Comparative Genome Analysis of the Yellow Fever Mosquito *Aedes aegypti* with *Drosophila melanogaster* and the Malaria Vector Mosquito *Anopheles gambiae*

D. W. SEVERSON, B. DEBRUYN, D. D. LOVIN, S. E. BROWN, D. L. KNUDSON, AND I. MORLAIS

From the Center for Tropical Disease Research and Training, Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556 (Severson, deBruyn, Lovin, and Morlais); and from the Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO 80523 (Brown and Knudson). This work was supported by the grants AI33127 and AI34337 from the National Institutes of Health.

Address correspondence to D. W. Severson at the above address, or e-mail: David.W.Severson.1@nd.edu.

Abstract

An *in silico* comparative genomics approach was used to identify putative orthologs to genetically mapped genes from the mosquito, *Aedes aegypti*, in the *Drosophila melanogaster* and *Anopheles gambiae* genome databases. Comparative chromosome positions of 73 *D. melanogaster* orthologs indicated significant deviations from a random distribution across each of the five *A. aegypti* chromosomal regions, suggesting that some ancestral chromosome elements have been conserved. However, the two genomes also reflect extensive reshuffling within and between chromosomal regions. Comparative chromosome positions of *A. gambiae* orthologs indicate unequivocally that *A. aegypti* chromosome regions share extensive homology to the five *A. gambiae* chromosome arms. Whole-arm or near-whole-arm homology was contradicted with only two genes among the 75 *A. aegypti* genes for which orthologs to *A. gambiae* were identified. The two genomes contain large conserved chromosome segments that generally correspond to break/fusion events and a reciprocal translocation with extensive paracentric inversions evident within. Only very tightly linked genes are likely to retain conserved linear orders within chromosome segments. The *D. melanogaster* and *A. gambiae* genome databases therefore offer limited potential for comparative positional gene determinations among even closely related dipterans, indicating the necessity for additional genome sequencing projects with other dipteran species.

The genomics era offers tremendous opportunities to improve our understanding of gene structure, function, and location, and to allow rapid dissections of the interactions among genes that produce a variety of complex phenotypes. As the availability of genome information for individual species grows, comparative genomics also will provide valuable insights into chromosomal and organismal evolution (Andersson et al. 1996; Nadeau and Sankoff 1998). Of particular interest are comparative relationships among orthologous genes and their associated regulatory elements across diverse taxa. *Synteny* refers to genes that reside on the same chromosome. *Conserved synteny* indicates that homologous genes are syntenic between species, regardless of gene

order, while conserved linkage indicates conservation of both synteny and linear orders of genes between species (Ehrlich et al. 1997). The more closely related two species are, the greater the expected size of chromosome segments containing conserved linkages. For example, exploitation of conserved linkages between human and mouse is proving invaluable for understanding genetic disorders in humans via their mouse orthologs (Boyd et al. 2000; McPeek 2000; Nadeau 1989).

Several members of the mosquito family, Culicidae, have been or likely will be targeted for complete genome analyses, because of their importance as disease vectors. These include *Anopheles gambiae*, *Aedes aegypti*, and *Culex pipiens*, the primary vectors for malaria, yellow fever and dengue, and lymphatic

filariasis transmission to humans, respectively. Indeed, the complete genome sequence was recently reported for *A. gambiae* (Holt et al. 2002), and an *A. aegypti* genome project has been initiated (Knudson et al. 2002). The Culicidae is divided into three subfamilies, Anophelinae, Culicinae, and Toxorhynchitinae, with all disease vectors found within the Anophelinae and Culicinae. *A. gambiae* and *A. aegypti* are representatives of the Anophelinae and Culicinae, respectively. Because the Culicidae contains about 3,500 species (Knight and Stone 1977), only a restricted group will become subjects for complete genome analysis. Using a comparative genomics approach, one may be able, however, to compare genome information from these species, with draft-level genome efforts for others, to determine gene locations and functions in other mosquito species of interest, rapidly and economically.

Until recently, only limited evidence for synteny conservation between the higher dipterans, such as *Drosophila melanogaster*, and the lower dipterans, like mosquitoes, has been reported (Fulton et al. 2001; Matthews and Munstermann 1994; Weller and Foster 1993). The availability of the complete sequence for the euchromatic regions of the *D. melanogaster* genome (Adams et al. 2000) and the *A. gambiae* genome (Holt et al. 2002) provides gene-order scaffolds that can be used for comparisons with other insects, particularly other dipterans, including other mosquitoes. The comparative analysis of *D. melanogaster* and *A. gambiae* identified considerable genome-wide synteny conservation, but only limited evidence for gene-order conservation (Bolshakov et al. 2002; Zdobnov et al. 2002). In this study, an *in silico* comparative genomics approach was used to identify orthologs to genetically mapped and sequenced *A. aegypti* genes (Severson et al. 2002) in the *D. melanogaster* and *A. gambiae* genome databases. We identified chromosome segments of long-range, or macrosynteny, conservation between these species, and some evidence for local or microsynteny conservation within chromosome segments, but limited evidence for gene-order conservation.

Materials and Methods

Molecular Markers

Marker data and assembly of the *A. aegypti* composite linkage map are described elsewhere (Severson et al. 2002). Most of these markers are cDNA clones for known genes from *A. aegypti* or *D. melanogaster*, or random expressed sequence tags (ESTs) from *A. aegypti*. Three established criteria (Lalley et al. 1987) were used to identify candidate cDNAs as orthologs between *A. aegypti*, *D. melanogaster*, and *A. gambiae*: (1) orthologous gene designations based on high sequence similarities; (2) use of cloned genes from *A. aegypti* that were isolated with degenerate primers to *D. melanogaster* gene sequences or with use of *D. melanogaster* clones as heterologous low-stringency probes to screen *A. aegypti* libraries; or by (3) use of *D. melanogaster* cDNA clones directly as probes to *A. aegypti* Southern blots under high-stringency

hybridizations in restriction fragment length polymorphism (RFLP) mapping efforts.

Orthology and Synteny Determination

All *A. aegypti* cDNA sequences were initially submitted to the BLASTX program (Altschul et al. 1997) for homology searches against the *D. melanogaster* database and are accessible in the GenBank database (Severson et al. 2002). Searches were performed with use of the PAM30 substitution matrix, with the gap penalty set to the maximum (10). The PAM30 matrix was selected because it is based on a simple evolutionary model and should identify sequences with a high degree of similarity (Altschul 1991). *A. aegypti* sequences were arbitrarily considered to be orthologous to *D. melanogaster* sequences when the expected number (E) was less than e^{-16}. *A. aegypti* sequences that could unambiguously be assigned to chromosome position in *D. melanogaster* were then submitted to the TBLASTN program for homology searches against the *A. gambiae* database, with use of the default parameters. *A. aegypti* sequences were arbitrarily considered to be orthologous to *A. gambiae* sequences when the expected number (E) was less than e^{-9}. The *D. melanogaster* database was screened first because of the extensive gene annotation available compared to *A. gambiae*. This also allowed for identification of the correct reading frame for the *A. aegypti* sequence and subsequent amino acid translation for TBLASTN searches. Two *A. aegypti* cDNA sequences with no orthologs in *D. melanogaster* were included in the *A. gambiae* comparison, because they represent the only markers located near the end of the *p* arm of chromosome 2 in *A. aegypti* (Severson et al. 2002).

