[Frontiers in Bioscience 14, 2666-2673, January 1, 2009]

Tick genomics – coming of age

Vishvanath Nene

Institute for Genome Sciences and Department of Microbiology and Immunology University of Maryland School of Medicine Baltimore, MD 21201, USA

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Classification of ticks
4. Gross features of tick genomes
4.1. The nuclear component
4.2. The mitochondrial component
4.3. Exploring gene content via EST databases
4.4. IGP – Ixodes scapularis genome project
5. Passenger genomes
6. Perspectives
7. Acknowledgements
8. References

1. ABSTRACT

Many challenges face tick genomics. Ticks have large genomes and their estimated sizes vary from 1.04~7.1 $x10^9$ bp, about one third to over two times the size of the human genome. Karyotype studies have revealed a range in chromosome number and the sex determining system seems to be primarily driven by a XY or a XO format. Re-association hybridization kinetics indicates that the bulk of the tick genome consists of repetitive sequences and only 30~35% of the genome consists of low copy number sequences. The former remain poorly characterized and most of what we know about the latter has been driven by gene discovery projects via generation of expressed sequence tags (ESTs). However, large scale EST data (>10,000 ESTs) are available for only three tick species. The only whole genome sequencing project for ticks is that on *Ixodes scapularis,* the primary tick vector of medical importance in the United States. Nevertheless, important advances are being made in developing genomics tools and these will stimulate research in tick and vector biology.

2. INTRODUCTION

The phylum Arthropoda consists of a highly diverse collection of taxa and it is estimated that there are more arthropod species than all the other animal phyla combined (1). Within arthropods, genome sequencing projects have concentrated on insects as some species are well developed model organisms or they represent organisms classified as beneficial insects, agricultural pests or global vectors of disease that affect humans, animals or plants (2). The large number of insect genome sequencing projects is providing valuable data for comparative genomics, the evolution of insect genomes and genes as well as clues in functional genomics. However, little is known about the genomes of arthropods outside the class Insecta. This status is changing for the Chelicerates. Due to the importance of ticks as vectors of disease (3), a project to generate a draft genome sequence of *Ixodes scapularis* is well under way (4, 5).

The blacklegged or deer tick as *I. scapularis* is more commonly known is the principal vector of human

tick-borne disease in the Unites States and transmits pathogens which cause Lyme disease (6), babesiosis (7) and anaplasmosis (8). Ticks are second to mosquitoes in order of global importance as vectors of human disease but of primary importance in veterinary medicine (3). Unlike mosquitoes, both sexes of all life-cycle stages of ticks are exclusive blood-feeders (9, 10). Male and female ticks co-operate on host (11) in the acquisition of a blood meal, a dynamic process which presents challenges at many different levels from host-seeking to overcoming innate and acquired host defense mechanisms during the many hours to days required to feed (12-14). The tick salivary glands play a central role in maintaining water and ion balance (15) and in contrast to mosquitoes digestion of the blood meal occurs within cells that line the gut wall (16, 17). This barrier as well as hemocyte and fat body activity represent other important tick organs/tissues that mount a defense to infection. Passive and active modes of transmission and trans-ovarial and trans-stadial routes of transmission add to the complexity of vector biology. It is expected that paradigm shifts that occur on acquisition of the genome sequence of an organism will stimulate research in tick and vector biology and help efforts directed towards controlling disease.

3. CLASSIFICATION OF TICKS

There are currently 889 species of ticks listed in Tickbase, a taxonomic catalog of ticks maintained by the Integrated Consortium on Ticks and Tick-borne Diseases (http://www.icttd.nl). Ticks have been classified into one of three subfamilies, Ixodidae (hard ticks), Argasidae (soft ticks) and Nuttalliellidae (18). Since live ticks in the latter family are not available many characteristics of *Nuttalliella namaqua*, the only member of the family, remain enigmatic. Based mainly on morphological detail hard ticks are further divided into two lineages, Prostriata and Metastriata. The former group consists of a single genus, *Ixodes,* which appears to contain two lineages, the Australasian *Ixodes* and other *Ixodes (*19). However, the evolutionary history and age of ticks remains controversial. Re-examination of the tick fossil records has not allowed definite conclusions to be made regarding tick evolution (20) and the phylogenetic sister group to ticks remains to be firmly established. Although substantial progress has been made in tick systematics the evolutionary relationship between many tick species is uncertain. Changes in tick phylogeny due to the use of molecular markers can create confusion within non-specialist audiences but are necessary in order to reflect the evolutionary history of different species. Since it is mentioned later in the text the genus *Boophilus* has been re-classified as a subgenus of the genus *Rhipicephalus* (21).

4. GROSS FEATURES OF TICK GENOMES

4.1.The nuclear component

Cytogenetic study of ticks indicates variety in the number of chromosome numbers present in different species and sex appears to be determined by different combinations of sex chromosomes (22). For example, *I. scapularis* contains 26 autosomes and X and Y sex chromosomes (2n=28); females are XX while males are XY (23). On the other hand *Rhipicephalus (Boophilus) microplus* contains 20 autosomes and lacks a Y sex chromosome (2n=22 or 21); females are XX while males are XO (22). In general, the X chromosomes are much bigger in size than other chromosomes (22) and recent data based on fluorescence intensity of fluorochrome stained nuclei as measured by flow cytometry indicate a measurable difference in DNA content between male and female tick cells (24). The XY and XO sex determination system seems to predominate in ticks with the former prevalent in soft ticks and the latter in hard ticks. More complex sex determination systems have been identified in some tick species and parthenogenesis has also been described (9, 22).

