-linked *Drosophila* kettin gene[17]. In both cases, the level of transcripts in males was two times greater than that in females unlike dosage compensated species such as mammals, in which one copy of the homogametic chromosome pair (XX) is inactivated in somatic tissues by heterochromatinization[18]. On the other hand, in many lepidopteran species, as well as in other taxa with heterogametic females, there is ample evidence that it is the W-chromosome that becomes inactivated. In silkworm and other lepidopteran insects, heterochromatic body is shown to be present only in females in somatic tissues such as silk glands, malphighian tissues, epidermal cells, etc[19]. The sex chromatin body has been correlated with the W-chromosome in endomitotic tissue such as silk gland[20], artificially induced polyploids[21] and in mutants carrying translocations between the W and various autosomes of silkworm.

In light of the findings that W-chromosome is strongly female determining and there is no dosage compensation in *Bombyx*, it would be of interest to investigate whether heterochromatin of W-chromosome specific sequences shows any features with well-studied phenomenon of X-chromosome inactivation in mammals. The identification of W-chromosome specific sequences would facilitate further investigation of this process at the molecular level.

B. mori was the first insect shown to have an interspersed pattern of repetitive and non-repetitive sequences typical of mammalian genomes[22,23]. Molecular characterization of the repetitive elements from *B. mori* has revealed the presence of transposable elements typical of the *Drosophila* genome as well as retroposons typical of the mammalian genomes[24]. The *B. mori* genome harbours abundant retroposons (SINEs, for short interspersed nucleotide elements repeats) like Bm_1 and Bm_2 elements which represent 5 to 10% of the total genomic mass, similar to the *Alu* and *Alu*-like elements found in the mammalian genomes. Interestingly, the fruitfly genome is devoid of such SINE elements. Unlike the mammalian genomes, which contain only a non-LTR (Long Terminal Repeat) LINE-1 elements (Long Interspersed Nucleotide Elements), the *Bombyx* genome contains several diverse elements some of which contain long terminal repeats like *Pao* and *Mag* and others like R_1*Bm*, R_2*Bm* and BMC_I which lack long terminal repeats similar to the many LTR (*Copia*, *Gypsy*)- and non-LTR elements (R_1*Dm*, R_2*Dm*, *Jockey*, F, G and I) described in *Drosophila*. *Bombyx* genome also harbours DNA-mediated mobile elements such as *mariner*[25,26] similar to fruitfly genome.

Telomeric DNA motifs have been characterized from the silkworm[27] and are made of a pentanucleotide repeat (TTAGG) characteristic of many insects instead of the usual hexanucleotides (TTNGGG) found in vertebrates. A similar telomeric repeat appears to be present in wild silk moths and other insects, although not in *Drosophila*, whose telomeric region contains a complex array of longer and more diverse sequence elements which are also found in heterochromatin[28,29].

A number of sequenced *B. mori* genes encoding proteins with diverse functions are now available (Table 1). In this regard, the silk gland and the chorion provided effective model systems to help unravel the molecular basis of gene expression. The silk gland has long been used as a model for studying gene regulation and cell differentiation[30]. An attractive feature of this

tissue is its simple construction with a succession of territories comprising a fixed number of cells each expressing a specific set of silk protein encoding genes. The single silk gland cell accumulates up to 400,000 to 800,000 haploid genomes by endomitosis, the highest known polyploidy ever reported. The silk proteins, fibroin and sericins are two distinct families of proteins. Fibroin is a complex of three associated polypeptides encoded by single genes, (the heavy and light chains) which are linked by disulphide bonds and p25, a co-secreted chaperone loosely attached to the two others. Sericins are the products of two distinct genes whose precursor mRNAs are the target of differential splicing leading to 4 and 2 sericin isoforms from the genes *Ser* 1 and *Ser* 2, respectively[31]. Functional assays for regulatory elements of chorion genes in follicular cells have been elegantly demonstrated using the heterologous transgenic system of *Drosophila*[32], and follicular cell regulatory factors and their genes are now well characterized[33].

Homeobox genes, *Bombyx ultrabithorax* (*BmUbx*), *Bombyx abdominal-A* (*Bmabd-A*) and *Bombyx Abdominal-B* (*BmAbd-B*), which constitute a part of the E-pseudoallelic complex that specifies the identity of body segments[34], have been cloned and characterized[35,36]. From the initial molecular analysis of the homeodomains the E-complex has been shown to be analogous to the Bithorax complex (BX-C) of *Drosophila*. However, some mutations of the E-complex reveal one interesting aspect that is different from those in *Drosophila*. E-mutant phenotypes such as E^{kp}/E^{kp} and $E^{D}/+$ cause shifts in one direction on the dorsal side and in the opposite direction on the ventral side. Such independent determinations of the dorsal and ventral sides is not observed in *Drosophila*. Differences between the *Bombyx* E-complex and the BX-C of *Drosophila* show that the structures and functions of the *cis*-regulatory and coding regions outside the homeodomains of the *Bombyx* E-complex may be different from those of *Drosophila* which remain to be investigated. Another homeotic complex, *Nc*-locus which harbours homeotic genes for thoracic segments, *Bombyx* Antennapedia (*BmAntp*) and sex comb reduced (*Scr*), which specifies identities of maxillae, labial and prothoracic segments has also been characterized. These two genes of *Nc*-locus lie in close proximity of approximately 100 kb and are located on the sixth chromosome at 1 cM distance from the E-complex. The *Nc*-locus corresponds to the *Drosophila* ANT-C[36]. The close proximity of *Nc* and E-complex in the sixth chromosome is similar to ANT-C and BX-C in *Drosophila*; likewise the adjacent location of *Bm Antp* and *Scr* genes within approximately 100 kb is as in *Drosophila,* where they are positioned at about only 70 kb distance.

