Adaptive Sampling as a New Tool for Pathogen Surveillance and Mitogenome Assembly in Blood-feeding Arthropod Vectors

University of Minnesota
Driven to Discover®

Evan J. Kipp[1], Laramie L. Lindsey[1], Ben Khoo[2], Marissa S. Milstein[1], Lexi E. Frank[1], Cristina M. Blanco[3], Julia P. Baker[3], Christopher Faulk[4], Jonathan D. Oliver[2], Peter A. Larsen[1]

[1] Department of Veterinary and Biomedical Sciences, University of Minnesota, St. Paul, USA
[2] Division of Environmental Health Sciences, School of Public Health, University of Minnesota, Minneapolis, USA
[3] College of Veterinary Medicine, University of Minnesota, St. Paul, USA
[4] Department of Animal Science, University of Minnesota, St. Paul, USA

INTRODUCTION

- Diseases spread by arthropod vectors (e.g., ticks, mosquitoes) are major global health threats and have been increasing considerably over recent decades[1]
- New molecular tools for genomic vector-borne pathogen surveillance and for characterizing a variety of arthropod vector taxa are urgently needed
- Nanopore sequencing supplemented with adaptive sampling represents an innovative strategy through which researchers may enrich or deplete their sequencing output in real-time based on near-instantaneous mapping of individual reads against a user-specified reference as sequencing is occurring[2,3]
- This approach has many exciting potential applications in vector-borne disease surveillance, pathogen discovery, and can be simultaneously leveraged for molecular identification of cryptic arthropods and their blood-meals through analysis of mitochondrial species barcoding genes[4,5]

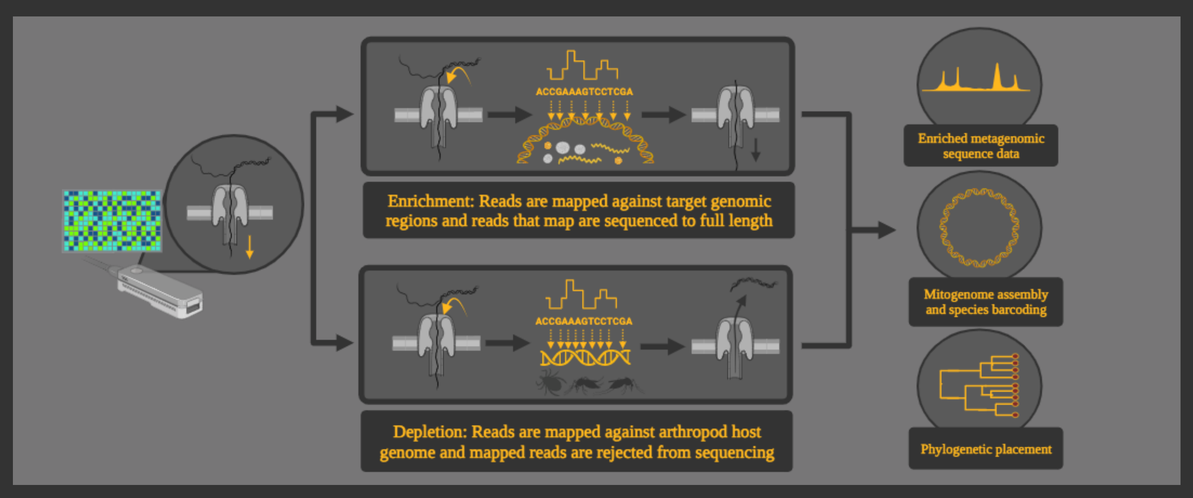

Figure 1. Adaptive sampling workflow using either target enrichment (e.g., pathogen genome assemblies, mitogenomes) or host depletion sequencing strategies.

Figure 2. Adult female *Ixodes scapularis* tick, vector of numerous bacterial, viral, and eukaryotic pathogens throughout North America. (Photo: E. Kipp)

METHODS

Metagenomic surveillance for tick-borne pathogens:

- Genomic DNA isolated from field-collected *Ixodes scapularis* ticks
- DNA sequenced on R9 MinION flow cells with adaptive sampling via two approaches:
  a. Enrichment for target tick-borne pathogen genomes
  b. Host depletion with *I. scapularis* whole genome assembly and endosymbiont *Rickettsia buchneri*
- Sequencing performed on Linux machine with live GPU-enabled basecalling
- Findings compared against a paired control run from same individual ticks, sequenced without adaptive sampling

Mitogenome assembly and species barcoding in insect vectors:

- Mosquitoes and other blood-feeding dipterans (e.g., deer fly, *Culicoides* midges) collected with CO_2-baited light traps; genomic DNA extracted
- Adaptive sampling performed with enrichment targeting invertebrate mtDNA for species barcoding and vertebrate mtDNA for blood-meal analysis
- Complete mitogenomes were assembled using Flye and target genes extracted for species barcoding; findings compared against control run