Little is known about the structural organization of tick chromosomes. The kinetics of re-association of genomic DNA (Cot analysis) provided the first estimates of the size of tick genomes. Such studies revealed a haploid genome size of ~1.04 Giga base pairs (Gbp - 10^9) for *Amblyomma americanum* (25), and 2.1 Gbp for *I. scapularis* and 7.1 Gbp for *R. (Boophilus) microplus* (26). As in other large genomes, these tick genomes appear to mostly consist of highly and moderately repeated sequences with 30~35% of the genome consisting of low copy number DNA. Flow cytometry data supports the genome size estimate of *I. scapularis* (2.26 Gbp) (24). However, this study estimated a ~3-fold larger genome size for *Am. americanum (*3.1 Gbp) and genome sizes of ~2.8 Gbp for *Am. cajennense* and ~2.9 Gbp for *Am. maculatum*. The reason for the discrepancy in genome size of *Am. americanum* is not clear, but it is important to resolve this issue as there is great interest in developing genomic resources for it (5). This tick species is associated with a number of emerging human diseases in the United States (27).

Flow cytometry data have also provided the first genome size estimates of soft ticks, ~1.47 Gbp for *Argas brevipes* and ~1.09 Gbp for *Ornithodoros turicata,* and suggest that Argasid ticks have smaller genomes than Ixodid ticks (24). The chromosome numbers in these soft ticks remains to be determined but in general most *Argas* species contain 26 chromosomes while in *Onrithodoros* the number ranges from 12 to 32 (24). Thus the C-value enigma (28), the huge variation in nuclear genome size between different species, is also discernible in ticks to the extent of intra-species variation in genome size (24). The repertoire of repeat families within tick genomes remains largely uncharacterized. The R2 element, a non-LTR retrotransposon inserted in large subunit rRNA genes has been found in both hard and soft ticks but is present in low copy numbers (29) and partial sequences of a reverse transcriptase-like protein found in ixodid ticks have been deposited in the National Center for Biotechnology Information (NCBI). Some of the more abundant repeat families present in ticks are beginning to be described (5).

Given the size and complexity of tick genomes it is not surprising that there are no physical maps of tick genomes. A great deal of effort has gone into production of

Table 1. The growth of tick nucleotide accessions in NCBI

Tick genus	11th July 2002	15th July 2004	5th July 2006	17th July 2007
Argasidae (soft ticks)				
Anticola	2	2	2	2
Argas	15	23	28	3,135
Carios	2	14	14	30
Otobios	4	4	4	6
Ornithodoros	30	1,610	1,751	3,480
Ixodidae (hard ticks)				
Amblyomma	6,037	6,175	10,666	12,491
R. (Boophilus)	761	21,205	43,392	43,406
Bothriocroton	1	1	1	33
Dermacentor	104	274	411	1,692
Ixodes	534	933	9,878	11,434*
Haemaphysalis	44	77	121	172
Hyalomma	24	122	133	137
Rhipicephalus	171	19,280	19,428	19,433
Anocentor	3	3	4	
Aponoma	16	30	30	11
Rhipicentor	5	6	6	6
Nosoma	2	2	2	2

The Broad Institute has deposited 201,600 ESTs at NCBI under the *Ixodes scapularis* genome project, sequence data not yet captured by the Taxonomy browser.

a preliminary genetic linkage map for *I. scapularis* based on different types of molecular markers (30). In agreement with karyotype data, the linkage of genotypes in an F_1 intercross family present in Mendelian ratios identified 14 linkage groups. Although the current map is of low resolution the re-assortment rates of markers provided evidence for a high rate of recombination within the progeny. This observation suggests that a higher density of markers could facilitate positional cloning of genes controlling quantitative trait loci (30) but indicates that tick population structure is likely to be highly complex.

4.2. The mitochondrial component

Because of their small size and the use of mitochondrial genes in inferring phylogenetic relationships a considerable amount of DNA sequence information has been derived from tick mitochondrial genomes, including complete genome sequences from both soft and hard ticks (31-33). The mitochondrial genomes of ticks are 14~15 kbp in length and they are circular in structure. Soft and prostriate ticks contain a typical complement of 37 metazoan mitochondrial genes and an ancestral arrangement of arthropod mitochondrial gene order (33). In contrast, metastriate mitochondrial gene order is rearranged for some genes and the genome contains two control regions (31, 32). Intriguingly, the latter is also a distinguishing feature of mitochondrial genomes of the Australiasian lineage of *Ixodes* ticks (18). As in other eukaryotes tick mitochondrial genome sequences have a low G+C content and protein coding genes exhibit a bias in using A+T rich codons (31, 32).

