https://doi.org/10.17221/59/2017-CJGPB

Structural and Functional Apple Genomics and its Application in Breeding

*Sylwia Elzbieta KELLER-PRZYBYŁKOWICZ**

Research Institute of Horticulture, Skierniewice, Poland

**Corresponding author: Sylwia.Keller@inhort.pl*

Abstract

Keller-Przybyłkowicz S.E. (2018): Structural and functional apple genomics and its application in breeding. Czech J. Genet. Plant Breed., 54: 147–153.

The present paper summarized some knowledge of modern technologies, applied in apple (*Malus domestica*) genome studies. New generation sequencing allowed single nucleotide polymorphism (SNP) chip technologies for genotyping, description of functional apple genes, characterization of the evolutionary results in apple genome fragment transition, as well as phylogenetic reconstruction of the genus *Malus*, being the confirmed progeny of *M. sieviersii*. Based on these technologies, newly developed putative markers may give the most important biological data such as age, geographical origins, tissue type determination, and external visible characters. The new generation genotyping platforms, representing very high efficiency, are now successfully applied for random apple genome-wide association (GWA) studies as well pedigree-based analysis and marker-assisted selection (MAS).

Keywords: genotyping; GWA; high-throughput sequencing; fine mapping; MAS

The techniques of DNA sequencing invented by F. Sanger (Sanger & Coulson 1975; Sanger *et al*. 1977) as well as A. Maxam and W. Gilbert (Maxam & Gilbert 1977) became fundamentals of promising tools applied in all research branches such as medicine, biology, agriculture, environment protection, evolution (Linnarsson 2010). From the historical point of view, discovery of the DNA structure by Watson and Crick (1953), DNA Polymerase I by Kornberg (1958) (Lehman *et al*. 1958; Lehman 2008), afterward the polymerase chain reaction (PCR) technique, developed by K. Mullis in 1983 (Stryer *et al*. 2003), allow evaluating different types of variable DNA sequences amplified in a single PCR test. On the basis of multiplication of such sequence variants, cloned in specific artificial libraries/bacterial plasmids, new technologies, so called next-generation sequencing (NGS), revealing the genome sequence order were developed. These technologies allow merging the regions of interest, identifying putative genes (Koboldt *et al*. 2013) and have become a milestone for all 'omics' areas such as genomics, transcriptomics, metabolomics and epigenomics (Gapper *et al*. 2014).

The analysis of samples, collected from different environmental sites and trials, devoted to genomic studies became very important for understanding the molecular background and impact of many biotic and abiotic factors and investigation of functional and biological diversity of all organisms (Ekblom & Galindo 2011; Shokralla *et al*. 2012).

Folding the genome fragments, firstly using BAC libraries sequencing (known as map-based cloning or positional cloning) (Patocchi *et al*. 1999; Han & Korban 2008), finally replaced by whole genome sequencing, allowed to construct densely saturated genetic maps (representing the precise gene localization) and to sequence the genome of many plant crops like apple, strawberry, blackcurrant and so on (Patocchi *et al*. 2009; Velasco *et al*. 2010; Shulaev *et al*. 2011; Russell *et al*. 2014).

The *Malus domestica* Borkh. genome is currently recognized as a reference in the study of the molecular basis of regulation of many important features in

https://doi.org/10.17221/59/2017-CJGPB

woody plants. A wide sequential database may now be a valuable source of molecular markers complementary to the target genes, successfully used in genetic breeding applications. The advantage of such markers is the ability to determine the genotype and its relationships with phenotype, and description of valuable seedling at early developmental stage, shortening the breeding process (Maliepaard *et al*. 1998; Leforestier *et al*. 2015).

The presented review is focused on the new-generation genome sequencing (NGS) technologies and their utilization in apple genomic and transcriptomic studies.

General apple genomics

Description of the whole genome sequence (~7 Mb) of *M. domestica* is of great importance in the current work of breeding and genetic selection of apple species. The sequence of diploid Golden Delicious available on a database for Rosaceae: http://www.rosaceae.org/species/malus/malus_x_domestica/genome_v 2.0 is now a reference for plant genome alignment and functional gene annotation (Velasco *et al*. 2010). It is very useful for an accurate molecular definition of many important agronomical traits, development of high-density molecular markers, and allows accelerating the apple breeding process.