FINDINGS

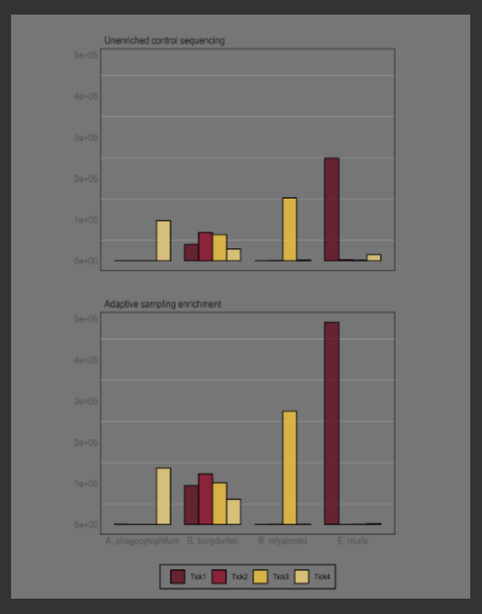

Figure 3. Total bases for individual *I. scapularis* ticks mapping to bacterial tick-borne pathogen genomes during control (top) and adaptive sampling experiments (bottom).

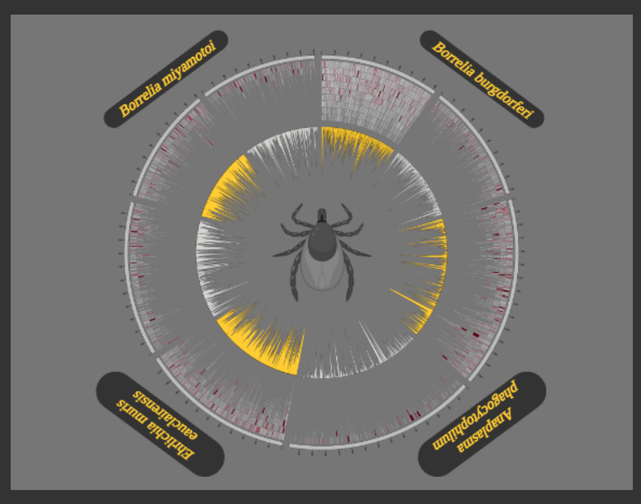

Figure 4. Detection of four bacterial tick-borne pathogens in paired adaptive sampling and control sequencing runs. Mean coverage generated through adaptive sampling enrichment is depicted in yellow and unenriched control nanopore sequence data is shown in gray. Individual reads are shown along the outer track, with reads > 5kb in length visualized in dark red.

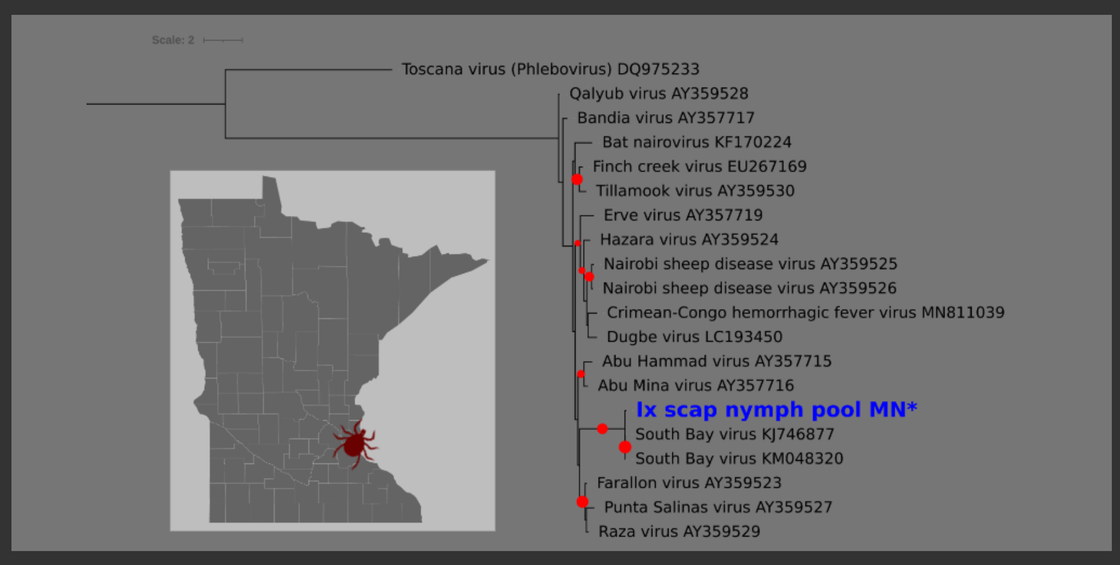

Figure 5. Maximum likelihood phylogeny of tick nairoviruses based on the RNA-dependent RNA polymerase gene. The consensus sequence for South Bay virus (blue) was generated through adaptive sampling targeting potential tick viruses from a pool of *I. scapularis* nymphs. Nodes with statistically supported bootstrap values of >70 are visualized in red.

