Assessing myBait Target Capture Sequencing Methodology using Short Read Sequencing for Variant detection in Oat Genomics and Breeding

Khalid Mahmood [1], **Pernille Sarup** [1], Lukas Oertelt [2], Ahmed Jahoor [1,2] and Jihad Orabi [1]

[1]Nordic Seed , Grindsnabevej 25, 8300 Odder, Denmark, [2]Nordic Seed Germany, Kirchhorster str. 16 ,31688 Nienstädt, Germany

Background

Oat (*Avena sativa* L.) is an allohexaploid. This complexity presents challenges in research and breeding. Hybridization-based target capture (myBait) is a cost-effective method for achieving high-depth coverage and identifying sequence variants in large genomes. This gene enrichment method has not been explored in oat research before.

Objective

We use myBait to identify genomic variants in oat and try to verify these variants using Sanger sequencing and high-resolution melting curve analysis (HRM).

Variant calling

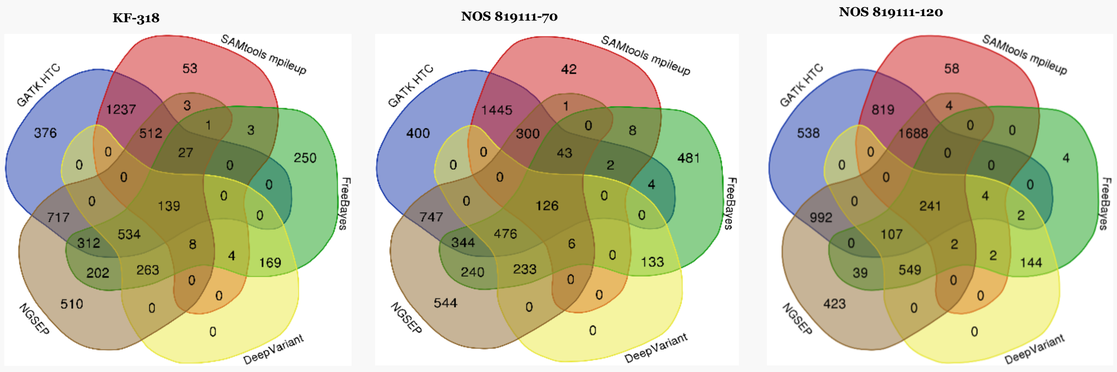

Fig1: Identified variants using different variant callers for each of three representative genotypes. DeepVariant had the highest propotion of variants that was also identified by at least one other variant caller.

Validation of indels

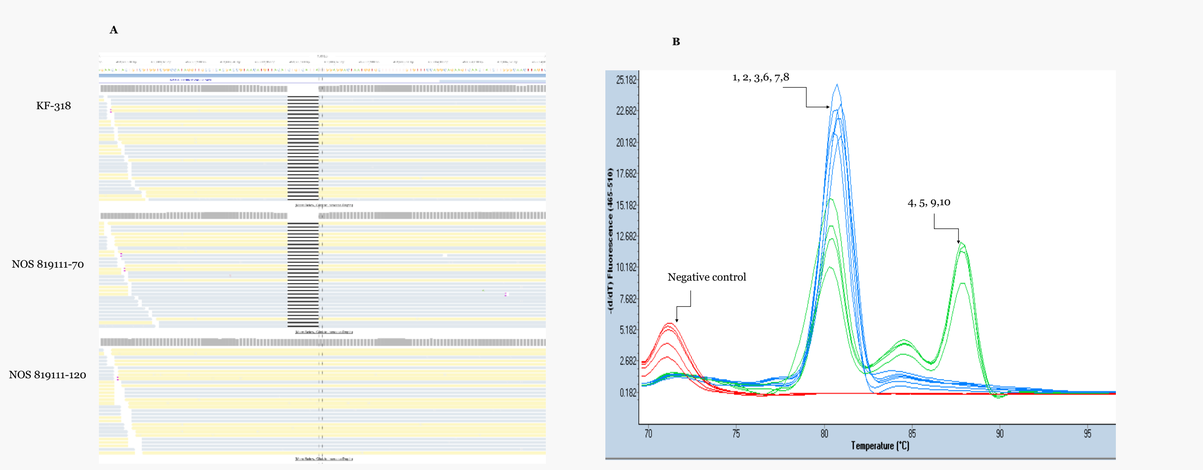

Fig2: A) BWA-MEM read alignment of a selected 9 bp deletion variant in genotypes KF-318 and NOS 819111-70 chromosome 2A at positions 453603957 the deletion is absent from the 8 contrasting genotypes. **B)** HRM curve chromatogram of all 10 oat genotypes. KF-318, NOS 819111-70, and NOS 819111-120 cluster together in HRM. Genotypes are represented as follows: (1) KF-318, (2) NOS 819111-70, (3) NOS 819111-120 (4) Symphony, (5) Delfin, (6) NOS 81920-15, (7) Mathilda, (8) WPB Oskar, (9) NOS 81937-11, (10) NOS 81950-13.