4.3. Exploring gene content via EST databases

Protein coding gene density in genome sequences is variable and generally low in higher eukaryotic organisms. Thus, genome sequence surveys (GSS) usually inform on some classes of DNA sequence repeats and may reveal partial open reading frames (ORFs) which are useful if reference genome sequences are available, but in general they are not very informative in revealing novel gene content. In contrast, an equivalent level of cDNA sequencing and creation of expressed sequence tag (EST) databases tends to be highly informative as sequencing is directed towards protein coding regions of the genome. Given the large size of tick genomes (>1.0 x 10^9 bp) it is not surprising that acquisition of EST data has out-stripped that of GSS data from different tick species.

Table 1 shows the rapid rise in tick DNA sequence deposition in NCBI since July 2002 as revealed by the NCBI Taxonomy browser (http://www.ncbi.nlm.nih.gov/) and in almost all cases where more than one hundred sequences are available is due to EST driven gene discovery projects. The EST data have been used to generate annotated databases which describe transcribed protein encoding tick genes. However, from a suborder perspective very few species of ticks have been sequence sampled: only six Ixodid species (*Am. americanum*, *Am. variegatum*, *Am. cajennense*, *R. (Boophilus) microplus*, *I. scapularis* and *R. appendiculatus*) and three Argasid ticks (*Argas monolakensis* and *Ornithodoros parkeri* and *O. porcinus*) have more than 1,000 sequence entries each. Due to the *I. scapularis* genome sequencing project (4), this species has been the one most extensively sampled and ~211,000 ESTs are available from it.

ESTs represent single pass DNA sequences. Thus, trimming of low quality sequences and use of consensus sequences derived from overlapping high confidence EST sequences represents a critical exercise when building a gene index. Ideally individual trimmed ESTs should be at least 100 bases in length and contain less than 3% N's to be of further use (34). "Gene indices" created from *Am. americanum*, *R. microplus*, *R. appendiculatus* and *I. scapularis* ESTs can be accessed via the internet (http://compbio.dfci.harvard.edu/tgi/tgipage.html). This site also harbors gene indices from a variety of different organisms. EST data have been clustered into unique datasets consisting of tentative consensus (TC) and singleton sequences and then assigned putative gene

function, gene ontology terms and those with enzyme function assigned to metabolic pathways. The relational database structure allows the user to query the gene index using key words and one can drill down, e.g., to view a map of the overlap of ESTs that constitute a TC, a six-frame translation, annotation statistics, a single nucleotide polymorphism report and unique oligonucloetide prediction. Users must be aware that the gene indices are not manually curated and the data is derived via a series of auto-annotation pipelines. This can lead to lead to some problems, e.g., in assigning gene names. Nevertheless, the gene indices offer an excellent repository of large amounts of otherwise unmanageable data and provides a convenient starting points for further studies.

Unfortunately, the use of different methods and stringencies for processing of raw sequence data, clustering and assembly can result in significant differences in EST database structure as well as in the auto-assignment of putative function to sequences. Genuine DNA sequence polymorphisms, due either to the presence of paralogous genes or allelic variants, as well as alternative splice forms could lead to the generation of mis-assembled sequences. Ideally, the outputs of different assembly parameters should be assessed in an effort to minimize mis-assembly of closely related DNA sequences. In addition, this exercise should be compared with assemblies derived using the Program to Assemble Spliced Alignments (PASA) software, a more stringent assembly tool as it maps ESTs to genome sequence data prior to assembly (35). Currently, the latter would only be feasible for *I. scapularis*. Resolving the technical issues due to differences in computational pipelines becomes important prior to carrying out global comparative genome studies based on EST data. A semi-manual curation of predicted tick proteomes based on clustered EST data has been recently reported by J. Ribeiro (http://exon.niaid.nih.gov/transcriptome/Page_tick.htm).

ESTs have been derived from whole life-cycle stages of ticks (36-38). Such projects are beginning to provide important base line data on tick genomes and represent essential knowledge building exercises which provide tools that may be used to study tick biology. Most EST projects have sampled the genes expressed in tick salivary glands (39-45) as this is the site of replication/maturation of many pathogens and tick saliva contains biologically active molecules which can modulate host inflammatory, hemostatic and immune responses (14, 46, 47). These glands play a prominent role in maintaining water and ion balance (15) and they also express proteins which are components of tick cement (48, 49). Besides supporting the identification of biologically active proteins in tick saliva (50) examination of EST data has revealed important temporal differences in genes expressed in *I. scapularis* salivary glands (51), a finding supported by transcript profiling using a mini-array consisting of ~150 genes (52). Such discoveries highlight the importance of genomic tools and data being derived from the ongoing *I. scapularis* genome sequencing project will facilitate the refinement of genomics platforms for research in tick and vector biology. Other tick organs and tissues have also been sampled and such studies begin to reveal tissue specific genes expressed in the mid-gut, ovaries and hemocytes (53). Comparative analyses of the temporal repertoire of genes expressed in ticks should yield interesting data on the common and species specific strategies that have evolved in different tick-host and tick-pathogen interactions. In addition, comparison of genomics data from ticks with other arthropods should yield interesting data on the evolution of vector biology and the adaptations to blood feeding.

