Insect Molecular Biology (2002) **11**(4), 371–378

Linkage map organization of expressed sequence tags and sequence tagged sites in the mosquito, *Aedes aegypti*

D. W. Severson, J. K. Meece, D. D. Lovin, G. Saha and I. Morlais

Center for Tropical Disease Research and Training, Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA

Abstract

A composite genetic linkage map for the yellow fever mosquito *Aedes aegypti* was constructed based on restriction fragment length polymorphism (RFLP), single nucleotide polymorphism (SNP) and single strand conformation polymorphism (SSCP) markers. The map consists of 146 marker loci distributed across 205 cM, and includes several morphological mutant marker loci. Most of the genetic markers are derived from random cDNAs or *Ae. aegypti* genes of known function. A number of markers are derived from random genomic DNAs, including several cloned RAPD-PCR fragments, and also several cDNAs from *Drosophila melanogaster*. Most of the random cDNAs (80.2%) have high BLASTX sequence identities to known genes, with the majority of matches to genes from *D. melanogaster*. Access to sequence data for all markers will facilitate their continued development for use in high-throughput SNP marker analyses and also provides additional physical anchor points for an anticipated genome sequencing effort.

Keywords: EST, genomics, RFLP, SNP, SSCP.

Introduction

The yellow fever mosquito *Aedes aegypti* remains one of the most intensively studied insects – which is largely a consequence of its direct role as vector for a variety of human pathogens – as well as being an excellent laboratory model for investigating the genetics of vector/pathogen interactions. The first detailed DNA marker-based genetic map for *Ae. aegypti* was constructed using restriction fragment length polymorphism, or RFLP markers (Severson *et al*., 1993). This map consisted of 50 DNA markers that identified 53 RFLP loci across the three linkage groups. Most of these RFLP markers were random cDNA clones or cDNAs of known genes. This map also included several morphological mutant marker loci that allowed its partial integration with an existing genetic map based on meiotic recombination among isozyme and morphological mutant marker loci (Munstermann & Craig, 1979). These and additional RFLP markers have since been applied toward the identification of quantitative trait loci (QTL), influencing *Ae. aegypti* competence to vector the malarial parasite, *Plasmodium gallinaceum* (Severson *et al*., 1995a), and the lymphatic filarioid nematode parasite, *Brugia malayi* (Beerntsen *et al*., 1995; Severson *et al*., 1994a). More recently, Fulton *et al*. (2001) have employed single strand conformation polymorphism (SSCP) analysis to identify single nucleotide polymorphisms (SNP) in cDNA sequences among segregating populations of *Ae. aegypti* and map them genetically. Finally, because *Ae. aegypti* cDNA sequences are highly conserved across mosquito species (Severson *et al*., 1994b), they have been used to construct comparative RFLP linkage maps for several culicine species, including *Aedes albopictus* (Severson *et al*., 1995b), *Armigeres subalbatus* (Ferdig *et al*., 1998), *Culex pipiens* (Mori *et al*., 1999), *Culex tritaeniorhynchus* (Mori *et al*., 2001) and *Ochlerotatus* (formerly *Aedes*) *triseriatus* (Anderson *et al*., 2001).

Because of its importance as a disease vector, *Ae. aegypti* is a good candidate for a systematic genome sequencing project (Knudson *et al*., 2002). Success in sequencing the *Drosophila melanogaster* genome demonstrated the feasibility of such projects, but also illustrated that while sequencing technology has advanced rapidly, genome projects involving large genomes remain dependent on the availability of a physical scaffold for global assembly (Adams *et al*., 2000). Integration of the *Ae. aegypti* genetic and physical maps using fluorescent *in situ* hybridization (FISH) technology represents an important step towards defining the basic genome architecture in terms of putative euchromatic and heterochromatic regions (Brown *et al*., 2001). To date, however, only a small number

Received 28 January 2002; accepted after revision 14 May 2002. Correspondence: Dr David Severson, Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA. Tel.: +1 574 631 3826, fax: +1 574 631 7413, e-mail: David.W.Severson.1@nd.edu

of the genetically mapped RFLP markers have been reported as physical landmark sequences (Severson & Zhang, 1996). Here we report an updated linkage map for *Ae. aegypti*, based largely on RFLP markers, but also including a small number of SNP and SSCP markers, and present an inventory of sequence information for the associated cDNA and genomic clones as expressed sequence tags (EST) and sequence tagged sites (STS), respectively. Putative identities for our random ESTs, based on BLAST analysis, are included.

