2002

A preliminary linkage map of the tick, *Ixodes scapularis*

Amy J. Ullmann
Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, P.O. Box 2087, Rampart Road, Foothills Campus, Fort Collins, CO 80522, USA

Joseph Piesman
Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, P.O. Box 2087, Rampart Road, Foothills Campus, Fort Collins, CO 80522, USA

M.C. Dolan
Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, P.O. Box 2087, Rampart Road, Foothills Campus, Fort Collins, CO 80522, USA

William C. Black, IV
Department of Microbiology, Colorado State University, Fort Collins, CO 80523-1677, USA

Experimental and Applied Acarology **28:** 107–126, 2002.

A preliminary linkage map of the tick, *Ixodes scapularis*

AMY J. ULLMANN[1,*], JOSEPH PIESMAN[1], M.C. DOLAN[1] and WILLIAM C. BLACK, IV[2]

[1]*Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, P.O. Box 2087, Rampart Road, Foothills Campus, Fort Collins, CO 80522, USA*

[2]*Department of Microbiology, Colorado State University, Fort Collins, CO 80523-1677, USA*

Abstract. A linkage map of the *Ixodes scapularis* genome was constructed based upon segregation amongst 127 loci. These included 84 random amplified polymorphic DNA (RAPD) markers, 32 Sequence-Tagged RAPD (STAR) markers, 5 cDNAs, and 5 microsatellites in 232 F_1 intercross progeny from a single, field-collected P_1 female. A preliminary linkage map of 616 cM was generated across 14 linkage groups with one marker every 10.8 cM. Assuming a genome size of $\sim 10^9$ bp, the relationship of physical to genetic distance is ~300 kb/cM in the *I. scapularis* genome.

Key words: *Ixodes scapularis*, linkage map, microsatellites, RAPD-SSCP, STARs, cDNA-SSCP

Introduction

Ticks and the pathogens they transmit have long been an important cause of morbidity and mortality. Ticks are second only to mosquitoes as vectors of human and animal diseases (Sonenshine, 1991). *Ixodes scapularis* has become an important vector species because of its ability to transmit *Borrelia burgdorferi*, the causative agent of Lyme disease, in eastern and midwestern North America (Spielman *et al*., 1985). *I. scapularis* is also a vector of the causative agent of human granulocytic ehrlichiosis (HGE) (Telford *et al*., 1996) and *Babesia microti* the causative agent of human babesiosis (Spielman, 1976).

A great deal of information is known about the development, physiology, anatomy and vector competence of ticks, however little is understood about the genetic basis of these traits. Karyotypes have been characterized for 103

*Author for correspondence (E-mail: aff1@cdc.gov)

of the ~830 known tick species (Oliver, 1977). However, a linkage map has never been produced for any of these or, for that matter, for any chelicerate arthropod. This is in part due to the minimal generation time required to establish a tick mapping family, which, at a minimum, over 1 year and 10 months are required.

Once an F_1 intercross family has been established, the modern paradigm for linkage mapping involves simultaneous estimates of recombination amongst hundreds (Antolin *et al*., 1996) or potentially thousands (Yasukochi, 1998) of highly polymorphic loci. This typically yields an intensive linkage map with one marker distributed every 5 cM. Many different genetic markers are currently in use for intensive linkage mapping. One class includes simple sequence repeats or microsatellites (MS). MS-based linkage maps have been generated for anopheline mosquitoes (Zheng *et al*., 1996), mice (Dietrich *et al*., 1992) and humans (Hudson *et al*., 1995). Characterization of MS markers in *I. scapularis* has been described (Fagerberg *et al*., 2001) but the abundance of and heterozygosity at MS loci was generally low.

Random Amplified Polymorphic DNA amplified by the Polymerase Chain Reaction (RAPD-PCR) provides another class of highly polymorphic markers. This technique uses a short (10 oligonucleotide) primer to amplify small regions of a genome (Williams *et al*., 1990). Alleles amplified by RAPD-PCR are typically dominant. Recessive alleles arise through mutations in or around the primer annealing site that cause the PCR process to fail. Genetic linkage maps utilizing RAPDs have been created for many insects including *Anopheles gambiae* (Dimopoulos *et al*., 1996), *Aedes aegypti* and a parasitic wasp, *Bracon hebetor* (Antolin *et al*., 1996), *Apis mellifera* (Hunt and Page, 1995), and *Bombyx mori* (Yasukochi, 1998). However, a great deal of the information needed to estimate recombination frequencies is lost when mapping loci with dominant alleles.

Single strand conformation polymorphism (SSCP) analysis of RAPD markers can reveal alternative, codominant alleles (Antolin *et al*., 1996). SSCP analysis is based on the principle that electrophoretic mobilities of single-strand DNA molecules in nondenaturing gels are dependent upon both the size and shape of the fragments. Several stable structures or conformations can be formed when secondary base pairing occurs among nucleotides on a single DNA strand. The SSCP technique detects 99–100% of point mutations in DNA molecules 100–300 base pairs (bp) in length and at least 89% of mutations in molecules 300–450 bp in length (Orita *et al*., 1989; Hayashi, 1991; Hiss *et al*., 1994; Vidal-Puig and Moller, 1994).

