Genome analysis of Major Tick and Mite Vectors of Human Pathogens

Submitted by Catherine A. Hill on behalf of the Tick and Mite Genomes Consortium
15 December 2010

Executive Summary

Ticks and mites (subphylum Chelicerata: class Acari) transmit a greater variety of human and animal pathogens than any other arthropod vector. Tick- and mite-borne diseases are global health problems caused by a variety of bacterial, viral and protozoan pathogens, which are responsible for significant morbidity and thousands of human deaths annually. The incidence of many tick-borne diseases is increasing worldwide - many are emerging zoonoses or exotic diseases that could be introduced to the U.S. The Acari are a diverse and basal group within the phylum Arthropoda, comprised of two lineages; the Acariformes or chigger mites and the Parasitiformes which includes the ticks (suborder Ixodida) and other medically important mites. Despite their impact on human health, little is known about the biology and genetic basis of vector competence in the Acari. The NIH funded *Ixodes scapularis* (Lyme disease tick) assembly represents the only available genome sequence of a medically important species within this entire diverse lineage. Unfortunately, application of this resource is limited by low sequence coverage and fragmentation of the assembly.

Here we propose immediate sequencing of the mite vector of scrub typhus, *Leptotrombidium deliense,* and additional sequencing of *I. scapularis* to expand genomic resources for this group (summarized in Table 1). These projects will provide the anchor for additional sampling of species at increasing evolutionary distance. We also identify six members of the Ixodida (*Dermacentor variabilis, Amblyomma americanum, I. pacificus, I. ricinus, I. persulcatus* and *Ornithodorus turicata*) that are considered high priority sequencing targets by the tick and mite research communities. The haploid genome size of these ticks is expected to exceed 1 Gbp. Recognizing the inherent challenges associated with *de novo* sequencing and assembly of such genomes and the likely rapid advances anticipated in genome sequencing and assembly technology, we propose extensive transcriptome sequencing of these species to position them for anticipated genome sequencing efforts. Should sequencing of *L. deliense* and *I. scapularis* prove successful, we identify two additional species of Ixodida (*D. variabilis* and *O. turicata*) as candidates for genome sequencing efforts.

Table 1. Summary of Proposed Tick and Mite Vectors Genome Projects

Classification	Genome Sequencing (no. species at 100X genome coverage)	Transcriptome Sequencing (no. RNAseq reads in millions)*	Priority
Leptotrombidium deliense	1	100M	Tier 1
Ixodes scapularis	1	100M	
Dermacentor variabilis	1	100M	Tier 2
Amblyomma americanum	-	100M	
Ixodes pacificus	-	100M	
Ixodes ricinus	-	100M	
Ixodes persulcatus	-	100M	
Ornithodorus turicata	1	100M	Tier 3
TOTAL	**4**	**800M**	**3**

*Illumina 100 bp average read length

Identifying new strategies to control tick and mite vectors and the pathogens they transmit is a central theme of tick and mite research programs worldwide. The genome sequence data needed to accomplish these goals are not yet in hand. This proposal represents the cooperative efforts of the international tick and mite research community to develop the critical resources to facilitate comprehensive genomic studies across this important, yet largely neglected group of arthropod vectors. It outlines our long term plan for implementing genomic research in major evolutionary groups and key vector species of the Acari, and identifies the steps the community will take to achieve this goal.

1. Introduction and Background

1.1 Biomedical significance of the Acari

The Acari: an evolutionarily diverse group comprising multiple mite and tick vectors of disease

Ticks and mites are members of the phylum Arthropoda, subphylum Chelicerata (spiders, scorpions, mites and ticks) and subclass Acari. The Chelicerates include a diverse assemblage of terrestrial and marine arthropods (Jeyaprakash and Hoy, 2008), and are the second largest group of arthropods after the insects. This lineage is thought to be ancient, having diverged from Trilobites during the Cambrian explosion (Brusca and Brusca, 1990). It has been an estimated 490-550 million years since arthropods in the subphylum Chelicerata shared a common ancestor with species in the subphylum Mandibulata, which contains the order Hexapoda (Insects) (Klompen et al., 1996). Not surprisingly, the Acari differ from insect vectors of disease in many aspects of their biology. The current understanding of the phylogeny within the subclass Acari is represented in Figure 1. The split between the Acariformes (which contains, among others, important disease-transmitting mites of the family Trombiculidae or chigger mites) and the Parasitiformes (ticks and all other mites) occurred approximately 400 million years ago (Jeyaprakash and Hoy, 2008). Species classified within both of these lineages are important disease vectors. The superorder Parasitiformes includes the sub-order Ixodida which is comprised by the families Argasidae (soft ticks), Ixodidae (hard ticks) and the Nuttallielidae (represented by only a single species and, consequently, which will not be discussed further). The Ixodidae is divided into two lineages, the Prostriata which contains the single genus *Ixodes* comprised of approximately 249 species, and the Metastriata (all other tick genera) which contains approximately 464 species. It has been proposed that species classified in the subclass Ixodida evolved from free-living, saprophytic mites (Klompen et al., 2000) and that hard ticks evolved from bird-feeding soft ticks (Black and Piesman, 1994; Ribeiro et al., 2006). However, there is much debate about the evolution of blood-feeding ability in the ticks which may have evolved multiple times. The scope of vector competence among Acarine species has not been thoroughly analyzed but the ability to transmit bacteria, rickettsia, protozoa and viruses appears to be widely distributed across all major groups of the Parasitormes (Table 4).

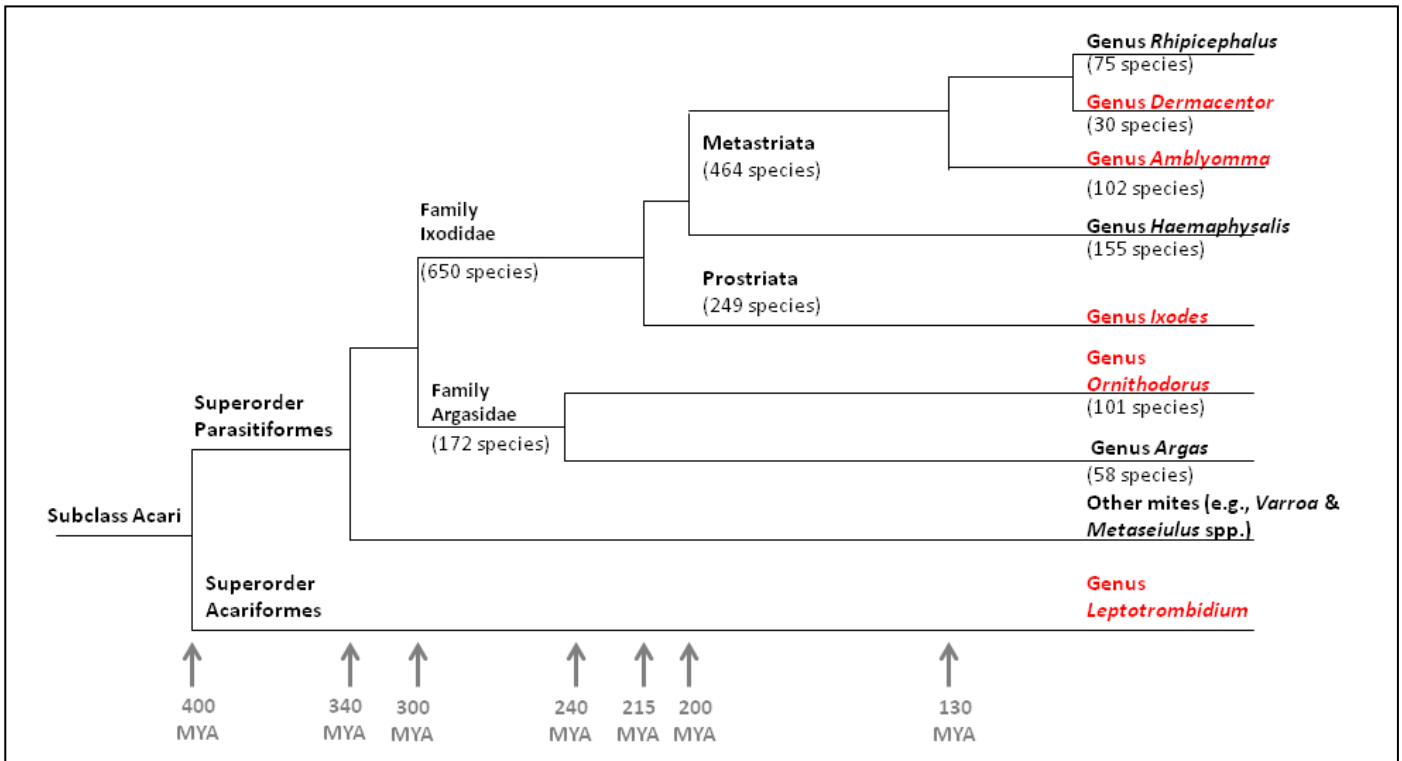

Figure 1. Phylogenetic relationships and first divergence time estimates of major groups in the Acari. After Balashov (1994), Klompen et al., (2000), Barker and Murrell (2004) and Jeyaprakash and Hoy (2008). Genera to be sampled are highlighted in red.