Based on the available sequence data, different DNA variants, characterizing the *Malus* genome, such as single nucleotide polymorphisms (SNPs) and copy number variations (CNVs) can be easily identified (Troggio *et al*. 2012; Bianco *et al*. 2014, 2016).

Single nucleotide polymorphism and apple genotyping

Recent apple genomics studies, supported by high-throughput DNA sequencing technologies, employing genome screening systems, helped to evaluate the DNA allelic variants, represented by single nucleotide insertions and deletions (InDels) (Khan *et al.* 2012; Kumar *et al*. 2012b). Such genomic single sequence rearrangements, representing genome structural variants (SVs), likely may differentiate phenotypic traits (Zhang *et al*. 2014; Chagne *et al*. 2012; Bianco *et al*. 2014, 2016). Application of SNP technology allowed to detect InDels (~5% of genome coverage) spanning all 17 chromosomes of the genomes of apple cultivars Indo and Su Shuai. Additionally, alignment of DNA sequences of both cultivars with the Golden Delicious genome confirmed the existence of numbers of modified genes (24 449 for Indo and 22 153 for Su Shuai). Those data were also a basis for new molecular marker development useful for the creation of pedigree/phylogenetic tree of analysed apple cultivars and predicted their ancestors (Golden Delicious and Golden Reinette × Grimes Golden – for Su Shuai, and White Winter Prarmain – for Indo) (Zang *et al*. 2014).

The techniques were successfully used in research on the apple resistance to Glomerella leaf spot and have made possible to determine the SNPs saturating the quantitative trait locus (QTL) fragments responsible for the trait, which, according to current knowledge, was considered to be regulated by a recessive allele at a single gene locus. In such approaches, the resequencing technique has a great impact on the determination of specific genome fragments that control quantitative traits (Liu *et al*. 2017).

In this venture the identification of different types of genome variation led to develop very precise apple genotyping platforms (GBS; genotyping by sequencing), such as TaqMan (Applied Biosystems, USA), GoldenGate, Infinium, Axion (Illumina; Termo-Fisher Scientific, USA) and AmpSeq (Chagne *et al*. 2012; Khan *et al*. 2012; Kumar *et al*. 2012a, 2012b; Bianco *et al.* 2014; Yang *et al*. 2016), improved by molecular marker multiplexing and automation, which allow to characterize up to one million samples in one assay. On such a platform, called 8K SNP Assay, totally 2 113 120 SNPs, corresponding to over 200 bp of the genome length, were detected. About 27 apple germplasm were aligned to the Golden Delicious reference genome. Finally the 8K array enabled to establish about 5 thousand SNPs, segregating in apple seedling collection and considered to be useful for genetic map saturation. Additionally, the application of this platform for alignment of almost 160 apple accessions allowed validating over 144 newly developed and selected SNPs, useful for the genotype description of other apple cultivars as well as newly developed apple breeding hybrids (Change *et al*. 2012).

Two arrays, 20K and 480K, developed by Bianco *et al.* (2014, 2016), were applied in further apple genome studies. Application of whole-genome genotyping based on SNP technology (20K array) allowed to characterize genomes of X9273 and X9748 apple accessions, derived from Golden Delicious and developed by INRA, France. Sequences available in the array covered about 70% of both double haploid (DH) apple genomes. For these genotypes (explored

https://doi.org/10.17221/59/2017-CJGPB

in parallel with additional 13 apple cultivars) over 2 million nucleotide pair reads, consisting of 60% of unique sequences, were obtained. The analysis of new apple genomes allowed to discover 16 330 other different SNPs, which were incorporated into the final test (FruitBreedomics SNP Chip). Regarding this 'apple sequence model', 257 223 high-resolution poly sequences of apple genome were revealed (Bianco *et al*. 2014).