Sequence-Tagged RAPD loci (STARs) (Bosio *et al*., 2000) are another useful modification of RAPD markers. A RAPD-PCR band from a SSCP gel is cloned into a plasmid vector and sequenced. The sequence is then used to

design targeted primers for PCR. STAR products are analyzed on a second SSCP gel, and usually segregate as codominant markers.

SSCP analysis of cDNA sequences provides another class of highly polymorphic markers for use in linkage mapping. Although these are unique, functional regions of a genome, polymorphisms are frequent, and can be used to generate linkage maps. This method has been used to map cDNA makers in *A. aegypti* (Fulton *et al*., 2001).

In this paper we describe the use of MS, RAPD-PCR, STAR, and cDNA markers to generate a preliminary linkage map in an F_1 intercross family of *I. scapularis.* We describe our successes and failures in developing polymorphic markers and provide a preliminary 616 cM linkage map across 14 linkage groups, the 1N number of chromosomes determined cytogenetically (Oliver, 1977). This provided a linkage map with one marker every 10.8 cM. However, counting markers that did not map to any of these 14 linkage groups, we estimate a much larger map size of 3165 cM with an average resolution of only one marker every 55 cM.

Materials and Methods

Mapping cross

All ticks were kept in Wheaton 8 ml sample vials (Wheaton, Millville, NJ) which were placed into glass desiccator jars with water in the bottom to maintain a high relative humidity. The desiccators were maintained in a Revco bioclimatic chamber (Revco Inc., Deerfield, MI) at 21°C, 95% humidity and received a photoperiod of 16:8 (light:dark). ICR outbred, pathogen-free mice were used for blood feedings (DVBID, CDC, Ft. Collins). Four fertilized P_1 females were collected from Bridgeport, Connecticut in the fall of 1997, brought back to DVBID, and fed on rabbits. Eggs were laid by the P_1 females and hatched ~4 weeks later to generate four initial families. After eggs were collected from each P_1 mother, each female was frozen in a labeled cryotube at −20°C in 70% ethanol. F_1 eggs from each female were hatched out and larvae were blood fed on mice. One hundred replete larvae were place in Wheaton vials and ~5 weeks after being fed, the larvae molted to nymphs. Once the cuticle hardened, the F_1 nymphs were blood fed. At this point the nymphs went into diapause for ~8 months due to an unintentional drop in incubator temperature to 16°C. When the incubator temperature was corrected, the ticks came out of diapause and development continued. When the fed nymphs molted to adults, eight single F_1 intercrosses were set up per family. Successful breeding pairs were established from P_1

female #2 (five pairs), and only a single pair from P_1 females #1, #3, and #4. For each F_1 intercross, females were isolated from the males, and single breeding pairs were made and subsequently fed individually in an ear bag on a rabbit ear.

Fed females were separated into labeled vials to collect F_2 eggs. These were laid from 4 to 6 weeks post-feeding. Each F_1 male was frozen in a labeled tube at −20°C in 70% ethanol, and given the same number as his F_1 mate. Each F_1 female was frozen after collecting eggs. F_2 eggs were hatched and when the cuticle had hardened, blood fed on mice. After the larvae had molted to nymphs, ∼4 weeks post-feed, and the cuticle had hardened, nymphs were fed on mice. The nymphs were reared to adults, and individually assigned a number that corresponded to the P_1 and F_1 female numbers. The F_2 adults were then frozen in a labeled tube at −20°C in 70% ethanol.

Markers

Microsatellite loci were derived as described by Fagerberg *et al*. (2001). RAPD loci were amplified and run on SSCP gels as described in Black and DuTeau (1997). STAR loci were derived from polymorphic RAPD loci following exactly the procedure in Bosio *et al*. (2000). Primers for cDNA sequences from an *I. scapularis* salivary gland genomic library (kindly provided by Jose Ribeiro) were developed following the methodology of Fulton *et al*. (2001).

Map data analysis

Map distances were converted from recombination fractions to map units (cM) using the Kosambi mapping function (Kosambi, 1944). Offspring genotypes of F_1 and F_2 *I. scapularis* were entered into JoinMap® 2.0 (Stam and van Oojin, 1995). Because the numbers of parental, recombinant, and uninformative genotypes, and the segregation ratios among the F_2 offspring may differ for each locus, JoinMap allows each marker to be analyzed according to parental genotypes. Initially, a threshold recombination fraction of 0.499 and a log odds density (LOD) score of 3.0 was used to group markers. The minimal LOD was then increased to 6.0 in increments of 0.1 to monitor the rate at which markers left individual linkage groups. Increasing the minimum LOD in increments tests how robust are the linkage groups in the genotype data set. DrawMap® (van Ooijen, 1994) was used to plot a linkage map from the recombination frequencies generated by JoinMap.

Results

The P_1 generation began as a field-collected, fertilized female. The genotype of the P_1 male had therefore to be inferred at all loci. Recombination frequencies among all marker loci were analyzed in 232 F_2 progeny.

Twenty RAPD primers (Table 1) amplified 84 polymorphic marker loci. Genotypes at 64 loci were in expected Mendelian ratios among the F_2 offspring. Sixty-three of these loci and D04.800a and b (bands ~1 mm apart on the SSCP gel that did not completely cosegregate) were extracted from SSCP gels, reamplified, cloned and sequenced to generate 65 primer pairs (Table 2). Of these, 52 STAR loci were polymorphic, and genotypes at 33 of these were in expected Mendelian ratios in the F_2 offspring. Alleles at 20 of the 33 STAR loci segregated as codominant markers.