Results and discussion

The composite linkage map for *Ae. aegypti* consists of 146 marker loci distributed across 205 cM (Fig. 1). The map also includes the Sex determination (Sex) locus on linkage

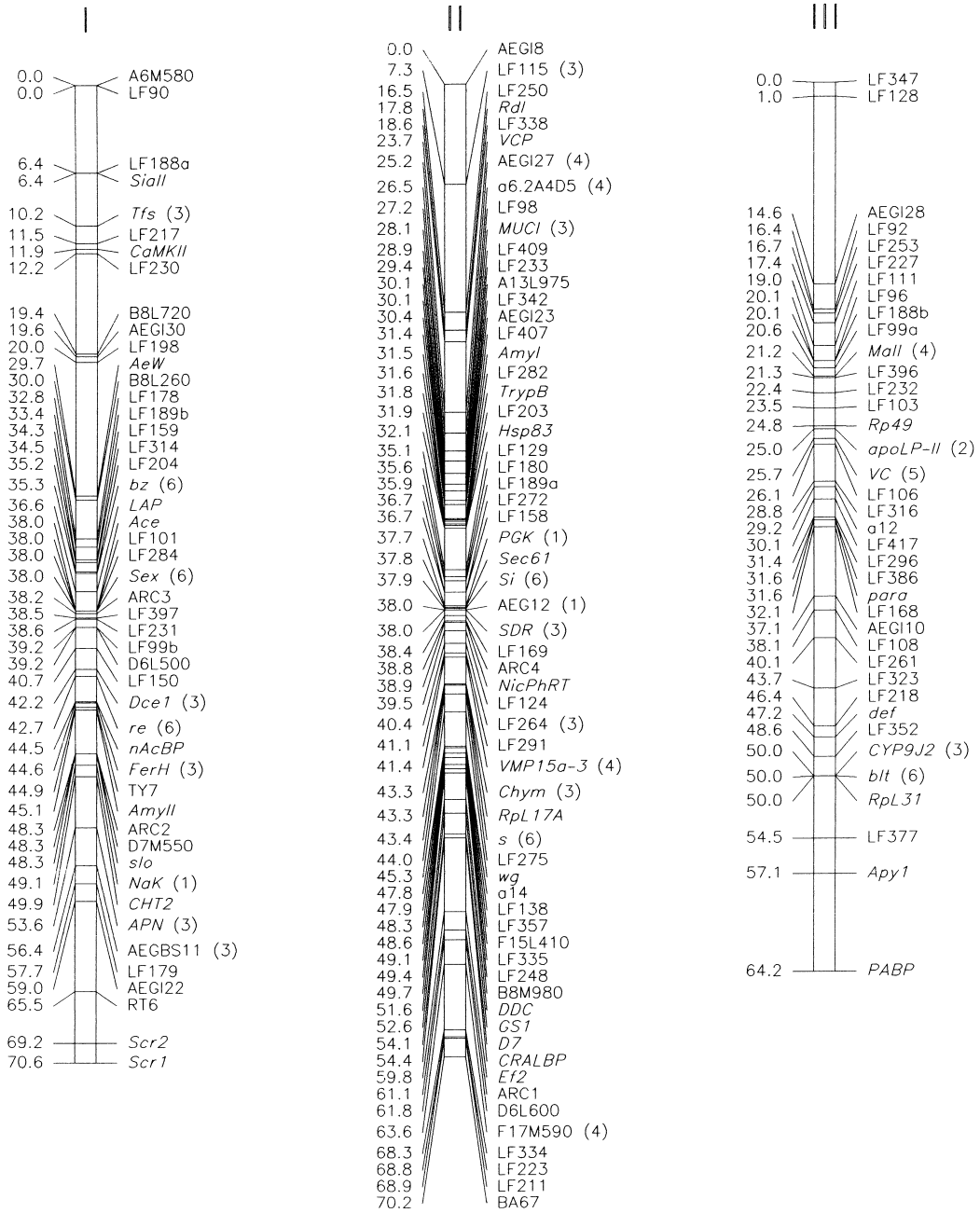

Figure 1. Composite RFLP, SNP and SSCP genetic linkage map for *Aedes aegypti*. Map distances are listed in Kosambi centiMorgans. All markers were mapped as RFLP loci except those indicated by '( )'. (1) = SNP only, (2) = SSCP only, (3) = SNP and RFLP, (4) = SSCP and RFLP, (5) = SSCP and SNP (6) morphological.

group (LG) I and 5 morphological mutant marker loci, red eye (*re*) and bronze (*bz*) on LG I, Silver mesonotum (*Si*) and spot abdomen (*s*) on LG II, and black tarsus (*blt*) on LG III (see Munstermann & Craig, 1979). The linear orders for many of these markers have repeatedly been confirmed across multiple segregating populations. Discrepancies have only been observed among tightly linked loci that most likely reflect the quality of the individual data sets. That is, for such loci, sample size within individual mapping populations or a single mis-typed individual can alter the outcome of multipoint linkage analyses. Furthermore, because the map was constructed from composite data, all