Evidence for the nonrandom distribution of *D. melanogaster* orthologs across the *A. aegypti* genome was tested with the binomial test (http://home.clara.net/sisa/binomial.htm). This test is based on the exact probability distribution of a single dichotomy and is particularly applicable to small sample sizes (Sokal and Rohlf 1981). The *A. aegypti* cDNAs were assessed with regard to their general orthologous chromosome locations in *D. melanogaster* (e.g., X, 2L, 2R, 3L, 3R, 4), and the observed frequencies were calculated for each location. Centromere positions on the *A. aegypti* chromosomes were estimated according to integrated genetic and physical map data (Brown et al. 2001), which allowed us to define 5 chromosome regions for comparison with *D. melanogaster*, including chromosome 1, the *p* arm and *q* arm of chromosome 2, and the *p* arm and *q* arm of chromosome 3. Note that *A. aegypti* (and other Culicinae) do not exhibit sex chromosome dimorphism, but instead sex determination is controlled by a single autosomal locus on chromosome 1 (Gilchrist and Haldane 1947), with relatively equal recombination rates in both males and females (Severson et al. 1993). The expected number of orthologs for each *A. aegypti* chromosome region was calculated with use of the observed frequencies of *D. melanogaster* orthologs and the total number of genes that mapped to each chromosome region in *A. aegypti*. We tested the null hypothesis that the genes within the

designated *D. melanogaster* chromosome locations are randomly distributed across the *A. aegypti* genome.

Microsynteny Determination

An *A. aegypti* region on chromosome 1 (FL*p*ter ~25%, where FL*p*ter represents the fractional length from the operationally defined *p* terminus) targeted in an ongoing contig construction effort and, where 134 cosmid or BAC end sequences were available, was compared to the *A. gambiae* predicted proteins by BLASTX, with use of the same parameters as described above. The *A. gambiae* predicted proteins that were used and their chromosome location came from the Ensembl build version 10.2.1 of the *A. gambiae* whole genomic sequence (www.ensembl.org).

Results

Synteny Conservation Between *A. aegypti* and *D. melanogaster*

We performed BLASTX analysis on 121 genetically mapped (Severson et al. 2002) *A. aegypti* genes to identify those with *D. melanogaster* orthologs. The mapped genes also included those from the direct screening (as RFLP markers) of three *D. melanogaster* cDNAs in *A. aegypti* segregating populations; these were included as identified orthologs. With a highly conservative substitution matrix and BLAST analysis conditions, a total of 87 genes (72%) with *D. melanogaster* orthologs were identified. A small number of these genes (14) had multiple possible orthologs and represent members of multigene families. For these, individual family members are distributed across different chromosomes or chromosome arms in *D. melanogaster* and could not unambiguously be evaluated for synteny conservation with *A. aegypti*. Unambiguous chromosome positions in *D. melanogaster* could be assigned for 73 genes (Table 1).

Comparative chromosome positions of the *D. melanogaster* orthologs indicated significant deviations from a random distribution across each of the five *A. aegypti* chromosome regions, suggesting that some ancestral chromosome elements have been conserved (Table 2). However, it is also very clear that the two genomes reflect extensive reshuffling within and between chromosome regions. *A. aegypti* chromosome 1 carries a greater than expected number of orthologs from the *D. melanogaster* X chromosome ($P = .0216$) and a slight but nonsignificant excess of orthologs from *D. melanogaster* 3R ($P = .0543$). The comparative ortholog positions for chromosomes 2 and 3 in both organisms are consistent with whole-arm translocations during chromosome evolution. With *A. aegypti* chromosome 2, 2*p* contains a significant excess of orthologs to *D. melanogaster* 3L ($P = .0021$), and 2*q* contains a significant excess of orthologs to *D. melanogaster* 2L ($P = .00005$). Conversely, chromosome 3*p* contains a significant excess of orthologs to *D. melanogaster* 3R ($P = .0139$), and 3*q* contains a significant excess of orthologs to *D. melanogaster* 2R ($P = .0477$). Only four orthologs to the dot chromosome 4 in

D. melanogaster were identified, providing little power for statistical comparisons.

Synteny Conservation Between *A. aegypti* and *A. gambiae*

We performed TBLASTN analysis of the *A. gambiae* genome database with the 73 *A. aegypti* genes that we were able to unambiguously assign to chromosome positions in *D. melanogaster*, and with two genes (*AEGI8* and *LF115*) that have no *D. melanogaster* homolog and represent the only genetic markers for the end of *A. aegypti* 2*p*. The 75 *A. aegypti* genes identified 78 orthologs in *A. gambiae* (Table 1). We used TBLASTN searches because gene annotation for *A. gambiae* was based largely on *in silico* gene-finding programs and preliminary BLASTX searches identified likely errors in annotation of some putative genes, usually use of the wrong reading frame. We also used a less stringent cutoff for the *A. gambiae* TBLASTN searches than that used for BLASTX *D. melanogaster* searches, because observed sequence similarities with TBLASTN would be impacted in some genes by the presence of introns and BLASTX matches with the conservative PAM30 matrix will give lower E values than TBLASTN with the default BLOSUM matrix for matches to the same sequence. Three *A. aegypti* genes (*LF90, LF178*, and *VCP*) each had essentially equally high scoring sequence identities to duplicate and physically independent genes in *A. gambiae*. It seems probable that these genes are also represented as duplications in the *A. aegypti* genome; our efforts to develop physical contigs of BAC clones around the *LF178* locus have confirmed the existence of two gene copies (Brown SE, deBruyn BS, Severson DW and Knudson DL, unpublished data). The physical distance between copies has not been determined (also see results for the microsynteny investigations), but our results from multiple genetic mapping experiments (Severson et al. 2002) indicate very low levels of recombination between them (less than 1 cM).

Comparative chromosome positions of the *A. gambiae* orthologs indicate unequivocally that the five *A. aegypti* chromosome regions share extensive homology with the five *A. gambiae* chromosome arms (Figure 1). The conserved synteny pattern is consistent with whole- or near-whole-arm homologies that generally correspond to break/fusion events and a reciprocal translocation with extensive paracentric inversions evident within chromosome segments. *A. aegypti* chromosome 1 contains all orthologs identified on the *A. gambiae* X chromosome and about half the orthologs identified on the 2R arm. Most of the orthologs to the *A. gambiae* X are located toward the telomere end of the *p* arm of *A. aegypti* chromosome 1. *A. aegypti* chromosomes 2 and 3 reflect a reciprocal translocation with *A. gambiae* chromosomes 2 and 3. *A. aegypti* 2*p* and 2*q* correspond to *A. gambiae* 2L and 3R, respectively. Conversely, *A. aegypti* 3*p* and 3*q* correspond to segments of *A. gambiae* 2R and the entire 3L arm, respectively. The *A. aegypti* chromosomes 1 and 3*p* conserved syntenies with the *A. gambiae* X and 2R suggest break and fusion events in their ancestral chromosome lineages. This phenomenon is supported by chromosome locations of the putative duplicate loci, *LF90* and *LF178*.