Fifty-six sequences from an *I. scapularis* salivary gland cDNA library (sequences kindly provided by Jose Ribeiro) were analyzed for primer design.

Table 1. List of RAPD primers and sequences utilized for mapping

Primer	Sequence
A09	GGGTAACGCC
A20	GTTGCGATCC
B15	GGAGGGTGTT
B18	CCACAGCAGT
B20	GGACCCTTAC
C19	GTTGCCAGCC
C01	TTCGAGCCAG
C04	CCGCATCTAC
D02	GGACCCAACC
D03	GTCGCCGTCA
D04	TCTGGTGAGG
D07	TTGGCACGGG
D08	GTGTGCCCCA
D12	CACCGTATCC
D13	GGGGTGAGCA
D16	AGGGCGTAAG
D17	TTTCCCACGG
D18	GAGAGCCAAC
D19	CTGGGGACTT
D20	ACCCGGTCAC

Table 2. List of 95 primer pairs for PCR amplification of 65 STAR, 20 cDNA and 10 microsatellite loci in *I. scapularis*

Locus name		T_a	Length (bp)	Primer sequences
STARs				
A09.306ST	Genbank Acc. #BZ385505	49	306	GGGTAACGCCAGGGTTTTCC
				GGGTAACGCCCGATGTATAG
A20.310ST	Genbank Acc. #BZ592381	51	310	GTTGCGATCCCTCAGAGCGA
				GTTGCGATCCAACGAAGTTT
A20.390ST	Genbank Acc. #BZ592382	57	390	GTTGCGATCCCAGGATATAC
				GTTGCGATCCTASAGCACAT
A20.517ST	Genbank Acc. #BZ592383	59	517	GTTGCGATCCCTGCGCCTAT
				GTTGCGATCCAGGCGATCAC
B15.874ST	Genbank Acc. #BZ385506	59	874	GGAGGGTGTTGCACAGTACA
				GGAGGGTGTTGGGGTGTGTG
B18.358ST	Genbank Acc. #BZ385507	62	358	CCACAGCAGTCAAACCTTCT
				CCACAGCAGTAGTGATACTC
B18.653ST	Genbank Acc. #BZ385508	62	653	CCACAGCAGTCGACCATGCG
				CCACAGCAGTAGTGATACTC
B20.361ST	Genbank Acc. #BZ385509	62	361	GGACCCTTACGAGCGAAAGA
				GGACCCTTACCCCAACCAAA
C13.357ST	Genbank Acc. #BZ385510	62	357	AAGCCTCGTCGGGGTAGAGA
				AAGCCTCGTCCTGCTGTTCT
C13.579ST	Genbank Acc. #BZ385511	60	579	AAGCCTCGTCGTTAGACGTT
				AAGCCTCGTCCACTTTCCTT
C16.362ST	Genbank Acc. #BZ385512	62	362	CACACTCCAGTCAGCATCAG
				CACACTCCAGGACCAATAGT

C16.432ST	Genbank Acc. #BZ385513	46	432	CACACTCCAGGCAAAAAAAC
				CACACTCCAGCATTAGGGAA
C16.458ST	Genbank Acc. #BZ385514	59	458	CACACTCCAGGCAGATGCAT
				CACACTCCAGCATCAAGTGC
C16.680ST	Genbank Acc. #BZ385515	48	680	CACACTCCACCAATCCCCTA
				CACACTCCACGATACAAACG
C01.169ST	Genbank Acc. #BZ385516	59	169	TTCGAGCCAGGGCAGGACCT
				TTCGAGCCAGCATTGATAAG
C01.170ST	Genbank Acc. #BZ385517	59	170	TTCGAGCCAGCATTGAGAAG
				TTCGAGCCAGGGCAGGACCT
C04.323ST	Genbank Acc. #BZ385518	60	323	CCGCATCTACGTACGATTGA
				CCGCATCTACGCCCGGTAGT
C04.331ST	Genbank Acc. #BZ385519	57	331	CCGCATCTACATATACAATA
				CCGCATCTACCCCATCGACG
C04.345ST	Genbank Acc. #BZ385520	57	345	CCGCATCTACGTTAGCAGTC
				CCGCATCTACACACAGTTTC
D02.328ST	Genbank Acc. #BZ385392	50	328	GGACCCAACCCTATACCGCT
				GGACCCAACCTACGGAGGAG
D02.330ST	Genbank Acc. #BZ385393	48	330	GGACCCAACCCTATACCGCT
				GGACCCAACCTACGGAGGAG
D02.460ST	Genbank Acc. #BZ385394	45	460	GGACCCAACCATGTTTTAAA
				GGACCCAACCCCAAGAAATC
D02.463ST	Genbank Acc. #BZ385395	48	463	GGACCCAACCCCCAAGAATC
				GGACCCAACCATGTTTTAAC

Table 2. (*continued*)