4.4. IGP – *Ixodes scapularis* genome project

Due to the medical importance of *I. scapularis* the National Institute for Allergy and Infectious Diseases has funded the production of a draft genome sequence of *I. scapularis* (4) at its Microbial Sequencing Centers (MSCs). The scope of the project includes whole genome shotgun (WGS) sequencing to 6-fold sequence coverage and generation of ~200,000 ESTs from a normalized, directional cDNA library.

The WGS method of genome sequencing and associated computational tools for assembly of shotgun sequence data has developed into a robust method which can be applied to genomes of different size and sequence complexity. However, this approach assumes "randomness" in many aspects of the protocol and assembly algorithms assume homogeneity in the starting DNA sample and produce a consensus genome sequence. The *I. scapularis* colony (referred to as the Wikel strain) was chosen for sequencing as it represents a "closed" colony where no new genotypes have been introduced into it since its founding >12 generations ago (5). Approximately 19.5 million random WGS reads from a variety of different sized genomic DNA libraries (small insert, medium insert, fosmids, BACs and BAC clones) and 201,600 ESTs have been generated and released to NCBI by the two MSCs (J. Craig Venter Institute and The Broad Institute) (5). The data can also be accessed via VectorBase (http://www.vectorbase.org), a resource center which is devoted to vector bioinformatics.

The project is in the assembly phase of the WGS reads and will be followed by auto-annotation of the assembled genome sequence data and is due to be released to the scientific community in mid-2008. A description of genome wide data derived from this project is the primary responsibility of IGP and will not be discussed here. However, given the level of random sequencing carried out and based on experience derived from other genome sequencing projects (54, 55) it is quite likely that the assembly will consist of a very large number of scaffolds. Ultimately the assembled sequence data will have to be assigned to chromosomes. The genetic linkage map of *I. scapularis* should provide useful markers for chromosomal assignments (30) and novel markers from the IGP could be used to increase the resolution of the current genetic map. It is likely that other physical mapping methods will have to be used for chromosomal assignments of sequence data and fluorescent *in situ* hybridization experiments are being explored with this in mind (5).

A striking feature of analyses of an *I. scapularis* EST dataset derived independently of the IGP has revealed that there are many sequence polymorphisms between cDNAs that are predicted to encode highly similar protein sequences (51). It is likely that some of the cDNAs are encoded by multi-gene families and it has been suggested that genome duplication (polyploidy) may have occurred in ticks providing a source of gene duplications (51), although some of the polymorphisms may be due to allelic variation. It remains to be seen whether the IGP will reveal the evolutionary history of the *I. scapularis* genome and its genes. In the interim clustering of ESTs derived from single ticks could begin to provide more definitive data on the presence of multi-gene families.

5. PASSENGER GENOMES

Symbiotic relationships are prevalent in arthropods (56) and the WGS approach has an added benefit, namely the discovery of microbial organisms and viruses associated with the organism being sequenced. This realization led to the reconstruction of high quality genome sequences of *Wolbachia* associated with *Drosophila ananassae*, *D. simulans*, and *D. mojavensis* (57) from WGS projects on *Drosophila*. Many arthropods which occupy highly restricted niches have evolved microbial relationships to compensate for deficiencies in their diet (56). It has been suggested that ticks could benefit from as association with bacteria predicted to contain intact biosynthetic pathways for the production of essential cofactors which are limiting in blood (e.g., biotin) (58) and ticks appear to be persistently infected with bacteria (59).

Ricketsiosis is not usually associated with *I. scapularis* so it is interesting to note that this tick harbors a *Rickettsia* (60), one which still remains to be characterized in detail. Bacterial sequences have been identified within the WGS sequence data generated by the IGP consortium and are most likely derived from this rickettsial symbiont. Thus a byproduct of the IGP project will include an analysis of the genomic properties of this organism. An intriguing symbiont of *I. ricinus* which invades mitochondria has been recently described (61). It will be of interest to screen the *I. scapularis* WGS reads and assembled sequence data for similar sequences as well as other microbial and viral sequences. Ticks may be co-infected with multiple microbes (62) raising interesting questions regarding the symbiotic relationships between co-habiting microbes and the tick host, the evolution of pathogenic organisms (56, 63) and horizontal gene transfer not only between microbes but also between microbes and their arthropod host (64).

6. PERSPECTIVES

Rapid progress has been made in building tick EST database, but there is clearly a long way to go in developing whole genome sequence data. Additional WGS data will have to be derived for *I. scapularis* and a deeper sampling of *I. scapularis* ESTs will undoubtedly aid the genome project. The growing importance of non-coding RNA in biology (65) indicates that analyses should be expanded to include such transcripts. The large genome sizes of ticks and sequence heterogeneity create technical problems which need to be overcome as the commitment of financial and laboratory resources is currently huge. Such considerations become critical as genome sequencing projects for other tick species are considered (5, 66). Tick cell lines (67) remain untapped as potential sources of large amounts of genomic DNA for WGS projects. Such lines would have to be established from single individuals and the benefits of early acquisition of genome sequence data may outweigh the risk of artifacts due to chromosomal aberrations. The latter could be reduced by monitoring karyotypes and chromosome banding patterns (23). The development of newer and cheaper DNA sequencing platforms, a rapidly evolving field, holds much promise but it is not yet clear how they will perform in *de novo* large genome sequencing projects. The ability to computationally reconstruct individual haplotype sequences rather than a consensus sequence raises new opportunities in exploring the true genetics of diploid genomes (68-70).