For validation of 480K, 1320 apple accessions, including 92 progenies of two mapping populations (Golden Delicous (ref) × Renetta Grigia di Torriana, and Fuji × Pinova) were directly tested. This enabled to validate totally 21 463 polymorphic SNPs, characterizing the apple genome (Bianco *et al*. 2016). The Axion 480K, built from the re-sequencing of 63 different apple cultivars is now the most promising and proper for commercialization. Generally, the application of available genotyping arrays allowed to distribute totally over 505 268 SNP sequences in apple genome (globally) (18 019 developed in 20K and 8K SNPs, and 487 249 SNPs developed by 480K), which could be successfully used as a genotype screening molecular marker set for genome-wide association (GWA) studies in apple (Bianco *et al*. 2014, 2016). Such a study gives the valuable information and enlarges the sequence database, applicable in further apple genotype description such as genetic map saturation and QTL localization (Kumar *et al*. 2013; Sun *et al*. 2015).

Copy number variation for apple genome structural analysis

Genomes of eukaryotic organisms exhibit about 4 to 15% of the copy number variations (CNVs) of sequences between individuals, which may influence the gene expression and their structure due to determination of changes in depth coverage of the genome, indicating changes of the CNV (Conrad *et al*. 2009; Munoz-Amatriain *et al*. 2013). The presence of such variable sequences was discovered within *ubiquitin* genes, representing the QTLs linked to the bud phenology of apple. Paralog sequences of scab (*Venturia inequalis*) and powdery mildew (*Podosphaera leucotricha*) resistance genes (*Rvi2*, *Rvi4*, *Rvi11* and *Pl-2*), represented by CNVs, were also confirmed in the region of chromosome 2 between positions 3 507 000 and 4 107 000 bp (Boocock *et al*. 2015).

In regard to apple chromosome studies CNV, consisting of different forms of DNA variability, has been discovered as repetitive elements (67%, 500Mb of total 742.3Mb genome size) and genome-wide duplications (Velasco *et al*. 2010; Chagne *et al*. 2013). It confirms the accidental allopolyploidization of genome originating from evolutionary events. This knowledge is useful for breeders to understand the chromosome fragment dissection in evaluated germplasm, and prediction of inheritance of important genes. The precise sequence analysis allowed to dissect the large segments of chromosomes 3 and 11, 5 and 10, 9 and 17, 13 and 17 representing high collinearity, resulting from the sequence copy number repetition (Velasco *et al.* 2010). Generally about 876 CNVs, with the average size over 16 kb covering about 3.5% of apple genome, were described (Boocock *et al*. 2015).

Microarray – fundamental approach to description of apple gene function

Reverse genetics, developed in the last decade, has an impact on the identification of functional genes controlling important agronomical traits. In this venture the cDNA microarray technique can be successfully applied without sequence description (Ward *et al*. 2012). Several years ago the microarray technology was widely used for monitoring the expression of thousands samples in a single experiment, but also to conduct a genome-wide analysis of variable expression of the transcripts in normal and treated / infected plant samples (Eisen & Brown 1999; Soglio *et al*. 2009).

The technology is based on laborious fabrication of microchips (slides) by probe coverage and relies on the specific cDNA/DNA hybridization of the target samples. Moreover, the microarray results have to be validated with quantitative reverse transcriptase performed by real-time PCR (qRT-PCR) technique, which allows more precise quantification of the gene expression level, following different criteria (Janssen *et al*. 2008; Soglio *et al*. 2009; Baldo *et al*. 2010). Regardless of labour consumption, the technology gives the valuable information on gene expression profiling, as well as on a comparison of gene transcript levels between genotypes (Eisen & Brown 1999; Clarke & Zhu 2006; Soglio *et al*. 2009).