Mite vectors of disease

It is estimated that as many as three million mite species exist on earth. The vast majority of the approximately 40,000 named mite species are harmless to humans or beneficial in ecosystems (Hoy, 2009). More than 250 species are recognized as the cause of health-related problems for humans and domestic animals. The most widely recognized mite problems affecting human health are mite-induced allergies. There are very few known human diseases that involve pathogens transmitted by mites; the most important is tsutsugamushi disease, or scrub typhus, caused by the bacterium (*Orientia tsutsugamushi*) that is transmitted by several members of the genus *Leptotrombidium* (Family Trombiculidae) in south eastern Asia, Australia and the pacific islands (Mullen and Durden, 2002). There are more than one million cases of scrub typhus every year and greater than one billion people are at risk for infection (Lerdthusnee et al., 2002). Trombiculid mites do not blood feed; these mites ingest epidermal cells and lymph that they digest largely externally through enzymes secreted in the saliva. Only the larvae of trombiculid mites are parasitic; they feed primarily on ground-dwelling rodents and humans are incidental hosts.

While scrub typhus is a major concern historically, recent evidence suggests that *Leptotrombidium* mites harbor (and may vector) other pathogens such as hantavirus (Houck et al. 2001) and *Bartonella* species such a *B. tamiae*, the causative agent of bartonellosis in humans (Kabeya et al. 2010). The latter study also shows that other mite genera (*Schoengastia* and *Blankarrtia*) belonging to the same family (Trombiculidae) as the genus *Leptotrombidium* can harbor *Bartonella* spp. (Kabeya et al. 2010). Furthermore, other mite genera belonging to the Trombiculid family (e.g., *Neotrombicula* and *Eushoengastia*) can harbor *O. tsutsugamushi* (B. Evans, pers. comm.). These observations suggest that mites may in fact play a greater role in disease transmission than is commonly assumed.

Tick vectors of disease

There are approximately 907 valid species of ticks; the majority of these are ectoparasites of wildlife and approximately 10% are recognized as vectors of disease to humans and animals or for their ability to cause direct damage through blood feeding (Jongejan and Uilenberg, 2004) and/or paralysis and toxic injury. Ticks transmit the greatest variety of human and animal pathogens of any arthropod vector and are second only to mosquitoes as vectors of human disease (Fivaz *et al.,* 1992; Sonenshine and Mather, 1994). Diseases transmitted by blood feeding ixodid and argasid ticks are global medical and veterinary health problems (Sonenshine, 1993) and include a wide variety of bacterial, rickettsial, viral and protozoan diseases. Other forms of pathogenesis attributed to ticks include anemia, dermatosis, toxemia and paralysis (Gothe, 1999; Roberts and Janovy, 1996; Sonenshine, 1991; Sonenshine, 1993). Important tick-borne diseases include Lyme disease (LD), tick-borne relapsing fever, babesiosis, anaplasmosis, Rocky Mountain spotted fever, Boutonneuse fever, Queensland tick typhus, Q fever, and numerous arboviruses (Sonenshine, 1993). The resurgence of LD and the emergence of other tick-borne diseases such as human granulocytic anaplasmosis (HGA) (Childs and Paddock, 2003; Gratz, 1999; Paddock and Childs, 2003) pose increasing public health concerns. Since the discovery of the causative agent of LD, *Borrelia burgdorferi*, fifteen previously unrecognized tick-borne bacterial pathogens have been described (Parola and Raoult, 2001). Furthermore, due to their efficiency as vectors of a wide variety of pathogens, broad vertebrate host range and worldwide distribution (Sonenshine, 1991), ixodid ticks are recognized as potential vectors of a number of pathogens considered to be possible bioterrorism agents for use against humans and livestock. These include Crimean-Congo hemorrhagic fever virus, *Rickettsia rickettsii* (Rocky Mountain Spotted Fever), the tick-borne encephalitis complex of flaviviruses (Central European tick-borne encephalitis, Far Eastern tick-borne encephalitis, Siberian tick-borne encephalitis, Kyasanur forest disease and Omsk hemorrhagic fever), *Coxiella burnetii* (Q Fever) and *Francisella tularensis* (tularemia) (Centers for Disease Control and Prevention Select Biological Agents and Toxins, 2004; http://www.cdc.gov/od/sap/docs/salist.pdf).

The Prostriata (genus *Ixodes*) represents an evolutionarily primitive phyletic line of the Acari that includes a number of medically significant tick species world-wide. Multiple *Ixodes* species are competent vectors of a variety of pathogens, most notably the *Borrelia* spp. spirochete that cause Lyme borreliosis in humans and animals. Lyme disease is arguably one of the most important vector borne diseases in the US, Europe and Asia. Over 33,000 positive LD cases were reported in the US in 2008 (Centers for Disease Control and Prevention, 2010). LD and other tick-borne diseases have important long term health consequences. In the United States, *I. scapularis* is the most important tick species from a human health perspective; *I. scapularis* transmits LD in the northeastern and north-central US, HGA and babesiosis, and the flavivirus that

causes Powassan encephalitis (POW), which is the North American representative of the TBE complex. Recent studies by Anderson *et al*., (2003) also suggest that West Nile virus can be transmitted trans-stadially by *I. scapularis* although vector competence has not yet been established. Of further concern is the fact that the incidence and geographic spread of LD and other tick-borne disease are increasing and many cases are suspected to be vastly under-reported or misdiagnosed (Walker, 1988). Other important species in this genus include *I. pacificus*, the vector of LD on the U.S. Pacific Coast, *I. ricinus* and *I. persulcatus,* the Eurasian *Ixodes spp*. vectors of LD and tick borne encephalitis, and *I. holocyclus*, an Australian ixodid species responsible for transmission of *Rickettsia* and *Borrelia* and human cases of tick paralysis. These factors, particularly the wide range of human diseases that it transmits, were the primary justification for selection of *I. scapularis* in 2004 for a genome sequencing effort seeking to have an ultimate impact on human welfare through development of novel vector suppression measures, therapeutics and vaccines. Already, this investment has begun to bear fruit (see below).

The Acari: status of genomic research and the need for additional genome sequence

The tick research community is developing genomic resources (summarized in Table 2) in anticipation of genome sequencing efforts for a number of vector species. The *I. scapularis* genome project has generated a draft genome assembly, trace data, ESTs and BAC clones that are used widely by the scientific research community. Other resources to support an expanded sequencing and assembly effort for *I. scapularis* include FISH-based physical mapping techniques (Meyer et al., 2010) and a preliminary genetic map (Ullmann et al., 2003). Transcript data are invaluable for identification of novel biopharmaceuticals and antimicrobial peptides, targets for acaricide and vaccine development, as well as determining genes important for vector competence. Large numbers of ESTs and RNAseqs have been generated for the ticks *I. scapularis* and *D. variabilis,* but additional sequence is necessary for comprehensive gene discovery in these species. The remaining sequencing targets identified in this proposal currently lack adequate transcript data and robust gene discovery efforts are essential to advance genomic research in these important vector species.