Earlier attempts at producing transgenic silkworms have resulted either in transient expression[37,38] or in rapid degradation and extensive rearrangement of the injected sequences[39]. Recently, the lepidopteran mobile element, *Piggy bac*[40], has been successfully put to the germline of silkworm (Pierre Couble, University of Lyon, France, pers. commun.). In another novel approach the recombinant *Autographa californica* nuclear polyhedrosis virus (Ac NPV) carrying chimeric fibroin light (L) chain green fluorescent protein (GFP) gene has been successfully used to target the chimeric gene to the L-chain region of the silkworm genome by homologous recombination. The integrated gene has been shown to stably transmit through generations[41].

Table 1. Selected genes cloned from *Bombyx mori*

Cloned genes	Encoded products	Ref.
Enzymes		
Sorbitol dehydrogenase	Sorbitol dehydrogenase	59
Trehalase (cDNA)	Trehalase	60
Glycyl-tRNA synthetase	Glycyl-tRNA synthetase	61
cdc2 and cdc2-related (cDNA)	Cell cycle kinases	62
Homeoproteins		
Bmftz-F1	Fujitaratzu-related	63
Bmantp	Antennapedia	36
BmUbx	Ultrabithorax	35
Bmcd1	Caudal	64
Bmabd-1	Abdominal	35
Bmen	Engrailed	65
Bmin	Invected	65
BmScr	Sex comb reduced	66
Bmdef	Deformed	66
Hormones and receptors PTTH	Prothoracicotropic hormone	67
DH/P-BAN	Diapause hormone/pheromone Biosynthesis activating neuropeptide	68, 69
Bombyxin A, B, C and D	Insulin related peptides	70,71
EcR	Ecdysone receptor	72
BmEcR-A and B1 (cDNA)	Ecdysone receptor isoforms	73
Immune proteins		
BmLeb	Lebocin	74
GNBP (cDNA)	Gram negative bacteria-binding protein	75
FPI-F (cDNA)	Serine protease inhibitor	76
BmAtt	Attacin	77
MORICIN (cDNA)	Moricin	78
CecB1 and 2	Cecropin	79
BmLyz	Lysozyme	80
Pro-PO (cDNA)	Prophenol oxidase	81
Regulatory proteins		
SCF (cDNA)	Super coiling factor	82
SGF-1	Silk gland cell regulatory protein	83
POU-M1/SGF3	Sericin gene regulatory protein	84
BmGATA a and b	GATA-like regulatory proteins	33
RNAs		
rDNA	Ribosomal RNA	85
$tRNA^{gly}$	Silk gland glycine tRNAs	86
$tRNA^{ala}$	Silk gland alanine tRNAs	87–89
Structural proteins		
BmColl (cDNA)	Collagen-like peptide	90
BmBtub (cDNA)	B-tubulin	91
A3	Cytoplasmic actin	92
A4	Cytoplasmic actin	93
A1 and A2	Muscle actins	94
ErA and ErB	Early chorion proteins	95
A and B	Middle chorion proteins	96
HcA and HcB	High cysteine late chorion proteins	97, 98
PCP-2	Pupal cuticle protein	99
H-Fib	Heavy fibroin chain (silk protein)	100, 101
L-Fib	Light fibroin chain (silk protein)	102
P25	Fibroin-associated peptide	103
Ser-1	Sericins (Silk gum protein)	104
Ser-2	Sericins (Silk gum protein)	105
ESP	Egg storage protein	106
SPI, SP2	Larval haemolymph storage proteins	107, 108

Molecular genetic maps of silkworm

Silkworm genetic stocks

The silkworm genetic stocks maintained around the world comprise geographical races, inbred lines, and mutants that carry numerous morphological, developmental, behavioural and biochemical features. It is estimated that more than 3000 silkworm genotypes are being maintained in Asia and Europe. The genetic diversity found in the present day inbred lines and elite stocks is mostly a result of cross breeding of geographical races which display a variety of qualitative, quantitative and biochemical traits. Broadly, four sets of geographical races have been identified: Japanese, Chinese, European and Tropical, which in addition to many visible characters, differ widely in qualitative and quantitative characters that affect silk yield. The genotypes of temperate origin produce high quantities of rated silk while the tropical races are poor producers, but hardy and able to survive under adverse climatic conditions.

The genetic stocks of the silkworm comprise more than 500 mutants for a wide variety of characters (serosa colours; larval and adult integument colours, skin markings and body shapes; cocoon colours and shapes; physiological traits such as diapause, number of larval moults and timing of larval maturity; food habits and biochemical features such as digestive amylase, blood amylase, blood and egg esterases, larval integument esterase, alkaline and acid phosphatases, haemolymph proteins; silk production and fibroin secretion; homeoproteins and body plan determination, etc.)[42] (Figure 1).*B. mori*, the only lepidopteran with well developed genetics, proves itself an ideal organism to apply the emerging molecular technologies to construct high density map. Such a map has manifold applications: positional or map-based cloning, localization of complex traits, comparative studies on genome evolution, marker-assisted breeding and selection and development of tools for fingerprinting. Further, information generated from *B. mori* can significantly benefit many economically and ecologically important lepidopteran species which include many wild silk moths and the most destructive agriculture pests.

Experimental strategies

The haploid genome of silkworm comprises 560 million base pairs per silk glandcell[22], which is approximately 3.5 ´ the size of the *D. melanogaster* genome (140 million base pairs)[43] and one-sixth the size of the human genome. Given the 560 Mb of the silkworm genome, more than 3000 molecular markers would be required to place one marker within every 250 kb, an average insert size which could be packaged in YAC (Yeast Artificial