Table 1. Known genes mapped as RFLP, SNP or SSCP marker loci

Clone ID	Chromosome	Gene*
***Aedes aegypti* clones:**		
CHT2	I	Chitinase (AF026492)
Ace	I	Acetylcholinesterase (AAB35001)
AEG12	II	Protein G12 (AY038041)
AeW	I	White-eye (U73826)
AmyI	II	Salivary gland amylase (AF000569)
AmyII	I	Midgut amylase (AF000568)
APN	I	Aminopeptidase N (AF378117)
apoLp-II	III	Apolipophorin II (AF038654)
Apy1	III	Apyrase 1 (L12389)
CaMKII	I	Calcium/calmodulin dependent protein kinase (AF311726)
CRALBP	II	Cellular retinaldehyde-binding protein (AF329893)
Chym	II	Chymotrypsin (AY038039)
		Ae. aegypti LF173 EST (BM005481)
CYP9J2	III	Cytochrome P450 CYP9J (AF329892)
D7	II	Salivary gland protein (M33156)
Dce1	I	Dopachrome conversion enzyme (AF288384)
DDC	II	Dopadecarboxylase (U27581)
def	III	Defensin (AF156088)
		Ae. aegypti LF283 EST (BM005517)
Ef2	II	Elongation factor 2 (AY040342)
FerH	I	Ferritin heavy chain (AF326341)
GS1	II	Glutamine synthetase (AF004351)
LAP	I	Lysosomal aspartic protease (M95187)
MalI	III	Maltase-like enzyme (M30442)
MUCI	II	Mucin-like protein (AF308862)
		Ae. aegypti LF398 EST (BM378048)
nAcBP	I	Nucleic acid binding protein (AY040341)
NaK	I	Sodium/potassium ATPase (AF393727)
NicPhRT	II	Nicotinate phosphoribosyltransferase (AF395330)
PABP	III	Polyadenylate-binding protein (AY038043)
para	III	DDT/pyrethroid resistance (AF468968)
PGK	II	Phosphoglycerate kinase (AY043171)
Rdl	II	Cyclodiene resistance (U28803)
RpL17A	II	Ribosomal protein L17A (AF315597)
		Ae. aegypti LF355 EST (BM005513)
RpL31	III	Ribosomal protein L31 (AF324863)
SDR	II	Short chain dehydrogenase/reductase (AY033621)
Sec61	II	Transport protein Sec61 alpha subunit (AF326338)
SiaI	I	Sialokinin I (AF108099)
slo	I	Slopoke (AF443282)
Tfs	I	Transferrin (AF019117)
		Immune-expressed polypeptide D6.12 (U11235)
		Ae. aegypti LF235 EST (BM005471)
TrypB	II	Trypsin (M77814)
VCP	II	Vitellogenic carboxypeptidase (L46594)
VC	II	Vitellogenin convertase (L46373)
VMP15a-3	II	Vitelline membrane protein (U91682)
***Drosophila melanogaster* clones:**		
Hsp83	II	Heat shock protein 82/83 (X03810)
Rp49	III	Ribosomal protein 49 (X00848)
Scr	I	Sex combs reduced (X05228)
wg	II	Wingless (M17230)

*A representative accession number is provided, also indicating in some cases, those genes initially mapped as unknown function ESTs. For some genes, multiple accessions are available in the GenBank database.