Table 1. Mapped *A. aegypti* clones with *D. melanogaster* and *A. gambiae* orthologs

A. aegypti clone ID	*A. aegypti* (cM) map position	*Drosophila* gene description	*Drosophila* chromosome	BLASTX E value	*A. gambiae* protein ID	*A. gambiae* chromosome	TBLASTN E value
Chromosome 1							
LF90|T58320	0.0	Ribosomal protein S14 (*RpS14*)|P14130	X (7C–D1)	6^{e-40}	EAA06897	X (1A–5A)	6^{e-23}
					EAA08220	2R (11B–14C)	2^{e-22}
Tfs|AF019117	10.2	Transferrin (*Tfs*)|AAC67389	X (17A10–11)	1^{e-128}	EAA06303	X (1D–4C)	1^{e-138}
CamKII|AF311726	11.9	Calcium/calmodulin-dependent protein kinase (*CamKII*)|AAF59388	4 (102F3–4)	1^{e-111}	EAA06500	X (1D–4C)	4^{e-66}
LF198|T58319	20.0	NADH-ubiquinone Oxidoreductase (*ND75*)|Q94511	X (7D18-20)	4^{e-42}	EAA00921	2R (7A–10D)	2^{e-12}
AeW|U73826	29.7	White eye (w)|X02974	X (3C2)	2^{e-36}	EAA06417[a]	X (1D–4C)	3^{e-96}
LF178|T58309	32.8	QM protein homology (*QM*)|O61231	3L (80A3–4)	1^{e-135}	EAA04923	X (5A–6)	6^{e-51}
					EAA08084[a]	2R (11B–14C)	6^{e-51}
LF314|BM005509	34.5	ADP-ribosylation factor 2 (*Arf102F*)|P40945	4 (102F1)	2^{e-98}	EAA08117	2R (11B–14C)	4^{e-54}
LF204|BM378050	35.2	CG15697 gene product|AE003732	3R (92F13–93A1)	9^{e-25}	EAA06802	X (1A–5A)	2^{e-23}
LAP|M95187	36.6	Cathepsin D (*cathD*)|AF220040	2R (43E18–F1)	0.0	EAA03535	2R (14C–14D)	0.0
Ace|AAB35001	38.0	Acetylcholinesterase (*Ace*)|X05893	3R (87E3–4)	0.0	EAA06531	X (1D–4C)	1^{e-150}
LF101|BM005475	38.0	CG9354 gene product|AAF54369	3R (85D17–19)	3^{e-74}	EAA05780	3R (33D–34A)	8^{e-38}
LF284|BM005502	38.0	Ribosomal protein L1 (*RpL1*)|AAG22173	3R (92B2)	2^{e-61}	EAA07484	2R (11B–14C)	2^{e-41}
LF397|BM378051	38.5	Small nuclear riboprotein (*DebB*)|Q24297	2R (48F1)	6^{e-48}	EAA07677	2R (11B–14C)	2^{e-42}
LF231|BM005478	38.6	Ribosomal protein L36 (*RpL36*)|P49630	X (1B10)	1^{e-28}	EAA08114	2R (11B–14C)	3^{e-20}
FerH|AF32634	44.6	Ferritin heavy chain (*Fer1HCH*)|U91524	3R (99F4–5)	3^{e-31}	EAA08169	2R (11B–14C)	3^{e-51}
TY7|R19560	44.9	Rab-protein 7 (*Rab7*)|AF079459	3R (95D1–3)	1^{e-43}	EAA03119	2R (8C–8C3)	2^{e-27}
slo|AF443282	48.3	Slowpoke (*slo*)|JH0697	3R (96A17–20)	0.0	EAA08854	2R (15D–16A)	3^{e-58}
NaK|AF393727	49.1	ATPase (*Atpα*)|AE003732	3R (93A7–B1)	1^{e-166}	EAA07474	2R (11B–14C)	1^{e-79}
LF179|BM005479	57.7	β-coatomer protein (β*Cop*)|AF145656	X (17A10–11)	1^{e-109}	EAA01097	2R (7A–10D)	1^{e-59}
Scr1/Scr2[b]	69.2/70.6	Sex combs reduced (*Scr*)|X77075	3R (84A5–B1)	—	EAA07260[c]	2R (19D)	6^{e-53c}
Chromosome 2							
AEGI8|AF326340	0.0	—			EAA03987	2L (25D–28D)	1^{e-153}
LF115|R67978	7.3	—			EAA04761	2L (25D–28D)	6^{e-14}
LF250|T58311	16.5	Ribosomal protein L14 (*RpL14*)|P55841	3L (66D8)	5^{e-29}	EAA11356	2L (22B–25B)	7^{e-9}
Rdl|U28803	17.8	Cyclodiene resistance (*Rdl*)|P25123	3L (67A1)	0.0	EAA11666	2L (22B–25B)	2^{e-69}
LF338|BM005508	18.6	Cuticle protein (*Lcp65Ac*)|U84745	3L (65A5)	1^{e-28}	EAA11701	2L (22B–25B)	3^{e-24}

Table 1. Continued

A. aegypti clone ID	A. aegypti (cM) map position	Drosophila gene description	Drosophila chromosome	BLASTX E value	A. gambiae protein ID	A. gambiae chromosome	TBLASTN E value
VCP|L46594	23.7	CG4572 gene product|AAF55705	3R (92B6)	1^{e-105}	EAA11956	2L (22B–25B)	1^{e-142}
					EAA04657	2L (25D–28D)	1^{e-135}
LF203|BM005503	31.9	CG18001 gene product|AE003048	2R (41C)	1^{e-52}	EAA13878	3R (36C)	1^{e-32}
Hsp83[b]	32.1	Heat shock protein 83 (Hsp83)|X03810	3L (63B13–C1)	—	EAA04769[a]	2L (25D–28D)	0.0^{c}
LF129|BM005504	35.1	CG5827 gene product|AAF52217	2L (25C3–4)	1^{e-58}	EAA13840	3R (34A–37D)	2^{e-41}
LF158|BM005485	36.7	CG3195 gene product|AAF47152	2R (60B2–3)	4^{e-76}	EAA13967	3R (34A–37D)	5^{e-29}
LF272|BM005484	36.7	Ribosomal protein S17 (RpS17)|P17704	3L (67B3–4)	5^{e-50}	EAA09708	2L (20D)	2^{e-34}
PGK|AY043171	37.7	Phosphoglycerate kinase (Pgk)|Q01604	2L (A6–B1)	0.0	EAA10208	3R (30E–32A)	1^{e-149}
Sec61|AF326338	37.8	SRP-dependent protein Transporter (Sec61alpha)|AB062670	2L (24D4–7)	0.0	EAA14690	3R (31D–33D)	0.0
NicPbRT|AF395330	38.9	Nicotinate Phosphoribosyltransferase (NicPbRT)|AE003577	2L (24D6–7)	0.0	EAA12203	3R (29A–31C)	1^{e-123}
LF124|T58324	39.5	CG10423 gene product|AAF56428	3R (96C3–5)	6^{e-50}	EAA04241[a]	2L (25D–28D)	3^{e-33}
LF264|BM005483	40.4	Ribosomal protein S4 (RpS4)|AAF49846	3L (69F2)	5^{e-98}	EAA04244	2L (25D–28D)	8^{e-45}
LF291|BM005482	41.1	Receptor of activated protein kinase C1 (Rack1)|U96491	2L (28D1–2)	3^{e-78}	EAA13872	3R (34A–37D)	1^{e-53}
RpL17A|AF315597	43.3	Ribosomal protein L17A (RpL17A)|JC1253	2R (59B1–2)	1^{e-97}	EAA13962	3R (34A–37D)	3^{e-69}
wg[b]	45.3	Wingless (wg)|J03650	2L (27F1–3)	—	EAA14180	3R (34A–37D)	1^{e-73c}
LF138|T58332	47.9	RNA helicase (Hel25E)|Q27268	2L (25E6–F1)	6^{e-55}	EAA14744	3R (31D–33D)	1^{e-24}
LF335|BM005505	49.1	CG10527 gene product|AAF46666	2R (57B17–20)	4^{e-56}	EAA14604	3R (34A–37D)	6^{e-47}
LF248|BM005480	49.4	Yippee interacting protein 6 (yip6)|AAF45418	2L (40B–D)	3^{e-76}	EAA14773	3R (31D–33D)	4^{e-63}
DDC|U27581	51.6	Dopa decarboxylase (DDC)|P05031	2L (37C1)	0.0	EAA14728	3R (31D–33D)	0.0
GS1|AF004351	52.6	Glutamine synthetase 1 (Gs1)|AAF51546	2L (21B1)	0.0	EAA14864	3R (31D–33D)	0.0
CRALBP|AF329893	54.4	CG5958 gene product|AAF52516	2L (27F–28A1)	1^{e-125}	EAA05798	3R (33D–34D)	1^{e-86}
Ef2|AY040342	59.8	Transcription elongation factor 2b (Ef2b)|P13060	2L (39E5–6)	0.0	EAA03632	3R (34A–34B)	0.0
LF334|BM005506	68.3	CG6105 gene product|AAF53041	2L (32C4–5)	4^{e-63}	EAA05846[a]	3R (34B–34D)	1^{e-39}
LF223|BM005515	68.8	Ribosomal protein L9 (RpL9)|JC6062	2L (32C5–D1)	2^{e-67}	EAA05902	3R (34B–34D)	9^{e-48}
LF211|BM005514	68.9	CG8360 gene product|AAF52621	2L (28F1–4)	9^{e-33}	EAA05898	3R (34B–34D)	8^{e-29}
BA67|AI561370	70.2	RNA polymerase II transcription factor (Sin3A)|AE003821	2R (49B2–3)	1^{e-125}	EAA12860	3R (29A–31C)	3^{e-96}