Locus name		T_a	Length (bp)	Primer sequences
D02.464ST	Genbank Acc. #BZ385396	48	464	GGACCCAACCCCAAGAATCC
				GGACCCAACCATGTTTTAAC
D02.466ST	Genbank Acc. #BZ385397	48	466	GGACCCAACCCCAAGAAATC
				GGACCCAACCATGTTTTAAC
D02.470ST	Genbank Acc. #BZ385398	50	470	GGACCCAACCATGTTTTAAC
				GGACCCAACCCCCAAGAATC
D02.472ST	Genbank Acc. #BZ385399	48	472	GGACCCAACCATGTTTTAAA
				GGACCCAACCCCAAGAATCC
D04.410ST	Genbank Acc. #BZ385400	45	410	TCTGGTGAGGGGGAAGAAGG
				TCTGGTGAGGCAACTACTGA
D04.534ST	Genbank Acc. #BZ385401	45	534	TCTGGTGAGGCGTTCAAGGC
				TCTGGTGAGGAGGAATGTTC
D04.800AST	Genbank Acc. #BZ85433-BZ85434	45	800	TCTGGTGAGGTTCGTGTCAA
				TCTGGTGAGGCGGGGCTCCA
D04.800BST	Genbank Acc. #BZ85435-BZ85436	45	800	TCTGGTGAGGCGGGACTTCA
				TCTGGTGAGGTTCGTGTCAA
D07.292ST	Genbank Acc. #BZ385402	45	292	TTGGCACGGGCACACAAATA
				TTGGCACGGGGCTGCAACAA
D07.434ST	Genbank Acc. #BZ385403	50	434	TTGGCACGGGCAATCATTGC
				TTGGCACGGGACGCGGAAGC

D07.457ST	Genbank Acc. #BZ385404	50	457	TTGGCACGGGCGCGTGGCTT
				TTGGCACGGGGGACAAAAGA
D07.649ST	Genbank Acc. #BZ385405	50	649	TTGGCACGGGCAGCAGCCGG
				TTGGCACGGGTGGAGGTCGT
D08.459ST	Genbank Acc. #BZ385406	50	459	GTGTGCCCCAGGACCTGGGT
				GTGTGCCCCACCGCAAAGAG
D08.463ST	Genbank Acc. #BZ385407	50	463	GTGTGCCCCACCGCAAAGAG
				GTGTGCCCCAGGACCTGGGT
D12.425ST	Genbank Acc. #BZ385408	50	425	CACCGTATCCAGGAAGTGTT
				CACCGTATCCTCGCAGAGTT
D12.617ST	Genbank Acc. #BZ385409	36	617	CACCGTATCCCCCTAGCGAA
				CACCGTATCCAACAAAGGAT
D12.821ST	Genbank Acc. #BZ385410	43	821	CACCGTATCCCGTTGGCCTA
				CACCGTATCCCCATGATTTC
D12.942ST	Genbank Acc. #BZ385411	50	942	CACCGTATCCATGTGTAATC
				CACCGTATCCCCAGACTCGG
D17.459ST	Genbank Acc. #BZ385412	45	459	TTTCCCACGGGCACGTACTA
				TTTCCCACGGCGTCATGACT
D17.984ST	Genbank Acc. #BZ385413	48	984	TTTCCCACGGCAGGATAACT
				TTTCCCACGGATTAATCAAA
D17.1200ST	Genbank Acc. #BZ385437-BZ85438	43	1200	TTTCCCACGGAACGACTCAC
				TTTCCCACGGCCAAACAAAG
D13.363ST	Genbank Acc. #BZ385414	50	363	GGGGTGACGACCGACTGGC
				GGGGTGACGAGACAACGAAA

Table 2. (*continued*)

Locus name		T_a	Length (bp)	Primer sequences
D13.443ST	Genbank Acc. #BZ385415	50	443	GGGGTGACGATGGGTTGTGG
				GGGGTGACGACCAACCCGGG
D16.330ST	Genbank Acc. #BZ385416	50	330	AGGGCGTAAGCAGAACCGTT
				AGGGCGTAAGCATCAGGTAT
D16.470ST	Genbank Acc. #BZ385417	50	470	AGGGCGTAAGGGCCGATAGC
				AGGGCGTAAGTGCGTGTGTG
D16.521ST	Genbank Acc. #BZ385418	50	521	AGGGCGTAAGCCTAACAAAA
				AGGGCGTAAGGGTAGTGTGT
D17.684ST	Genbank Acc. #BZ385419	50	684	TTTCCCACGGCACCCTTTGT
				TTTCCCACGGGCCATTCCGC
D18.265ST	Genbank Acc. #BZ385420	41	265	GAGAGCCAACACGTCAAATA
				GAGAGCCAACCTCACTGACC
D18.266ST	Genbank Acc. #BZ385421	43	266	GAGAGCCAACCTCACTGACCT
				GAGAGCCAACACGTCAAAT
D18.284ST	Genbank Acc. #BZ385422	39	284	GAGAGCCAACCTACTTGGCA
				GAGAGCCAACATGGAGAGGA
D18.353ST	Genbank Acc. #BZ385423	39	353	GAGAGCCAACGGGTAAATAG
				GAGAGCCAACAAGAGACTAC
D18.366ST	Genbank Acc. #BZ385424	50	366	GAGAGCCAACCCACCTTCAT
				GAGAGCCAACGCGCGGGCGA
D18.367ST	Genbank Acc. #BZ385425	50	367	GAGAGCCAACGCGCGGGCGA
				GAGAGCCAACCCACCTACAT