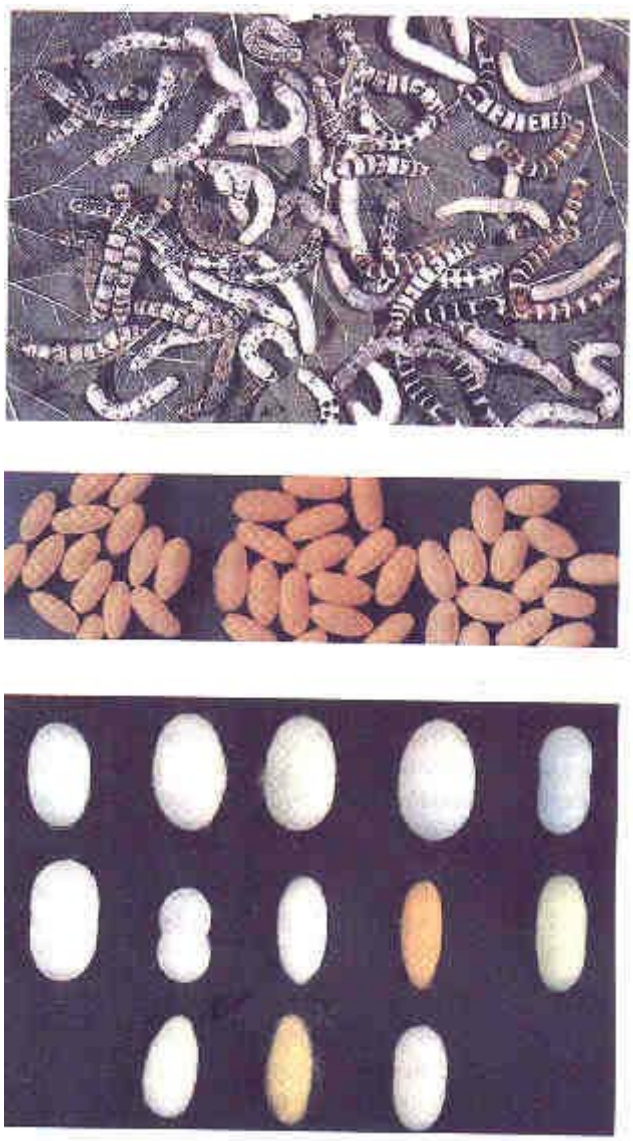

Figure 1. Genetic resources of silkworm. An array of larval mutants (upper panel); different ecotypes and inbred lines showing various cocoon shapes, sizes and colours (middle and lower panels).

Chromosome) or BAC (Bacterial Artificial Chromosome) if one plans to use the genetic map for positional cloning using YAC or BAC Library. In order to pursue this scale of molecular map one requires high volume marker technology.

The recent explosion of knowledge on the use of molecular techniques and remarkable progress made in the large-scale genomics of mammalian systems have resulted in the high volume marker assays that promise precision, speed, and cost effectiveness. These methods have accelerated the progress in high density mapping of various genomes, map-based cloning of genes of interest for which no prior information was available, identification of quantitative trait loci (QTL) for economic characters and utilization of molecular markers for 'marker assisted selection' and estimation of genetic diversity. Such an integrated approach would form the basis for the analysis and manipulation of the silkworm genome, and provide important raw material for comparative genomics and germline transformation.

Mapping population

Selection of suitable parents and synthesis of appropriate genetic populations are important steps in the molecular mapping strategy. The two parental strains, C_{108} and p50 have been used for molecular linkage map[44–46]. The former is a diapausing strain (bivoltine – 2 generations per year) and the latter a non-diapausing strain (polyvoltine – 5–6 generations per year) and they exhibit high phenotype diversity for such complex characters such as size, growth rate, diapause, morphology, nutritional requirements, general vigour and cocoon properties suggesting that considerable polymorphism exists at DNA level. These strains have undergone a high degree of inbreeding and are relatively homozygous. Both the strains are of Chinese origin. The F_2 progeny raised from a single pair mating of F_1 sibs have been used for molecular map construction and various molecular markers (restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), simple sequence repeat polymorphism (SSR) and inter-simple sequence repeat (ISSR)) are being integrated to this common F_2 mapping population.

It is common to use backcross population for linkage mapping in *B. mori* because of lack of crossing over in females. These backcrosses have differing advantages depending on the sex of the informative F_1 parent. If the F_1 is female, rapid detection of linkage is possible because of the complete absence of recombinants but the estimation of map distance is not possible. If the F_1 is male, the estimation of map distance is possible because of crossing over, but the detection of linkage is more difficult. Clearly, one approach to molecular mapping is to conduct the eight backcrosses obtainable by all possible combinations of the initial cross (C108 ´ p50 or p50 ´ C108), sex of the parent, and recurrent strain for the backcross (C108 or p50). This would permit mapping of co-dominant loci or dominant loci originating from either strain making data analysis straightforward. For maximum efficiency in the initial phases of linkage mapping, an F_2 cross if used, would yield eight-fold savings in labour with only about a two-fold decrease in information content.

Molecular markers

Molecular markers have several advantages over the traditional phenotypic markers that are available in silkworm genetic stocks. They are unaffected by environment, detectable in all stages of development and ubiquitous in number covering the entire genome. Many techniques have been employed in silkworm to reveal the genome-wide polymorphism. These are RFLP, RAPD, SSR and ISSR anchored PCR.

A preliminary RFLP map has been developed in silkworm using 52 progeny of an F_2 cross from a pair mating of two parental genotypes, C_{108} and p50. The RFLP probes included 15 characterized single copy sequences, 36 anonymous sequences derived from a follicular cDNA library and 10 loci corresponding to a low copy number retrotransposon. The RFLP map covers a total recombination length of 413 cM dispersed over 15 linkage groups. Out of the 15 characterized single copy sequences, 10 (which have already been localized on the conventional linkage map using biochemical and morphological variants) could be correlated with the molecular map[44].

Although RFLP analysis offers several advantages, it requires large quantities of DNA, is labour intensive, time consuming and expensive. The advent of PCR technology increased the approaches available for the detection of DNA polymorphisms, which offer effective alternatives to the hybridization methods of RFLP. The PCR-based methods meet the wide spectrum of requirements for large-scale genetic studies since they require only small quantities of DNA, avoid DNA blotting and radioactivity and are

amenable to automation. Some of the methods used in silkworm genetic analysis are described briefly here.

RAPD marker technology involves the amplification of genomic DNA template primed by a 10 base-oligonucleotide primer of a random sequence but with a minimum guanidine–cytosine content of 50%. A low annealing temperature (37–40° C) is maintained throughout the cycles (35–40 cycles). Polymorphisms occur as indicated by the presence or absence of a specific amplification product among individuals.