Table 1. Continued

A. aegypti clone ID	A. aegypti (cM) map position	Drosophila gene description	Drosophila chromosome	BLASTX E value	A. gambiae protein ID	A. gambiae chromosome	TBLASTN E value
Chromosome 3							
LF128|BM005494	1.0	CG7998 gene product|AAF55516	3R (90F1–3)	1^{e-46}	EAA01572	2R (7A–10B)	5^{e-47}
LF253|T58331	16.7	CG10652 gene product|AAF53738	2L (37B9)	1^{e-49}	EAA05968	2R (16B–16D)	4^{e-20}
LF227|T58323	17.4	CG11522 gene product|AAF57167	3R (100D1)	2^{e-23}	EAA01025	2R (7A–10D)	4^{e-33}
LF111|BM005492	19.0	Ribosomal protein L13 (RpL13)|P41126	2L (30F2–4)	5^{e-57}	EAA01175	2R (7A–10D)	2^{e-43}
LF96|BM005491	20.1	Opsin (*ninaE*)|K02320	3R (92B7–9)	2^{e-44}	EAA01516	2R (7A–10D)	2^{e-48}
LF232|BM005489	22.4	CG2099 gene product|AAF52027	3R (83A4–5)	1^{e-23}	EAA07823	2R (11B–14C)	9^{e-35}
LF103|BM005488	23.5	Ribosomal protein L11 (*RpL11*)|AAF57560	2R (56D6–7)	1^{e-57}	EAA05139	3L (40A–42B)	1^{e-32}
Rp49[b]	24.8	Ribosomal protein L32 (*RpL32*)|U92431	3R (99D5–6)	—	EAA00946[a]	2R (7A–10D)	5^{e-56c}
apoLpII|AF038654	25.0	Retinoid and fatty acid binding protein (*RfaBp*)|AAF59387	4 (102F3–4)	1^{e-55}	EAA01116	2R (7A–10D)	1^{e-152}
VC|L46373	25.7	Furin 1 (*Fur1*)|L12372	3R (96D1–3)	0.0	EAA01182	2R (7A–10D)	6^{e-86}
LF106|BM005490	26.1	Ribosomal protein S25 (*RpS25*)|P48588	3R (86D6–7)	1^{e-45}	EAA09243	2R (18C–19C)	4^{e-31}
LF316|BM005516	28.8	Ubisnap (*usnp*)|AAF47071	2R (60A3–5)	6^{e-19}	—[d]	3L (40A–42B)	2^{e-33}
LF417|BM005499	30.1	Ribosomal protein L27A (*RpL27A*)|Q94530	2L (24F4–6)	1^{e-37}	EAA00079	3L (43D–46D)	1^{e-31}
LF296|BM005501	31.4	Ribosomal protein S3A (*RpS3A*)|P55830	4 (102A1)	3^{e-63}	EAA08803	2R (15D–16A)	5^{e-50}
LF386|BM005497	31.6	CG4759 gene product|AAF56495	3R (96E10–12)	5^{e-62}	EAA00091	3L (43D–46D)	1^{e-46}
para|AF468968	31.6	Sodium channel protein (*para*)|M32078	X (14D1–E1)	3^{e-24}	EAA13231	2L (20C–20D)	9^{e-12}
LF168|R47184	32.1	Ribosomal protein S26 (*RpS26*)|P13008	2L (36F7)	7^{e-74}	EAA00291	3L (43D–46D)	4^{e-28}
LF108|T58322	38.1	Ribosomal protein S19 (*RpS19*)|P39018	X (14F4–5)	2^{e-62}	EAA05616	3L (40B–41A)	1^{e-49}
LF323|BM005507	43.7	Proteasome β2 subunit (*Prosβ2*)|AAB82570	3L (71A3)	4^{e-98}	EAA13087	3L (42A–43C)	6^{e-73}
LF218|BM005487	46.4	Ribosomal protein S20 (*RpS20*)|P55828	3R (92F12–13)	8^{e-84}	EAA09966	3L (38C–39A)	2^{e-17}
def|AF156088	47.2	Defensin (*Def*)|P36192	2R (46D7–9)	3^{e-16}	EAA05234	3L (40A–42B)	1^{e-23}
LF352|T58330	48.6	Ribosomal protein L46 (RpL46)|O16130	2R (60B2–3)	1^{e-36}	EAA00736[a]	3L (43D–46D)	2^{e-19}
RpL31|AF324863	50.0	CG1821 gene product|AE003832	2R (45F4–6)	7^{e-79}	EAA00150	3L (43D–46D)	1^{e-57}
Apy1|L12389	57.1	CG1961 gene product|AAF47996	X (10A4)	8^{e-47}	EAA05427	3L (40B–41A)	1^{e-158}
PABP|AY038043	64.2	PolyA-binding protein (*pAbp*)|P21187	2R (55B5–8)	1^{e-141}	EAA05186	3L (40A–42B)	1^{e-118}

[a] Probable error in automated annotation with this protein.