7. ACKNOWLEDGEMENTS

I would like to thank my “tick” colleagues Richard Bishop, John Quackenbush, Felix Guerrero, Cate Hill and Don Knowles. The *Ixodes scapularis* genome project is funded via contracts by the National Institute for Allergy and Infectious Diseases at its Microbial Sequencing Centers (J. Craig Venter Institute and The Broad Institute) and VectorBase, a bioinformatics resource center (HHSN266200309D266030071, HHSN266200400001C, and HHSN266200400039C).

8. REFERENCES

1. E. O. Wilson: The diversity of life. *Penguin* London. (1992)

2. J. D. Evans, & D. Gundersen-Rindal: Beenomes to *Bombyx*: future directions in applied insect genomics. *Genome Biology* 4, 107 (2003)

3. F. Jongejan & G. Uilenberg: The global importance of ticks. *Parasitology* 129 Suppl, S3-14 (2004)

4. C. A. Hill & S. K. Wikel: The *Ixodes scapularis* Genome Project: an opportunity for advancing tick research. *Trends Parasitol* 21, 151-153 (2005)

5. P. vanZee, N. S. Geraci, F. D. Guerrero, S. K. Wikel, J. J. Stuart, V. M. Nene & C. A. Hill: Tick genomics: The *Ixodes* genome project and beyond. *Int. J. Paratiol* 37, 1297-1305 (2007)

6. J. Piesman & L. Gern: Lyme borreliosis in Europe and North America. *Parasitology* 129 Suppl, S191-220 (2004)

7. J. F. Anderson, E. D. Mintz, J. J. Gadbaw & L. A. Magnarelli: *Babesia microti*, human babesiosis, and

Borrelia burgdorferi in Connecticut. *J Clin Microbio.* 29, 2779-2783 (1991)

8. S.R. Telford 3rd, J.E. Dawson, P. Katavolos, C.K. Warner, C.P. Kolbert & D.H. Persing: Perpetuation of the agent of human granulocytic ehrlichiosis in a deer tick-rodent cycle.*Proc Natl Acad Sci U S A* 93, 6209-6214 (1996)

9. D.E.Sonenshine: Biology of ticks Vol. 1. *Oxford University Press*, Oxford. (1991)

10. D.E. Sonenshine: Biology of ticks Vol. 2. *Oxford University Press*, Oxford. (1993)

11. W.R. Kaufman: Assuring paternity in a promiscuous world: are there lessons for ticks among the insects?. *Parasitology* 129 Suppl, S145-160 (2004)

12. J.M. Ribeiro: Role of saliva in blood-feeding by arthropods. *Annu Rev Entomol* 32, 463-478 (1987)

13. S.K. Wikel: Host immunity to ticks. *Annu Rev Entomol* 41, 1-22 (1996)

14. S.K. Wikel: Tick modulation of host cytokines. *Exp Parasitol* 84, 304-309 (1996)

15. A.S. Bowman & J.R. Sauer: Tick salivary glands: function, physiology and future. *Parasitology* 129 Suppl, S67-81 (2004)

16. F.A. Lara, U. Lins, G. Paiva-Silva, I.C. Almeida, C.M. Braga, F.C. Miguens, P.L. Oliveira & M. Dansa-Petretski: A new intracellular pathway of haem detoxification in the midgut of the cattle tick *Boophilus microplus*: aggregation inside a specialized organelle, the hemosome. *J Exp Biol* 206, 1707-1715 (2003)

17. F.A.Lara, U. Lins, G.H. Bechara & P.L.Oliveira: Tracing heme in a living cell: hemoglobin degradation and heme traffic in digest cells of the cattle tick *Boophilus microplus*. *J Exp Biol* 208, 3093-3101 (2005)

18. S.C. Barker & A.Murrell: Systematics and evolution of ticks with a list of valid genus and species names. *Parasitology* 129 Suppl, S15-36 (2004)

19. J.S.H. Klompen, W.C. Black, J.E. Keirans & D.E. Norris: Systematics and biogeography of hard ticks, a total evidence approach. *Cladistics* 16, 79-102 (2000)

20. J. de la Fuente: The fossil record and the origin of ticks (Acari: Parasitiformes: Ixodida). *Exp Appl Acarol* 29, 331-344 (2003)

21. S.C. Murrell & A. Barker: Synonymy of *Boophilus* Curtice, 1891 with *Rhipicephalus* Koch, 1844 (Acari: Ixodidae). *Syst Parasitol* 56, 169-172 (2003)

22. J.H. Jr. Oliver: Cytogenetics of mites and ticks. *Annu Rev Entomol* 22, 407-429 (1977)

23. C. Chen, U.G. Munderloh & T.J. Kurtti: Cytogenetic characteristics of cell lines from *Ixodes scapularis* (Acari: Ixodidae). *J Med Entomol* 31, 425-434 (1994)

24. N.S. Geraci, J. Spencer Johnston, J. Paul Robinson, S.K. Wikel & C.A. Hill: Variation in genome size of argasid and ixodid ticks. *Insect Biochem Mol Biol* 37, 399-408 (2007)