Considering the apple study, hybridization between DNA probes and cDNA from samples of interest allowed to collect about 150 thousand of expressed sequence tags (ESTs) from 43 different cDNA libraries, representing about 34 plant treatments and tissues of

https://doi.org/10.17221/59/2017-CJGPB

Royal Gala. Annotation of the investigated genes was transferred from the *Arabidopsis thaliana* database (Newcomb *et al*. 2006). After analysis of the Royal Gala fruit samples collected at different developmental stages, the differentially expressed genes (validated by qRT-PCR), activated in early (61 putative and annotated genes), middle (8 putative and annotated genes) and ripe (42 putative and annotated genes) stages of apple fruits were identified (Janssen *et al*. 2008). Moreover, the data collected on the basis of cDNA hybridization of transcripts of 1 536 cDNA clones from apple libraries (PRI, The Netherlands) were used for clustering genes as up- or downregulated in regard to fruit development in Fiesta. The validated array dataset has indicated 17 putative genes involved in the primary and secondary metabolism of apple fruit development (Soglio *et al*. 2009).

The cDNA-AFLP strategy which relies on the RNA transcript amplification, performed after its restriction enzyme digestion and their assembling (BLASTn) to apple EST sequence available in NCBI database, was applied to identify several differentially expressed sequence tags in regard to response to *Erwinia amylovora* (FB, fire blight agent) infection of two apple rootstocks (G.41 – FB resistant, M.26 – FB susceptible). The sequences associated with the disease resistance were then verified by quantitative PCR technique which allows to determine over 90 putative ESTs, specifically involved in the interaction between *Malus* and *E. amylovora* (Baldo *et al*. 2010).

The reverse technologies of gene expression profiling were first applied to gain the linkage of gene function and to explore molecular mechanisms of variability of plant physiological processes. They represent the prototype of a high-throughput assay for the identification of differentially expressed genes and produce suitable insights into different plant metabolic pathways (Wullschleger & Difazio 2003; Goulao & Oliveira 2007).

Reading the apple gene sequences – RNA-seq

The transcriptome analysis provided by massive throughput RNA sequencing is now the most effective in the identification of annotated functional genes as well as small RNA molecules introduced from pathogens. It was successfully used for detection of pathogenic RNA transcripts, regulating apple rubbery wood disease resistance in the genome of Lord Lambourne cv. and gives valuable information about interaction between other viroid sequence coverage of plant genome, as well allowed to perform the expression analysis of genes involved in lignin synthesis. Such knowledge gives the huge impact for further breeding program directions (Jakovljevic *et al*. 2017).

RNA-seq technology contributes to the production of various sequence tag collections from different types of plant tissues, environmental conditions, stress treatment and so on. The selected differentially expressed sequences are then compared with available plant genome databases, annotated and applied as functional molecular markers of traits of interests (Xu *et al*. 2012; Bai *et al*. 2014; Jakovljevic *et al*. 2017).

RNA deep-sequencing technology is based on qualitative and quantitative measurements of sequence copy numbers and enables the identification of changes in transcript variants. This can be done by profiling of poly-A RNA template capturing, performed directly in cells (Ozsolak *et al*. 2010) or fractionation of the RNA molecule converted to cDNA (cDNA libraries obtained after adapter attachment to one or both ends of the RNA molecule) (Wang *et al*. 2009; Xu *et al*. 2012).

Generally, transcriptome studies generate the big amount of read sequences (range of about 30–400 pair base length), which must be mapped and annotated by comparison with well described genome sequences. In case of non-model plants, for which no reference sequence of the genome is available, a novel transcript annotation can be assembled to the already published ones (in e.g. GeneBank, UniProt, Ensembl, TrEMBL, SwisProt) as *de novo* sequencing genome approach (Ekblom & Galindo 2011; Lopez-Maestre *et al*. 2016).

For *M. domestica*, functional annotation of genes was performed based on the available sequences of *A. thaliana*, papaya, rice, maize, grape, *Sorghum bicolor, Oryza sativa*, *Vitis vinifera* and *Cucumis sativus*. This allowed to group the sequences of genes involved in biochemical pathways responsible for the synthesis of volatiles, antioxidants, pigments and to predict their chromosome position as well to determine their origin (Velasco *et al*. 2010; Micheletti *et al*. 2011). Combination of *de novo* and Golden Delicious genome assembly of the new apple transcriptomes allowed to enrich: over 53 500 predicted functional genes involved in the fruit development of Golden Delicious (Bai *et al*. 2014), 1080 apple transcripts expressed after open flower inoculation by *E. amylovora* (Kamber *et al*. 2016), group of genes activated under abiotic

https://doi.org/10.17221/59/2017-CJGPB

stress such as MYB and DREB transcription factors (Chagne *et al*. 2007, 2012; Zhao *et al*. 2012), genes regulating fruit ripening (ERF-ethylene transcription factors) (Girard *et al*. 2013), cell cycles, flavour and colour biosynthesis, photosynthesis (Soria Guerra *et al*. 2011) as well as gene families responsible for columnar apple trees (Krost *et al*. 2012, 2013).