[b] *D. melanogaster* cDNA.

[c] E value based on *D. Melanogaster* and *A. Gambiae* comparisons.

[d] Not called as a protein by automated annotation; E value based on *D. melanogaster* and *A. gambiae* comparisons.

Table 2. Chromosomal distribution of putative
D. melanogaster orthologs of *A. aegypti* genes

A. aegypti chromosome regions	*D. melanogaster* polytene chromosome arms					
	X	2L	2R	3L	3R	4
1 (20 genes)						
Observed	6	0	2	1	9	2
Expected	2.47	5.21	3.29	2.19	5.75	1.1
P value	.0216	.0024	.2026	.2416	.0543	.2069
2p (14 genes)						
Observed	0	4	1	6	3	0
Expected	1.73	3.64	2.3	1.53	4.03	0.77
P value	.1584	.2254	.2223	.0021	.2176	.4543
2q (14 genes)						
Observed	0	11	3	0	0	0
Expected	1.73	3.64	2.3	1.53	4.03	0.77
P value	.1584	.00005	.2243	.1969	.0087	.4543
3p (11 genes)						
Observed	0	2	1	0	7	1
Expected	1.36	2.86	1.81	1.21	3.16	0.6
P value	.2352	.2471	.3001	.2789	.0139	.3431
3q (14 genes)						
Observed	3	2	5	1	2	1
Expected	1.73	3.64	2.3	1.53	4.03	0.77
P value	.1604	.1655	.0477	.3393	.1285	.3687

Figure 1. Comparative genome positions of orthologous genes identified between *Aedes aegypti* (Ae) and *Anopheles gambiae* (Ag). Linkage map positions for *A. aegypti* loci are from Severson et al. (2002) and lines link each ortholog to the corresponding cytological position for the *A. gambiae* ortholog according to data in Table 1. Dashed lines identify putative duplicate loci or transpositions. Putative centromere positions in *A. aegypti* extrapolated from Brown et al. (2001) are indicated by arrows, with *p* and *q* referring to individual arms on chromosomes 2 and 3.

Both map genetically to single loci on *A. aegypti* chromosome 1, but each has homologues on the *A. gambiae* X and 2R. As discussed above, we have verified that *LF178* represents a duplication in *A. aegypti*. We note that the third putative duplicate gene, *VCP*, maps genetically to a single locus on *A. aegypti* chromosome 2 and to two loci on *A. gambiae* 2L. Whole-arm or near-whole-arm correspondence between conserved synteny segments was contradicted by only two of the genes we examined, *LF101* (Ae:1 versus Ag:3R ortholog positions) and *para* (Ae:3q versus Ag:2L ortholog positions), which is consistent with transpositions. Although caution is warranted regarding the exact centromere position on *A. aegypti* chromosome 2, the comparative orientations reflected in gene-order conservation around the centromere region suggest that a pericentric inversion occurred in the *A. aegypti* and *A. gambiae* chromosome lineages (also see the discussion section).

Synteny conservation is also consistently greater in *A. aegypti*–*A. gambiae* comparisons than in comparisons of either mosquito species with *D. melanogaster* (Table 3). In addition, these results indicate that, though extensive chromosome rearrangements are evident among all three species, *A. aegypti* and *A. gambiae* have largely maintained whole-arm gene associations, whereas comparisons of either with *D. melanogaster* indicate considerable interarm exchanges.

Microsynteny Comparison Between *A. aegypti* and *A. gambiae*

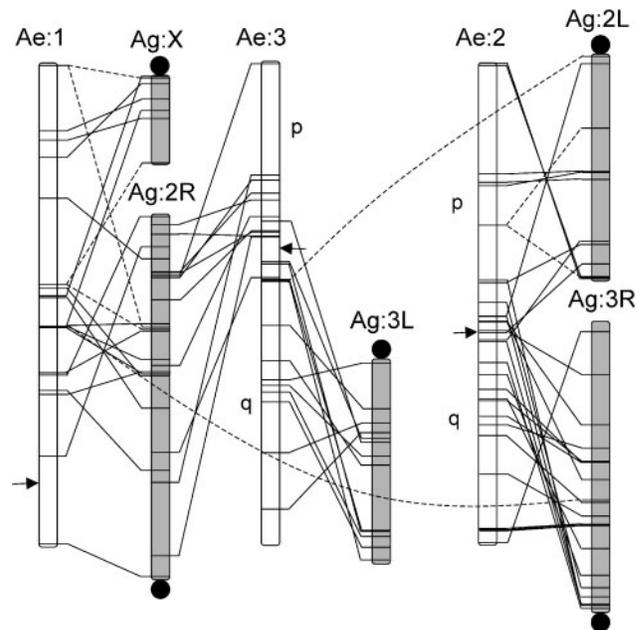We performed BLASTX analysis of the *A. gambiae* proteins with 134 *A. aegypti* cosmid or BAC end sequences produced during a contig building effort for a region on chromosome 1. A single cosmid (*LSBC29.1*) was isolated as containing the *LF178* gene sequence and used initially as an anchor point for contig construction. As previously indicated, we have determined that there are two tightly linked copies of this gene in *A. aegypti* and two copies (located on X and 2R) in *A. gambiae*. For the current analysis, we included sequences from cosmid and BAC clones that are within that contig or physically map near the *LSBC29.1* cosmid (Brown SE and Knudson DL, unpublished data). Fourteen *A. aegypti* end sequences yielded significant BLASTX results, with four representing unique, unambiguous locations on the *A. gambiae* genome (Table 4). Comparative genome positions of these sequences indicated that they remain microsyntenic in *A. gambiae*. The LSBC29.1 clone is about 30 kb and contains one copy of *LF178* and a *wapl* ortholog; if we assume the corresponding *LF178* ortholog is on the X in *A. gambiae*, these genes are about 4 Mb apart. The LSBC29.1 clone overlaps partially with the K6.157E6 clone that also is about 30 kb; LSBC29.1 and the *Yippee* ortholog are, therefore, less than 60 kb apart in *A. aegypti* and about 11.1 kb in *A. gambiae*. Two orthologs within clones that physically map near LSBC29.1 also reflect microsynteny conservation in *A. gambiae*; the K6.2C1 clone contains a *CG3822* ortholog

Table 3. Proportion of orthologs shared among chromosome regions

	Ag:X	Dm:X	Ag:X + Dm:X	Ag:X to Dm:X[a]
Ae:1	0.32	0.30	0.15	0.27
	Ag:2R	Dm:3R	Ag:2R + Dm:3R	Ag:2R to Dm:3R
	0.64	0.45	0.30	0.29
	Ag:2L	Dm:3L	Ag:2L + Dm:3L	Ag:2L to Dm:3L
Ae:2p	0.65	0.43	0.43	0.25
	Ag:3R	Dm:2L	Ag:3R + Dm:2L	Ag:3R to Dm:2L
Ae:2q	1.00	0.79	0.79	0.33
	Ag:2R	Dm:3R	Ag:2R + Dm:3R	Ag:2R to Dm:3R
Ae:3p	0.91	0.64	0.64	0.29
	Ag:3L	Dm:2R	Ag:3L + Dm:2R	Ag:3L to Dm:2R
Ae:3q	0.86	0.36	0.36	0.17

[a] Data extracted from Zdobnov et al. (2002).

that is about 364.5 kb from the *wapI* ortholog, and the K20.3A10 clone contains a *Aats-ser* ortholog that is about 631.4 kb from the *LF178* ortholog on the X.