A medium density linkage map of RAPDs has been constructed using 102 F_2 individuals of the same parents (p50 and C_{108}) used for the RFLP map. The screening of 320 decanucleotide-primers resulted in 243 clear polymorphic products between C_{108} and p50, of which 168 bands followed Mendelian inheritance. The MAPMAKER program sorted 168 bands into 29 linkage groups and 10 unlinked loci covering the 892.4 cM map distance[45].

Recently, a high density linkage map containing 1018 genetic markers representing all 27 autosomes and the Z-chromosome has been constructed[47]. Most of the markers were RAPDs with primer pairs in combinations of 140 commercially available decanucleotides. This is in contrast to the earlier studies[45,46] which used only single random primers. The use of random primer pairs has greater advantages in map construction because of exponential increase of combinations of primer pairs using a limited number of primers and higher reproducibility when compared to using single primer. Furthermore, this map is an improvement over previously reported map distances of 593 cM based on 60 RFLP markers, 897.4 cM based on 168 RAPD markers and 900.2 cM based on 207 phenotypic markers. Considering the 560 Mb genome size of the silkworm, the average interval of markers works out to be ~560 kb and the relationship between physical and genetic distances is ~250 kb/cM. This value is intermediate between the honey bee (52 kb/cM)[48] and the fruit fly (575 kb/cM)[49]. These values agree well with the idea that crossing over rates are inversely correlated with chromosome size, because average chromosome sizes are 11.1 Mb for the honey bee, 17.1 Mb for the silkworm and 47 Mb for the fruit fly. The overall distance of Yasukochi's map is nearly 2 cM/marker.

The routine application of RAPD analysis for genetic mapping has limitations. First, the efficiency of amplification of fragments using decamers at low annealing temperatures (37° C) is very sensitive to reaction conditions and may be inconsistently reproducible in different experiments. This is particularly problematic if one is interested in integrating mapping data from different crosses by comparison to common reference markers. Second, RAPD markers are inherited as dominant markers. Lack of allelism is a disadvantage as the heterozygous and homozygous loci are not distinguishable.

Three criteria define ideal polymorphic loci: (i) they should be highly polymorphic; (ii) they should be easy to identify and type; and (iii) they should be amenable to automation. Microsatellite based genetic markers, which are distributed across most of the eukaryotic genomes, fulfill these criteria. They have proved to be extrordinarily efficient means to pursue genetic mapping. Microsatellites are short stretches of DNA which consist of an array of simple di-, tri-, or tetranucleotide repeats such as $(CA)_n$, $(GA)_n$, $(GTA)_n$, $(ATT)_n$, $(GATA)_n$. A unique oligonucleotide on each side of the repeat region is chosen for the production of a primer pair for each of the microsatellite array (locus). Allelic variations are based mostly on differences in the number of tandem repeats in a microsatellite array. Thus, the only way in which alleles can be distinguished is by measuring the total length of the microsatellite array. This is most readily accomplished through PCR amplification of microsatellite itself along with a small amount of defined flanking sequences

on both sides with designed primer pairs for each locus. A higher level of cost effficiency is achieved by combining two or more loci for simultaneous analysis through multiplex PCR. By using the fluorescent labelled primers and automated sequencers, the entire process of microsatellite analysis can be automated.

Recent studies have shown that the silkworm genome is abundantly interspersed with CA/GT and GA/CT repeats. The $(GT)_n$ repeats occur at every 49 kb while $(CT)_n$ repeats occur, on an average, at every 104 kb in the silkworm genome[50] (Table 2). These frequencies compare favourably with the well-characterized mouse and human genomes where intensive maps of microsatellite markers have been accomplished. Initial studies show that the primers for the flanking microsatellite loci revealed distinct allelic differences between different strains of silkworm and inherit in Mendelian pattern (Figure 2). In the thirteen silkworm populations analysed using 14 microsatellite loci, the number of alleles ranged from 3 to 17 with PIC values of 0.66 to 0.90. The microsatellite markers can thus provide high volume markers for intensive mapping of the silkworm genome. As many microsatellite loci could be multiplexed based on allelic sizes in an automated sequencer, the microsatellite markers could be mapped quickly using the appropriate silkworm mapping population. Efforts are now underway to utilize these markers to integrate with the existing RFLP and RAPD maps.

Examination of conservation of microsatellites among distantly related saturniid silkmoths such as *Antheraea mylitta, A. pernyi, Philosamia cynthia ricini, A. yamamai* and *A. polyphemus* and patterns of cross-species polymorphism revealed that almost one-third of the 60 *Bombyx* microsatellite loci tested are conserved in the heterologous species[51]. These results show that microsatellite loci provide another potential source of markers for comparative genome work in lepidoptera.

As PCR technology is finding increased use in genetic analysis, more variations of this technique are emerging. One such variation takes advantage of the ubiquitous distribution of di, tri and tetra nucleotide repeats in the silkworm genome using primers designed for the repeat regions to amplify the inter repeat genomic regions.

Table 2. Average distance in kb between $(GT)_n$ and between $(CT)_n$ microsatellite repeats in the genomes of silkworm, other insects and vertebrates[50]

Repeat	Silkworm	Honey bee	Bumble bee	Salmon	Pig	Rat	Human
$(GT)_n$	49	34	500	12	47	15	30
$(CT)_n$	104	15	40	ND	ND	50	113

ND, Not determined.