CONCLUSIONS

As apple (*Malus* × *domestica* Borkh.) represents self-incompatibility, outbreeding mode of reproduction systems, as well as long generation cycle hampering the breeding process, the new generation technologies, based on genome characterization, can accelerate time-consuming field evaluations.

The present paper summarized some knowledge of modern technologies, applied in apple genome studies, leading to characterization of the genome segments as well as numbers of precisely characterized genes. The new generation genotyping platforms, representing very high efficiency (few gaps between SNPs over 100 kpb were noted), seems to be an ideal for random apple genome-wide association (GWA) studies as well pedigree-based analysis.

Recent description of apple genome and transcriptome as well as chromosome structures are powerful deliverables improving the efficiency of production of new varieties (apple and other species of the family Rosaceae) by bridging the gap between scientific genetic research and its application in breeding. These new technologies proved to be useful as assembly template for transcriptome re-sequencing, new molecular marker development, creation of the targeted sequencing assays such as microarrays/genotyping platforms, as well for gene expression profiling surveys of the plant genetic variation. Finally, they are proper technologies also for gene alternative splicing studies as a phenomenon of phylogenetic, plant speciation and adaptation processes.

References

Bai Y., Dougherty L., Xu K. (2014): Towards an improved apple reference transcriptome using RNA-seq. Molecular Genetics and Genomics, 289: 472–438.

Baldo A., Norelli J.L., Farrell Jr. R.E., Basset C.L., Aldwinckle H.S., Malnoy M. (2010): Identification of genes differentially expressed during interaction of resistance and susceptible apple cultivars (*Malus* × *domestica*) with *Erwinia amylovora*. BMC Plant Biology, 10: 1–10.

Bianco L., Cerstaro A., sergeant D.J., Banchi E., Derdak S., Di Guardi M., Salvi S., Jansen J., Vila R., Gut I., Laurence F., Chagne D., Velasco R., Van de Weg E., Troggio M. (2014): Development and validation of a 20K single nucleotide polymorphism (SNP) whole genome genotyping array for apple (*Malus* × *domestica* Borkh.). PLoS ONE, 9: e110377.

Bianco L., Cestaro A., Linsmith G., Muranty H., Denace C., Theron A., Poncet C., Micheletti D., Kerschbamer E., Di Piero E.A., Langer S., Pindo M., Van de Weg E., Davassi A., Laurens F., Velasco R., Durel C-E., Troggio M. (2016): Development and validation of Axiom Apple 480K SNP genotyping array. The Plant Journal, 86: 62–74.

Boocock J., Chagne D., Marriman T.R., Balck M.A. (2015): The distribution and impact of common copy-number variation in the genome of the domesticated apple, *Malus* × *domestica* Borkh. BMC Genomics, 16: 848.

Chagne D., Carlisle C.M., Blond C., Volz R.K., Withworth C.J., Oraguzie C.N., Crowhurst R.N., Allan A.C., Espley R.V., Hellens R.P., Gardiner S.E. (2007): Mapping a candidate gen (*MdMYB10*) for flesh and foliage colour in apple. BMC Genomics, 8: 212.

Chagne D., Crowhurst R.N., Troggio M., Davey M.W., Gilmore B., Lawley C., Vanderzande S., Hellens R.P., Kumar S., Cestaro A., Velasco R., Main D., Rees J.D., Iezzoni A., Mockler T., Wilhelm L., Van de Weg E., Gardiner S.E., Bassil N., Peace C. (2012): Genome – wide SNP detection, validation, and development of an 8K SNP array for apple. PLoS ONE, 7: e31745.