D18.369ST	Genbank Acc. #BZ385426	50	369	GAGAGCCAACCCACCTTCAT
				GAGAGCCAACCCGCGGGCGA
D18.826ST	Genbank Acc. #BZ385427	50	826	GAGAGCCAACAATCAACCTC
				GAGAGCCAACGCCAAAGAAA
D19.382ST	Genbank Acc. #BZ385428	50	382	CTGGGGACTTCAAACAACCC
				CTGGGGACTTGCCACGTCAG
D19.833ST	Genbank Acc. #BZ385429	50	833	CTGGGGACTTCACAGAGGCA
				CTGGGGACTTCCGTTTGGAT
D20.284ST	Genbank Acc. #BZ385430	50	284	ACCCGGTCACCAATTTGTAG
				ACCCGGTCACAGATGGAGAA
D20.294ST	Genbank Acc. #BZ385431	50	294	ACCCGGTCACACGAGCGCTT
				ACCCGGTCACGTCTTCACTG
D20.817ST	Genbank Acc. #BZ385432	50	817	ACCCGGTCACCAGCAGATTT
				ACCCGGTCACATCATTGTGT
cDNA[a]				
Is218 (APOP-IN (AF333765) DAD-1-like protein [*Schistosoma japonicum*] (3.00E−41)*)	Genbank Acc. #CA763761	60	209	GCCAGCGTAAACGAAACC
				GATGAAGCCCGACAGGAA
Is155 (ATP-S (mitochondrial ATP synthase [*Drosophila melanogaster*] (1.00E−25))	Genbank Acc. #CA763762	52	195	CTGAAAGGGTCCCAGAAGA
				GGAATGGCACCGTAAAGC

Table 2. (*continued*)

Locus name		T_a	Length (bp)	Primer sequences
Is108 (EF1-A (AF378368 elongation factor 1-alpha [*Coccidioides immitis*] (3.00E−34))	Genbank Acc. #CA763763	62	196	TGTGGGCGTCATCAAGTC AAACAGAGTAACCAAACGAAAG
Is119 (ELF-2B (CAC08449 eukaryote initiation factor 2 beta [*Gallus gallus*] (5.00E−57))	Genbank Acc. #CA763764	58	301	CCTTCGCCAACTTCCTCG AAGCCAGACTTGATGCTCAC
Is211 (GST-1 (AF366931 glutathione S-transferase [*Boophilus microplus*]) (1.00E−30))	Genbank Acc. #CA763765	53	218	GAAATAACGGTTGAGGGC GTAGGGAAGGTTGGGAAA
Is75 (GST-2 (JX0095 glutathione transferase (EC 2.5.1.18) b – guinea pig) (1.00E−30))	Genbank Acc. #CA763766	53	164	TCCTCGTTGAGATCCAGT GTTCCAGCAGTAATAGTCG
ISAC (AF270496 *I. scapularis* anticomplement protein)		56	188	TTTTGGCGATTTCGTTTC AATTCCTTCCCAGGTTGC
Is152 (ND(CG7580 gene product [*D. melanogaster*] putative NADH dehydrogenase) (6.00E−27))	Genbank Acc. #CA763767	53	251	GCACTTTGGGAACCTGAT TGGGCTGCTTCTACTTGTC

Is136 (QM (XP 141003 similar to ribosomal protein L10 [*Rattus norvegicus*]) (2.00E−49))	Genbank Acc. #CA763768	62	184	AGGCCAAGGTGGATGAGTT TTGTTGATGCGGATGACG
Is30 (RPS12 (AF470687 40S ribosomal protein S12 [*Branchiostoma belcheri*]) (1.00E−49))	Genbank Acc. #CA763769	62	214	GCTCATCTGTGCGTCCTG TCTCCTTGCCGTAGTCCTT
Is140 (RPS13 (AF116857 40S ribosomal protein S13 [*Cricetulus griseus*]) (3.00E−48))	Genbank Acc. #CA763770	60	177	CCCAGTCGGCTCTGCCATAC TCTTGTTGCCCGTCACCC
Is45 (RPS14B (AF402822 40S ribosomal protein S14 [*Ictalurus punctatus*]) (3.00E−52))	Genbank Acc. #CA763771	53	183	TTTGGCGTAGCCCACATA GACCTGCTTGCATTTCTCG
Is10 (RPS18) (AAN52390 ribosomal protein S18 [*B. belcheri*]) (1.00E−65))	Genbank Acc. #CA763772	55	206	TCGCCCTGACAGCCATTA CTGAACTTGCCGTCCTTG
Is66 (RPL12 (AAH08230 ribosomal protein L12 [*Homo sapiens*]) (1.00E−49))	Genbank Acc. #CA763773	61	310	CCCAAGTTTGATCCCACG ATGTCAGGTTGCCGCTGT

Table 2. (*continued*)