Table 2. Mapped random ESTs with best BLASTX identity matches

Clone ID	Best match GENBANK ID	Gene description	Organism	%AA identity	%AA similarity	BLASTX *E*-value
Chromosome I						
AEGI22|BI099650	None					
AEGI30|BI096856	None					
AEGbS11|AY033622	None					
LF90|T58320	sp.|P14130	Ribosomal protein S14	*Drosophila*	87% (49/56)	92% (52/56)	7e-22
LF99b|BM005477	gb|AF194819	Heat shock cognate protein 79 protein	*Manduca sexta*	72% (113/155)	78% (123/155)	8e-55*
LF101|BM005475	sp.|P45842	Ribosomal protein L34	*Aedes albopictus*	85% (102/120)	90% (109/120)	1e-52
LF150|BM005476	dbj|AB032721	*Pol* protein	*Bombyx mori*	54% (31/57)	76% (44/57)	3e-16†
LF159|T58315	sp.|P29692	Translation elongation factor 1-delta	Human	59% (43/72)	84% (61/72)	7e-10*,†
LF178|T58309	sp.|O61231	QM protein	*Drosophila*	94% (136/144)	96% (139/144)	1e-77
LF179|BM005479	gb|AF145656	Beta-coat protein	*Drosophila*	92% (102/110)	96% (107/110)	1e-61
LF188a|BM005472	gb|AF148814	Translation elongation factor 1-gamma	*Drosophila*	86% (32/37)	94% (35/37)	1e-30
LF189b|BM005474	sp.|Q24154	Ribosomal protein L29	*Drosophila*	64% (49/76)	77% (59/76)	2e-20
LF198|T58319	sp.|P28331	NADH-ubiquinone oxidoreductase	Human	87% (34/39)	87% (34/39)	3e-27
LF204|BM378050	gb|AE003732	CG15697 gene product	*Drosophila*	89% (51/57)	92% (53/57)	7e-16*
LF217|BM005473	gb|AF161709	Ferritin G subunit precursor	*Calpodes ethlius*	45% (20/44)	60% (27/44)	5e-04
LF230|T58326	dbj|AB030307	Haustellum specific protein	*Sarcophaga peregrina*	38% (35/91)	56% (52/91)	5e-09†
LF231|BM005478	sp.|P49630	Ribosomal protein L36	*Drosophila*	100% (36/36)	100% (36/36)	1e-12
LF284|BM005502	sp.|P09180	Ribosomal protein L4	*Drosophila*	68% (74/108)	84% (92/108)	6e-40
LF314|BM005509	sp.|P40945	ADP-ribosylation factor 2, *Arf2*	*Drosophila*	90% (102/113)	96% (109/113)	2e-56
LF397|BM378051	sp.|Q15356	Small nuclear ribonucleoprotein F	Human	86% (56/65)	95% (62/65)	1e-28
Chromosome II						
BA67|AI561370	gb|AF024603	Transcriptional co-repressor SIN3A	*Drosophila*	65% (143/218)	81% (180/218)	1e-80
AEGI8|AF326340	gb|AF309553	Meiotic recombination protein REC14	Human	53% (175/330)	70% (234/330)	5e-86*
AEGI23|AY033624	None					
AEGI27|BG937399	None					
ARC1|R19561	None					
LF98|T58313	gb|AF062077	Syntaxin 8	Human	41% (27/65)	71% (47/65)	1e-11†
LF115|R67978	gi|548749	Ribosomal protein L18	Human	72% (36/50)	90% (45/50)	3e-15
LF124|T58324	sp.|P47904	Ribosomal protein S27	*Xenopus*	84% (53/63)	88% (56/63)	8e-28
LF129|BM005504	sp.|O61462	Ribosomal protein L37A	*Cryptochiton stelleri*	85% (76/89)	90% (81/89)	2e-33*
LF138|T58332	sp.|Q27268	RNA helicase	*Drosophila*	93% (57/61)	97% (60/61)	1e-27
LF158|BM005485	sp.|P23358	Ribosomal protein L12	Rat	77% (80/103)	88% (92/103)	7e-44
LF169|BM378049	None					
LF180|BM005486	gb|AE003559	CG8562 gene product	*Drosophila*	35% (21/59)	62% (37/59)	3e-04
LF189a|BM005474		See chromosome 1 data				
LF203|BM005503	gb|AE003048	CG18001 gene product	*Drosophila*	88% (62/70)	93% (66/70)	1e-28
LF211|BM005514	gb|U80980	Cytidine deaminase	*Brugia malayi*	50% (42/83)	70% (59/83)	1e-17
LF223|BM005515	sp.|P50882	Ribosomal protein L9	*Drosophila*	83% (81/97)	92% (90/97)	1e-39
LF233|T58327	None					
LF248|BM005480	sp.|O44248	Ribosomal protein L5	*Anopheles gambiae*	94% (116/123)	94% (117/123)	6e-53*
LF250|T58311	sp.|P55841	Ribosomal protein L14	*Drosophila*	71% (37/52)	90% (47/52)	1e-16
LF264|BM005483	gb|AE003539	Ribosomal protein S4	*Drosophila*	79% (106/134)	86% (116/134)	5e-62
LF272|BM005484	gb|AF164153	Ribosomal protein S17	*Anopheles gambiae*	92% (73/79)	95% (76/79)	1e-35
LF275|BM005500	None					