Table 2. Genomic resources for the Acari

Species	No. ESTs[1]	RNAseqs (in million bp)	BAC Library	Other
Superorder Acariformes				
*Leptotrombidium deliense**	3	-	-	-
Superorder Parasitiformes				
Family Ixodidae (Lineage Prostriata) *Ixodes scapularis**	193,773	50M[2]	Yes	Genetic map[6], FISH-based physical mapping[7]
*Ixodes pacificus**	149	-	-	-
*Ixodes ricinus**	1,969	-	-	-
*Ixodes persulcatus**	73	-	-	-
(Lineage Metastriata) *Dermacentor variabilis**	23,051	1M[2]	-	Microarray[2]
Dermacentor andersoni	124	-	-	-
*Amblyomma americanum**	6,480	-	-	Microarray[8]
Rhipicephalus appendiculatus	18,422	-	Yes[4]	-
Rhipicephalus microplus	52,838	7.2M[3]	Yes[5]	Microarray[9]
Family Argasidae *Ornithodorus spp.**	532	-	-	-

[1]from NCBI dbEST database (13 November, 2010); [2]by 454 and Illumina sequencing, D. Sonenshine & M. Roe, pers. comm.; [3]F. Guerrero, Pers. comm.; [4]R. Bishop, pers. comm.; [5]Guerrero et al., 2006; [6]Ullmann et al., 2003; [7]Meyer et al., 2010; [8]Aljamali et al., 2009; [9]Saldivar et al., 2008; BAC, bacterial artificial chromosome clone; *sequencing target identified in this proposal.

2. Rationale

2.1 Sequencing of multiple tick and mite vectors will permit identification of the genetic basis for disease transmission and the development of novel strategies for disease control

Genome sequence will be used to understand the role of ticks and mites in the transmission of emerging and re-emerging infectious disease. Sequence data offer an opportunity to study vector-host-pathogen relationships, to identify novel biologically active molecules and to develop new vaccine and acaricide targets for tick and mite pests. Numerous studies have investigated the molecular cross-talk of the tick-host-pathogen interface (Anguita *et al.,* 2002, Leboulle *et al*., 2002) and many laboratories have established the necessary animal and pathogen models for these studies (T. Mather, pers. comm.). The competence of ixodid ticks to vector many diseases has been the focus of detailed studies (Edlow, 2002). Ticks have a complex innate immune system; nevertheless, many infectious microbes have evolved mechanisms for evading these immune defenses and are able to be transmitted when their competent hosts' blood feed (Sonenshine and Hynes, 2008). There is strong evidence that components of ixodid tick saliva can promote pathogen transmission (reviewed by Nuttall, 1998; Schoeler and Wikel, 2001) but the specific components of saliva responsible for this have not been identified. Vector competence for many tick-borne diseases is presumed to be under genetic control (Ochanda et al., 1998; Young et al., 1995). Many studies are aimed at elucidating these mechanisms (Estrada-Pena et al., 2009) and a significant number are directed toward understanding the ability of ticks to vector newly identified species of spirochetes and potential agents of bioterrorism (Azad and Radulovic, 2003). Despite the fact that genes determining vector competence are obvious targets for control strategies, the genetic basis of vector competence in ticks has received limited research attention. Tick and mite genome sequence will help to unravel the complicated molecular mechanisms that underpin pathogen acquisition and transmission of a wide range of emerging and resurging tick- and mite-borne diseases.

Ticks remain attached to the host for long time periods and have evolved unique blood feeding and osmo-regulation strategies, and mechanisms to modulate and avoid host immune responses. Tick salivary glands play a critical role in survival and maintenance of water balance and supply a pharmacopoeia of compounds that promote blood feeding and pathogen transmission. Various laboratories are currently focused on the genomic and proteomic analysis of ticks and in particular, tick salivary gland proteins (the sialome) as a source of targets for vaccine and acaricide development (de la Fuente and Kocan, 2003; Ribeiro and Francischetti, 2003; Trimnell *et al.,* 2002; Valenzuela, 2002; Wikel *et al*., 2003). Patents have already been issued for application of ixodid cement as a surgical adhesive and ixodid histamine binding proteins and a mast cell tryptase inhibitor as novel anti-inflammatory drugs, with the former already at the stage of phase IIb human clinical trials for treatment of allergic conjunctivitis (P. Nuttall, pers. comm.). The TickGARD vaccine commercialized for control of *B. microplus* in Australia is based on a tick gut glycoprotein and is one of the few successful recombinant vaccines to be developed for tick control (reviewed by Willadsen *et al*., 1995). The recombinant vaccine against *B. burgdorferi* based on a tick-expressed antigen (Fikrig *et al*., 1990; Sigal *et al*., 1998) for protection against Lyme disease in endemic areas of the United States was recently withdrawn from the market for reasons of safety and efficacy. Tick and mite genome sequence will enable expanded discovery efforts directed at the development of new vaccines and acaricides for tick/mite and disease control.

2.2 Tick and mite population genetics

One of the fields of tick and mite research that has been affected most by lack of genome data is that of tick and mite population genetics. Very few markers have been developed to date for tick and mite genetics research and only one genetic map, namely a preliminary *I. scapularis* linkage map, has been published for a member of the Acari (Ullmann et al., 2003). Historically, efforts to develop molecular markers for ticks have proved challenging, and have presumably been confounded by the high repeat content of tick genomes. Many studies of the Acari are based on mitochondrial 12S and 16S rDNA markers and small numbers of microsatellite markers. With the exception of *I. scapularis*, analyses at a genome-wide scale are simply not possible. Consequently, studies of tick and mite population structures at both the micro- and macrogeographic scale, and the resolution of systematic relationships of the Acari have been limited. This has prevented efforts aimed at understanding the genetic basis of host preference, adaptation to ecological niches, vector competence, acaricide resistance and other traits of interest. It has also hindered the development of

diagnostic tools for species identification and monitoring of important phenotypes such as acaricide resistance, which is wide spread in many populations of ticks and mites. The availability of the *I. scapularis* genome assembly is rapidly changing the research landscape as it is facilitating the discovery of large numbers of SNPs for population genetics studies (L. Beati & C. Hill, pers. comm.). Additional tick and mite genome sequence will provide a major boost to current efforts to develop added markers for phylogenetic analysis to help solve outstanding questions regarding the evolution of vectoring capabilities of the Acari.

2.3 Comparative and evolutionary genomic studies

The *I. scapularis* genome was the first genome sequence produced for the evolutionarily diverse sub-phylum Chelicerata; this genome sequence has facilitated valuable comparative and evolutionary eukaryotic genomics studies across the phylum Arthropoda. The sequencing of the targets identified in this proposal will expand the scope and comparative power of these analyses. Following approval and development of the *I. scapularis* project in 2004, genome sequencing efforts have been established for four additional members of the Acari. These are the plant parasitic two-spotted spider mite (*Tetranychus urticae*; Grbic et al., 2007), the varroa mite, *Varroa destructor* parasite of honey bees (Cornman et al., 2010), the predatory phytoseiid mite, *Metaseiulus occidentalis* (Hoy, 2009), and the southern cattle tick, *Rhipicephalus microplus* (Guerrero et al., 2006). Data from these projects will be invaluable for the assembly and annotation of the tick and mite species described in this proposal. However, it should be noted that of these five species, only *I. scapularis* is a vector of pathogens to humans. 454 pyrosequencing has been used to generate a draft (8X coverage) *T. urticae* genome assembly, and to successfully sample the *V. destructor* (2.4 Gbp sequence reads, 4.3X coverage; N_{50} contig length: 2,262 bp; Cornman et al., 2010) and *R. microplus* (1.8 Gbp sequence reads, 1.6X coverage of the gene-rich fraction of the genome, N_{50} contig length: 624 bp; Guerrero et al., 2010) genomes. Sequencing of *V. destructor* is ongoing; the release of a draft genome assembly (50X coverage) and transcriptome data is anticipated in 2011 (J. Evans, pers. comm., see letter of support attached). Sequencing of the *M. occidentalis* genome and transcriptome data is expected within the next two to three years (M. Hoy, pers. comm.). There is considerable interest from the veterinary entomology community to develop a genome sequence for a tick species of animal health significance. Dr. F. Guerrero of the USDA-ARS Livestock Insect Research Laboratory in Kerrville, TX is leading an international effort to sequence the *R. microplus* genome. *Rhipicephalus microplus* is a one-host tick that causes significant losses in animal production systems in tropical and sub-tropical regions of the world (Guerrero et al., 2006). A BAC library has also been generated from *R. appendiculatus*, an important vector of East Coast fever to cattle, as part of collaborations between ILRI and USDA. Both *B. microplus* and *R. appendiculatus* are members of the Metastriata, a tick lineage comprising 17 genera of ticks that vector many pathogens of medical and veterinary significance (Ahmed and Mehlhorn, 1999; Kocan et al., 2002; McQuiston et al., 2003; Wagner et al., 2002). The sequencing targets proposed here will complement these efforts and permit valuable comparative studies between highly divergent ixodid tick species and between the acariform and parasitiform lineages.