Chagne D., Lin Wang K., Espley R.V., Volz R.K., How N.M., Rouse S., Brendolise C., Carlisle C.M., Kumar S., De Silva N., Micheletti D., McGhie T., Crowhurst R.N., Storey R.D., Velasco R., Hellens R.P., Gardiner S.E., Allan A.C. (2013): An ancient duplication of apple MYB transcription factors is responsible for novel red fruit flesh phenotypes. Plant Physiology, 161: 225–239.

Clarke J.D., Zhu T. (2006): Microarray analysis of the transcriptome as a stepping stone towards understanding biological systems: practical considerations and perspectives. Plant Journal, 45: 630–650.

Conrad D.F., Pinto D., Redon R., Feuk L., Gokcumen O., Zang Y., Aerts J., Andrews T.D., Barnes C., Campbell P.,Fitzgerald T., Hu M., Ihm C.H., Kristiansson K., MacArthu D.G., MacDonald J.R., Onyiah I., Pang A.W.C, Robson S., Stirrups K., Valsesia A., Walter K., We J.I., Tyler-Smith C., Carter N.P., Lee C., Scherer S.W., Hurles M.E. (2009): Origin and the functional impact of copy number variation in human genome. Nature, 464: 704–712.

Eisen M.B., Brown P.O. (1999): DNA arrays for analysis of gene expression. Methods in Enzymology, 303: 179–205.

https://doi.org/10.17221/59/2017-CJGPB

Ekblom R., Galindo J. (2011): Application of next generation sequencing in molecular ecology of non-model organisms. Heredity, 107: 1–15.

Gapper N.E., Giovannoni J.J., Watkins C.B. (2014): Understanding development and ripening of fruit crops in an 'omics' era. Horticulture Research, 1: 14034.

Girard C.L., Rombaldi C.V., Dal Cero J., Nobile P.M., Laurens F., Bouzayen M. (2013): Genome-wide analysis of the AP2/ERF superfamily in apple and transcriptional evidence of ERF involvement in scab pathogenesis. Scientia Horticulturae, 151: 112–121.

Goulao L.F., Oliveira C.M. (2007): Molecular identification of novel differentially expressed mRNAs up regulated during ripening of apples. Plant Science, 172: 306–318.

Han Y., Korban S.S. (2008): An overview of apple genome through BAC end sequence analysis. Plant Molecular Biology, 67: 581–588.

Jakovljevic V., Otten P., Berwarth C., Jelkman W. (2017): Analysis of the apple rubbery woody disease by next generation sequencing of total RNA. European Journal of Plant Pathology, 148: 637–646.

Janssen J.B., Thodey K., Schaffer R.J., Alba R., Bala-krishnan L., Bishop R., Bowen J.H., Crowhurst R.N., Gleave A.P., Lendger S., McArtney S., Pichler F.B., Snowden K.C., Ward S. (2008): Global gene expression analysis of apple fruit development from the floral bud to ripe fruit. BMC Plant Biology, 8: 16.

Kamber T., Buchmann J.P., Porthier J.F., Smits T.H.M., Wicker T., Duffy B. (2016): Fire blight disease reactome: RNA seq transcriptional profile of apple host plant defense response to *Erwinia amylovora* pathogen infection. Nature: Scientific Reports, 6: 21600.

Khan M.A., Han Y., Zhao Y., Korban S.S. (2012): A high-throughput apple SNP genotyping platform using GoldenGate assay. Gene, 494: 196–201.

Koboldt D.C., Steinberg K.M., Larson D.E., Wilson R.K., Mardis E.R. (2013): The next-generation sequencing revolution and its impact on genomics. Cell, 155: 27–38.

Krost C., Petersen R., Schmidt E.R. (2012): The transcriptomes of columnar type apple trees (*Malus domestica*) – a comparative study. Gene, 498: 223–230.

Krost C., Petersen R., Lokan S., Brauksiepe B., Braun P., Schmidt E.R. (2013): Evaluation of the hormonal state of columnar apple trees (*Malus* × *domestica*) base on high throughput gene expression studies. Plant Molecular Biology, 81: 211–220.