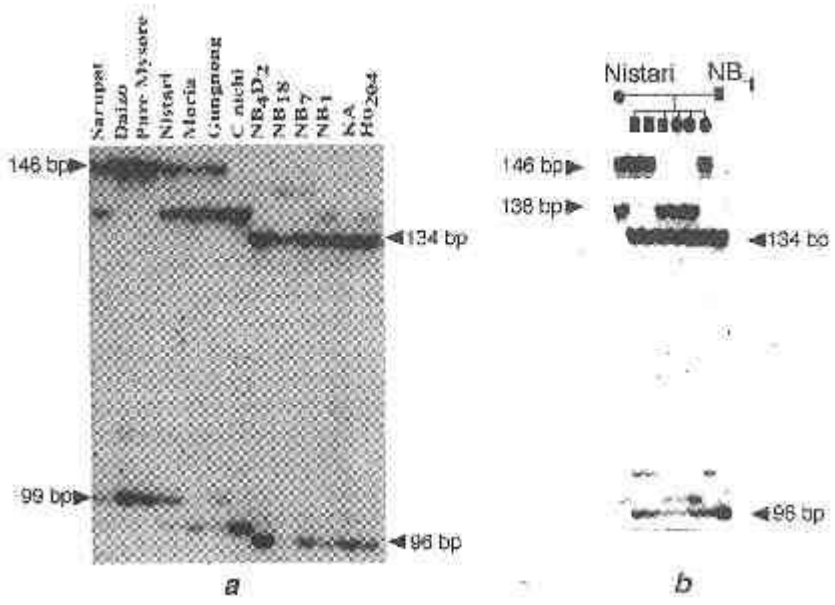

Figure 2. ***a***, Allelic variations detected by a microsatellite locus sat 211 in 13 silkworm strains. Arrows indicate diapause and non-diapause strain specific alleles; ***b***, Co-dominant inheritance of microsatellite alleles detected by sat 211 in the F_1 offspring of Nistari and NB_1 silkworm strains.

The primers consisted of $(CA)_n$ or $(GA)_n$ repeats extended into the 5¢ or 3¢ termini by 2 to 4 purine or pyrimidine residues to initiate PCR amplification of genomic segments flanked by inversely oriented, closely repeated sequences. The radioactively labelled primers are used for PCR amplification of template genomic DNA. The amplified DNA when resolved under non-denaturing polyacrylamide gel electrophoresis reveal information on multiple genomic loci in a single lane (Figure 3 *a*). Such markers are dominant and are shown to inherit in Mendelian fashion (Figure 3 *b*)[52]. Although most ISSR loci are dominant, rather than co-dominant, they offer several advantages over RFLPs and RAPDs for genotyping, the major one being the rapid production of a large number of markers in a cost-effective manner. The ISSR markers offer higher reproducibility due to the use of longer primers and higher annealing temperature than those used for RAPDs. Due to these advantages we believe that the ISSR technique has great potential in silkworm breeding and germplasm evaluation. The ability to screen so many polymorphisms in a single assay makes the ISSR-PCR a technique of choice for a large-scale screening of silkworm germplasm and makes it useful to augment the marker resources for the silkworm genome mapping programme.

The banded krait minor satellite DNA Bkm-2(8) containing a 545 bp sequence consisting mainly of 22 GATA repeats was also used as a probe to reveal polymorphisms in different silkworm races. Although polymorphism level was relatively lower than SSR anchored PCR, the Bkm probe revealed very clear fingerprint profiles. Such DNA fingerprint-based information will be useful to estimate genetic diversity of silkworm germplasm[53,54].

Two bacterial artificial chromosome (BAC) libraries have been constructed from the two *B. mori* reference strains, p50 and C_{108}. The libraries contain a total of 36,864 clones and are approximately 5.8 and 3.2 genome

equivalents, respectively. The average insert sizes in the libraries are 134.5 kb and 120.8 kb, respectively[55]. The previously reported close linkage between the *Bombyx* homologues of the *invected* (*Bm in*) and *engrailed* (*Bm en)* genes was confirmed in the newly constructed BAC contigs that contained both. This is the first report on large-insert libraries of non-dipteran insects. Such libraries are available only for *D. melano-gaster*[56,57]. With the ongoing efforts on construction and integration of high density linkage maps, the availability of BAC libraries and mutant genetic resources of the silkworm, map-based cloning will be feasible within a few years, and light will be thrown on molecular mechanisms that control quantitative trait loci (QTL), various mutant phecotropes. These libraries will be used for map-based cloning of genes and projected large-scale genomic sequencing projects. The construction of a complete physical map has been initiated by a hybridi-

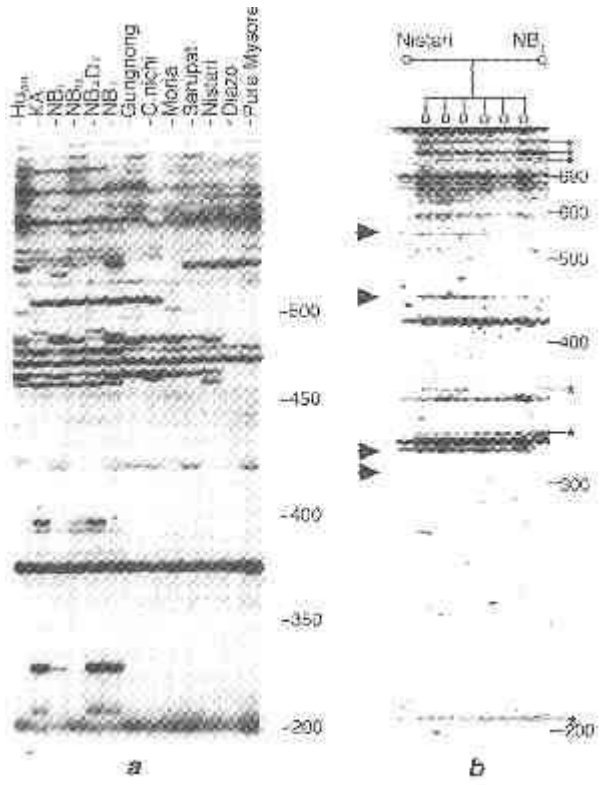

Figure 3. ***a***, Inter-SSR PCR profiles of 13 silkworm strains with 5¢ -anchored primer RRY RYY RRYR $(CA)_7$; ***b***, Mendelian inheritance of polymorphic bands in F_1 (Nistari ´ NB_1) offspring generated by the 5¢ -anchored primer RYY RRY RYY $(CA)_7$ Y. Arrows and asterisks indicate Nistari and NB_1 strain specific amplification products, respectively. Squares and circles represent males and females, respectively.

zation method using independent ESTs as probes[58]. In a complementary approach based on PCR, RAPD markers derived from the high density maps and STSs are being used to construct contig islands encompassing known silkworm genes and gene complexes such as *Hox* genes[55]; this will provide information about the chromosomal organization of regions likely to be conserved in other species of insects.