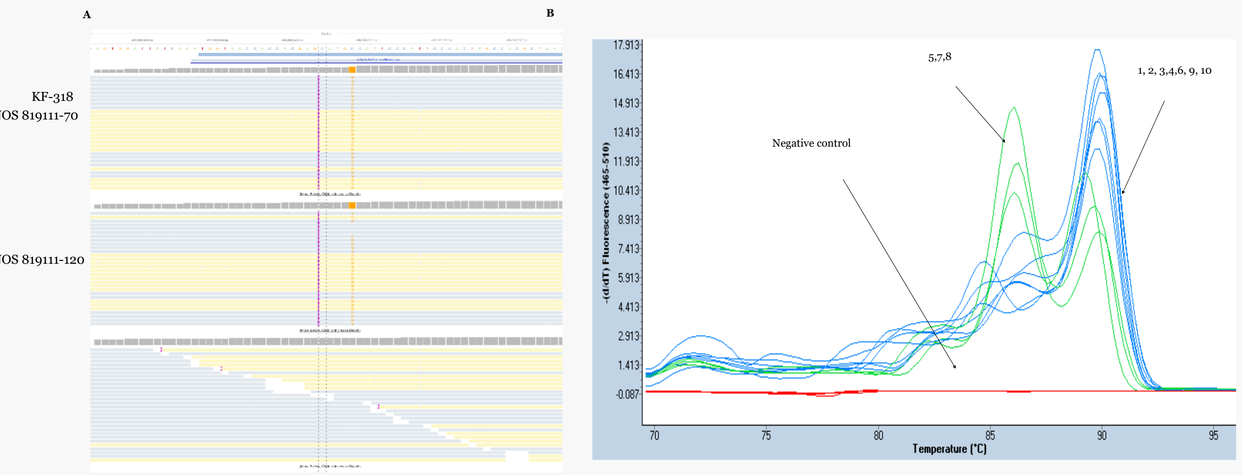

Fig3: A) BWA-MEM read alignment of an insertion variant (2A-456585644), the purple line show the insertion of 3b and orange line show the SNP in KF-318 and NOS 819111-70 but absent in the contrasting 8 genotypes. **B)** HRM chromatogram of all 10 oat genotypes. KF-318, NOS 819111-70, and NOS 819111-120 cluster together in HRM. Genotypes are represented as follows: (1) KF-318, (2) NOS 819111-70, (3) NOS 819111-120 (4) Symphony, (5) Delfin, (6) NOS 81920-15, (7) Mathilda, (8) WPB Oskar, (9) NOS 81937-11, (10) NOS 81950-13.

Methodology

1-Plant Material

Symphony, Delfin, NOS 81920-15, KF-318, Mathilda, WPB_Oskar, NOS 81937-11, NOS 81950-13, NOS 819111-70, and NOS 819111-120.

2-Target Capture Sequencing

A genomic region on chromosome 2A encompassing 98 annotated genes were target captured using short-read sequencing.

3-Mapping and Variant Calling

High quality reads were mapped using three different read aligners, BWA MEM , Bowtie2 , and NGSEP.

Variant calling was done using five different callers, FreeBayes , GATK HaplotypeCaller (HC), SAMtools-mpileup, DeepVariant , and NGSEP.

4-Validation of Targeted Variants

Of the variants identified by all variant callers, we selected two SNP variants, and two indels found homozygotely in two genotypes (KF-318 and NOS 819111-70) but absent in all the other eight genotypes for validation (results of NOS 819111-120 representing the eight genotypes). The SNP variants were validated by sanger sequencing, indels were validated by HRM.

