G107 Developing Genomic Tools for Forensically Important Flies to Improve Forensics

Aaron M. Tarone, PhD, Texas A&M University, Department of Entomology, College Station, TX 77843; Sing-Hoi Sze, PhD, Texas A&M University, Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843; Joseph P. Dunham, BS, University of Southern California, Molecular and Computational Biology, Los Angeles, CA 90089; Brandi-Lee Carey, BS, University of Massey, Institute of Molecular BioSciences, Palmerston North, NC, NEW ZEALAND; Peter C. Chang, BS, University of Southern California, Molecular and Computational Biology, Los Angeles, CA; Fang Li, PhD, North Carolina State University, Department of Genetics, Raleigh, NC 27695; Courtney Fjeldsted, BS, University of Southern California, Molecular and Computational Biology, Los Angeles, CA 90089; Maxwell J. Scott, PhD, North Carolina State University, Department of Genetics, Raleigh, NC 27695; and Sergey V. Nuzhdin, PhD, University of Southern California, Molecular and Computational Biology, Los Angeles, CA 90089*

After attending this presentation, attendees will learn that developmental and quantitative genetics has an improving forensic entomology. Attendees will also gain an appreciation of how to develop genomic tools and implement genomic techniques to improve forensic entomology.

This presentation will impact the forensic science community by showing how recent technological advances have allowed non-model organism researchers to conduct genomic research which means that genomics can be used to improve problems in forensic entomology, which is dominated by non-model organisms. The data presented here represent the first set of genomic tools for a common forensically important blowfly species.

Recent research indicates that quantitative and functional genetic principles will be useful in improving the accuracy and precision of arthropod derived postmortem interval estimates. However, such endeavors require genomic tools, which are lacking in non-model species. Next-generation sequencing provides the necessary capabilities to develop genomic tools for blowflies and other forensically relevant species. For many species whole genome assembly is not feasible, but it is possible to assemble sequences that represent the transcriptome: the subset of the genome represented by transcribed genes. However, transcripts possess alternative splices, which must be accounted for in the assembly process (or ignored at the cost of losing information). The *de novo* assembly of the *Lucilia sericata* (Diptera: Calliphoridae) transcriptome is described here, outlining computational and molecular efforts to identify and confirm alternative splice, allelic, and gene expression information in a species important to forensic entomology.

An algorithm that enables the identification of alternative splices in assembled transcripts through the use of de bruijn networks was developed. This tool, ASplice, also identifies putative SNPs and reports library specific gene expression estimates (expressed as reads per million mapped reads per thousand bases of transcript; or RPKM). Before assembling blowfly data, *Drosophila melanogaster* data was assembled in this manner, revealing false positive and false negative rates for transcript and splice identification. The performance of ASplice was also compared to other algorithms, demonstrating the ability of the program to be conservative (unlikely to identify false positives) and/or less memory intensive than competing *de novo* assembly software packages.

After validation of the assembly methods, *Lucilia sericata* RNA derived from embryonic, larval, pupal, and adult samples was sequenced using a combination of Illumina and 454 sequencing. All developmental stages were sequenced using Illumina, producing >6 billion bases of raw useable sequence. In addition, reciprocal subtractive hybridizations were performed between larval and pupal samples in a manner that would enrich for transcripts that are differentially expressed between: (1) feeding and postfeeding third instars; and, (2) early and mid pupation. Samples derived from subtractive hybridizations and salivary gland RNA were sequenced using 454 sequencing, yielding tens of millions of bases of sequence data.

Results indicate that the authors have identified hundreds of transcripts (and isoforms) that are strong candidates for use in predicting blowfly developmental age in a manner that increases precision compared to traditional forensic entomology approaches. Many of these loci also have *Drosophila* homologs that are expressed in developmentally regulated patterns commensurate with expression patterns observed in *Lucilia*, while others are unique to the species. Thousands of putative SNPs and splices have been identified, which will be useful in quantitative and population genetic studies. Sanger sequencing has been used to confirm assembled sequences in seven genes, including several alternatively spliced sex determination genes. Results from the analysis of the transcriptome will be presented and the audience will be exposed to several lines of research that can be employed once genomic tools are available for forensically important species.

Postmortem Interval, Genomics, Gene Expression

