June 2016

Faculty of Engineering Sciences

Avram and Stella Goldstein-Goren

Department of Biotechnology Engineering

אוניברסיטת בן-גוריון בנגב
Ben-Gurion University of the Negev

Project number: 369-16-58

Locating Alleles Responsible to Heat Conditions Tolerance in Recombinant Barley by Lab Methods and Statistical Analysis by ComSeq Algorithm

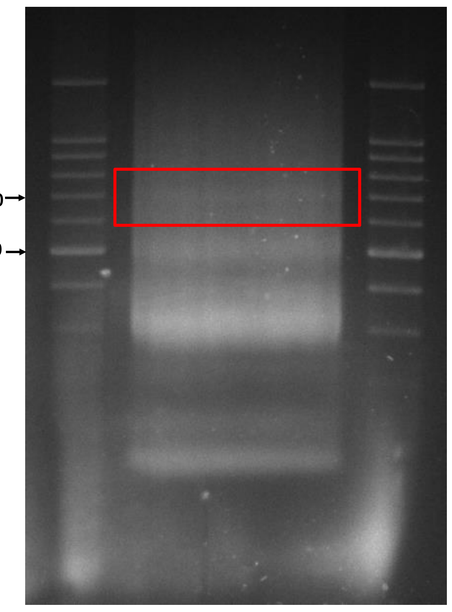
Rotem Sarig

Supervisors: Dr. Eyal Fridman & Prof. David Katoshevski

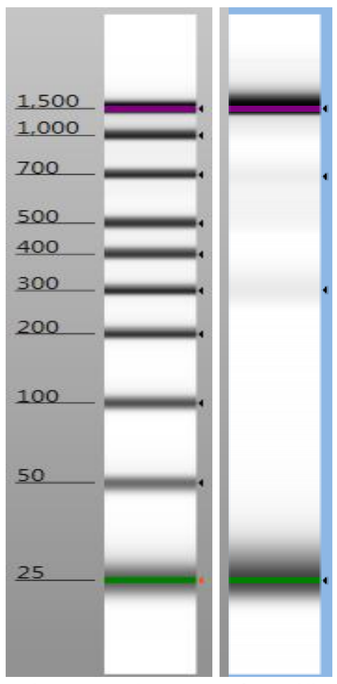
ABSTRACT

The projects purpose is to develop a protocol for locating rare recombination events in recombinant barley plants in large amounts of barley plants using DNA extraction and smart DNA tailoring and analysis of the data through statistical methods and advanced sequencing methods. The project is trying to locate recombination events through two snips (single nucleotide polymorphism) in the Gigenta gene, a gene coding for a protein responsible to the circadian clock and heat tolerance, and I-select, a genetic marker who was located through genetic scanning as a sequence with large significance on the genome, the prospect is to understand weather a recombination occurred in long sequence of the genome through mapping of only two snips that will be inserted to the same amplicon, the process will occur in many plants and analyses will be performed on large quantities of amplicons. According to the suggested protocol[1] nine hundred and sixty barley seeds where chosen due to the fact that their parents had unique combination of alleles from both wild and domesticated barley allowing them to have high heat tolerance, DNA was extracted through grinding and NaOH based extraction and a unique PCR, using amplification of two snips and merging of the two amplicon through another PCR based on complementary tail from non plant source, because of unexplained reason there was not enough DNA in some of the samples ,according to the suggested protocol in order to reduce the number of sequences that will be sent to sequencing a pooling to 48 pools was supposed to be performed using compressed sequencing algorithm – ComSeq[2], in order to distinguish between the pools a barcoding PCR was supposed to be performed, in order to avoid side products magnetic purification with Beckman Coulter beads was supposed to be performed before the PCR, the pools were supposed to poled into one pool , have another magnetic purification and be sent to next generation sequencing in MiSeq machine by Illumina, after decoding of the sequencing barley recombinant plants were supposed to be located.
because of the fact that there was not enough DNA in the first PCR a control experiment was performed in order to check if the compressed sequencing method works on a two snips containing amplicon, high quality DNA from the domesticated (Barke) and wild type (Hid) was the template for the control experiment, 956 samples where imaging regular events and this samples contained 1:1 amplicons containing snips who had Barke:Barke snips or Hid:Hid snips ,another four amplicons where rare recombination events ,two contained Barke:Hid snips and another two contained Hid:Barke snips, the 960 wells where barcoded and pooled according to the suggested protocol, the last pool was purified from gel and not purified by magnetic beads, the purified sample was sent to NGS sequencing, the sequencing showed no results probably due to human error.