Efforts are being made by different groups to construct a high density molecular map of the silkworm that would integrate RFLPs (M.R. Goldsmith, USA), RAPDs (T. Shimada, Japan; J. Nagaraju, India; Y. Yasukochi, Japan), microsatellites (J. Nagaraju, India) and expressed sequence tags (ESTs) (K. Mita, Japan).

Conclusion

The molecular genetics of silkworm genome is poised on the edge of a new era. The large collections of carefully maintained silkworm mutations and practical breeding stocks are valuable resources that have only begun to be exploited to their full potential. The well-studied genetic resources of *B. mori* make it an ideal reference for the lepidoptera, where comparative genetics and genomics can work together to elucidate conserved evolutionary pathways and their diversification, identify new genes and gene systems as targets for transgenesis. Besides, it leads to basic research on new genome-based approaches for the exploitation of other economically important species such as *Antheraea* and in the control of pest species. An intensive genomics programme for the silkworm should enhance its usefulness for strain improvement and aid in the development of tools for targeted knockouts.

1. Aratake, Y. and Kayamura, T., *J. Seric. Sci. Jpn*., 1973, **42**, 331–339.
2. Murakami, A. and Imai, H. T., *Chromosoma*, 1974, **47**, 167–178.
3. Maekawa, H., Takada, N., Mikitani, K., Ogura, T., Miyajima, N., Fujiwara, H., Kobayashi, M. and Ninaki, O., *Chromosoma,* 1988, **96**, 263–269.
4. Kusuda, J., Tazima, Y., Onimaru, K., Ninaki, O. and Suzuki, Y., *Mol. Gen. Genet*., 1986, **203**, 359–364.
5. Shimada, T., Kurimoto, Y. and Kobayashi, M., *Mol. Phylogenet. Evol*., 1995, **4**, 223–234.
6. Friedlander, M. and Wahrman, J., *J. Cell Sci*., 1970, **7**, 65–89.
7. Rasmussen, S. W. and Holm, P. B., *Insect Ultrastructure* (eds King, R. C. and Akai, H.), Plenum Press, New York, 1982, pp. 61–85.
8. Puttaraju, H. P. and Nagaraju, J., *Curr. Sci.*, 1985, **54**, 471–472.
9. Nagaraju, J. and Jolly, M. S., *Theor. Appl. Genet*., 1985, **72**, 269–273.
10. Nagaraju, J., Kleminkob, V. and Couble, P., *Encyclopedia of Genetics* (ed. Reeve, E. C. R.), Fitzcoy Dearbon Publishers, London, 1999, in press.
11. Toyama, K., *Breeding Methods of Silkworm*, Sango, Shimpo, 1906, vol. 158, pp. 282–286.
12. Tanaka, Y., *J. Coll. Assoc. Sapporo*, 1916, **7**, 129–255.
13. Tazima, Y., *The Genetics of the Silkworm,* Logos Press, London and Englewood Cliffs, N.J., 1964.
14. Ito, T. and Kobayashi, M., *The Silkworm – An Important Laboratory Tool* (ed. Tazima, Y.), Kodansha, Tokyo, 1978, pp. 83–102.
15. Goldsmith, M. R. and Wilkins, A. S. (eds), *Molecular Model Systems in the Lepidoptera*, Cambridge University Press, 1995.
16. Suzuki, M. G., Shimada, T. and Kobayashi, M., *Heredity*, 1998, **81**, 275–283.
17. Suzuki, M. G., Shimada, T. and Kobayashi, M., *Heredity*, 1999, **82**, 170–179.
18. Lyon, M. F., *Annu. Rev. Genet*., 1992, **26**, 17–28.
19. Traut, W., *Chromosoma* (*Berl*), 1976, **58**, 275–284.
20. Perdix-Gillot, S., *Biochime*, 1979, **61**, 171–204.
21. Ito, S., *Jpn. J. Genet*., 1977, **52**, 327–340.
22. Gage, L. P., *Chromosoma*, 1974, **45**, 27–42.
23. Gage, L. P., Friedlander, E. and Manning, R. F., *Mechanisms in the Control of Gene Expression* (eds Nierlich, D. P., Rutter, W. J. and Fox, C. F.), Academic Press, New York, 1975, pp. 593–598.
24. Eickbush, T. H., *Molecular Model Systems in the Lepidoptera* (eds Goldsmith, M. R. and Wilkins, A. S.), Cambridge University Press, 1995, pp. 77–106.
25. Robertson, H. M. and Asplund, M. L., *Insect Biochem. Mol. Biol*., 1996, **26**, 945–954.
26. Ueda, H., Mizuno, S. and Shimura, K., *J. Mol. Biol*., 1986, **190**, 319–327.
27. Okazaki, S., Tsuchida, K., Maekawa, H., Ishikawa, H. and Fujiwara, H., *Mol. Cell. Biol*., 1993, **13**, 1424–1432.
28. Young, B. S., Pession, A., Traverse, K. L., French, C. and Pardue, M. L., *Cell*, 1983, **34**, 85–94.
29. Valgeirsdottir, K., Traverse, K. L. and Pardue, M. L., *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 7998–8002.
30. Hui, C. C. and Suzuki, Y., *Molecular Model Systems in the Lepidoptera* (eds Goldsmith, M. R. and Wilkins, A. S.), Cambridge University Press, 1995, pp. 249–272.
31. Couble, P., Michaille, J. J., Garel, A., Couble, M. L. and Prudhomme, J. C., *Dev. Biol*., 1987, **124**, 431–440.
32. Mitsialis, S. A. and Kafatos, F. C., *Nature*, 1985, **317**, 453–456.
33. Drevet, J. R., Seiky, Y. A. and Iatrou, K., *J. Biol. Chem*., 1994, **269**, 10660–10667.
34. Hasimoto, H., *Bull. Sericult. Exp. Stn. Jpn*, 1941, **10**, 328–363.
35. Ueno, K., Hui, C. C., Fukuta, M. and Suzuki, Y., *Development*, 1992, **114**, 555–563.
36. Ueno, K., Nagata, K. and Suzuki, Y., *Molecular Model Systems in the Lepidoptera* (eds Goldsmith, M. R. and Wilkins, A. S.), Cambridge University Press, 1996, pp. 165–180.
37. Coulon, M., Mounier, N., Couble, P. and Prudhomme, J. C., *Arch. Roux's Dev. Biol*., 1993, **202**, 123–127.
38. Tamura, T., Kanda, T., Takiya, S., Okano, K. and Maekawa, H., *Jpn. J. Genet*., 1990, **65**, 401–410.