Table 2. *continued*

Clone ID	Best match GenBank ID	Gene description	Organism	%AA identity	%AA similarity	BlastX *E*-value
LF282|T58328	None					
LF291|BM005482	sp.|O18640	Guanine nucleotide-binding protein β subunit	*Drosophila*	88% (86/97)	92% (90/97)	1e-49
LF334|BM005506	sp.|Q28852	ATP synthase G chain, mitochondrial	Bovine	51% (49/96)	70% (68/96)	4e-23
LF335|BM005505	gb|AF249871	Farnesoic acid o-methyltransferase	*Panulirus interruptus*	41% (46/112)	58% (66/112)	9e-20
LF338|BM005508	gb|U84745	Cuticle protein Lcp65Ac	*Drosophila*	68% (40/58)	81% (48/58)	5e-19
LF342|BM005512	gb|AF051780	Chymotrypsin 2	*Anopheles darlingi*	35% (53/148)	57% (86/148)	5e-20
LF357|BM005495	None					
LF407|BM005510	gb|AE003550	CG11989 gene product	*Drosophila*	78% (22/28)	95% (27/28)	3e-06†
LF409|BM005511	None					
Chromosome III						
AEGI10|BI096854	None					
AEGI28|BI096849	None					
LF92|BM005493	pir|S23988	Ribosomal protein CEP52	*Drosophila*	100% (91/91)	100% (91/91)	4e-49
LF96|BM005491	dbj|AB007424	Opsin	*Papilio xuthus*	77% (91/117)	85% (101/117)	7e-46
LF99a|BM005477		See chromosome 1				
LF103|BM005488	gb|AE003796	Ribosomal protein L11	*Drosophila*	84% (67/79)	89% (71/79)	1e-31
LF106|BM005490	gb|AE003690	Ribosomal protein S25	*Drosophila*	90% (60/66)	90% (60/66)	9e-28
LF108|T58322	sp.|P39018	Ribosomal protein S19	*Drosophila*	72% (81/111)	84% (95/111)	5e-44
LF111|BM005492	sp.|P41126	Ribosomal protein L13	*Drosophila*	75% (61/81)	88% (72/81)	5e-33
LF128|BM005494	gb|AF218064	Malate dehydrogenase precursor	*Nucella lapillus*	65% (74/113)	74% (85/113)	3e-35
LF168|R47184	sp.|P13008	Ribosomal protein S26	*Drosophila*	90% (74/82)	96% (79/82)	
LF188b|BM005472		See chromosome 1				
LF218|BM005487	sp.|P55828	Ribosomal protein S20	*Drosophila*	87% (93/106)	93% (100/106)	7e-48
LF227|T58323	gb|AI003779	CG11522 gene product	*Drosophila*	50% (73/144)	60% (88/144)	5e-22*
LF232|BM005489	sp.|P18077	Ribosomal protein L35a	Human	64% (29/45)	75% (34/45)	7e-11
LF253|T58331	sp.|P04645	Ribosomal protein L30	Human	80% (62/74)	89% (67/74)	4e-29
LF261|BM378052	sp.|P35035	Trypsin 1	*Anopheles gambiae*	57% (83/145)	69% (101/145)	1e-39*
LF296|BM005501	sp.|P52813	Ribosomal protein S3a	*Anopheles gambiae*	86% (74/86)	91% (79/86)	1e-37
LF316|BM005516	sp.|O95721	Synaptosomal associated protein 29, *SNAP-29*	Human	30% (25/82)	55% (46/82)	6e-13
LF323|BM005507	gb|AF025791	20S proteasome beta2 subunit	*Drosophila*	73% (117/160)	83% (133/160)	6e-61
LF347|T58329	None					
LF352|T58330	gb|AF400202	Ribosomal protein L39	*Spodoptera frugiperda*	93% (44/47)	97% (46/47)	2e-20
LF377|BM005496	None					
LF386|BM005497	sp.|P08526	Ribosomal protein L27	Human	60% (75/124)	76% (95/124)	9e-36
LF396|BM005498	gb|AE003471	CG9134 gene product	*Drosophila*	28% (14/49)	52% (26/49)	7e-08
LF417|BM005499	sp.|Q94530	Ribosomal protein L27a	*Drosophila*	98% (56/57)	100% (57/57)	9e-25*
TY7|R19560	gb|AF079459	Ras-like GTPase, *Rab7*	*Drosophila*	76% (55/92)	84% (61/72)	4e-25

*Filtered *E*-value but unfiltered percentage matches.
†Best ungapped identities.

linked markers have not been compared directly within the same segregating population, and as such, some local misplacements are anticipated. Care was taken during the JoinMap assembly to confirm that known linear orders were maintained. The data for most markers were completely informative, although for a small group (LF99b, LF101, LF150, LF179, LF231 and *AeW*) the data were only partially informative and their map positions are therefore likely to be tentative.