Discussion

We combined genetic linkage data and gene and EST sequence data for the mosquito *A. aegypti* with the public domain *D. melanogaster* and *A. gambiae* genome databases to perform an *in silico* genome comparison of *A. aegypti* to the other well-described dipterans. We initially submitted 121 *A. aegypti* sequences for BLASTX comparison with the *D. melanogaster* whole-genome database because of the extensive gene annotation available. Significant matches to putative *D. melanogaster* orthologs were observed with 87 (72%) of these sequences, indicating a high level of gene peptide sequence conservation between the two species. Fourteen of these sequences were excluded from comparative genome analyses because they represented multigene families that could not unambiguously be assigned to specific loci in *Drosophila*. All 73 genes with clear orthologs in *D. melanogaster* also had significant TBLASTN matches to *A. gambiae* orthologs. Because these 73 orthologs are distributed across each of the four *D. melanogaster* and the three *A. gambiae* chromosomes, they provided an adequate basis for a comparative genome analysis with *A. aegypti*.

Extent of Synteny Conservation Between *A. aegypti* and *D. melanogaster*

Genome-wide, conserved syntenies between *A. aegypti* and *D. melanogaster* are clearly evident yet less extensive than those observed in, for example, mammals (Ehrlich et al. 1997; Wiltshire et al. 1999) or even among more distant vertebrates (Barbazuk et al. 2000). For example, 80% of putative orthologous genes belong to conserved synteny groups in the human and zebrafish genomes, despite a divergence time of ca. 420 million (Mya) ago (Barbazuk et al. 2000). We observed significant departures from an expected random distribution of the individual genes. That is, for each of the *A. aegypti* chromosomes, we observed a greater than expected number of orthologs to genes from particular chromosome regions in *D. melanogaster*. *A. aegypti* chromosome 1 shows clear homology to the *D. melanogaster* X and some evidence for homology to 3R. In general, chromosomes 2 and 3 for the two species suggest whole-arm translocation events in the ancestral chromosome lineages. The *A. aegypti* 2p and 2q arms have significant homology with *D. melanogaster* 3L and 2L, respectively. Conversely, the *A. aegypti* 3p and 3q arms have significant homology with *D. melanogaster* 3R and 2R, respectively. We note that *D. melanogaster* 3R has homology with both *A. aegypti* chromosome 1 and the 3p arm, suggesting a possible break and fusion event in the ancestral chromosome lineages. Only four *A. aegypti* homologues to *D. melanogaster* chromosome 4 (the dot chromosome) were observed; these genes are distributed on chromosomes 1 and 3. Overall, our *A. aegypti* and *D. melanogaster* comparisons indicate that some orthologs show conserved syntenies that have retained ancestral evolutionary associations, yet these genome segments reflect the effects of considerable intra- and interchromosomal reorganization.

It seems likely that as more *A. aegypti* genes become available for comparison the individual *A. aegypti* chromosomes will at best reflect a mosaic of very short homology segments to *D. melanogaster* chromosomes or individual chromosome arms, as was observed with *D. melanogaster–A. gambiae* comparisons (Zdobnov et al. 2002). Further, it seems likely that *A. aegypti–D. melanogaster* homology segments represent the results of uniform random genome rearrangement events that are independent of functional or mechanistic constraints, given that linkage associations across *Drosophila* species are apparently not based on functional interactions between chromosomal segments (Hilliker and Trusis-Coulter 1987).

From an evolutionary viewpoint, the observation of limited synteny conservation between *A. aegypti* and *D. melanogaster* is not unexpected. That is, the main dipteran lineages were evident in the fossil record during the Upper Triassic (Rohdendorf 1974), and, therefore, the two lineages diverged about 250 million years ago. Further and perhaps most significantly, the two most distant lineages within the genus *Drosophila*, which diverged 40 to 60 million years ago (Russo et al., 1995), exhibit synteny conservation corresponding to individual chromosome arms but extensive interspecific gene-order reshuffling within arms (Ranz et al. 2001). Indeed, the *Drosophila* genome shows one of the highest rates of chromosome evolution observed for any eukaryote, and as such only tightly linked genes in *D. melanogaster* would be expected to reflect linkage conservation in nondrosophilid insects (Ranz et al. 2001).

Extent of Synteny Conservation Between *A. aegypti* and *A. gambiae*

Our results demonstrate that the *A. aegypti* and *A. gambiae* genomes share extensive syntenic associations that corre-

Table 4. Microsynteny comparison between *Aedes aegypti* and *Anopheles gambiae*

A. aegypti clone ID	Genbank accession	A. aegypti map position (FLpter)[a]	A. gambiae Ensembl protein ID	A. gambiae chromosome boundary ba	D. melanogaster gene ortholog	D. melanogaster chrosomone:gene boundary bases
K6.157E6.t7	CC144562	In region	ENSANGP00000017573	X:14238572..14239413	*Yippee*	X:13016562..13218053
LSBC29.1.t7	CC144866	25.6 +/− 6.3%	ENSANGP00000017525	X:14250534..14253933	*wapl*	X:1806874..2015755
K6.2C1.t7	CC144560	23.1 +/− 3.7%	ENSANGP00000001667	X:14618455..14702074	CG3822	3R:16585926..16791968
LF178[b]	T58309	—	ENSANGP00000023750	X:18339733..18342693	*Qm*	3L:22910908..23112752
K20.3A10.t3	CC144563	24.3 +/− 4.3%	ENSANGP00000018742	X:18974130..18975506	*Aats-ser*	3R:11066425..11267957

Orthologs identified among *A. aegypti* cosmid and BAC end sequences during a contig building effort around *LF178* locus on chromosome 1.

[a] FLpter represents the fractional length from the operationally defined *p* terminus.

[b] The LSBC29.1 cosmid contains one copy of the *LF178* gene, and this clone overlaps the K6.157E6.t7 cosmid clone.

spond largely with a whole chromosome-arm conservation pattern that is consistent with the occurrence of paracentric inversions in the ancestral chromosome lineages. Whole-arm or near-whole-arm correspondence between conserved syntenies was contradicted with only two genes among the 75 *A. aegypti* genes for which orthologs to *A. gambiae* were identified. Whole-arm homology is clearly represented by Ae:2p with Ag:2L, Ae:2q with Ag:3R, and Ae:3q with Ag:3L. A likely chromosome arm break and fusion event is suggested by the Ag:2R homology with Ae:3p and Ae:1, and by Ag:X with Ae:1. Further, the *A. gambiae* gene homologies to the general centromere position on Ae:2 suggest a likely pericentric inversion in the chromosome lineages. This interpretation is consistent with our previous comparative mapping data for several culicine mosquito species (Anderson et al. 2001), and particularly with respect to *A. aegypti* comparisons with *Armigeres subalbatus* (Ferdig et al. 1998); the linear order of genes mapped across the centromere region in *A. subalbatus* chromosome 2 is inverted relative to the that of their orthologs in *A. aegypti* chromosome 2, whereas linear orders are conserved between the two species for genes distal to the respective centromere regions. *Armigeres* is considered phylogenetically basal to *Aedes*, and the Anophelinae are basal to the Culicinae (Ross 1951).