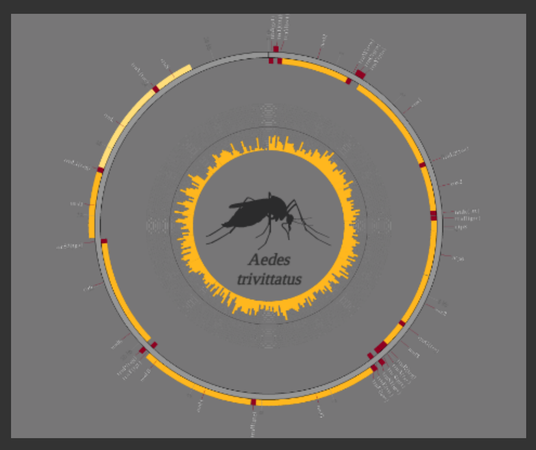

Figure 6. Mitochondrial genome map depicting annotated genes (outer track) and GC content (inner track) for the mosquito *Aedes trivittatus*, sequenced through adaptive sampling with enrichment targeting mtDNA.

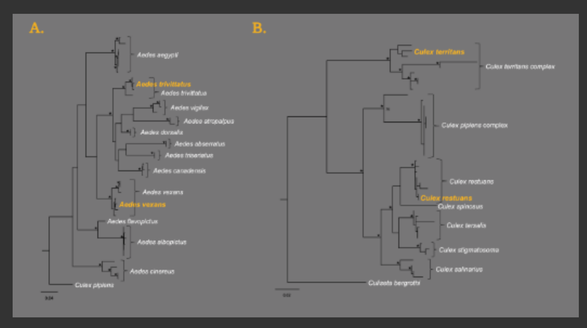

Figure 7. Maximum likelihood phylogenies generated based on COI gene consensus sequences for select *Aedes* (panel A) and *Culex* (panel B) mosquito species sequenced using adaptive sampling with mtDNA enrichment. Statistically supported nodes (≥ 75 bootstrap value) are depicted with an asterisk (*); trees generated using 1000 bootstrap replicates.

DISCUSSION

- Adaptive sampling serves as an innovative and effective method for rapid and streamlined metagenomic vector-borne pathogen surveillance and arthropod species barcoding
- In comparison to control libraries, adaptive sampling enrichment and depletion strategies enabled a roughly two-fold improvement for target tick-borne pathogen reads
- In these preliminary experiments, adaptive sampling successfully detected and sequenced at low coverage, genomes for four important emerging bacterial pathogens: *Anaplasma phagocytophilum*, *Borrelia burgdorferi* s.s., *Borrelia miyamotoi*, and *Ehrlicha muris eauclairensis*
- Furthermore, adaptive sampling enrichment targeting mtDNA enabled sequencing of high-coverage mitogenomes, while also enabling identification of host blood-meals
- Future experiments will further evaluate adaptive sampling to better understand its sensitivity and specificity against traditional surveillance approaches (e.g., rt-PCR)
- Ongoing work by our team is also aimed at utilizing the approach in remote field settings for real-time vector species barcoding and biosurveillance efforts

REFERENCES

1. Rosenberg, R., Lindsey, N. P., Fischer, M., Gregory, C. J., Hinckley, A. F., Mead, P. S., ... & Petersen, L. R. (2018). Vital signs: trends in reported vectorborne disease cases—United States and Territories, 2004–2016. *Morbidity and Mortality Weekly Report*, *67*(17), 496.
2. Payne, A., Holmes, N., Clarke, T., Munro, R., Debebe, B. J., & Loose, M. (2021). Readfish enables targeted nanopore sequencing of gigabase-sized genomes. *Nature biotechnology*, *39*(4), 442-450.
3. Wanner, N., Larsen, P. A., McLain, A., & Faulk, C. (2021). The mitochondrial genome and Epigenome of the Golden lion Tamarin from fecal DNA using Nanopore adaptive sequencing. *BMC genomics*, *22*, 1-11.
4. De Meulenaere, K., Cuypers, W. L., Rosanas-Urgell, A., Laukens, K., & Cuypers, B. (2022). Selective whole-genome sequencing of Plasmodium parasites directly from blood samples by Nanopore adaptive sampling. *bioRxiv*, 2022-11.
5. Kipp, E. J., Lindsey, L. L., Milstein, M. S., Blanco, C. M., Baker, J. P., Faulk, C., ... & Larsen, P. A. (2023). Nanopore adaptive sampling for targeted mitochondrial genome sequencing and bloodmeal identification in hematophagous insects. *Parasites & Vectors*, *16*(1), 68.

ACKNOWLEDGEMENTS

We are grateful to Suzanne Stone for logistical support and assistance with molecular experiments. Funding for ongoing tick-borne pathogen experiments is provided through the National Institutes of Health (NIH) under award R01AI155472-01A1. The Minnesota Supercomputing Institute (MSI) at the University of Minnesota provided resources used for data storage and computational analysis. Financial support for C.M.B. and J.P.B. was provided by the University of Minnesota College of Veterinary Medicine Summer Scholars program, with funds from the Office of the Director of the NIH under award T35OD0111118. We are also thankful to the Wildlife Rehabiliation Cener of Minnesota and luciano Caixeta for access to sites used for arthropod collections.