3. Sequencing targets, priorities and considerations

3.1 Whole genome assembly of *I. scapularis* and the need for an improved assembly.

NIAID approved the sequencing of the *I. scapularis* genome in May 2004. The genome was sequenced using a whole genome shotgun sequencing approach. The assembly (IscaW1; ABJB010000000) was performed at JCVI and consists of 369,495 scaffolds built from 570,637 contigs with an estimated 3.8X coverage of the assembly, representing a combined assembly length of 1.4 Gb. The *I. scapularis* assembly and annotation were released in 2008 and statistics are summarized in Table 3.

The *I. scapularis* assembly is highly fragmented; multiple short scaffolds (~370,000; N_{50} scaffold length ~ 50 Kb) limit the accurate annotation of gene models and the identification of gene families and other features of interest. Genome analyses are also limited by the fact that approximately one-third of the *I. scapularis* genome is not represented by the assembled sequence. The assembly and annotation of the *I. scapularis* genome proved challenging for several reasons. Firstly, fold coverage (3.8X) of the *I. scapularis* genome is lower than anticipated because the haploid genome is large (2.1 Gb or nearly 8 fold larger than the *Anopheles gambiae* genome and almost double the size of the *Aedes aegypti* genome). Secondly, the genome comprises approximately 70% repetitive DNA (Geraci et al., 2007; Ullman et al., 2005). Over 20,000 gene models were

predicted by automated annotation of the *I. scapularis* assembly (scaffolds shorter than 10 kb were excluded from automated analyses); many gene models are incomplete and some were undoubtedly missed. The vector community relies heavily on the *I. scapularis* assembly and annotation and there is a desperate need to improve these resources. However, flow cytometry and reassociation kinetics studies of multiple hard and soft tick species suggest that large, repeat-rich genomes are a common feature of species of Ixodida (Geraci et al., 2007; Palmer et al., 1994; Ullmann et al., 2005). These challenges must be considered when designing strategies for sequencing the genomes of these ticks.

Table 3. Summary of the *I. scapularis* genome assembly and annotation statistics

IscaW1 Assembly Statistics	
Total no. sequence reads	17.4 million
Estimated fold coverage of the assembly	3.8 fold
Number of scaffolds	369,495
N_{50} scaffold length	51,551 bp
Total length of combined contigs	1.4 Gb
Total length of combined scaffolds (including gaps)	1.8 Gb
Annotation Release 1.0 Statistics	
Total no. genes	20,486
Mean gene length	10,589 bp
Mean CDS length	855 bp

4.2 Choice of sequencing targets

Table 4 summarizes both the immediate and longer term sequencing priorities identified by the community of researchers working on medically important ticks and mites. The species proposed for genome sequencing are important vectors of multiple human and animal pathogens, and each is representative of a clade of vectors within the class Acari. *Leptotrombidium deliense* (tier 1) is the mite vector of scrub typhus in southeast Asia where there are more than a million cases every year and one billion people at risk (Lerdthusnee et al., 2002). *Ixodes scapularis* (tier 1) is a prostriate tick vector of Lyme disease (LD), the most common vector-borne disease in North America and Europe; more than 33,000 cases (28,921 confirmed; 6,277 probable) were reported in the U.S. in 2009 (MMWR, 2010), and more than 100,000 cases were reported in Europe in 2002 (WHO, 2002). *Ixodes* species ticks are also competent vectors of the bacteria and protozoa that cause human granulocytic anaplasmosis (HGA) and babesiosis, respectively; 1,009 cases of HGA were reported in 2009, and more than 450 cases of *Babesia microti* infections have been confirmed in the U.S. The actual prevalence of human babesiosis is unknown because most individuals are asymptomatic and are not tested for the pathogen. *Dermacentor variabilis* (tier 2) vectors the bacterium, *Rickettsia rickettsiae* that causes the potentially fatal Rocky Mountain Spotted Fever (RMSF); 2,563 human cases were reported in the U.S. in 2009. *Amblyomma americanum* (tier 2) vectors an emerging bacterial zoonosis - human monocytic erhlichiosis (HME) in the U.S; 957 cases were reported to the CDC in 2009. These metastriate ticks are also potential vectors of the bacterium *Francisella tularensis*, which is recognized as a potential bioterrorism agent. *Ornithodorus turicata* (tier 3) is a soft tick (family Argasidae) vector of the bacterium that causes tick-borne relapsing fever (TBRF); there are multiple cases of TBRF in the U.S. every year. This argasid tick will serve as an outgroup to the ixodid ticks. Having one representative of the superorder Acariformes, lineage Prostriata, lineage Metastriata and family Argasidae is a minimum on which to build a framework for future tick/mite genome research, and promote studies of genome evolution and vector competence among ticks and mites.

In the preparation of this document, members of the Tick and Mite Genomes Consortium considered multiple species of ticks and mites as sequencing targets. Two criteria governed the selection of species. These were: (1) vector status and (2) availability of a colony suitable for sequencing. There is considerable interest amongst community scientists in the identification and comparative analyses of closely related vector/non-vector ticks and mites as has been pioneered for mosquitoes that vector malaria (Besansky, 2008). However, the consortium was unable to identify acarine species incapable of pathogen transmission. Three parasitiform mite species associated with serious allergic conditions in humans were also considered as an approach to increase sequence representation of medically important mites. These were the scabies mite (*Sarcoptes scabei*), the house dust mite (*Dermataphagoides pteronyssinus*), and the straw itch mite (*Pyemotes ventricosis*). The scabies mite is a highly contagious ectoparasite of humans that causes severe dermatitis,

particularly in immune compromised individuals. The house dust mite is a ubiquitous, cosmopolitan pest and is responsible for severe rhinitis in a significant percentage of the human population. Although these three mite species are not recognized as vectors of disease causing organism, the generation of genomic resources will be essential to advance research in these mites. They are identified here by the Tick and Mite Genomes Consortium as important targets for future genome sequencing efforts.

Table 4. Proposed Tick and Mite Genomes, Clinical Significance and Sequencing Priority

Acari Classification	Species/Geographic Region	Diseases Transmitted[1]	Priority
Superorder Acariformes	*Leptotrombidium deliense* Asia	Scrub typhus	**Tier 1**
Superorder Parasitiformes	*Ixodes scapularis* Nth. America	LD, HGA, babesiosis, POW	
Family Ixodidae (hard ticks)	*Ixodes pacificus* Nth. America	LD, HGA	**Tier 2**
Lineage Prostriata	*Ixodes ricinus* Africa/Eurasia	LD, TBE, babesiosis, HGA	
	Ixodes persulcatus Eurasia	LD, TBE	
Lineage Metastriata	*Dermacentor variabilis* Nth. & Central America	RMSF, tularemia, anaplasmosis, tick-induced paralysis	
	Amblyomma americanum Nth., Central & Sth. America	HME, STARI, tularemia	
Family Argasidae (soft ticks)	*Ornithodorus turicata* Nth. America	TBRF	**Tier 3**

[1]From Jongejan and Uilenberg (2004) and Pagel VanZee et al., 2007; human babesiosis (*Babesia microti*); HGA, human granulocytic anaplasmosis (*Anaplasma phagocytophilum*); HME, human monocytic erhlichiosis (*E. chaffeensis*); LD, Lyme disease (*Borrelia burgdorferi)*; POW, Powassan virus; RMSF, Rocky Mountain spotted fever (*Rickettsia rickettsii*); scrub typhus (*Orientia tsutsugamushi*), STARI, southern tick associated rash illness (*Borrelia lonestari)*; TBRF*,* tick-borne relapsing fever (*Borrelia turicatae*.); ND, not determined.

4.3 Suggested sequencing strategy

The two-phase sequencing strategy proposed by the Tick and Mite Genomes Consortium is outlined below and was developed in consultation with Dr. E. Caler, Assistant Professor, Eukaryotic Genomics, J. Craig Venter Institute (JCVI).

Table 5. Proposed sequencing strategy

Priority	Species	Colony	Genome Size	Genome Coverage	Transcriptome Sequencing*
Tier 1	*L. deliense*	AFRIMS	ND	24X 454 100X Illumina[†]	100M
	I. scapularis	WIKEL	2.1 Gb[1]	100X Illumina[‡]	100M
Tier 2	*I. pacificus*	WIKEL	ND	-	100M
	I. ricinus	WIKEL	ND	-	100M
	I. persulcatus	WIKEL	ND	-	100M
	D. variabilis	WIKEL	2.9 Gb[1]	100X Illumina[‡]	100M
	A. americanum	WIKEL	3.3 Gb[1]	-	100M
Tier 3	*O. turicata*	Sonenshine	1.0 Gb[1]	100X Illumina[‡]	100M

[†]Hybrid 454-Illumina sequencing & *de novo* assembly; [‡]Illumina (supplemented with 454 paired-end runs if necessary) and *de novo* assembly; * Illumina 100bp average reads in millions; [1]after Geraci et al., 2007; AFRIMS, Armed Forces Research Institute for Medical Sciences; ND, not determined.