Despite the apparent extensive conservation of large syntenic chromosome segments between *A. aegypti* and *A. gambiae*, we observed only limited evidence for linear order conservation within these genome segments. In most instances, orthologs to putative *A. gambiae* genes physically located within the same scaffold accession are clearly not conserved in the same discrete chromosome segments in *A. aegypti*. Obviously, our results largely represent a broad-scale, or macrosynteny, survey, and additional efforts are needed to adequately address the issue of fine-scale, or microsynteny, and linear gene orders. Although our data set for microsynteny analysis is limited, we observed that *A. aegypti* genes identified within or very near a contig construction effort on chromosome 1 are represented by orthologs that generally remain within a relatively small genome region on the X in *A. gambiae*. The same gene associations were not evident in *A. aegypti* and *D. melanogaster* comparisons, suggesting that microsynteny conservation is likely to be greater with

A. gambiae. Of note, conservation of gene order recently has not been found evident in genome comparisons of *A. gambiae* and *Anopheles funestus*, despite an estimated divergence time between them of only 5 million years (Sharakhov et al. 2002). Gene-order conservation between the Anophelinae and Culicinae is, therefore, unlikely to be extensive even at the micro-scale, given a probable divergence time of at least 95 million years (Krzywinski et al. 2001).

Genome Size, Genetic Length, and Evolution in Mosquitoes

A. aegypti and *A. gambiae*, as representatives of the Culicinae and Anophelinae, reflect interesting dichotomies in mosquito chromosome evolution. For example, genome size varies up to eightfold among mosquito species with strong delineation between the subfamilies (Knudson et al. 1996; Rai and Black 1999). The Anophelinae have smaller genome sizes; for example, the 278 Mb *A. gambiae* genome is only about 1.6 times larger than the 170 Mb *D. melanogaster* genome (Holt et al. 2002). Genome size in the Culicinae is typically much larger, with *A. aegypti* at 813 Mb (Warren and Crampton 1991) or about 2.9 and 4.8 times the *A. gambiae* and *D. melanogaster* genome sizes, respectively. Despite such differences in genome size, the basic chromosome number ($2n = 6$) has remained constant within the Culicidae (the only known exception is the primitive anopheline *Chagasia bathana*, with $2n = 8$), with the differences being attributed to an increase in repetitive DNA among the Culicinae (Rai and Black 1999).

Although the basic chromosome complement is highly conserved, genome evolution in mosquitoes reflects considerable chromosome reorganization due to inversions and translocations (Matthews and Munsterman 1994). Still, however, whole-arm or near-whole-arm conservation seems to be the rule. Our *A. aegypti* and *A. gambiae* comparisons clearly demonstrate that, despite extensive rearrangement largely reflecting numerous inversions, chromosome arm syntenies between the two species remain remarkably intact. However, the Anophelinae appear to have followed the *Drosophila* paradigm, in that genome evolution in both groups is orders of magnitude higher than that reported for any other eukaryote (Ranz et al. 2001; Sharakhov et al. 2002). In

contrast, the Culicinae reflect extensive synteny and evidence for broad-scale linear order conservation (Matthews and Munstermann 1994). Indeed, comparative analyses using cDNA-based genetic markers for *A. aegypti* (Severson et al. 1993, 2002) indicate that they not only have high sequence identities with other mosquito species (Severson et al. 1994), but also provide direct evidence that large-scale inter- and intrachromosome rearrangements may not disrupt gene colinearity within relatively large conserved chromosome segments across a diverse group of culicines, including *Aedes albopictus* (Severson et al. 1995), *Ochlerotatus triseriatus* (formerly *Aedes triseriatus*) (Anderson et al. 2001), *A. subalbatus* (Ferdig et al. 1998), *C. pipiens* (Mori et al. 1999), and *Culex tritaeniorhynchus* (Mori et al. 2001).

The basis for the extreme differences in evolutionary rates between drosophilids/anophelines and culicines is unknown, though three factors have been suggested for the high rate in drosophilids: the short developmental time, a greater mutation rate, and a less detrimental effect on fertility of inversions (Ranz et al. 2001). Although developmental times are similar among the three groups, the other factors are likely applicable, though the exact mechanisms remain to be determined. Of note, there are interesting sex-specific differences in chiasmata formation between the three groups. Crossing over is suppressed in male but not in female drosophilids, but it occurs in relatively equal frequencies in autosomes of males and females of most mosquito species. However, the Anophelinae are characterized by typical X and Y sex chromosomes that reflect limited crossovers in males, whereas the Culicinae show no sex chromosome dimorphism and in general crossovers occur at about equal frequencies in males and females with the sex-determining chromosome (Clements 1992). Curiously, however, crossing over appears to be much reduced in the Culicinae: despite a considerably larger genome size, the genetic map size for *A. aegypti* is only 205 cM (Severson et al. 2002), compared to 215 cM for *A. gambiae* (Zheng et al. 1996) and 290 cM for *D. melanogaster* (O'Brien 1993).

Conclusions

We have performed a broad-scale, genome-wide analysis of synteny conservation of the yellow fever mosquito, *A. aegypti*, with *D. melanogaster* and the primary malaria vector, *A. gambiae*. An *in silico* comparative genomics approach allowed us to identify orthologs to genetically mapped *A. aegypti* genes. Although linkage estimates of gene order in *A. aegypti* could reflect some positional errors, it seems likely that our overall conclusions for residual whole-arm synteny conservation with *D. melanogaster* and whole-arm synteny conservation but limited linear order conservation with *A. gambiae* are appropriate. Indeed, accurate determinations of patterns of chromosome rearrangements in mammals have been produced with quite limited genetic linkage data sets (Ehrlich et al. 1997). Further, our interpretations of whole-arm synteny conservation of *A. aegypti* with both *D. melanogaster* and *A. gambiae* are completely concordant with previously reported comparisons of *A. gambiae* to *D. melanogaster*

(Bolshakov et al. 2002; Zdobnov et al. 2002). Another potential for error in our comparative analysis is the misclassification of paralogous genes as orthologs. However, given our stringent criteria for sequence comparisons and the exclusion of genes representing obvious gene families, most of the identified orthologous pairs are likely valid.

The *D. melanogaster* and *A. gambiae* genome databases, though clearly offering tremendous potential for comparative gene identity and function, seem to offer limited value for comparative positional gene determinations among even closely related dipterans. That is, to provide useful positional information, reasonably sized genome segments must reflect not only synteny conservation, but also conservation of linear gene orders. The apparent rapid rate of chromosome rearrangement observed in *Drosophila*, combined with the drosophilid-mosquito divergence time, suggests that even genome segments on the order of 50 kb have a less than .5 likelihood of linear gene order conservation (extrapolated from Ranz et al. 2001, figure 3). Further, the pattern of rapid chromosome rearrangement seems to be an even more striking feature of anopheline chromosome evolution (Sharakhov et al. 2002). These results and other factors, including differences in genome organization (Knudson et al. 2002), indicate that it remains paramount that genome sequencing projects be conducted for other dipteran species, including representative culicine mosquito species such as *A. aegypti*.

References

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, et al., 2000. The genome sequence of *Drosophila melanogaster*. Science 287:2185–2195.

