Adaptive sequencing in crop species

René Hogers, Alexander Wittenberg & Dick Roelofs

While most genome sequences are elucidated at an ever increasing pace, post-genome research focuses mostly on particular target loci in the genome. But how to select these target sequences, while rejecting the ones that are not of interest? KeyGene is among the first Agro Biotech companies that has implemented READ UNTIL, a tool enabling selective sequencing, while rejecting sequences that are not of interest.

Here, we provide the first evidence that adaptive Oxford Nanopore-based sequencing enables enrichment of targets in the melon genome.

Background and approach

Resequencing of targeted regions in the genome is of continued great interest to the plant breeding community. It promises to be a fast, efficient and cost-effective way to screen for genetic variation among loci of interest. Currently, several resequencing methods have been developed to target these specific loci of interest.

Wet-lab approaches can be subdivided into PCR-based methods and methods based on selection of native DNA sequence regions. PCR-based targeted resequencing has now been well established, due to the fact that many loci can be targeted and the work flow is relatively easy to implement.

Figure 1. the MinION device applied for adaptive sequencing (©Oxford Nanopore Technologies)

Still, amplification-based methods suffers from allele-specific preferential amplification, hampering insights in allelic distribution, especially in heterozygous or polyploid species. Also, the designed amplicons are of restricted lengths, so that larger and/or complex regions cannot be addressed.

These problems are mainly solved, when applying targeted resequencing on native DNA coupled to long read sequencing platforms such as those from Oxford Nanopore Technologies.

While wet-lab methods are continued to be optimized, bioinformatics tools became available for selecting target loci on the Oxford Nanopore Technology sequence platform [1]. The principle is simple and elegant at the same time. The MinION sequencer (Figure 1) is able to selectively sequence DNA in real-time by rejecting DNA molecules that do not match the target sequences through reversing the voltage of pores that are bound with off-target DNA molecules.

The voltage reversal causes the DNA strand to be rejected and pushed out of the pore. In turn, the pore is ready to bind a new DNA molecule for selective sequencing (Figure 2). Despite of its elegance, the method exhibits some severe computational challenges related to real-time base calling and mapping to a reference, hampering it from being used in more complex genomes.

With the introduction of graphical processing units (GPU) this challenge has been solved to a large extent. Indeed, Payne et al. used direct GPU base calling to selectively sequence the odd numbered chromosomes in the human genome, to show that a gigabase sized genome is not a constraint for Read Until [2-4].

At KeyGene we asked whether Read Until could be applied on plant crop genomes, which provide their own challenges with regard to for instance highly repetitive genomic regions.

About the authors of this White Paper

Ing. R.C.J. Hogers is scientist at KeyGene and responsible for innovations in DNA/RNA detection technologies.

Dr. Ir. Dick Roelofs is Program Scientist at KeyGene and team leader of technology innovations and next generation sequencing support.

Dr. Ir. A.H.J. Wittenberg is next generation sequencing expert at KeyGene and responsible for technology development and scouting