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

26. A.J.Ullmann, C.M. Lima, F.D. Guerrero, J. Piesman & W.C.Black: Genome size and organization in the blacklegged tick, *Ixodes scapularis* and the Southern cattle tick, *Boophilus microplus*. *Insect Mol Biol* 14, 217-222 (2005)

27. J.E. Childs & C.D. Paddock: The ascendancy of *Amblyomma americanum* as a vector of pathogens affecting humans in the United States. *Annu Rev Entomol* 48, 307-337 (2003)

28. T.R. Gregory, Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. *Biol Rev Camb Philos Soc* 76, 65-101 (2001)

29. J. Bunikis & A.G. Barbour, Ticks have R2 retrotransposons but not the consensus transposon target site of other arthropods. *Insect Mol Biol* 14, 465-474 (2005)

30. A.J. Ullmann, J. Piesman, M.C. Dolan & W.C.i. Black: Preliminary linkage map of the hard tick, *Ixodes scapularis*. *Insect Mol Biol* 12, 201-210 (2003)

31. W.C.t. Black & R.L.Roehrdanz: Mitochondrial gene order is not conserved in arthropods: prostriate and metastriate tick mitochondrial genomes. *Mol Biol Evol* 15, 1772-1785 (1998)

32. N.J. Campbell & S.C. Barker: The novel mitochondrial gene arrangement of the cattle tick, *Boophilus microplus*: fivefold tandem repetition of a coding region. *Mol Biol Evol* 16, 732-740 (1999)

33. R. Shao, Y. Aoki, H. Mitani, N. Tabuchi, S.C. Barker & M. Fukunaga: The mitochondrial genomes of soft ticks have an arrangement of genes that has remained unchanged for over 400 million years. *Insect Mol Biol* 13, 219-224 (2004)

34. G. Pertea, X. Huang, F. Liang, V. Antonescu, R. Sultana, S. Karamycheva, Y. Lee, J. White, F. Cheung, B. Parvizi, J. Tsai & Quackenbush, J. TIGR: Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* 19, 651-652 (2003)

35. Haas, B. J., Delcher, A. L., Mount, S. M., Wortman, J. R., Smith, R. K., Jr., Hannick, L. I., Maiti, R., Ronning, C.

M., Rusch, D. B., Town, C. D., Salzberg, S. L. & White, O. Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res* 31, 5654-5666 (2003)

36. A.L. Crampton, C. Miller, G.D. Baxter & S.C. Barker: Expressed sequenced tags and new genes from the cattle tick, *Boophilus microplus*. *Exp Appl Acarol* 22, 177-186 (1998)

37. C.A. Hill & J.A. Gutierrez: Analysis of the expressed genome of the lone star tick, *Amblyomma americanum* (Acari: Ixodidae) using an expressed sequence tag approach, *Microb Comp Genomics* 5, 89-101 (2000)

38. F.D.Guerrero, R.J. Miller, M.E. Rousseau, S. Sunkara, J. Quackenbush, Y. Lee & V. Nene: BmiGI: a database of cDNAs expressed in *Boophilus microplus*, the tropical/southern cattle tick. *Insect Biochem Mol Biol* 35, 585-595 (2005)

39. Valenzuela, J. G., Francischetti, I. M., Pham, V. M., Garfield, M. K., Mather, T. N. & Ribeiro, J. M. Exploring the sialome of the tick *Ixodes scapularis*. *J Exp Biol* 205, 2843-2864 (2002)

40. G. Leboulle, C. Rochez, J. Louahed, B. Ruti, M. Brossard, A. Bollen. & E. Godfroid: Isolation of *Ixodes ricinus* salivary gland mRNA encoding factors induced during blood feeding. *Am J Trop Med Hyg* 66, 225-233 (2002)

41. J.G. Valenzuela: Exploring the messages of the salivary glands of *Ixodes ricinus*. *Am J Trop Med Hyg* 66, 223-224 (2002)

42. V. Nene, D.Lee, J. Quackenbush, R. Skilton, S. Mwaura, M.J.Gardner & R. Bishop: AvGI, an index of genes transcribed in the salivary glands of the ixodid tick *Amblyomma variegatum*. *Int J Parasitol* 32, 1447-1456 (2002)

43. V. Nene, D.Lee, S. Kang'a, R. Skilton, T. Shah, E. de Villiers, S. Mwaura, D. Taylor, J. Quackenbush & R. Bishop: Genes transcribed in the salivary glands of female *Rhipicephalus appendiculatus* ticks infected with *Theileria parva*. *Insect Biochem Mol Biol* 34, 1117-1128 (2004)

44. I.M. Francischetti, V. My Pham, B.J. Mans, J.F. Andersen, T.N. Mather, R.S. Lane & J.Mribeiro: The transcriptome of the salivary glands of the female western black-legged tick *Ixodes pacificus* (Acari: Ixodidae). *Insect Biochem Mol Biol* 35, 1142-1161 (2005)

45. F.J. Alarcon-Chaidez, J. Sun. & S.K. Wikel: Transcriptome analysis of the salivary glands of *Dermacentor andersoni* Stiles (Acari: Ixodidae). *Insect Biochem Mol Biol* 37, 48-71 (2007)