Altschul SF, 1991. Amino acid substitution matrices from an information theoretic perspective. J Mol Biol 219:555–565.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, and Lipman DJ, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402.

Anderson JR, Grimstad PR, and Severson DW, 2001. Chromosomal evolution among six mosquito species (Diptera: Culicidae) based on shared restriction fragment length polymorphisms. Mol Phylogenet Evol 20: 316–321.

Andersson L, Ashburner M, Audun S, Barendse W, Bitgood J, Bottema C, Broad T, Brown S, Burt D, Charlier C, et al., 1996. Comparative genome organization of vertebrates: the first international workshop on comparative genome organization. Mamm Genome 7:717–734.

Barbazuk WB, Korf I, Kadavi C, Heyen J, Tate S, Wun E, Bedell JA, McPherson JD, and Johnson SL, 2000. The syntenic relationship of the zebrafish and human genomes. Genome Res 10:1351–1358.

Bolshakov VN, Topalis P, Blass C, Kokoza E, della Torre A, Kafatos FC, and Louis C, 2002. A comparative genomic analysis of two distant Diptera, the fruit fly, *Drosophila melanogaster*, and the malaria mosquito, *Anopheles gambiae*. Genome Res 12:57–66.

Boyd Y, Blair HJ, Cunliffe P, Masson WK, and Reed V, 2000. A phenotypic map of the mouse X chromosome: models for human X-linked disease. Genome Res 10:277–292.

Brown SE, Severson DW, Smith LA, and Knudson DL, 2001. Integration of the *Aedes aegypti* genetic linkage and physical maps. Genetics 157:1299–1305.

Clements AN, 1992. Aspects of genetics. In: The Biology of Mosquitoes. London: Chapman & Hall; 29–33.

Ehrlich J, Sankoff D, and Nadeau JH, 1997. Synteny conservation and chromosome rearrangements during mammalian evolution. Genetics 147:289–296.

Ferdig MT, Taft AS, Severson DW, and Christensen BM, 1998. Development of a comparative genetic linkage map for *Armigeres subalbatus* using *Aedes aegypti* RFLP markers. Genome Res 8:41–47.

Fulton RE, Salasek ML, DuTeau NM, and Black WC IV, 2001. SSCP analysis of cDNA markers provides a dense linkage map of the *Aedes aegypti* genome. Genetics 158:715–726.

Gilchrist BM and Haldane JBS, 1947. Sex linkage and sex determination in a mosquito, *Culex molestus*. Hereditas 33:175–190.

Hilliker AJ and Trusis-Coulter SN, 1987. Analysis of the functional significance of linkage group conservation in *Drosophila*. Genetics 117:233–244.

Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, Wincker P, Clark AG, Ribeiro JM, Wides R, et al., 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. Science 298:129–149.

Knight KL and Stone A, 1977. A catalog of the mosquitoes of the world (Diptera: Culicidae), vol. 2, 2nd ed. Baltimore, Maryland: Thomas Say Foundation.

Knudson DL, Brown SE, and Severson DW, 2002. Culicine genomics. Insect Biochem Mol Biol 32:1193–1197.

Knudson DL, Zheng L, Gordon SW, Brown SE, and Kafatos FC, 1996. Genome organization of vectors In: the biology of disease vectors (Beaty BJ and Marquardt WC, eds). Niwot: University Press of Colorado; 175–214.

Krzywinski J, Wilkerson RC, and Besansky NJ, 2001. Toward understanding Anophelinae (Diptera, Culicidae) phylogeny: insights from nuclear single-copy genes and weight of evidence. Syst Biol 50:540–556.

Lalley PA, O'Brien SJ, Creau-Goldberg N, Davisson MT, and Roderick TH, 1987. Report of the committee on comparative mapping. Cytogenet Cell Genet 46:367–389.

Matthews TC and Munstermann LE, 1994. Chromosomal repatterning and linkage group conservation in mosquito karyotypic evolution. Evolution 48:146–154.

McPeek MS, 2000. From mouse to human: fine mapping of quantitative trait loci in a model organism. Proc Natl Acad Sci USA 97:12389–12390.

Mori A, Severson DW, and Christensen BM, 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 DW, 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.

Nadeau JH, 1989. Maps of linkage and synteny homologies between mouse and man. Trends Genet 5:82–86.

Nadeau JH and Sankoff D, 1998. Counting on comparative maps. Trends Genet 14:495–501.

O'Brien SJ, 1993. Genetic maps: locus maps of complex genomes. Plainview, New York: Cold Spring Harbor Laboratory Press.

Rai KS and Black WC IV, 1999. Mosquito genomes: structure, organization, and evolution. Adv Genet 41:1–33.

Ranz JM, Casals F, and Ruiz A, 2001. How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. Genome Res 11:230–239.

Rohdendorf B, 1974. The Historical Development of Diptera. Edmonton: University of Alberta Press.

Ross HH, 1951. Conflict with *Culex*. Mosq News 11:128–132.

Russo CAM, Takezaki N, and Nei M, 1995. Molecular phylogeny and divergence times of drosophilid species. Mol Biol Evol 12:391–404.

Severson DW, Meece JK, Lovin DD, Saha G, and Morlais I, 2002. Linkage map organization of expressed sequence tags and sequence tagged sites in the mosquito, *Aedes aegypti*. Insect Mol Biol 11:371–378.

Severson DW, Mori A, Zhang Y, and Christensen BM, 1993. Linkage map for *Aedes aegypti* using restriction fragment length polymorphisms. J Hered 84:241–247.

Severson DW, Mori A, Zhang Y, and Christensen BM, 1994. The suitability of restriction fragment length polyporphism markers for evaluating genetic diversity among and synteny between mosquito species. Am J Trop Med Hyg 50:425–432.

Severson DW, Mori A, Kassner VA, and Christensen BM, 1995. Comparative linkage maps for the mosquitoes, *Aedes albopictus* and *Aedes aegypti*, based on common RFLP loci. Insect Mol Biol 4:1–45.

Sharakhov IV, Serazin AC, Grushko OG, Dana A, Lobo N, Hillenmyer ME, Westerman R, Romero-Severson J, Costantini C, Sagnon N, et al., 2002. Inversions and gene order shuffling in *Anopheles gambiae* and *Anopheles funestus*. Science 298:182–185.

Sokal R and Rohlf FJ, 1981. Biometry, 2nd ed. New York: Freeman.

Warren AM and Crampton JM, 1991. The *Aedes aegypti* genome: complexity and organization. Genet Res Camb 58:225–232.

Weller GL and Foster GG, 1993. Genetic maps of the sheep blowfly *Lucilia cuprina*: linkage-group correlations with other dipteran genera. Genome 36:495–506.

Wiltshire T, Pletche M, Cole SE, Villanueva M, Birren B, Lehoczky J, Dewar K, and Reeves RH, 1999. Perfect conserved linkage across the entire mouse chromosome 10 region homologous to human chromosome 21. Genome Res 9:1214–1222.

Zdobnov EM, von Mering C, Letunic I, Torrents D, Suyama M, Copley RR, Christophides GK, Thomasova D, Holt RA, Subramian GM, et al., 2002. Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. Science 298:149–159.

Zheng L, Benedict MQ, Cornel AJ, Collins FH, and Kafatos FC, 1996. An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. Genetics 143:941–952.