Session A: Next generation methods and chances for medicinal and aromatic plants

APL 1: Next-generation sequencing in MAP breeding – efficient and affordable?

Johannes Novak

Institute for Animal Nutrition and Functional Plant Compounds, University of Veterinary Medicine Vienna, Veterinaerplatz 1, 1210 Wien, Austria, e-mail: Johannes.Novak@vetmeduni.ac.at

DOI 10.5073/jka.2016.453.001

Abstract

'Next-generation sequencing' (NGS, syn. 'high-throughput sequencing', 'massively parallel sequencing') are terms comprising different new DNA sequencing technologies allowing sequencing DNA and RAN much more quickly and cheaply than the 'classical' Sanger sequencing (Metzker, 2010). The most widely used NGS technology is Illumina, which has short reading lengths of up to 2x 150 base pairs (bp) (Sanger sequencing: 600 – 1000 bp). However, in contrast to maximum 96 sequences per run in Sanger sequencing, Illumina is able to generate up to 3 billion sequences per run. Although the price per million base pair has decreased from 1.700.-€ (Sanger sequencing) to 0.07.-€ (Illumina sequencing), the price for one NGS run is high. Therefore, it depends on a clever strategy to include NGS affordably into MAP breeding. The most widely used strategy is 'sample barcoding'. The primers used in sequencing are extended by some base pairs unique for each sample ('barcode'). The samples are mixed and sequenced in one NGS run. Afterwards, the bulk of sequences are separated into sequences per sample by sorting them according to their barcode.

Molecular markers are valuable tools in plant breeding used to study genetic relationships, following crosses and finding markers linked to specific phenotypes in order to efficiently select for wanted traits as early as possibly, shortening breeding time ('marker assisted selection', MAS). NGS can be used to develop efficiently molecular markers. Instead of weeks and months necessary to develop around 10 microsatellites by the classical approach of cloning and sequencing, hundreds of microsatellites can be identified in one NGS run today (Takayama et al., 2011). In a similar approach, microsatellites were identified in the plastome of oregano, useful to follow specifically matrilineal relationships (Lukas and Novak, 2013). This fast development of a plethora of molecular markers creates the possibility to increase the number of molecular markers used in breeding. An approach to use NGS in both, SNP marker discovery and genotyping is called 'genotyping by sequencing' (GBS) (Kim et al., 2016).

NGS and the fast development of bioinformatics driven by NGS have revolutionized the fields of genomics and transcriptomics, thus increasing the number of genomes and transcriptomes publicly available as information resources. Posivitely for MAP breeding, genome and transcriptome data are not only limited to model organisms any more, but already extended to some MAP plants as well. Instead of indirect markers for specific traits, NGS offers the possibility to speed up identification of the mutations directly responsible for a specific phenotype. Especially metabolic pathway analysis of medicinal and aromatic plants by plant transcriptome analysis ('RNA-sequencing' RNA-seq) opens the way to functional plant breeding of MAPs (Hao et al., 2012).

Keywords: next generation sequencing, medicinal and aromatic plants, molecular markers, genomics, RNA sequencing

References

HAO, D.C., CHEN, S.L., XIAO, P.G. UND LIU, M. 2016: Application of high-throughput sequencing in medicinal plant transcriptome studies. Drug Development Research **73,** 487-498.

KIM C., GUO H., KONG, W., CHANDNANI R., SHUANG L.S., PATERSON, A.H. 2016: Application of genotyping by sequencing technology to a variety of crop breeding programs. Plant Science **242,** 14-22.

LUKAS, B. und NOVAK, J. 2013: The complete chloroplast genome of *Origanum vulgare* L. (Lamiaceae). Gene **528,** 163-169.

METZKER, M.L. 2010: Sequencing technologies – the next generation. Nature Reviews Genetics **11,** 31-46.

TAKAYAMA, K., LOPEZ P.S., KÖNIG, C., KOHL, G., NOVAK, J. und T.F. STUESSY, 2011. A simple and cost-effective approach for microsatellite isolation in non-model plant spcies using small-scale 454 pyrosequencing. Taxon **60**, 1442–1449.