46. J.M. Ribeiro: How ticks make a living. *Parasitol Today* 11, 91-93 (1995)

47. S.K. Wikel, R.N.Ramachandra & D.K. Bergman: Tick-induced modulation of the host immune response. *Int J Parasitol* 24, 59-66 (1994)

48. R. Bishop, B.Lambson, C.Wells, P. Pandit, J. Osaso, C. Nkonge, S. Morzaria, A. Musoke & V. Nene: A cement protein of the tick *Rhipicephalus appendiculatus*, located in the secretory e cell granules of the type III salivary gland acini, induces strong antibody responses in cattle. *Int J Parasitol* 32, 833-842 (2002)

49. A.R. Trimnell, G.M. Davies, O. Lissina, R.S. Hails & P.A. Nuttall: A cross-reactive tick cement antigen is a candidate broad-spectrum tick vaccine. *Vaccine* 23, 4329-4341 (2005)

50. J.G.Valenzuela: Exploring tick saliva: from biochemistry to 'sialomes' and functional genomics. *Parasitology* 129 Suppl, S83-94 (2004)

51. J.M.Ribeiro, F. Alarcon-Chaidez, I.M. Francischetti, B.J. Mans, T.N. Mather, J.G. Valenzuela, J. G. & S.K.Wikel: An annotated catalog of salivary gland transcripts from *Ixodes scapularis* ticks. *Insect Biochem Mol Biol* 36, 111-129 (2006)

52. S. Narasimhan, K. Deponte, N. Marcantonio, X. Liang, T.E. Royce, K.F. Nelson, C.J. Booth, B. Koski, J.F. Anderson, F. Kantor & E. Fikrig: Immunity against *Ixodes scapularis* Salivary Proteins Expressed within 24 Hours of Attachment Thwarts Tick Feeding and Impairs *Borrelia* Transmission. *PLoS ONE* 2, e451 (2007)

53. I.K. Santos, J.G. Valenzuela, J.M. Ribeiro, M. de Castro, J.N. Costa, A.M. Costa, E.R. da Silva, O.B. Neto, C. Rocha, S. Daffre, B.R.Ferreira, J.S. da Silva, M.P. Szabo & G.H. Bechara: Gene discovery in *Boophilus microplus*, the cattle tick: the transcriptomes of ovaries, salivary glands, and hemocytes. *Ann N Y Acad Sci* 1026, 242-246 (2004)

54. J.P. Vinson, D.B. Jaffe, K. O'Neill, E.K. Karlsson, N. Stange-Thomann, S. Anderson, J.P. Mesirov, N. Satoh, Y. Satou, C. Nusbaum, B. Birren, J.E. Galagan & E.S. Lander: Assembly of polymorphic genomes: algorithms and application to *Ciona savignyi*. *Genome Res* 15, 1127-1135 (2005)

55. V. Nene, J.R. Wortman, D. Lawson, B. Haas, C. Kodira, Z.J. C. Tu, B. Loftus, Z. Xi, K. Megy, M. Grabherr, Q.Ren, E.M. Zdobnov, N.F. Lobo, K.S. Campbell, S.E. Brown, M.F. Bonaldo, J. Zhu, S.P. Sinkins, D.G. Hogenkamp, P. Amedeo, P. Arensburger, P.W. Atkinson, S. Bidwell, J.Biedler, E. Birney, R.V. Bruggner, J. Costas, M.R.Coy, J. Crabtree, M. Crawford, B. Debruyn, D. Decaprio, K. Eiglmeier, E. Eisenstadt, H. El-Dorry, W.M. Gelbart, S.L. Gomes, M. Hammond, L.I. Hannick, J.R. Hogan, M.H. Holmes, D. Jaffe, J.S. Johnston, R.C. Kennedy, H. Koo, S. Kravitz, E.V. Kriventseva, D. Kulp, K. Labutti, E. Lee, S. Li, D.D. Lovin, C. Mao, E. Mauceli, C.F. Menck, J.R. Miller, P. Montgomery, A. Mori, A.L. Nascimento, H.F. Naveira, C. Nusbaum, S.O'Leary, J.

Orvis, M. Pertea, H. Quesneville, K.R. Reidenbach, Y.H. Rogers, C.W. Roth, J.R. Schneider, M. Schatz, M. Shumway, M. Stanke, E.O. Stinson, J.M. Tubio, J.P. Vanzee, S. Verjovski-Almeida, D. Werner, O. White, S.Wyder, Q. Zeng, Q. Zhao, Y. Zhao, C.A. Hill, A.S. Raikhel, M.B. Soares, D.L. Knudson, N.H. Lee, J. Galagan, S.L. Salzberg, I.T. Paulsen, G. Dimopoulos, F.H.Collins, B. Birren, C.M. Fraser-Liggett & D.W. Severson: Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* 316, 1718-1723 (2007)

56. C.Dale & N.A. Moran: Molecular interactions between bacterial symbionts and their hosts. *Cell* 126, 453-465 (2006)

57. S.L. Salzberg, J.C. Hotopp, A.L. Delcher, M. Pop, D.R. Smith, M.B. Eisen & W.C. Nelson: Serendipitous discovery of *Wolbachia* genomes in multiple *Drosophila* species. *Genome Biol* 6, R23 (2005)