Phase I: Developing reference mite and tick genome assemblies

Developing a reference genome of a medically important trombiculid (acariform) mite

Our immediate goal is to develop a high quality reference genome for a medically important mite. The predicted genome characteristics of *L. deliense* make this mite the best option to achieve this goal. The expected haploid genome size of *L. deliense* is approximately 500 Mbp based on that of the closely related acariform mite, *Tetranychus urticae* (75 Mbp; Grbic et al., 2007) and the parasitiform mites, *Metaseiulus occidentalis* (88 Mbp; Hoy 2008) and *Varroa destructor* (~450 Mbp, Cornman et al., 2010), and is relatively small compared to the genomes of ixodid ticks. We propose deep sequencing of the *L. deliense* genome to approximately 125X using a combination 454-Illumina approach (Table 5). This will accomplish multiple objectives: it will provide (1) the first genome sequence of a medically important trombiculid mite and increase genomic representation of species in the super-order Acariformes (*Tetranychus urticae* is the only acariform species assembly to date; M. Grbic and M. Navajas pers. comm.), (2) an important reference genome to assist the re-assembly and re-annotation effort proposed for *I. scapularis* below, and (3) the framework for expanding sequencing efforts to key vectors identified as tier 2 and 3 priorities.

The AFRIMS has established several *L. deliense* colonies from single-pair matings of mites collected from the field and one of these colonies will be expanded for this project. It will be important to determine the haploid genome size of *L. deliense* prior to sequencing using flow cytometry or possibly other approaches suitable for mites. The small physical size of *L. deliense* will necessitate DNA extraction from a single embryo batch, but deep sequencing should overcome heterozygosity issues. Female *L. deliense* typically produce one to several thousand embryos. Based on work with other mite species, we anticipate obtaining nanogram to microgram quantities of DNA from a single embryo batch, sufficient for library production and sequencing. We note that 454 pyrosequencing of the *V. destructor* genome was conducted using DNA extracted from a single batch of approximately 1,000 eggs (Cornman et al., 2010) and this is also the approach proposed for the *M. occidentalis* project. An alternative we will consider is whole-genome amplification from a single adult female mite. The successful assembly and annotation of the *T. urticae* genome suggests that our proposed approach is feasible, and this assembly provides an invaluable tool for assembly and annotation of *L. deliense*. The *L. deliense* sequence will facilitate important comparative analyses of acariform and parasitiform vectors, and in particular, comparison of a non-blood feeding mite vector with blood-feeding tick vectors of disease.

Developing a reference genome of an ixodid (parasitiform) tick: improving the draft I. scapularis "core" genome

An additional and immediate goal is to develop a high quality reference genome of an ixodid tick to promote more comprehensive identification of genes and non-coding DNA of interest such as transcriptional regulatory elements, and transposable elements and other classes of repetitive DNA in ticks. Several characteristics recommend *I. scapularis* for this role: (1) *Ixodes* spp. ticks are some of the most important vectors of human disease worldwide and *I. scapularis* is the most important tick vector in the U.S., (2) *I. scapularis* has the smallest haploid genome of any ixodid tick investigated to date, and (3) an existing assembly and multiple genomic resources (e.g., ESTs, BAC library, cell lines, physical mapping techniques and a preliminary genetic map) are available and provide a foundation for this effort. Physical mapping and genome sequence analysis has shown that the heterochromatic telomeric, sub telomeric and centromeric regions of *I. scapularis* chromosomes comprise kilobase stretches of extremely low complexity tandem repeats (Meyer et al., 2010); we recognize that it may not be feasible to assemble these regions in *I. scapularis* and other ixodid ticks. Our goal is to generate an improved "core" assembly of the euchromatic genome to enable more thorough identification and annotation of coding and non-coding elements.

To achieve this goal, we are considering several approaches. We propose deep sequencing of *I. scapularis* using an Illumina approach (100 bp average read length), possibly supplemented with 454 paired end reads of 8 kb and 20 kb clones to close gaps and link scaffolds. We recognize that repeats will complicate the assembly of *I. scapularis*, but we still expect to obtain useful isochores and resolve small repeat regions by this method. If possible, we suggest sequencing of a single tick to minimize heterozygosity. We will attempt Illumina, and possibly low level 454 coverage to generate large contigs for gene rich regions. We note that 454 has been used to sequence multiple individual *Anopheles* mosquitoes (M. Donnelly, pers. comm.) but this approach is likely prohibitive in the case of larger tick genomes. An alternative we are considering is sequencing of a high $C_o t$ "euchromatin enriched" fraction from a single embryo batch of *I. scapularis* to minimize the repetitive fraction. This approach has recently proved successful in combination with 454

pyrosequencing for generating a preliminary assembly of the ixodid tick, *R. microplus* (Guerrero et al., 2010) which has a genome size (~7.1 Gb) more than twice that of *I. scapularis* (Ullmann et al., 2005). Successful *de novo* assembly of the *L. deliense* genome will provide proof of concept for this effort and an important reference genome for an improved *I. scapularis* assembly. We will also consider targeted sequencing of *I. scapularis* BACs via a 454 or traditional Sanger approach to improve the core genome assembly. An. *I. scapularis* BAC library (10X clone coverage, 120 kb average insert size) was produced as part of the NIH funded *I. scapularis* genome project and is available for this purpose. Forty five *I. scapularis* BACs have been sequenced and successfully assembled and annotated to date and there is precedent for expanding this BAC-based sequencing effort to complement the next gen sequencing proposed here.

Ixodes scapularis may be a candidate for sequencing using Solexa (10 kb libraries) or Pacific Biosciences technologies should these become available within the timeframe of this project. We will investigate the possibility of combining sequence data with optical mapping techniques to improve assemblies. High resolution optical (physical) mapping is an attractive approach for ordering and joining sequence data. It has been used successfully to guide and validate assemblies of the rice and maize genomes (Zhou et al., 2007; 2009). Recently, the NIH/NIAID approved a Driving Biological Project through VectorBase (http://www.vectorbase.org/Other/News/?id=125) to build optical maps for several vectors; *I. scapularis* was one of five vector species identified as a candidate for assembly improvement by this method. Optical mapping will be performed in collaboration with Dr. David Schwartz, University of Wisconsin, Madison, who is the PI of the optical mapping effort funded through VectorBase.

Phase II: Developing reference metastriate and argasid genome assemblies

The long-term objective of the tick/mite research community is to generate a reference genome assembly for a metastriate and an argasid tick vector. The metastriate lineage includes multiple genera of ticks that vector disease to humans and animals. The consortium has identified two metastriate species as high priority (tier 2) sequencing targets based on their importance to human health. These are *Dermacentor variabilis* (American dog tick), the vector of Rocky Mountain Spotted Fever, tularemia and anaplasmosis, and *Amblyomma americanum* (lone star tick), a vector of ehrlichiosis and Southern tick-associated rash illness (STARI). The argasid tick, *Ornithodorus turicata*, has also been identified by the consortium as an important (tier 3) sequencing target. We acknowledge that tier 2 and 3 sequencing targets represent a challenge for assembly and annotation because these species have large (> 1Gbp), repeat rich genomes (based on Geraci et al., 2007, Palmer et al., 2004 and Ullmann et al., 2005). The estimated haploid genome size of *D. variabilis*, *A. americanum* and *O. turicata* as determined by flow cytometry analysis is 2.9 Gbp, 3.3 Gbp and 1.0 Gbp, respectively (Geraci et al., 2007). If we are able to generate an improved core reference *I. scapularis* genome assembly, we recommend sequencing of one tier 2 metastriate tick (*D. variabilis* or *A. americanum*) and one tier 3 argasid tick (*O. turicata*) to 100X genome coverage using paired end Illumina reads and sequencing of single ticks to minimize heterozygosity. This will generate much needed sequence data for multiple tick vectors and may be useful to improve the *I. scapularis* reference assembly. *Dermacentor variabilis* is particularly well positioned for a genome sequencing effort as multiple genome resources (ESTs, RNAseq and a microarray) have been established by the *Dermacentor* research community (Table 2).

