Technological University, Singapore, (4)University of Wuerzburg, Germany, (5)J. Craig Venter Institute, (6)Singapore Botanic Gardens, Singapore, (7)Penn State University, State College, PA, (8)University of Helsinki, Helsinki, Finland, (9)University of Lausanne, Switzerland

The development of long read sequencing technologies has entirely changed the landscape of possibilities for sequencing and assembling complex plant genomes. While Pacific Biosciences SMRT sequencing has served admirably for a number of years now, Oxford Nanopore technology is highly portable and requires much less up-front investment. Concerns have arisen over Nanopore error rates compared to PacBio or Illumina, but at least using current versions of flow cells, chemistry, and base-calling, we no longer find these misgivings tenable regarding the construction of a highly contiguous genome. Still, a combined approach is required, at minimum including polishing using low-error - but cheap - Illumina reads. Assemblies at the chromosome scale often require further efforts, such as HiC scaffolding - but the work flow is now democratized to the extent that any university lab should be able to generate a high-quality genome of its choice. We use case-by-case workflows for generating chromosome-scale assemblies of various-sized plant genomes. Unfortunately, no one assembly approach (e.g., De Bruijn graph or overlap-layout-consensus method) works best for all species, given the various and sundry nature of their heterozygosities, ploidy levels, and transposable element blooms (the latter two also in terms of their event ages). Despite not yet achieving a truly pipeline approach, we are satisfied with our ability to generate excellent de novo genomes on unprecedentedly low time and cost scales. We will describe several of our recent projects and the individual challenges encountered and how they were overcome.

W952: Sequencing Complex Genomes

The Importance of Sequencing Depth for a Complex Transcriptome

Adhini Sudhindra Kumar Pazhany[1], Virginie Perlo[2], Frederik Botha[3], Agnelo Furtado[4], Angela O'Keeffe[5], Ardy Kharabian Masouleh[5], Robert Henry[6], Karen Aitken[7], Angelique d'Hont[8], Adam Healey[9], Jane Grimwood[10], Kerrie Barry[11] and Jeremy Schmutz[10], (1)Queensland Alliance for Agriculture and Food Innovation,The University of Queensland, St Lucia, QLD, Australia, (2)University of Queensland, Brisbane, Australia, (3)QAAFI Queensland Alliance for Agriculture and Food Innovation - UQ University, St Lucia, QLD, Australia, (4)University of Queensland/QAAFI, Brisbane, QLD, Australia, (5)QAAFI (Queensland Alliance for Agriculture and Food Innovation), The University of Queensland, Brisbane, QLD, Australia, (6)Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Brisbane, QLD, Australia, (7)CSIRO Agriculture and Food, St Lucia, Australia, (8)CIRAD (Centre de Coopération Internationale en Recherche Agronomique pour le Développement), Montpellier, France, (9)HudsonAlpha Institute For Biotechnology, Huntsville, AL, (10)HudsonAlpha Institute for Biotechnology, Huntsville, AL, (11)Department of Energy Joint Genome Institute, Walnut Creek, CA

Sugarcane, a unique biological system which has evolved over time to suit the changing human needs is now being manipulated as a source of alternate energy and an array of platform chemicals. Sugarcane is a crop with diverse end uses and applications, a wealth of genetic resources and a rich breeding history. The daunting size and extreme complexity of the genome together with high heterozygosity and variable chromosome numbers has long hampered genomic research in sugarcane. Analysis of the transcriptome using long read sequencing has been reported. Normalization of libraries before sequencing has been widely employed in transcriptome analysis. In the complex sugarcane transcriptome, normalization was found to both reveal more rare sequences and result in the loss of many sequence variants. This suggested the need for deep sequencing to capture the diversity of sequences in sugarcane. Recently, a monoploid genome sequence was generated for the cultivar R570. We now report analysis of the transcriptome of this genotype and the impact of sequencing at greater depth on the recovery of transcripts. The results may guide future transcriptomic studies in sugarcane to make informed decisions on the required depth of sequencing.

W953: Sequencing Complex Genomes