39. Nagaraju, J., Kanda, T., Yukuhiro, K., Chavancy, G., Tamura, T. and Couble, P., *Appl. Entomol. Zool*., 1996, **31**, 587–596.
40. Fraser, M., Ciszczon, T., Elich, T. and Bauser, C., *Insect Biochem. Mol. Biol*., 1996, **5**, 141–151.
41. Yamao, M., Katayama, N., Nakazawa, H., Yamakawa, M., Hayashi, Y., Hara, S., Kamei, K. and Mori, H., *Genes Dev*., 1999, **13**, 511–516.
42. Chikushi, H., *Genes and Genetical Stocks of the Silkworm*, Keigaku Publishing Co., Tokyo, 1972.
43. Rasch, E. M., Barr, H. J. and Rasch, R. W., *Chromosoma* (*Berl*), 1971, **33**, 1–18.
44. Shi, J., Heckel, D. G. and Goldsmith, M. R., *Genet. Res*., 1995, **66**, 109–126.
45. Promboon, A., Shimada, T., Fujiwara, F. and Kobayashi, M., *Genet. Res*., 1995, **66**, 1–7.
46. Nagaraja, G. M. and Nagaraju, J., *Electrophoresis*, 1995, **16,** 1633–1638.
47. Yasukochi, Y., *Genetics*, 1998, **150**, 1513–1525.
48. Hunt, G. J. and Page, R. E., *Genetics*, 1995, **139,** 1371–1382.
49. Merriam, J., Ashburner, M., Hartl, D. L. and Kafatos, F. C., *Science*, **254**, 221–225.
50. Reddy, K. D., Abraham, E. G. and Nagaraju, J., *Genome*, 1999, in press.
51. Goldsmith, M., Shimada, T., Yasukochi, Y., Mita, K. and Nagaraju, J., Intl. Conf. Inv. Transgenesis, Crete, Greece, 29–31 August 1999.
52. Reddy, K. D., Nagaraju, J. and Abraham, E. G., *Heredity*, 1999, in press.
53. Nagaraju, J., Sharma, A., Sethuraman, B. N., Rao, J. V. and Singh, L., *Electrophoresis*, 1995, **16**, 1639–1642.
54. Sharma, A., Niphadkar, M. P., Kathirvelu, P., Nagaraju, J. and Singh, L., *J. Hered.*, 1999, **90**, 315–319.
55. Wu, C., Asakawa, S., Shimizu, N., Kawa Baki, S. and Yasukochi, Y., *Mol. Gen. Genet*., 1999, **261**, 698–706.
56. Ajioka, J. W., Smoller, D. A., Jones, R. W., Carulli, J. P., Vellek, A. E. C., Garza, D., Link, A. J., Duncan, D. W. and Hartl, D. L., *Chromosoma*, 1991, **100**, 495–509.
57. Smoller, D. A., Petrov, D. and Hartl, D. L., *Chromosoma*, 1991, **100**, 487–494.
58. Mita, K., Morimyo, M., Okano, K., Shimada, T. and Maeda, S., *RIKEN Rev.*, 1999, **22**, 63–67.
59. Niimi, T., Yamashita, O. and Yaginuma, T., *Insect Mol. Biol*., 1996, **5**, 269–280.
60. Su, Z. H., Sato, Y. and Yamashita, O., *Biochem. Biophys. Acta*, 1996, **1173**, 217–224.
61. Nada, S., Chang, P. K. and Dignam, J. D., *J. Biol. Chem*., 1993, **268**, 7660–7667.
62. Iwasaki, H., Takahashi, M., Niimi, T., Yamashita, O. and Yaginuma, T., *Insect Mol. Biol*., 1997, **6**, 131–141.
63. Sun, G. C., Hirose, S. and Ueda, H., *Dev. Biol*., 1994, **162**, 426–437.
64. Xu, X., Xu, P. X. and Suzuki, Y., *Development*, 1994, **120**, 277–285.
65. Hui, C. C., Matsuno, K., Ueno, K. and Suzuki, Y., *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 167–171.
66. Kokubo, H., Ueno, K., Amanai, K. and Suzuki, Y., *Dev. Biol*., 1997, **186**, 46–57.
67. Kawakami, A., Kataoka, H., Oka, T., Mizoguchi, A., Kimura-Kawakami, M., Adachi, T., Iwami, M., Nagasawa, H., Suzuki, A. and Ishizaki, H., *Science*, 1990, **247**, 1333–1335.
68. Sato, Y., Oguchi, M., Menjo, N., Imai, K., Saito, H., Ikeda, M., Isobe, M. and Yamashita, O., *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 3251–3255.
69. Kawano, T., Kataoka, H., Nagasawa, H., Isogai, A. and Suzuki, A., *Biochem. Biophys. Res. Commun*., 1992, **189**, 221–226.
70. Kawakami, A., Iwami, M., Nagasawa, H., Suzuki, A. and Ishizaki, H., *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 6843–6847.
71. Kondo, H., Ino, M., Suzuki, A., Ishizaki, H. and Iwami, M., *J. Mol. Biol*., 1996, **259**, 926–937.
72. Swevers, L., Drevet, J. R., Lunke, M. D. and Iatrou, K., *Insect Biochem. Mol. Biol*., 1995, **25**, 857–866.
73. Kamimura, M., Tomita, S., Kiuchi, M. and Fujiwara, H., *Eur. J. Biochem*., 1997, **248**, 786–793.
74. Furukawa, S., Taniai, K., Ishibashi, J., Hara, S., Shono, T. and Yamakawa, M., *Biochem. Biophys. Res. Commun*., 1997, **238**, 769–774.
75. Lee, W. J., Lee, J. D., Kravchenko, V. V., Ulevitch, R. J. and Brey, P. T., *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 7888–7893.
76. Pham, T. N., Hayashi, K., Takano, R., Nakazawa, H., Mori, H., Ichida, M., Itoh, M., Eguchi, M., Matsubara, F. and Hara, S., *J. Biochem*., 1996, **119**, 1080–1085.
77. Taniai, K., Ishii, T., Sugiyama, M., Miyanoshita, A. and Yamakawa, M., *Biochem. Biophys. Res. Commun*., 1996, **220**, 594–599.
78. Hara, S. and Yamakawa, M., *Biochem. Biophys. Res. Commun*., 1996, **220**, 664–669.
79. Taniai, K., Kodano-Okuda, K., Kato, Y., Yamamoto, M., Shimabukuro, M., Chowdhury, S., Xu, J., Kotani, E., Tomino, S. and Yamakawa, M., *Gene*, 1995, **163**, 215–219.
80. Lee, W. J. and Brey, P. T., *Gene*, 1995, **161**, 199–203.
81. Kawabata, T., Yasuhara, Y., Ochiai, M., Matsuura, S. and Ashida, M., *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 7774–7778.
82. Ohta, T., Kobayashi, M. and Hirose, S., *J. Biol. Chem*., 1995, **270**, 15571–15575.
83. Mach, V., Takiya, S., Ohno, K., Handa, H., Imai, T. and Suzuki, Y., *J. Biol. Chem*., 1995, **270**, 9340–9346.
84. Xu, P. X., Fukuta, S., Takiya, S., Matsuno, K., Xu, X. and Suzuki, Y., *J. Biol. Chem*., 1994, **269**, 2733–2742.
85. Lecanidou, R., Eickbush, T. H. and Kafatos, F. C., *Nucleic Acids Res*., 1984, **12**, 4703–4713.
86. Fournier, A., Taneja, R., Gopalkrishnan, R., Prudhomme, J. C. and Gopinathan, K. P., *Gene*, 1993, **134**, 183–190.