Development of EST datasets to improve gene identification and annotation in the Acari

Very few acarine ESTs have been produced to date (Table 2), and even the ~195,000 *I. scapularis* ESTs are insufficient for thorough gene annotation. In parallel with Phase I activities, we propose deep (100M reads) transcriptome sampling by Illumina (100bp reads) to support discovery and annotation of the estimated 20,000 genes in the genomes of *L. deliense* and *I. scapularis*. We also propose transcriptome sampling of multiple tier 2 and tier 3 species during this phase, to position these tick species for genome sequencing efforts. This will facilitate comparative analyses of transcriptomes from closely related *Ixodes* vectors, and between the more divergent metastriate and argasid ticks. To increase gene discovery, sampling of pooled tissue libraries is recommended for each target species identified in this proposal. Transcriptome analyses will be performed using cDNA from whole female, whole male, immature developmental stages (larvae and nymphs), and where possible, infected and uninfected ticks and mites. We are confident that this approach will be successful; two consortium members, Drs. Sonenshine and Roe have recently produced RNAseq data from an *I. scapularis* synganglia (brain) and a *D. variabilis* brain library. We suggest that transcriptome sampling be

performed early in the project to generate data for community scientists and facilitate functional studies ahead of genome sequencing and assembly.

Annotation strategy

Comprehensive genome annotation is an essential component of the genome sequencing efforts described in this proposal. Multiple curated arthropod genomes and other genomic resources (e.g., ESTs, RNAseqs and proteomics data) are available and will assist the annotation of the new genomes we propose, and the improvement of the genome annotation for *I. scapularis*. RNAseq datasets from pooled tick and mite stages will be useful in several ways. In the case of *I. scapularis*, these sequences will be added to the current gene set to validate existing gene models and determine stage/tissue specific transcripts, and to promote gene discovery. RNAseq data for *L. deliense,* a metastriate tick (*D. variabilis* or *A. americanum)* and the argasid tick, *O. turicata* will be mapped onto *de novo* genome assemblies as these become available and used to generate gene models. This will provide a training set that together with protein homology derived gene structures, will be used for *ab initio* gene identification in ticks and mites. Gene annotation will be greatly benefited by community input. Workshops and jamborees are great avenues to capture expert gene annotations. After genome assembly and preliminary automated annotation have been validated in Phase I of this project, we propose to invite community scientists to participate in a collaborative effort to manually curate the *L. deliense* and *I. scapularis* annotations. We will identify and invite key investigators to participate in a 3 to 4 day workshop, possibly to be held in the selected sequencing institute, where all computer and human resources and pipelines are in place to perform high throughput annotation.

Sample Availability:

In-bred (>20 generations) colonies suitable for genome sequencing are available for each sequencing target identified (see Table 5). The *L. deliense* colony is maintained by Dr. Ratree Takhampunya at the AFRIMS in Bangkok, Thailand. Dr. Stephen Wikel maintains the Wikel strain used for the *I. scapularis* genome project and colonies of *I. pacificus, I. ricinus, I. persulcatus*, *A. americanum* and *D. variabilis* at the University of Texas Medical Branch (UTMB), Galveston, TX. An *O. turicata* colony and back-up colony of the Wikel strain is maintained by Dr. Dan Sonenshine at Old Dominion University, Norfolk, VA. All sequencing targets suggested in this proposal are diploid (Oliver, 1977), and all ixodid tick targets are expected to have a GC content (56% coding regions; 32% intergenic regions) and gene architecture (approx. 4.4 exons/gene) similar to that of *I. scapularis*.

4. Management of the project

The Tick and Mite Genomes Consortium, comprising members of the tick and mite research communities has been established to oversee this project:

Tick and Mite Genomes Consortium

Catherine Hill, Purdue University, West Lafayette, IN
Abdu Azad, University of Maryland, MD
Lorenza Beati, Georgia Southern University, Statesborough, GA
Nathalie Boulanger, Universite de Strasbourg, France
Alan Bowman, University of Aberdeen, Scotland, United Kingdom
Jose de la Fuente, University of Castilla La Mancha, Spain
Frank Collins, University of Notre Dame, Notre Dame, IN
Brian Evans, AFRIMS, Bangkok, Thailand
Felix Guerrero, USDA-ARS, Kerrville, TX
Marjorie Hoy, University of Florida, Gainesville, FL
Jason Meyer, Purdue University, West Lafayette, IN
Uli Munderloh, University of Minnesota, MN
Patricia Nuttall, Centre for Ecology and Hydrology, National Environmental Research Council, Swindon, Wiltshire, England

Joseph Piesman, Centers for Disease Control and Prevention, Fort Collins, CO
Jose Ribeiro, National Institutes of Health, Bethesda, MD
Mike Roe, North Carolina State University, Raleigh, NC
Dan Sonenshine, Old Dominion University, Norfolk, VA
Ratree Takhampunya, AFRIMS, Bangkok, Thailand
Stephen Wikel, University of Texas Medical Branch, Galveston, TX

Consortium members have broad expertise in the biology of tick and mite vectors and tick- and mite-borne disease research. Members of the consortium also have considerable experience in the sequencing and analysis of arthropod genomes. Drs. Hill and Wikel are the lead community scientists for the *I. scapularis* genome effort. Dr. Frank Collins was the lead community scientist of the malaria mosquito (*Anopheles gambiae*) genome project and is the PI of the NIH-funded VectorBase. Dr. Marjorie Hoy is the PI of the *M. occidentalis* genome project and Dr. F. Guerrero leads the USDA effort to sequence *R. microplus*.

5. Community input and support

Members of the tick and mite research communities represent an international consortium of scientists that will benefit from, and participate in, the analysis and annotation of the genomes proposed here. The entire community of scientists working on arthropod vectors and vector-borne diseases will also benefit from these projects. More than 100 scientists have collaborated on the analysis of the *I. scapularis* genome sequence and are interested to expand research to other tick and mite systems. The tick and mite research community represents a mature community that has conducted many decades of research on various aspects of tick and mite biology, physiology, genetics, population biology, ecology, pathogen transmission and control. Multiple scientific meetings dedicated to tick and mite research serve as evidence of this fact; notable examples include the *International Congress of Acarology* and the *International Conference on Ticks and Tick-borne Pathogens (TTP),* as well as dedicated sessions at the annual meetings of both the *American Society for Tropical Medicine and Hygiene* and *Entomological Society of America*. In addition, the tick community has met several times to mobilize for tick genome research (e.g., the 2003 meeting "*Working Group on Tick Genomics"* and the 2004 meeting "*Tick-Borne Diseases: Genomics and Proteomics Approaches"*).

6. Data release and relevant repositories for strains and sequence data

All sequence data generated by this project will be released to GenBank and other public databases in accordance with NHGRI and NIAID data release policies described at:

NHGRI: http://www.genome.gov/25521732
NIAID: http://www.niaid.nih.gov/labsandresources/resources/gsc/pages/data.aspx

Drs. Wikel and Sonenshine currently maintain the *I. scapularis* WIKEL colony sequenced by NIAID in 2008. Both have facilities to support tick culture and will maintain reference colonies of the strains selected for sequencing as part of this project. The AFRIMS laboratory in Bangkok, Thailand will maintain the reference *L. deliense* colony.

The NIH-NIAID funded VectorBase (www.vectorbase.org) is the logical repository for the genome sequence and EST data generated by this project. VectorBase is a web-based resource that houses, displays and manages sequence and related data from arthropod vector genome projects (Lawson et al., 2009). VectorBase provides initial annotation of new genome sequence and re-annotation of existing sequence, and it maintains and updates reference data sets for multiple vector species, including *I. scapularis*.

7. Conclusion

One of the broad objectives of vector research programs around the world is to develop an improved understanding of tick/mite vector biology and to identify new approaches to control disease vectors among this important, but largely neglected arthropod lineage. Two factors limit our ability to achieve this goal. First, the community lacks genome resources for multiple species of medically important disease vectors, including

trombiculid mite vectors, and ixodid and argasid ticks. Second, the global research community has need of a high quality *I. scapularis* assembly to advance research among the *Ixodes* species complex which vector of LD. This proposal outlines the immediate and long-term strategies we propose to overcome these limitations.

Acknowledgements: The consortium gratefully acknowledges the assistance provided by Dr. E. Caler, Assistant Professor, Eukaryotic Genomics, JCVI, with the development of a sequencing strategy for the species recommended in this proposal. We also thank the Pathogens and Vectors Working Group for review of this proposal and the provision of constructive criticism.