Locus name		T_a	Length (bp)	Primer sequences
Is149 (RPL17A) (AF395586 ribosomal protein L17/23 [*Spodoptera frugiperda*]) (1.00E−57))	Genbank Acc. #CA763774	44	293	GCATCTCCCTCGGTCTTC ATTTCGCCTTTGTTGTTGA
Is171 (RPL40Ribosomal Protein L40 ubiquitin extension protein [*D. melanogaster*] (2.00E−41))	Genbank Acc. #CA763775	68	154	GACTGGTGGGTGGTGTCATT GGTCGCAGGTTGTTGGAG
Is28 (RPL44) (AAM94276 ribosomal protein L44 [*Chlamys farreri*]) (3.00E−49))	Genbank Acc. #CA763776	45	250	ATGATAAGAAGGGCAAGGAT CTGGTGGATTAGAACTGGAT
Is27 (PRP (AF400199 ribosomal protein L37 [*S. frugiperda*] (6.00E−30))	Genbank Acc. #CA763777	58	152	GCAGCTACCACATCCAAA CTGAAACGCCTCCAGACC
SALP16 (AF061845 Ixodes scapularis salivary gland 16 kDa protein (salp16))		63	371	CAGTGAGACGGGAGCATC GTTCGGCACCTTCCTTTA
Is79 (UBQ (CAA44453 ubiquitin-conjugating enzyme [*D. melanogaster*] (2.00E−43))	Genbank Acc. #CA763778	47	196	ACAGCAATGGCAGCATCT TTCCTGGTCCACTCCCTA

Microsatellites				
ISAG4	Genbank Acc. #AF331743	41	250	AAGAAAATAAAGCGAACAAG
				ATAAGCAATTCATACGAGATAGT
ISAG25	Genbank Acc. #AF331742	57	171	AAATGTCCGAACAGCCTTAT
				GCCCTTGAGTCTACCCACTA
ISAC4	Genbank Acc. #AF331736-AF331739	49	170	GGGTCCCAACGATTGCTAAACCAG
				AAGCGTATCCGATTTGCCCTTCAT
ISAC8	Genbank Acc. #AF331740	42	170	TCTTCCCGCTGCTGTCTCGTATTC
				GAGCTACCCCTTTCATCGTCTTCG
ISAC20	Genbank Acc. #AF331734	42	150	AGAAACACGGAAGGAGAAAGGAGA
				AACTGTGCCAGATGGGAAAGAAGA
ISAC22	Genbank Acc. #AF331735	44	200	CAGCTGGGCCCCTCCTTTTAATCC
				TATTGTAAGGCCAGTCGCCGCTGC
ISCAG12	Genbank Acc. #AF331755	44	170	GAAGAACACCGAGCGAACCAGAAC
				GTGAGCTGAGGGTTGCTGTTGATG
ISGATA3	Genbank Acc. #AF331752	47	150	AGTCCCCTCAGAGCGATTTTCA
				GGCCGCCAGTTTGATGGATA
ISGATA4	Genbank Acc. #AF331753	41	200	CAGACAATGTCATTCAATCGCA
				CGCACAATGCAAAACAAATCTA
ISCTGY17	Genbank Acc. #AF331745-AF331747	41	150	TTCTTGTTTTATTGGTTGGGTG
				AATGCAGGGTAAGTTGAGATTG

[a] ∗designates E score for cDNAs of unknown function. Information following the IS# designations on cDNAs indicate the results of a BLASTX search on Genbank.

Primers were designed to amplify 250–400 bp fragments (optimal size for SSCP) in 20 of the genes (Table 2) with high matches to existing genes. However, ultimately only five of these could be mapped because the genomic DNA amplified by the primers was much larger than the anticipated size based upon the cDNA sequence. Amplified products from these loci were extracted from the polyacrylamide gels, reamplified and cloned into a plasmid vector for subsequent sequencing. As expected we found that the cDNA sequence was interrupted by large introns, some up to 800 bp in size. SSCP analysis can only reliably reveal polymorphisms in fragments <500 bp. We did not design internal primers from the genomic sequences.

Primers for 10 MS loci (Fagerberg *et al*., 2001) are also listed in Table 2. However, only six of these could be mapped, one of which was not in expected Mendelian ratios, consequently only five MS markers have been mapped. The remaining four were monomorphic, further supporting the observations of low heterozygosity found at *I. scapularis* MS loci (Fagerberg *et al*., 2001).

A linkage map was generated using JoinMap (Figure 1) (Ullmann *et al*., 2003). The linkage map contains only those loci at which genotypes were in Mendelian ratios. Fourteen linkage groups were identified by JMGRP at an LOD of 3.0. Above this LOD, markers began splitting away either singly or into separate groups. A total of 57 markers mapped across 616 cM on the 14 linkage groups with one marker every 10.8 cM. A total map distance of 3166 cM was estimated by making the assumption that the remaining 51 markers that did not cosegregate onto any of the 14 linkage groups were $\geq$50 cM from mapped markers (51 markers $\times$ 50 cM/marker) + 616 cM = 3166 cM).

Linkage groups on the map are temporarily assigned numbers according to their relative lengths. These lengths could change if more markers are added to the map. Ultimately numbers will be assigned to chromosomes *via* physical mapping when a physical map for *I. scapularis* exists. When this is accomplished then chromosome numbers will be assigned according to overall chromosome length. Currently, the map shows that linkage group 1 is four times longer than any of the other 13 linkage groups. We suspect that linkage group 1 corresponds to the sex chromosome, because it is known from cytogenetic work that in *Ixodes* the sex chromosome is three to four times longer than the autosomes (Oliver, 1977), and when sex was mapped it was most closely linked to A09.583 on linkage group 1 with an LOD of 2.667.