Validation of SNPs

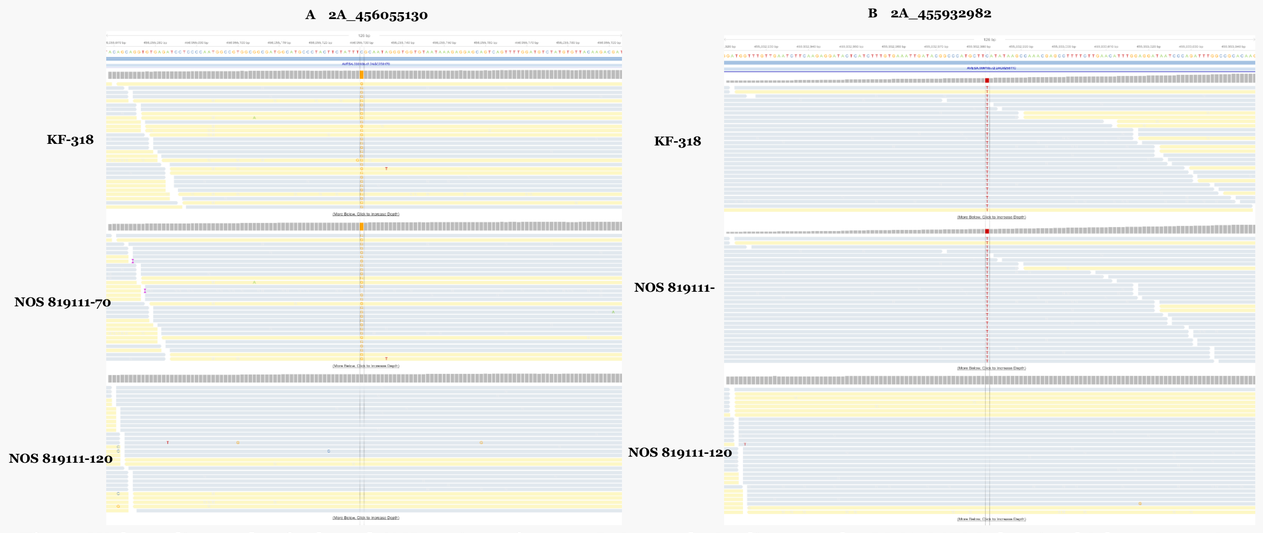

Fig4: High quality reads of selected SNP variants from myBait annalysis alignment by BWA-MEM showing high coverage in KF-318, NOS 819111-70, and NOS 819111-120 flanking the selected SNPs on chromosome 2A at positions 456055130 and 455932982. **A)** 2A_456055130: KF-318 and NOS 819111-70 are homozygote for G. B) 2A_455932982: KF-318 and NOS 819111 -70 are homozygote for T.

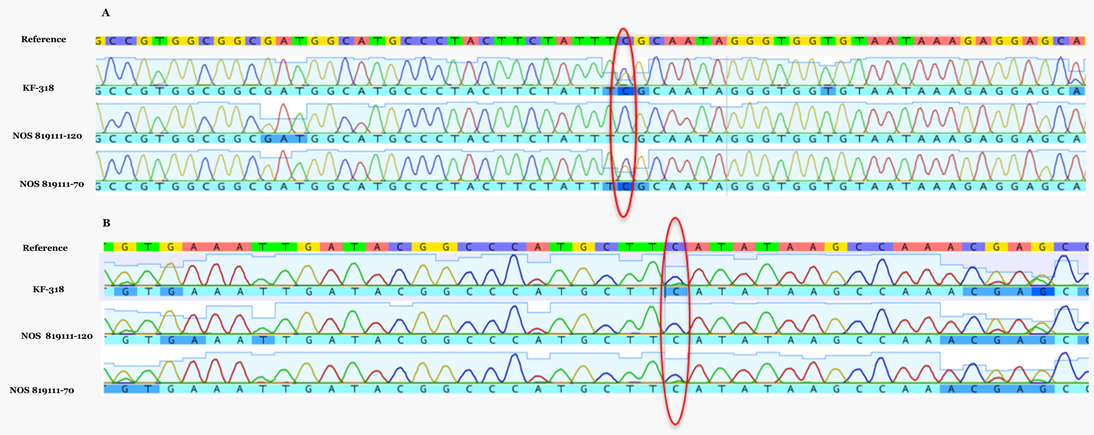

Fig5: Sanger sequencing flanking the myBait selected SNPs (highlighted by a red circle). **A)** 2A_456055130 , alignment clearly show a heterozygous SNP in KF-318 and NOS 819111-70 acording to myBait they were homozygote for the alternative allele, as expected NOS 819111-120 is homozygot for the reference allele. **B)** 2A_455932982. There is no polymorphism.

Conclusion

- Clear SNP and Indel variants was found in the myBait results.
- None of the 4 selected variants could be verifyed.
- We recommend caution when using myBait in oat.

For detailed information: Khalid, et al. (2024) "Assessing myBaits Target Capture Sequencing Methodology Using Short-Read Sequencing for Variant Detection in Oat Genomics and Breeding." Genes 15.6: 700.