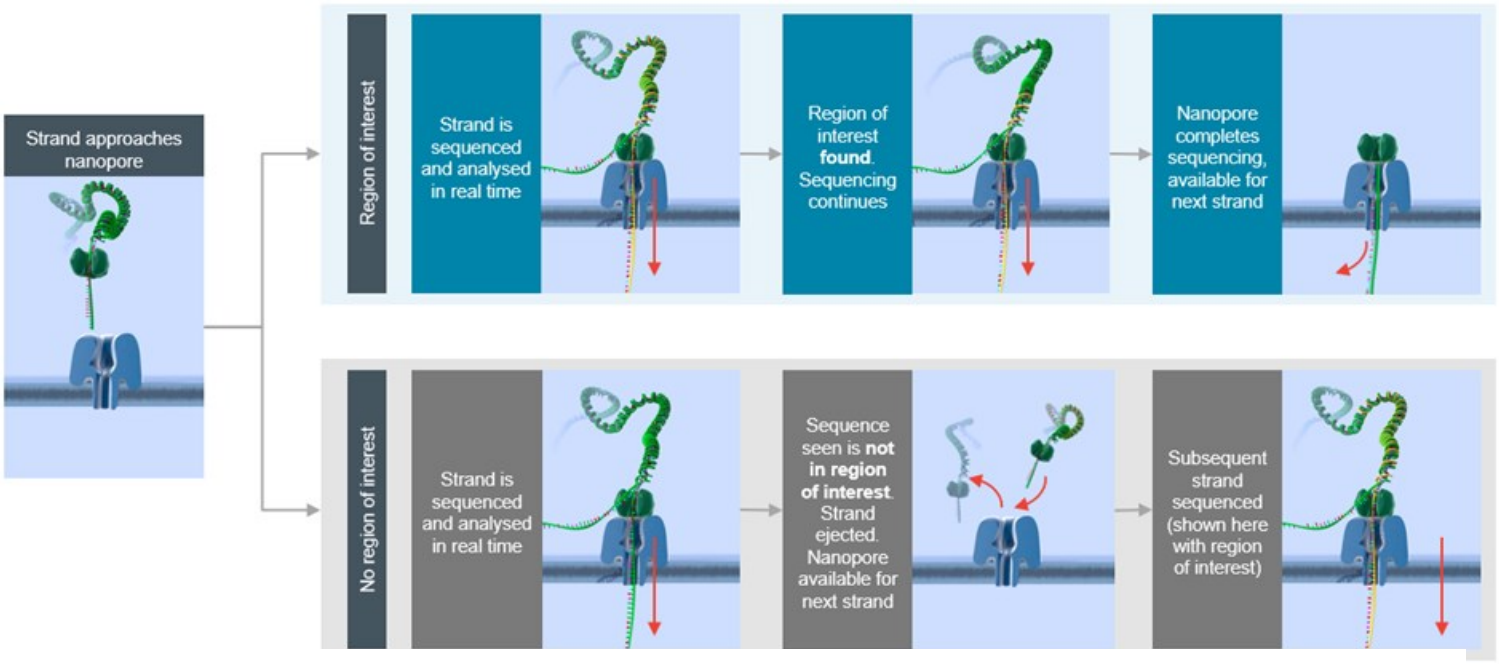

Figure 2. principle of adaptive sequencing (©Oxford Nanopore Technologies)

Approach

Melon cultivar Vedrantais was taken as proof of concept for Read Until sequencing in plants, for which we had previously generated a golden standard chromosome level reference genome sequence for this line using the very long reads from the PromethION platform. Moreover, ~800 regions of ~5-7 Kb, representing a total of ~4.7Mbp of the ~450Mbp melon genome and already selected and characterized in a previous experiment, were taken as targets for Read Until (RU) adaptive sequencing. High molecular weight DNA was isolated and used for Oxford Nanopore library construction and subsequent sequencing.

A Nvidia GeForce RTX 2080 Ti GPU was installed in order to facilitate real-time base calling of the MinION device. The Read Until Application (RU API), developed by the Loose lab was downloaded and installed from GitHub.

After successful testing of the API, we first performed a standard Whole Genome Sequencing (WGS) run using a R9.4.1 flow cell, without the Read Until API. Subsequently, a 2nd aliquot of the same WGS library was run along with Read Until adaptive sequencing.

Reads from both MinION runs were mapped against our Melon Vedrantais reference genome.

Results and discussion

During sequencing of the MinION flow cell with the RU API, we observed generation of short reads, which were not observed when sequencing without the RU API. This indicates that the RU API was indeed rejecting library fragments from being sequenced in full length.

After high accuracy basecalling of the resulting sequencing runs and mapping the sequence reads to the reference genome, we observed a moderate increase of ~1.7 fold when sequence depth on targets of WGS-RU was compared to WGS.

As expected, a more pronounced decrease of sequence depth of ~4.5 fold was observed throughout the genome when WGS-RU was compared to the regular WGS run (Table 1). This clearly indicates successful rejection of non-target sequences representing 99% of all sequences in the WGS library.

Overall, we observe an 8 fold increase in enrichment of the targeted regions. Figure 3 shows an overview of the coverage across the targeted loci.

.

Table 1. Summary statistics of the sequencing runs

	WGS MinION run	WGS MinION run Read Until
Sequencing output (Gb)	23	7
Average depth on target	21.38	36.7
Av. depth whole genome	20.48	4.53
Relative frequency target regions	1.04	8.18
Enrichment with Read Until		**7.86**