The size of this linkage maps suggests a very high recombination rate for the *I. scapularis* genome. The physical size of another ixodid tick genome, *Amblyomma americanum*, was previously determined to be 1.08 pg or 1.04×10^9 bp, and repetitive elements were in a long-period interspersion pattern (Palmer *et al*., 1994). Assuming this is an accurate estimate of the size of the

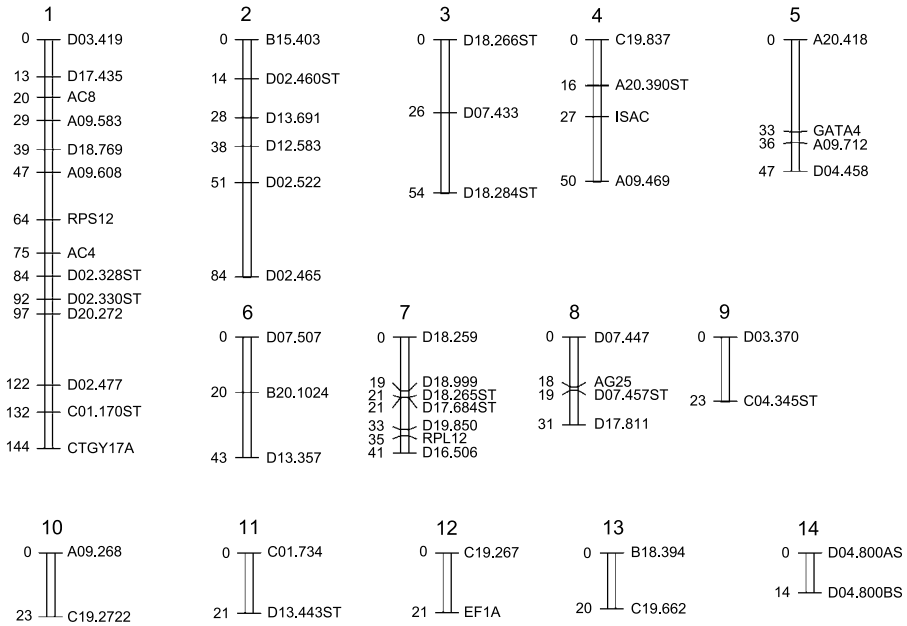

Figure 1. A preliminary linkage map of the *I. scapularis* genome that includes only those loci at which genotypes were in Mendelian ratios. Fourteen linkage groups were identified by JMGRP at an LOD of 3.0. A total of 57 markers map across 616 cM on 14 linkage groups with one marker every 10.8 cM. Linkage groups on the two maps are temporarily assigned numbers according to their relative linkage distances. These do not necessarily correspond to chromosomes numbered according to length cytogenetically (Oliver, 1977). RAPD markers are indicated by the name of the Operon primer, a period followed by the size of the amplified fragment. A STAR locus is indicated by a RAPD label followed by an 'ST'. cDNA loci follow the labels designated in Valenzuela *et al*. (2002). Microsatellite loci are indicated by AC, AG, CTGY, or GATA and the clone number.

I. scapularis genome, then the relationship of physical to genetic distance is ∼267–329 kb/cM. Work is currently underway to determine the physical size and organization of the *I. scapularis* genome.

Discussion

Our experience in generating this preliminary linkage map demonstrates both promising aspects to and pitfalls for future work in tick genetics and genomics. Having tried conventional RAPDs, STARs, MS, and cDNAs, it appears that STARs are likely to be the marker of choice for expanding the *I. scapularis* linkage map and eventually integrating the linkage map into a physical map. Of the 65 STAR primer sets designed (Table 2), 52 were polymorphic and genotypes at 37 of these conformed to expected Mendelian ratios. The

numbers of RAPD loci and the amount of heterozygosity at these loci were not as great as those detected in *A. aegypti* (Antolin *et al*., 1996). From the 20 RAPD primers, a total of 63 markers were mapped to the *I. scapularis* genome while 94 markers were generated with only 10 RAPD primers in an *Ae. aegypti* F_1 intercross family (Antolin *et al*., 1996). Reassociation kinetic studies of another ixodid tick, *A. americanum,* showed that the organization of repetitive DNA was of the long-period interspersion type (Palmer *et al*., 1994). We suspect that this is also the case for *I. scapularis.*

Fagerberg *et al*. (2001) had to use a specialized capture technique to identify a few MS loci in *I. scapularis*. Furthermore, the heterozygosity at these few loci was low. We have no explanation for the low abundance and variability at MS loci in *I. scapularis*. Sequences obtained from cDNA libraries were useful in generating PCR primers that amplified single nucleotide polymorphisms (SNPs) for linkage mapping in an *A. aegypti* F_1 intercross family (Fulton *et al*., 2001). In most cases the amplified regions were of the same size as the cDNA. However, this was not the case with cDNAs in the *I. scapularis* genome. In only 5 of the 20 cDNAs, for which primers were designed, were we able to recover a fragment that was sufficiently short for SSCP analysis. This appears to be due to the presence of large intervening introns. Taking the additional steps required to identify the locations and sequences of intron/exon boundaries will be very laborious and expensive.