References:

Ahmed, J.S. and Mehlhorn, H. 1999. Review: the cellular basis of the immunity to and immunopathogenesis of tropical theileriosis. *Parasitology Research*. 85, 539-549.

Aljamali, M.N., Ramakrishnan, V.G., Weng, H., Tucker, J.S., Sauer, J.R. and Essenberg, R.C. 2009. Microarray analysis of gene expression changes in feeding female and male lone star ticks, *Amblyomma americanum* (L). *Arch. Insect Biochem. Physiol*. 71, 236-253.

Anderson, J.F., Main, A.J., Andreadis, T.G., Wikel, S.K. and Vossbrinck, C.R. 2003. Transstadial transfer of West Nile virus by three species of ixodid ticks (Acari: Ixodidae). *Journal of Medical Entomology*. 40, 528-533.

Anguita, J., Ramamoorthi, N., Hovius, J. W.R., Das, S., Thomas, V., Persinski, R., Conze, D., Askenase, P.W., Kantor, F.S. and Fikrig, E. 2002. Salp15, an *Ixodes scapularis* Salivary Protein, Inhibits CD4-T Cell Activation. *Immunity.* 16, 849–859.

Azad, A. and Radulovic, S. 2003. Pathogenic rickettsiae as bioterrorism agents. *Annals of the New York Academy of Sciences*. 990, 734-738.

Brusca, R.C. and Brusca, G.J. 1990. Invertebrates. Sinauer Assoc., Sunderland, MA, U.S.A.

Bell-Sakyi, L., Zweygarth, E., Blouin, E., Gould, E.A. and Jongejan, F. 2008. Tick cell lines: tools for tick and tick-borne disease research. *Trends Parasitol*. 9, 450-457.

Besansky, N. J. 2008. Genome analysis of vectorial capacity in major *Anopheles* vectors of malaria parasites. White paper submitted to the Eurkaryotic Pathogens and Disease Vectors Target Selection Group on behalf of the *Anopheles* Genomes Cluster Committee. August 2008.

Black, W.C. IV. And Peisman, J. 1994. Phylogeny of hard- and soft-tick taxa (Acari:Ixodida) based on mitochondrial 16S rDNA sequences. *Proc. Natl., Acad. Sci. U.S.A*. 91, 10034-10038.

Childs, J.E. and Paddock, C.D. 2003. The ascendancy of *Amblyomma americanum* as a vector of pathogens affecting humans in the United States. *Annual Review of Entomology*. 48, 307-337.

Cornman, S.R., Schatz, M.C., Johnston, S.J., Chen, Y.P., Pettis, J., Hunt, G., Bourgeois, L., Elsik, C., Anderson, D., Grozinger, C.M. and Evans, J.D. 2010. Genomic survey of the ectoparasitic mite *Varroa destructor*, a major pest of the honey bee *Apis mellifera. BMC Genomics*. 11, 602.

Dearden, P.K., Donly, C., Grbic, M. 2002. Expression of pair-rule gene homologues in a chelicerate: early patterning of the two-spotted spider mite *Tetranychus urticae*. *Development*. 129, 5461–5472.

de la Fuente, J. and Kocan, K. M. 2003. Advances in the identification and characterization of protective antigens for development of recombinant vaccines against tick infestations. *Expert Review of Vaccines.* 2, 583-593.

Edlow, J.A. (Ed.). 2002. Tick-borne diseases. *Medical Clinics of North America*. 86, 205-453.

Estrada-Pena, A., Naranjo, V., Acevedo-Whitehouse, K., Mangold, A.J., Kocan, K.M. and de la Fuente, J. 2009. Phylogeographic analysis reveals association of tick-borne pathogen, *Anaplasma marginale*, MSP1a sequences with ecological traits affecting tick vector performance. BMC Biol. 7, 57-69.

Fikrig, E., Barthold, S.W., Kantor, F.S. and Flavell, R.A. 1990. Protection of mice against the Lyme disease agent with recombinant OspA. *Science*. 250, 563-566.

Fivaz, B., Petney, T. and Horak, I. 1992. "Tick Vector Biology: Medical and Veterinary Aspects". *Springer-Verlag, pp. 191.*

Geraci, N.S., Johnston, J.S., Robinson, J.P., Wikel, S.K. and Hille, C.A. 2007. Variation in genome size of argasid and ixodid ticks. *Insect Biochem. Mol.* Biol. 37, 399-408.

Gothe, R. 1999. Zeckentoxikosen – Tick Toxicoses. *Hieronymus Buchreproduktions GMBH, München.*

Gratz, N.G. 1999. Emerging and resurging vector-borne diseases. *Annual Review of Entomology.* 44, 51-75.

Guerrero, F.D., Moolhuijzen, P., Peterson, D.G., Bidwell, S., Caler, E., Bellgard, M. Nene, V.M. and Djikeng, A. 2010. Reassociation kinetics-based approach for partial genome sequencing of the cattle tick, *Rhipicephalus* (*Boophilus*) *microplus*. *BMC Genomics*. 11, 374-382.

Guerrero, F.D., Vishvanath, N., Neneb, M., George, J.E., Barker, S.C. and Willadsen. 2006. Sequencing a New Target Genome: The *Boophilus microplus* (Acari: Ixodidae) Genome Project. *J. Med. Entomol.* 43, 9-16.

Grbic, M., Khila, A., Lee, K-Z., Bjelica, A., Grbic, V., Whistlecraft, J., Verdon, L., Navajas, M. and

Nagy, L. 2007. Mity model: *Tetranychus urticae*, a candidate for chelicerate model organism. *BioEssays*. 29, 489-496.
Hill, C.A. and Wikel, S.K. 2005. The *Ixodes scapularis* genome project: an opportunity for advancing tick research. *Trends Parasitol*. 21, 151-153.
Houck, M.A., Qin, H. and Roberts, H.R. 2001. Hantavirus transmission: potential role of ectoparasites. *Vector-Borne Zoonotic Dis*. 1, 75-79.
Hoy, M.A. 2009. The predatory mite *Metaseiulus occidentalis*: mitey small and mitey large genomes. *BioEssays*. 31, 581-590.
Jeyaprakash, A. and Hoy, M.A. 2008. First divergence time estimates of spiders, scorpions, mites and ticks (subphylum: Chelicerata) inferred from mitochondrial phylogeny. *Exp. Appl. Acarol*. 47, 1-18.
Jongejan, F. and Uilenberg, G. 2004. The global importance of ticks. *Parasitology.* 129, S3-S14.
Kabeya, H., *et al.,* 2010. Detection of *Bartonella tamiae* DNA in ectoparasites from rodents in Thailand and their sequence similarity with bacterial cultures from Thai patients. *Vector-borne Zoonotic Dis*. 10, 429-434.
Klompen, J.S.H., Black, W.C. IV., Keirans, J.E. and Oliver, J.H. Jr. 1996. Evolution of ticks. *Ann. Rev. Entomol.* 41, 141-161.
Klompen, J.S.H., Black, W.C. IV., Keirans, J.E. and Norris, D.E. 2000. Systematics and biogeography of hard ticks, a total evidence approach. *Cladistics*. 16, 79-102.
Kocan, K.M., de La Fuente, J., Blouin, E.F. and Garcia-Garcia, J.C. 2002. Adaptations of the tick-borne pathogen, *Anaplasma marginale*, for survival in cattle and ticks. *Experimental and Applied Acarology*. 28, 9-25.
Leboulle, G., Crippa, M., Decrem, Y., Mjri, N., Brossard, M., Bollen, A. and Godfroid, E. 2002. Characterization of a novel salivary immunosuppressive protein from *Ixodes ricinus* ticks. *Journal of Biological Chemistry*. 277, 1083-1089.
Lerdthusnee, K., et al., 2002. *J. Med. Ent*. 39, 521-525.
McQuiston, J.H., McCall, C.L. and Nicholson, W.L. 2003. Ehrlichiosis and related infections. *Journal of the American Veterinary Medical Association*. 223, 1750-1756.
Meyer, J.M., Kurtti, T.J., Van Zee, J.P. and Hill, C.A. 2010. Genome organization of major tandem repeats in the hard tick, *Ixodes scapularis. Chrom. Res*. 18, 357-370.
MMWR, Morbidity and Mortality Weekly Report. CDC, November 5, 2010. 59, 1416-1429. www.cdc.gov/mmwr
Mullen, G. and Durden, L. (Eds.). 2002. Medical and Veterinary Entomology. *Academic Press, San Diego, CA*.
Nuttall, P.A. 1998. Displaced tick-parasite interactions at the host interface. *Parasitology.* 116, S65-S72.
Ochanda, H., Young, A.S., Medley, G.F. and Perry, B.D. 1998. Vector competence of seven rhipicephalid tick stocks in transmitting two *Theileria parva* parasite stocks from Kenya and Zimbabwe. *Parasitology*. 116, 539-545.
Oliver, J. 1977. Cytogenetics of mites and ticks. *Ann. Rev. Entomol.* 22, 407-429.
Paddock, C.D. and Childs, J.E. 2003. *Ehrlichia chaffeensis*: a prototypical emerging pathogen. *Clinical Microbiology Reviews*. 16, 37-64.
Pagel Van Zee, J., Geraci, N.S., Guerrero, F.D., Wikel, S.K., Stuart, J.J., Nene, V.M. and Hill, C.A. 2007. Tick genomics: The *Ixodes* genome project and beyond. *Int. J. Parasitol.* 37, 1297-1305.
Palmer. M.J., Bantle, J.A., Guo, X. and Fargo, W.S. 1994. Genome size and organization in the ixodid tick *Amblyomma americanum. Insect Mol. Biol*. 3, 57-62.
Ribeiro, J.M., Alarcon-Chaidez, F., Francischetti, I.M., Mans, B.J., Mather, T.N. et al., 2006. An annotated catalog of salivary gland transcripts from *Ixodes scapularis* ticks. *Insect Biochem. Mol. Biol*. 36, 111-129.
Ribeiro, J.M. and Francischetti, I.M. 2003. Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. *Annual Review of Entomology*. 48, 73-88
Roberts, L.S. and Janovy, J. 1996. "Foundations of Parasitology". *WCB Publishers, Dubuque, IA.*
Saldivar, L., Guerrero, F.D., Miller, R.J., Bendele, K.J., Gondro, C. and Brayton, K.A. 2008. Microarray analysis of acaricide-inducible gene expression in the southern cattle tick, *Rhipicephalus* (*Boophilus*) *microplus*. *Insect Mol. Biol*. 17, 597-606.
Schoeler, G.B. and Wikel, S.K. 2001. Modulation of host immunity by haematophagous arthropods. *Annals of Tropical Medicine and Parasitology*. 95, 755-771.

