

Workshop: The Utility of Next Generation Sequencing for Genome Scale Studies

Rachel J. O'Neill, James Lindsay, Dawn Carone, Sahar Al Seesi, Thomas Heider, Ion Mandoiu, and Andrew Pask

University Of Connecticut

Storrs, CT, USA

e-mail: rachel.oneill@uconn.edu

Molecular biological enquiry has entered a new age in the last few years with the development of high-throughput, massively parallel data acquisition platforms. Since it's advent with the introduction of gene microarrays early in this century, the technology has continued to evolve at an astonishing pace into sequencers capable of millions to billions of sequencing reads in one instrument run. After introducing the platforms, applications and advances in next generation sequencing technologies, I will discuss our involvement with several genome sequencing initiatives, focusing more closely on the tammar wallaby genome project.

The Baylor College of Medicine, Australian Government and University of Connecticut teamed together (under the KanGO consortium) as part of an international effort to generate a de novo genome sequence for the marsupial model species, the tammar wallaby (*Macropus eugenii*). This important model organism is part of a large group of mammals (Metatherian) that harbor unique life history traits and genome features. Our work involved assembling this 2.7 GB mammalian genome from sequence produced from Sanger sequencing reads (Baylor and RIKEN), Illumina

sequencing reads (Australian Genome Research Foundation, AGRF), and 454 reads (AGRF and the CAGT at UConn) using unique assembly tools we have developed. The genome sequencing initiative for the tammar wallaby has also afforded an opportunity to explore the evolution of repeat classes as well as each of the major classes of small RNAs, siRNAs, miRNAs, piRNAs, and the newest class of small RNAs, crasiRNAs (centromere repeat associated short interacting RNAs), first discovered in the tammar wallaby.

Although only recently discovered, small RNAs have proven to be essential regulatory molecules encoded within eukaryotic genomes. These molecules, represented by four major class sizes ranging from 20nt to 42nt, are participants in a diverse array of cellular processes including gene regulation, chromatin dynamics and genome defense. Our analysis includes the detailed examination of these small RNAs, salient features that define their canonical members and the constitution of species-specific members derived from mobile genetic elements. Understanding the evolution of these important regulators and surveyors of the genome in this unique marsupial model brings valuable insights to the community of small RNA researchers.