We are in the process of performing the reassociation kinetics of the *I. scapularis* genome. However, assuming that the physical size of the *I. scapularis* genome is approximately the same size as the *A. americanum* genome (1.04×10^9 bp) (Palmer *et al*., 1994), perhaps an inappropriate assumption to make, our results suggest that there is a great deal of recombination in *I. scapularis*. A resolution of ~300 kb/cM compares favorably to the average 1100 kb/cM resolution in *A. gambiae* (Dimopoulos *et al*., 1996) or the 1000–3400 kb resolution for *A. aegypti* (Brown *et al*., 2001). A map of this resolution has a higher probability of being successfully used in mapped-based positional cloning of candidate genes for vector competence to pathogens, host preference, insecticide resistance and other important characters in *I. scapularis*.

References

Antolin, M.F., Bosio, C.F., Cotton, J., Sweeney, W., Strand, M.R. and Black, W.C.t. 1996. Intensive linkage mapping in a wasp (*Bracon hebetor*) and a mosquito (*Aedes aegypti*) with single-strand conformation polymorphism analysis of random amplified polymorphic DNA markers. Genetics 143: 1727–1738.

Black, W.C. and DuTeau, N.M. 1997. In: The Molecular Biology of Insect Disease Vectors: A Methods Manual, pp. 361–373. Chapman & Hall, New York.

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

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

Dietrich, W., Katz, H., Lincoln, S.E., Shin, H.S., Friedman, J., Dracopoli, N.C. and Lander, E.S. 1992. A genetic map of the mouse suitable for typing intraspecific crosses. Genetics 131: 423–447.

Dimopoulos, G., Zheng, L., Kumar, V., della Torre, A., Kafatos, F.C. and Louis, C. 1996. Integrated genetic map of *Anopheles gambiae*: use of RAPD polymorphisms for genetic, cytogenetic and STS landmarks. Genetics 143: 953–960.

Fagerberg, A.J., Fulton, R.E. and Black, W.C. 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.

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

Hayashi, K. 1991. PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. PCR Meth. Appl. 1: 34–38.

Hiss, R.H., Norris, D.E., Dietrich, C.H., Whitcomb, R.F., West, D.F., Bosio, C.F., Kambhampati, S., Piesman, J., Antolin, M.F. and Black, W.C.t. 1994. Molecular taxonomy using single-strand conformation polymorphism (SSCP) analysis of mitochondrial ribosomal DNA genes. Insect Mol. Biol. 3: 171–182.

Hudson, T.J., Stein, L.D., Gerety, S.S., Ma, J., Castle, A.B., Silva, J., Slonim, D.K., Baptista, R., Kruglyak, L., Xu, S.H. et al. 1995. An STS-based map of the human genome. Science 270: 1945–1954.

Hunt, G.J. and Page Jr., R.E. 1995. Linkage map of the honey bee, *Apis mellifera*, based on RAPD markers. Genetics 139: 1371–1382.

Kosambi, D.D. 1944. The estimation of map distances from recombination values. Ann. Eugen. 12: 172–175.

Oliver Jr., J.H. 1977. Cytogenetics of mites and ticks. Annu. Rev. Entomol. 22: 407–429.

Orita, M., Suzuki, Y., Sekiya, T. and Hayashi, K. 1989. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics 5: 874–879.

Palmer, M.J., Bantle, J.A., Guo, X. and Fargo, W.S. 1994. Genome size and organization in the ixodid tick *Amblyomma americanum* (L.). Insect Mol. Biol. 3: 57–62.

Sonenshine, D.E. 1991. Biology of Ticks, Vol. 1, New York.

Spielman, A. 1976. Human babesiosis on Nantucket Island: transmission by nymphal *Ixodes* ticks. Am. J. Trop. Med. Hyg. 25: 784–787.

Spielman, A., Wilson, M.L., Levine, J.F. and Piesman, J. 1985. Ecology of *Ixodes dammini*-borne human babesiosis and Lyme disease. Annu. Rev. Entomol. 30: 439–460.

Stam, P. and van Oojin, J.W. 1995. CPRO-DLO, Wageningen, The Netherlands.

Telford III, S.R., Dawson, J.E., Katavolos, P., Warner, C.K., Kolbert, C.P. and Persing, D.H. 1996. Perpetuation of the agent of human granulocytic ehrlichiosis in a deer tick-rodent cycle. Proc. Natl. Acad. Sci. USA 93: 6209–6214.

Ullmann, A.J., Piesman, J., Dolan, M.C. and Iv, W.C. 2003. A preliminary linkage map of the hard tick, *Ixodes scapularis*. Insect Mol. Biol. 12: 201–210.

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

Vidal-Puig, A. and Moller, D.E. 1994. Comparative sensitivity of alternative single-strand conformation polymorphism (SSCP) methods. Biotechniques 17: 490–492, 494, 496.

Williams, J.G., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl. Acids Res. 18: 6531–6535.

Yasukochi, Y. 1998. A dense genetic map of the silkworm, *Bombyx mori*, covering all chromosomes based on 1018 molecular markers. Genetics 150: 1513–1525.

Zheng, L., Benedict, M.Q., Cornel, A.J., Collins, F.H. and Kafatos, F.C. 1996. An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. Genetics 143: 941–952.