Most of the markers were derived from random cDNAs, although a fair number (45) were derived from known genes of diverse function (Tables 1 and 2). We were also able to successfully directly use four *Drosophila melanogaster* cDNA clones as probes to *Ae. aegypti* DNA, and subsequently, to map their position. These clones were included, as they hybridized to Southern blots of *Ae. aegypti* DNA under high stringency conditions, indicating that they recognized the *Ae. aegypti* orthologue of the *D. melanogaster* gene. A relatively small number (16) of the markers were derived from random genomic DNAs, including several cloned RAPD-PCR fragments (Table 3). The marker distribution across the three linkage groups reflects an average map distance of 70.6, 70.2 and 64.2 cM for LG I, II and III, respectively. It is of interest that the largest linkage group (LG I) is physically the smallest chromosome (McDonald & Rai, 1970; Brown *et al.*, 2001).

Sequence data have been obtained and were submitted to the GenBank EST database for all mapped random cDNA clones and to the GSS database for all mapped random genomic clones. Most of the cDNA clones had high BlastX sequence identities to known genes; only 16 (19.8%) had no matches in the public databases (Tables 1 and 2). Not surprisingly, most best matches were to genes from *Drosophila melanogaster*, including three clones with relatively weak matches to gene products of unknown function. A small number of the clones had high sequence identities to previously described mosquito genes, including both *Aedes* and *Anopheles* spp. This largely reflects the general lack of mosquito sequences in the public databases. Many of the random cDNA clones ($n = 28$) had a high sequence identity to various ribosomal proteins. The remainder are genes involved in a variety of general cell activities.

Table 3. Mapped random genomic clones

Clone ID	Chromosome	Accession no.
a6.2A4D5	II	BH614629
a12	III	BH214530
a14	III	BH214531
ARC2	I	BH214538
ARC3	I	BH214542
ARC4	II	BH214543
A6M580	I	BH214540
A13L975	II	BH214533
B8L260	I	BH214532
B8L720	II	BH214539
B8M980	II	BH214534
D6L500	I	BH214541
D6L600	II	BH214535
F15L410	II	BH214536
F17M590	II	BH214537
RT6	I	BH214544

Our composite linkage map seems to provide a dense coverage of the euchromatic regions of the *Ae. aegypti* genome (Brown *et al.*, 2001). Apparently, few genes are located in the centromeric regions of each chromosome and little recombination occurs across them. Given the present density of genetic markers across all three chromosomes, it also seems likely that few, if any, genes are located on the distal portion of the *q*-arm of chromosome III. It is also of interest that all the genomic DNA sequences that we and others (Bosio *et al.*, 2000; Fulton *et al.*, 2001) have mapped are interspersed within the euchromatic regions of the genome.

Our results indicate that the *Scr* gene is located on the *q*-arm of chromosome I (Fig. 1), and is apparently separated from genes on the *p*-arm by an extensive heterochromatic region (Brown *et al.*, 2001). Because independent studies by Fulton *et al.* (2001) also placed the *AbdA* gene in the same general genome location, it seems likely that Hox gene members of the Antennapedia complex and Bithorax complex are clustered in this region. The available data is insufficient to determine whether the Hox gene cluster is tightly organized in a single cluster as was recently reported for the mosquito, *Anopheles gambiae*, or if they are split into two widely separated but linked groups, as in *D. melanogaster* (Devenport *et al.*, 2000; Powers *et al.*, 2000).

Success in the positional cloning of genes that influence complex phenotypes of interest in *Ae. aegypti* will depend heavily on our abilities to develop PCR-based genetic markers. That is, only a limited number of Southern blot-based RFLP markers can be evaluated for individual mosquitoes. The development of microsatellite markers has proven difficult for *Ae. aegypti* (Fagerberg *et al.*, 2001), indicating that alternative marker systems must be developed. The successful use of SSCP analysis for the detection of SNPs demonstrates that these polymorphisms are abundant and useful for *Ae. aegypti* (Bosio *et al.*, 2000; Fulton *et al.*, 2001). Access to EST and STS sequence data for our RFLP markers will facilitate their development as SNP-based markers. Indeed, we have screened 32 cDNA sequences and determined that 24 (75%) contained at least one SNP within the regions compared in laboratory strain mosquitoes, and are successfully developing them as genetic markers (Meece *et al.*, unpublished data). Because SNPs are highly abundant, and can be adapted for high-throughput analysis, it seems clear that they will become the genetic marker of choice for most organisms, including *Ae. aegypti* (Landegren *et al.*, 1998; Brookes, 1999).