Kumar S., Chagne D., Bink M.C.A., Volz R.K., Whitwirth C., Carlisle C. (2012a): Genomic selection for fruit quality traits in apple (*Malus* × *domestica* Borkh.). PLoS ONE, 7: e36674.

Kumar S., Banks T.W., Cloutier S. (2012b): SNP discovery through next-generation sequencing and its application. International Journal of Plant Genomics, 831460: 1–15.

Kumar S., Garrick D.J., Bink M., Whitworth C., Chagne D., Volz R.K. (2013): Novel genomic approaches unravel genetic architecture of complex traits in apple. BMC Genomics, 14: 393.

Leforestier D., Ravon E., Muranty H., Cornille A., Lemaire C., Giraud T., Durel C.E., Branca A. (2015): Genomic basis of the differences between cider and dessert apple varieties. Evolutionary Applications, 8: 650–661.

Lehman R.I. (2008): Historical perspective: Arthur Kornberg, a giant of 20th century biochemistry. Trends in Biochemical Science, 33: 291–296.

Lehman I.R., Bessman M.J., Simms E.S., Kornberg A. (1958): Enzymatic synthesis of deoxyribonucleic acid. (I). Preparation of substrates and partial purification of an enzyme from *Escherichia coli*. The Journal of Biological Chemistry, 233: 163–170.

Linnarsson S. (2010): Recent advances in DNA sequencing methods-general principles of sample preparation. Experimental Cell Research, 316: 1339–1343.

Liu Y., Lan J., Li Q., Zhang Y., Wang C., Dai H. (2017): Rapid location of Glomerella leaf spot resistance gene locus in apple by whole genome resequencing. Molecular Breeding, 37: 96.

Lopez-Maestre H., Brinza L., Marchet C., Kielbassa J., Bastien S., Boutigny M., Monnin D., Filali A.E, Carareto C.M., Vieira C., Picard F., Kremer N., Vavre F., Sagot M.F, Lacroix V. (2016): SNP calling from RNA-seq data without a reference genome: identification, quantification, differential analysis and impact on the protein sequence. Nucleic Acid Research, 1: 1–13.

Maliepaard C., Aston F.H., van Arkel G., Brown L.M., Chevereau E., Dunemann F., Evans K.M., Gardiner S., Guilford P., van Heusden A.W., Janse J., Laurens F., Lynn J.R., Manganaris A.G., den Nijs A.P.M., Periam N., Rikkerink E., Roche P., Ryder C., Sansavini S., Schmidt H., Tartarini S., Verhaegh J.J., Vrielink-van Ginkel M., King G.J. (1998): Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. Theoretical and Applied Genetics, 97: 60–73.

Maxam A.M., Gilbert W. (1977): A new method for sequencing DNA. Proceedings of the National Academy of Science, 74: 560–564.

Micheletti D., Troggio M., Zharkikh A., Costa F., Malnoy M., Velasco R., Salvi S. (2011): Genetic divesresity of the genus *Malus* and implication of linkage mapping with SNPs. Tree Genetics and Genomes, 7: 857–868.

Munoz-Amatriain M., Eichten S.R., Wicker T., Richmond T.A., Mascher M., Steuernagel B., Scholz U., Ariyadasa R., Spannagl M., Nussbaumer T., Mayer K.F., Taudien S., Platzer M., Jeddeloh J.A., Springer N.M., Muehlbauer G.J., SteinN. (2013): Distribution, functional impact and origin mechanisms of copy number variation in the barley genome. Genome Biology, 14: R58.

Newcomb R.D., Crowhurst R.N., Gleave A.P., Rikkerink E.A.H., Allan A.C., Beuning L.L., Bowen J.H., Gera E.,

https://doi.org/10.17221/59/2017-CJGPB

Jamieson K.R., Janssen B.J., Laing W.A., McCartney S., Nain B., Ross G.S., Snowden K.C., Soulyere E.J.F,. Walton E.F., Yauk Y.-K. (2006): Analyses of expressed sequence tags from apple. Plant Physiology, 141: 147–166.