RESULTS



100 micro liter of barcoding PCR product loaded on 1.5% agarose gel , DNA was extracted only from the 695 bp bend with NGE gel purification Kit



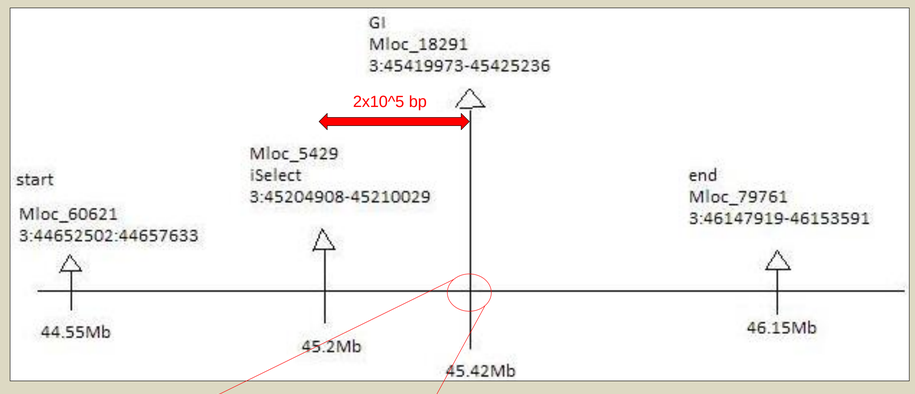
Tape station image of post gel purification of 695 bp DNA bend, concentrations: 302 bp - 0.373 ng/microL, 695 bp 0.243 ng/microL, 695 bp segment has molarity of 0.54 nano-molar

Probably due to wrong defrosting of the MiSeq reagents the NGS didn't create any clusters and the machine showed an error message after three ours of running (full run – 39 hours).

CONCLUSIONS

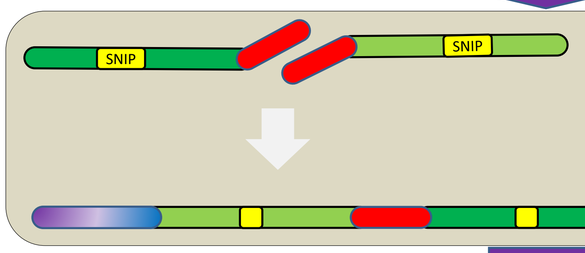
*Creating an amplicon 695 base pairs long that will contain the two snips and Illumina tails is possible.
*Further experiments should use smaller than 200 base pairs amplicons in order to avoid side products and allow ultimate use of the magnetic beads.
*Better PCR on fast NaOH based DNA extraction should be considered ,possible suggestions are use of other TAQ polymerase and use of different DNA extraction protocol.

THE PROTOCOL - METHODOLOGY

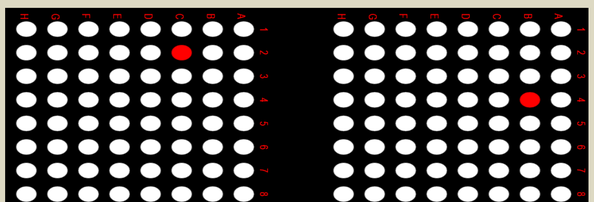


Choosing of an important sequence that we want to locate recombination evets at

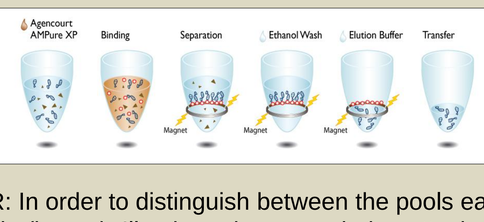
Choosing of a seeds whom are suspected as recombinants, seeding, irrigating and, grinding the plants and extraction of DNA with NaOH based extraction



First PCR on plant DNA amplifies 960 x 2 amplicons with non plant tail (red) and snip (yellow).
Second PCR on two amplicons from each plant with external Illumina tails (purple-blue).

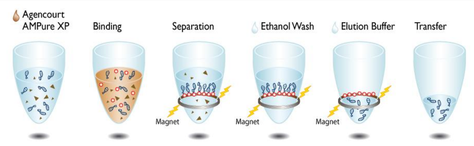


Pooling with iPipet application:
In order to reduce the number of samples that will be sent to sequencing a smart pooling application was developed ,the app is based on ComSeq algorithm , each well is being pipeted to six different pools, locating of recombinants is based on different frequencies of each pool.



In order to avoid side products in future NGS and PCR the pools pass magnetic purification with Becman Coulter magnetic beads.

Barcoding PCR: In order to distinguish between the pools each pool is barcoded with different barcode (brown), Illumina primers are being attached in the same PCR to their complementary tails.
The pools are being pooled to one pool that will be sent to NGS



In order to avoid side products in future NGS final pools pass magnetic purification with Becman Coulter magnetic beads.

Next generation sequencing by Illumina (NGS) – the purified pool (695bp) is inserted to synthesis based sequencing in MySeq machine using nano kit, 1.2 million reads ,250bp from each side, the sample is being inserted together with other samples of other members of the lab in ratio of 1:6 favoring the other lab member.

BIBLIOGRAPHY

1. Nida.H et al, Highly efficient de novo mutant identification in a Sorghum bicolor TILLING population using the ComSeq approach, The Plant Journal, published May 2016, vol. 86, pages: 349-359.

2. Noam Shental, Amnon Amir and Or Zuk, Identification of rare alleles and their carriers using compressed se(que)nsing, Nucleic Acids Research - Advance Access, published August 2010, vol. 10, pages: 1-22.