#048

Exploiting Long Read Sequence Technology to Resolve the Hidden Genomic Landscape of Brassica Species

Andrew Sharpe
Isobel Parkin, Sampath Perumal, Chu Shin Koh, Erin Higgins, Lingling Jin, Miles Buchwald, Tricia Bender, Steve Robinson

Global Institute for Food Security, Saskatoon, Canada

Plant genome assembly has been developing rapidly with costs declining and scaffold size and genome coverage improving; however, with short read technologies, underlying contig size remains limited and it is inevitable that some genomic regions will not be captured and duplicated or repetitive regions are often collapsed. Concomitant with these improvements there is a growing appreciation that copy number variants, presence/absence variants and structural rearrangements have played an important role in the adaptation of phenotype. Long read sequencing technologies offer a unique opportunity to capture these often elusive genome differences. In order to study a large number of lines the technology needs to be both cost effective and preferably accessible to many labs. To test its applicability to polyploid species, two de novo genome assemblies were generated for Brassica nigra, a paleohexaploid, using Oxford Nanopore Technologies (ONT) sequence reads. The resultant assemblies were error corrected using Illumina short reads, and HiC and genotype data was added to generate pseudomolecules. The resulting assemblies were compared to each other and to a previously available assembly of one of the genotypes generated using an Illumina short-read sequencing based approach. The ONT assemblies extended the original reference assembly by ~90 Mb, covering ~94% of the expected genome size. The majority (85%) of the additional assembled sequence represented repetitive DNA, yet ~6,000 additional genes were added to the new assembly. The long read assemblies provided novel insights into the repetitive genome structure, access to previously hidden genes, and could span non-recombinant regions. This technology is advancing rapidly and offers many opportunities for accessing some of the important structural variation in Brassica species.