58. J.C. Hotopp, M. Lin, R. Madupu, J. Crabtree, S.V. Angiuoli, J. Eisen, R. Seshadri, Q. Ren, M. Wu, T.R. Utterback, S. Smith, M. Lewis, H. Khouri, C. Zhang, H. Niu, Q. Lin, N.Ohashi, N. Zhi, W. Nelson, L.M.Brinkac, R.J. Dodson, M.J. Rosovitz, J. Sundaram, S.C. Daugherty, T. Davidsen, A.S. Durkin, M. Gwinn, D.H. Haft, J.D. Selengut, S.A. Sullivan, N. Zafar, L. Zhou, F. Benahmed, H. Forberger, R. Halpin, S. Mulligan, J. Robinson, O. White, Y. Rikihisa & H. Tettelin: Comparative genomics of emerging human ehrlichiosis agents. *PLoS Genet* 2, e21 (2006)

59. H. Noda, U.G. Munderloh & T.J. Kurtti: Endosymbionts of ticks and their relationship to *Wolbachia* spp. and tick-borne pathogens of humans and animals. *Appl Environ Microbiol* 63, 3926-3932 (1997)

60. S.J.Weller, G.D. Baldridge, U.G. Munderloh, H. Noda, J. Simser & T.J. Kurtti: Phylogenetic placement of rickettsiae from the ticks *Amblyomma americanum* and *Ixodes scapularis*. *J Clin Microbiol* 36, 1305-1317 (1998)

61. L. Sacchi, E. Bigliardi, S. Corona, T. Beninati, N. Lo & A.A. Franceschi: symbiont of the tick *Ixodes ricinus* invades and consumes mitochondria in a mode similar to that of the parasitic bacterium *Bdellovibrio bacteriovorus*. *Tissue Cell* 36, 43-53 (2004)

62. S. Varde, J. Beckley & I. Schwartz: Prevalence of tick-borne pathogens in *Ixodes scapularis* in a rural New Jersey County. *Emerg Infect Dis* 4, 97-99 (1998)

63. T.J. Kurtti, J.A. Simser, G.D. Baldridge, A.T. Palmer & U.G. Munderloh: Factors influencing *in vitro* infectivity and growth of *Rickettsia peacockii* (Rickettsiales: Rickettsiaceae), an endosymbiont of the Rocky Mountain wood tick, *Dermacentor andersoni* (Acari, Ixodidae). *J Invertebr Pathol* 90, 177-186 (2005)

64. J.C. Hotopp, M.E. Clark, D.C. Oliveira, J.M. Foster, P. Fischer, M.C. Torres, J.D. Giebel, N. Kumar, N. Ishmael, S. Wang, J.Ingram, R.V. Nene, J. Shepard, J. Tomkins, S. Richards, D.J. Spiro, E. Ghedin, B.E. Slatko, H.Tettelin & J.H. Werren: Widespread Lateral Gene Transfer from Intracellular Bacteria to Multicellular Eukaryotes. *Science* (2007)

65. F. F. Costa: Non-coding RNAs: lost in translation?. *Gene* 386, 1-10 (2007)

66. F.D. Guerrero, V.M. Nene, J.E. George, S.C. Barker & P. Willadsen: Sequencing a new target genome: the *Boophilus microplus* (Acari: Ixodidae) genome project, *J Med Entomol* 43, 9-16 (2006)

67. L. Bell-Sakyi, E. Zweygarth, E. Blouin, E.A. Gould & F. Jongejan: Tick cell lines: tools for tick and tick-borne disease research. *Trends in Parasitol* 23, 450-457 (2007)

68. J.H. Kim, M.S. Waterman & L.M. Li: Diploid genome reconstruction of *Ciona intestinalis* and comparative analysis with *Ciona savignyi*. *Genome Res* 17, 1101-1110 (2007)

69. S. Levy, G. Sutton, P.C. Ng, L. Feuk, A.L. Halpern, B.P. Walenz, N. Axelrod, J. Huang, E.F. Kirkness, G. Denisov, Y. Lin, J.R. Macdonald, A.W. Pang, M. Shago, T.B. Stockwell, A. Tsiamouri, V. Bafna, V. Bansal, S.A. Kravitz, D.A. Busam, K.Y. Beeson, T.C. McIntosh, K.A. Remington, J.F. Abril, J. Gill, J. Borman, Y.H. Rogers, M.E. Frazier, S.W. Scherer, R.L. Strausberg & J.C. Venter: The Diploid Genome Sequence of an Individual Human. *PLoS Biol* 5, e254 (2007)

70. K.S. Small, M. Brudno, M.M. Hill & A.A. Sidow: A haplome alignment and reference sequence of the highly polymorphic *Ciona savignyi* genome. *Genome Biol* 8, R41 (2007)

Key Words: Tick, Genome, EST, Sequence, Arthropod, Review

Send correspondence to: Vishvanath Nene, Institute for Genome Sciences and Department of Microbiology and Immunology University of Maryland School of Medicine HSF-II, Room S447 20 Penn Street Baltimore, MD 21201, USA, Tel: 410-706-3860 Fax: 410-706-1482**,** E-mail: vnene@som.umaryland.edu

http://www.bioscience.org/current/vol14.htm