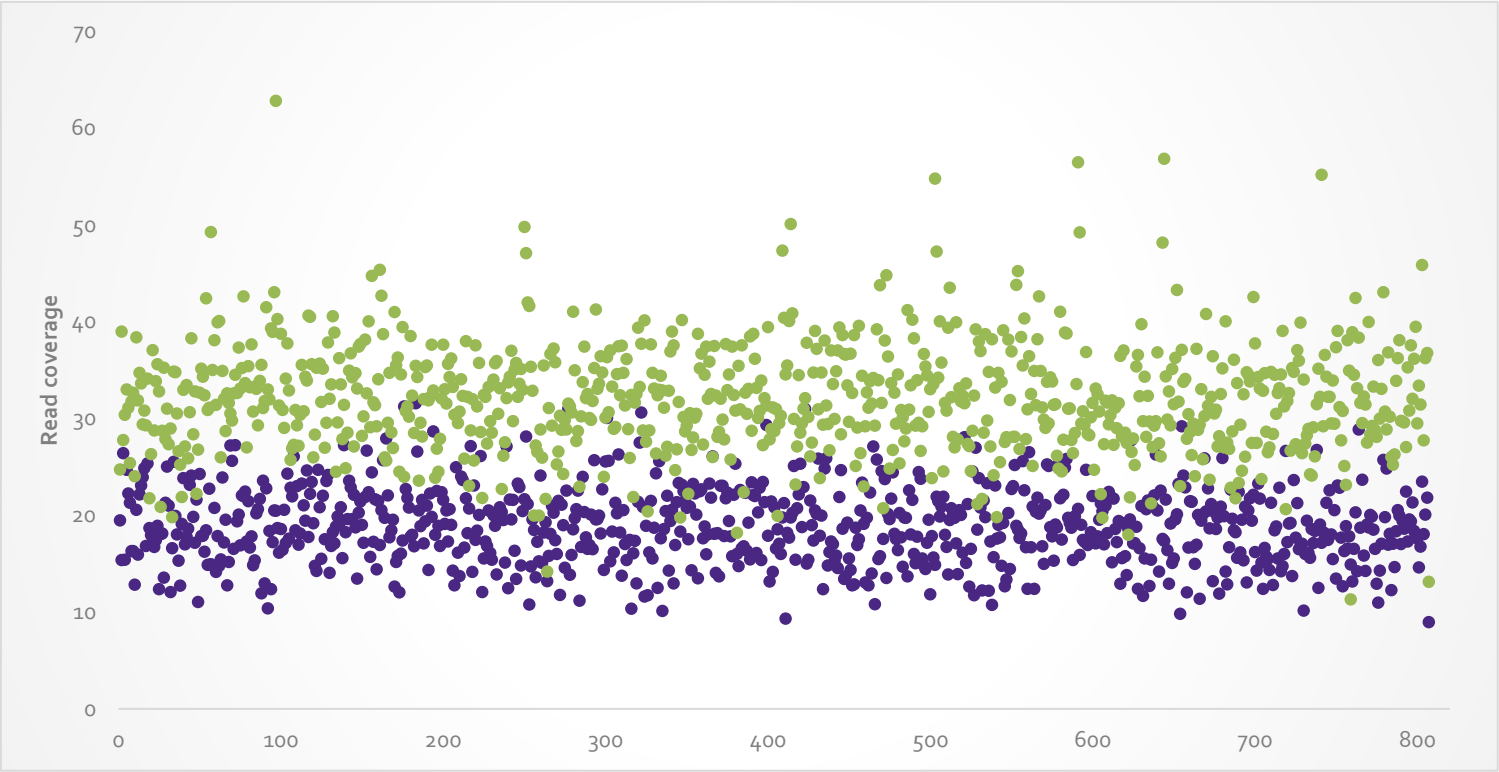

Figure 3. average read coverage across the target loci (x-axis) with (green) and without (purple) Read Until enrichment

When considering all ~800 targeted loci, we observe a highly homogeneous enrichment; with the RU API at least 20X coverage was achieved for almost all loci. This is in strong contrast to the WGS run without RU API, where only ~300 loci showed a coverage of at least 20X.

To our knowledge, this is for the first time that Read Until is applied for target enrichment in a plant crop genome. We show a very homogeneous enrichment distribution over ~800 investigated target loci, mostly in the range of 30 to 40X coverage.

We are currently validating adaptive sequence for single nucleotide variant calling in the enriched sequence regions. As long reads fully span the regions of interest, this application also allows for structural variant analysis among samples. Moreover, we expect to be able to call base-modifications, since the data is generated from native DNA.

As such, we are looking into ways to integrate this tool into DNA innovations for crop breeding research. For instance, the method could be highly instrumental to select loci associated with shelf life, an important trait in melon breeding. Also, Read Until can aid in elucidating genetic variation in R-gene family clusters, which are essential in breeding for enhanced pathogen resistance.

References

1. *Loose M, Malla S, Stout M. Real-time selective sequencing using nanopore technology. Nat Methods. 2016;13:751–4*
2. *T. Payne A, Holmes N, Clarke T, Munro R, Debebe B, Loose MW. Nanopore adaptive sequencing for mixed samples, whole exome capture and targeted panels. bioRxiv. 2020;:2020.02.03.926956.*
3. *Oxford Nanopore Technologies webpage on read until: https://nanoporetech.com/about-us/news/towards-real-time-targeting-enrichment-or-other-sampling-nanopore-sequencing-devices*
4. *Matt Loose on "Read Until" or Adaptive Sequencing on Mendelspod: https://mendelspod.com/podcasts/matt-loose-adaptive-sequencing/*

Contact details
KeyGene
Alexander Wittenberg

+31 (0)317 466 866
alexander.wittenberg@keygene.com