Experimental procedures

Sources of clones and linkage analysis

The clones reported in this study were obtained from a wide variety of sources, including a previously described random cDNA library (Severson *et al*., 1993), cDNAs for known genes from *Ae. aegypti* and *D. melanogaster* isolated in our laboratory or received as gifts from other investigators or via PCR from genomic DNA using GenBank sequence data, genomic clones generated by targeted marker development (Severson *et al*., 1999), and random genomic clones serendipitously obtained in our laboratory. Map positions for some clones have been reported previously (Beerntsen *et al*., 1995; Grossman *et al*., 1997; Severson *et al*., 1993, 1994a, 1995a, 1997, 1999), but have not been collectively integrated. RFLP analysis on individual segregating populations from a variety of independent crosses was performed using our standard protocol (Severson, 1997). Twenty-one crosses provided two-point linkage estimates for markers on chromosomes I and III, while 31 crosses provided two-point linkage estimates for markers on chromosome II. The cDNA sequences for SSCP analysis were selected from Fulton *et al*. (2001) and the SSCP analysis followed Bosio *et al*. (2000). To identify informative SNP loci, a target sequence was amplified from genomic DNA from each parent of a segregating population, and subjected to direct cycle sequencing (see below). Thereafter, SNP analysis followed the McSNP protocol (Akey *et al*., 2001). McSNP products from individual segregants were visualized and scored used a gel-free system (Hybaid DASH System) that identifies alleles based on melting temperature calculations of the PCR products directly in 96-well microplates. A composite linkage map was assembled based on all our previously reported mapping efforts, as well as on unpublished data using the JoinMap computer program (Stam, 1993). The JoinMap output files containing all composite two-point recombination frequencies and associated standard errors are available by request (contact D.W. Severson). The linkage map was drawn using the DrawMap computer program (van Ooijen, 1994). Map distances are reported in Kosambi centiMorgans.

Sequence analysis and database comparisons

Recombinant plasmids were purified using the alkaline lysis technique (Sambrook *et al*., 1989) and subjected to cycle sequencing using the ABI Prism Big Dye Terminator kit (PE Applied Biosystems), using sequencing primers specific to the plasmid vectors. Clones from orientated libraries were sequenced from the 5′ ends, although for some clones the sequence data for 3′ ends also was obtained. DNA sequences were submitted to the BlastX program (Altschul *et al*., 1997) for homology searches with the non-redundant sequence database using the default BLOSUM62 matrix and other default settings. Selected sequences were re-analysed using the gapped or the unfiltered settings. Data for the expressed sequences have been submitted to the GenBank EST database, and data for the random genomic sequences have been submitted to the GenBank GSS database.

Acknowledgements

This work was supported by grants AI33127, AI34337 and AI45123 from the National Institutes of Health. We thank Jenny Walerak, Sara Cywinski, Jill Williams and Kristen Jones for technical assistance, and William Black for assistance with the SSCP analysis.

References

Adams, M.D., Celniker, S.E., Holt, R.A., *et al*. (2000) The genome sequence of *Drosophila melanogaster*. *Science* **287**: 2185–2195.

Akey, J.M., Sosnoski, D., Parra, E., Dios, S., Hiester, K., Su, B., Bonilla, C., Jin, L. and Shriver, M.D. (2001) A melting curve analysis of SNPs (McSNP): a gel-free and inexpensive approach for SNP genotyping. *Biotechniques* **30**: 358–367.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl Acids Res* **25**: 3389–3402.

Anderson, J.R., Grimstad, P.R. and Severson, D.W. (2001) Chromosomal evolution among six mosquito species (Diptera: Culicidae) based on shared restriction fragment length polymorphisms. *Mol Phylogenet Evol* **20**: 316–321.

Beerntsen, B.T., Severson, D.W., Klinkhammer, J.A., Kassner, V.A. and Christensen, B.M. (1995) *Aedes aegypti*: a quantitative trait locus (QTL) influencing filarial worm intensity is linked to QTL for susceptibility to other mosquito-borne pathogens. *Exp Parasitol* **81**: 355–362.

Bosio, C.F., Fulton, R.E., Salasek, M.L., Beaty, B.J. and Black, W.C.I.V. (2000) Quantitative trait loci that control vector competence for dengue-2 virus in the mosquito *Aedes aegypti*. *Genetics* **156**: 687–698.

Brookes, A.J. (1999) The essense of SNPs. *Gene* **234**: 177–186.

Brown, S.E., Knudson, D.L. and Severson, D.W. (2002) Culicine genomics. *Insect Biochem Mol Biol*, in press.

Brown, S.E., Severson, D.W., Smith, L.A. and Knudson, D.L. (2001) Integration of the *Aedes aegypti* genetic linkage and physical maps. *Genetics* **157**: 1299–1305.

Devenport, M.P., Blass, C. and Eggleston, P. (2000) Characterization of the Hox gene cluster in the malaria vector mosquito, *Anopheles gambiae*. *Evol Dev* **2**: 326–339.

Fagerberg, A.J., Fulton, R.E. and Black, W.C.I.V. (2001) Microsatellite loci are not abundant in all arthropod genomes: analyses in the hard tick, *Ixodes scapularis* and the yellow fever mosquito, *Aedes aegypti*. *Insect Mol Biol* **10**: 225–236.