Sigal, L.H., Zahradnik, J.M., Lavi, P., Patella, S.J., Bryant, G., Haselby, R., Hilton, E., Kunkel, M., Aderklein, D., Doherty, T., Evans, J. and Malawista, S.E. 1998. A vaccine consisting of recombinant *Borrelia Burgdorferi* outer surface protein A to prevent Lyme disease. *New England Journal of Medicine.* 339, 216-222.

Sonenshine, D. 1991. "Biology of Ticks". *Volume 1. Oxford University Press, New York*.

Sonenshine, D. 1993. "Biology of Ticks"*. Volume 2. Oxford University Press, New York.*

Sonenshine, D. and Mather, T.N. 1994. "Ecological Dynamics of Tick-Borne Zoonoses"*. Oxford University Press, New York. pp. 447.*

Sonenshine, D.E. and Hynes, W.L. 2008. Molecular characterization and related aspects of the innate immune response in ticks. *Front. Biosci.* 13:7046-7063.

Teague, B., Waterman, M.S., Goldstein, S., Potamousis, K., Zhou, S., Reslewic, S., Sarkar, D., Valouev, A., Churas, C., Kidd, J., Kohn, S., Runnhein, R., Lamers, C., Forrest, D., Newton, M.A., Eichler, E.E., Kent-First, M., Surti, U., Livny, M., and Schwartz, D.C. High-resolution human genome structure by single molecule analysis. 2010. *Proc. Natl. Acad. Sci. USA.* 107, 10848-53.

Trimnell, A.R., Hails, R.S. and Nuttall, P.A. 2002. Dual action ectoparasite vaccine targeting exposed and concealed antigens. *Vaccine*. 20, 3560-3568.

Ullmann, A.J., Lima, C.M., Guerrero, F.D., Piesman, J. and Black, W.C., IV. 2005. Genome size and organization in the black-legged tick, *Ixodes scapularis* and the southern cattle tick, *Boophilus microplus. Insect Mol. Biol*. 14, 217-222.

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

Valenzuela, J.G. 2002. Exploring the messages of the salivary glands of *Ixodes ricinus*. *American Journal of Tropical Medicine and Hygiene.* 66, 223-234.

Wagner, G.G., Holman, P. and Waghela, S. 2002. Babesiosis and heartwater: threats without boundaries. *Veterinary Clinics of North America, Food Animal Practice*. 18, 417-430.

Walker, D. 1998. Tick-transmitted infectious diseases in the United States. *Annual Review of Public Health*. 19, 237-269.

WHO. 2006. Lyme borreliosis in Europe: Influences of climate and climate change, epidemiology, ecology and adaptation measures*. World Health Organization, Regional Office for Europe, Copenhagen, Denmark.*

Wikel, S.K., Alarcon-Chaidez and Mueller-Doblies, U. 2003. Immunological control of vectors. In: "Biology of Disease Vectors". J. Hemingway, (Ed.). *In press. Academic Press, San Diego.*

Willadsen, P., Bird, P., Cobon, G.S. and Hungerford, J. 1995. Commercialization of a recombinant vaccine against *Boophilus microplus*. *Parasitology*. 110, S43-S50.

Young, A.S., Dolan, T.T., Mwakima, F.N., Ochanda, H., Mwaura, S.N., Njihia, G.M., Muthoni, M.W. and Dolan, R.B. 1995. Estimation of heritability of susceptibility to infection with *Theileria parva* in the tick *Rhipicephalus appendiculatus*. *Parasitology*. 111, 31-38.

Zhou, S., Bechner, M.C., Place, M., Churas, C.P., Pape, L., Leong, S.A., Runnheim, R., Forrest, D.K., Goldstein, S., Livny, M., and Schwartz, D.C. 2007. Validation of rice genome sequence by Optical Mapping. *BMC Genomics* 8: 278.

Zhou, S., Wei, F., Nguyen, J., Bechner, M., Potamousis, K., Goldstein, S., Pape, L., Mehan, M., Churas, C., Pasternak, S., Forrest, D.K., Wise, R., Ware, D., Wing, R., Waterman, M.S., Livny, M., and Schwartz, D.C. 2009. A single molecule scaffold for the maize genome. *PLoS Genetics* 5: e1000711.

USDA

United States Department of Agriculture

Research, Education, and Economics
Agricultural Research Service

Prof. Catherine Hill
Tick and Mite Genomes Consortium
Dept. of Entomology
Purdue University
West Lafayette, IN

Dec. 23, 2010

Dear Professor Hill and Consortium Members,

I wanted to express my enthusiasm for your efforts to expand sequencing efforts for mites and ticks of human importance. These proposed resources will provide a helpful meeting point not just for the targeted species but for all of those studying Acari of medical, veterinary, and agricultural importance. Your arguments for a mix of genomic and transcriptomic sequencing are sound and doable, and you propose a great model for how such resources can be strengthened by including taxonomically diverse organisms.

We are engaged in genomic and transcript sequencing projects with the honey bee parasitic mite, *Varroa destructor*, and are addressing many similar questions to those posed by you and your Consortium members. As one example, we are quite interested in contrasting our discoveries with those found through genomic and wet-lab work in the vector tick *Ixodes scapularis*. We have enjoyed the excellent public access to resources, tools and analysis generated by the *Ixodes* genome project and will benefit much more from the expanded comparative resources you are proposing. *Varroa* mites are significant vectors of RNA viruses and other bee disease agents, and the use of common databases to understand host-parasite-vector interactions should help us move forward with *Varroa* research and hopefully contribute complimentary insights from our system.

In short, expanded genomic and transcriptomic efforts for this important group seem sure to generate insights and tools for medical and veterinary solutions. Secondarily, your efforts will influence research on a wide range of ticks and mites of agricultural importance, including the bee mites we are studying. I hope that the NIH-NHGRI sees fit to enable this project and will, of course, be happy to share any insights we generate from our focal species in order help with computational and experimental efforts from your project .

Yours sincerely,

Dr. Jay D. Evans
Research Scientist
USDA-ARS Bee Research Lab
BARC-East Bldg. 476
Beltsville, MD 20705
301-504-5143 (ph)
Jay.evans@ars.usda.gov http://www.coloss.org/partners/evans

ars