87. Young, L. S., Takahashi, N. and Sprague, K. U., *Proc. Natl. Acad. Sci. USA*, 1986, **83**, 374–378.
88. Sullivan, H. S., Young, L. S., White, C. N. and Sprague, K. U., *Mol. Cell. Biol.*, 1994, **14**, 1806–1814.
89. Hagenbuchle, O., Larson, D., Hall, G. I. and Sprague, K. U., *Cell*, 1979, **18**, 1217–1229.
90. Chareyre, P., Besson, M. T., Fourche, J. and Bosquet, G., *Insect Biochem. Mol. Biol*., 1996, **26**, 677–685.
91. Mita, K., Nenoi, M., Morimyo, M., Tsuji, H., Ichimura, S., Sawai, M. and Hamana, K., *Gene*, 1995, **38**, 583–592.
92. Mounier, N. and Prudhomme, J. C., *Biochimie*, 1986, **68**, 1053–1061.
93. Mangé, A., Couble, P. and Prudhomme, J. C., *Gene*, 1996, **183**, 191–199.
94. Mounier, N., Gaillard, J. and Prudhomme, J. C., *Nucleic Acids Res*., 1987, **6,** 2781.
95. Hibner, B. L., Burke, W. D. and Eickbush, T. H., *Genetics*, 1991, **128**, 595–606.
96. Spoerel, N., Nguyen, H. T. and Kafatos, F. C., *J. Mol. Biol.,* 1986, **190**, 23–35.
97. Iatrou, K., Tsitilou, S. G. and Kafatos, F. C., *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 4452–4456.
98. Rodakis, G. C., Lecanidou, R. and Eickbush, T. H., *J. Mol. Biol*., 1984, **20**, 265–273.
99. Nakato, H., Izumi., S. and Tomino, S., *Biochem. Biophys. Acta*, 1992, **1218**, 64–74.
100. Suzuki, Y. and Brown, D. D., *J. Mol. Biol*., 1972, **63**, 409–429.
101. Mita, K., Ichimura, S. and James, T. C., *J. Mol. Evol*., 1994, **38**, 583–592.
102. Yamaguchi, K., Kikuchi, Y., Takagai, T., Kikuchi, A., Oyama, F., Shimura, K. and Mizuno, S., *J. Mol. Biol*., 1989, **210**, 127–139.
103. Couble, P., Chevillard, M., Moine, A., Ravel-Chapuis, P. and Prudhomme, J. C., *Nucleic Acids Res*., 1985, **13**, 1801–1814.
104. Michaille, J. J., Couble, P., Prudhomme, J. C. and Garel, A., *Biochimie*, 1986, **68**, 1165–1173.
105. Michaille, J. J., Garel, A. and Prudhomme, J. C., *Gene*, 1990, **86**, 177–184.
106. Sato, Y. and Yamashita, O., *Insect Biochem. Mol. Biol*., 1991, **21**, 495–503.
107. Sakurai, H., Fujii, T., Izumi, S. and Tomino, S., *J. Biol. Chem*., 1988, **263**, 7876–7880.
108. Fujii, T., Sakurai, H., Izumi, S. and Tomino, S., *J. Biol. Chem*., 1989, **264**, 11020–11025.

ACKNOWLEDGEMENTS. I thank the Department of Biotechnology, New Delhi and Central Silk Board, Bangalore for financial assistance.

Received 30 August 1999; revised accepted 11 October 1999

CONTENTS

Volume 78 Number 2, 25 JANUARY 2000