Ferdig, M.T., Taft, A.S., Severson, D.W. and Christensen, B.M. (1998) Development of a comparative genetic linkage map for *Armigeres subalbatus* using *Aedes aegypti* RFLP markers. *Genome Res* **8**: 41–47.

Fulton, R.E., Salasek, M.L., DuTeau, N.M. and Black, W.C.I.V. (2001) SSCP analysis of cDNA markers provides a dense linkage map of the *Aedes aegypti* genome. *Genetics* **158**: 715–726.

Grossman, G.L., Campos, Y., Severson, D.W. and James, A.A. (1997) Evidence for two distinct members of the amylase gene family in the yellow fever mosquito, *Aedes aegypti*. *Insect Mol Biol Biochem* **27**: 769–781.

Landegren, A., Nilsson, M. and Kwok, P.-Y. (1998) Reading bits of genetic information: methods for single-nucleotide polymorphism analysis. *Genome Res* **8**: 769–776.

McDonald, P.T. and Rai, K.S. (1970) Correlation of linkage groups with chromosomes in the mosquito, *Aedes aegypti*. *Genetics* **66**: 475–485.

Mori, A., Severson, D.W. and Christensen, B.M. (1999) Comparative linkage maps for the mosquitoes, *Culex pipiens* and *Aedes aegypti*, based on common RFLP loci. *J Hered* **90**: 160–164.

Mori, A., Tomita, T., Hidoh, O., Kono, Y. and Severson, D.W. (2001) Comparative map development and identification of an autosomal locus for insensitive acetylcholinesterase-mediated insecticide resistance in *Culex tritaeniorhynchus*. *Insect Mol Biol* **10**: 197–203.

Munstermann, L.E. and Craig, G.B. Jr (1979) Genetics of *Aedes aegypti*: updating the linkage map. *J Hered* **70**: 291–296.

Powers, T.P., Hogan, J., Ke, Z., Dymbrowski, K., Wang, X., Collins, F.H. and Kaufman, T.C. (2000) Characterization of the Hox cluster from the mosquito *Anopheles gambiae* (Diptera: culicidae). *Evol Dev* **2**: 311–325.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Press, Cold Spring Harbor, NY.

Severson, D.W. (1997) RFLP analysis of insect genomes. In *The Molecular Biology of Insect Disease Vectors: a Methods Manual* (Crampton, J.M., Beard, C.B. and Louis, C., eds), pp. 309–320. Chapman & Hall, London.

Severson, D.W., Anthony, N.M., Andreev, O. and ffrench-Constant, R.H. (1997) Molecular mapping of insecticide resistance genes in the yellow fever mosquito *Aedes aegypti*. *J Hered* **88**: 520–524.

Severson, D.W., Mori, A., Kassner, V.A. and Christensen, B.M. (1995b) Comparative linkage maps for the mosquitoes, *Aedes albopictus* and *Aedes aegypti*, based on common RFLP loci. *Insect Mol Biol* **4**: 41–45.

Severson, D.W., Mori, A., Zhang, Y. and Christensen, B.M. (1993) Linkage map for *Aedes aegypti* using restriction fragment length polymorphisms. *J Hered* **84**: 241–247.

Severson, D.W., Mori, A., Zhang, Y. and Christensen, B.M. (1994a) Chromosomal mapping of two loci affecting filarial worm susceptibility in *Aedes aegypti*. *Insect Mol Biol* **3**: 67–73.

Severson, D.W., Mori, A., Zhang, Y. and Christensen, B.M. (1994b) The suitability of RFLP markers for evaluating genetic diversity among and synteny between mosquito species. *Am J Trop Med Hyg* **50**: 425–432.

Severson, D.W., Thathy, V., Mori, A., Zhang, Y. and Christensen, B.M. (1995a) Restriction fragment length polymorphism mapping of quantitative trait loci for malaria parasite susceptibility in the mosquito *Aedes aegypti*. *Genetics* **139**: 1711–1717.

Severson, D.W., Zaitlin, D. and Kassner, V.A. (1999) Targeted identification of markers linked to malaria and filarioid nematode parasite resistance genes in the mosquito *Aedes aegypti*. *Genet Res* **73**: 217–224.

Severson, D.W. and Zhang, Y. (1996) Generation of expressed sequence tags and sequence-tagged sites as physical landmarks in the mosquito, *Aedes aegypti*, genome. *Genome* **39**: 224–229.

Stam, P. (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. *Plant J* **3**: 739–744.

van Ooijen, J.W. (1994) DrawMap: a computer program for drawing genetic linkage maps. *J Hered* **85**: 66.