Ozsolak F., Ting D.T., Wittner S.B., Branningan B.W., Suchismita P., Bardeesy N., Ramaswamy S., Milos P.M., Haber D.A. (2010): Amplification-free digital gene expression profiling from minute cell quantities. Nature Methods, 7: 619–621.

Patocchi A., Vinatzer B.A., Gianfranceschi L., Tartarini S., Zhang H.B., Sansavini S., Gessler C. (1999): Construction of a 550Kb BAC contig spanning the genomic region containing the apple scab resistance gene *Vf*. Molecular and General Genetics, 262: 884–891.

Patocchi A., Fernandez-Fernandez F., Evans K., Gobbin D., Rezzonico F., Boudichevskaia A., Dunemann F.A., Stankiewicz-Kosyl M., Mathis-Jeanneteau F., Durel C.E., Gianfranceschi L., Costa F., Toller C., Cova V., Mott D., Komjanc M., Barbaro E., Kodde L., Rikkerink E., Gessler C., van de Weg E. (2009): Development and test of 21 multiplex PCRs composed of SSR spanning most of the apple genome. Tree Genetics and Genomes, 5: 211–223.

Russell J., Hackett C., Hedley P., Liu H., Milne L., Bayer M., Marshall D., Jorgensen L., Gordon S., Brennan R. (2014): The use of genotyping by sequencing in black currant (*Ribes nigrum*): developing high-resolution linkage maps in species without reference genome sequences. Molecular Breeding, 33: 835–849.

Sanger F., Coulson A.R. (1975): A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. Journal of Molecular Biology, 94: 441–448.

Sanger F., Nicklen S., Coulson A.R. (1977): DNA sequencing with chain terminating inhibitors. Proceedings of the National Academy of Science, 74: 5463–5467.

Shokralla S., Spall J.L., Gibson J.F., Hajibabaei M. (2012): Next-generation sequencing technologies for environmental DNA research. Molecular Ecology, 21: 1794–1805.

Shulaev V., Sargent D.J., Crowhurst R.N., Mockler T.C., Folkerts O., Delcher A.L., *et al*. (2011): The genome of woodland strawberry (*Fragaria vesca*). Nature Genetics, 43: 109–116.

Soglio V., Costa F., Molthoff J.W., Mieke W., Weemen-Hendriks J., Schouten H.J., Gianfranceschi L. (2009): Transcription analysis of apple fruit development using cDNA microarrays. Tree Genetics and Genomes, 5: 685–698.

Soria-Guerra R.E., Rosales-Mendoza S., Gasic K., Wiśniewski M., Band M., Korban S.S. (2011): Gene expression is highly regulated in early developing fruit of apple. Plant Molecular Biology and Reposiotory, 29: 885–897.

Stryer L., Augustyniak J., Michejda J. (2003): Biochemia. Warszawa, Wydawnictwo Naukowe PWN. (in Polish)

Sun R., Chagne Y., Yang F., Wang Y., Li H., Zhao Y., Chen D., Wu T., Zhang X., Han Z. (2015): A dense SNP genetic map constructed using restriction site associated DNA sequencing enables detection of QTLs controlling apple fruit quality. BMC Genomics, 16: 747.

Troggio M., Gleave A., Salvi S., Chagne D., Cestaro A., Kumar S., Crowhurst R.N., Gardiner S. (2012): Apple from genome to breeding. Tree Genetics and Genomes, 8: 509–529.

Velasco R., Zharkikh A., Affourtit J., Dhingra A., Cestaro A., Kalyanaraman A., Fontana P., Bhatnagar S.K., Troggio M., *et al*. (2010): The genome of the domesticated apple (*Malus* × *domestica* Borkh.). Nature Genetics, 42: 833–841.

Wang Z., Gerstein M., Snyder M. (2009): RNA-seq: a revolutionary tool for transcriptomics. Nature Review Genetics, 10: 57–63.

Ward J.A., Ponnala L., Weber C.A. (2012): Strategies for transcriptome analysis in nonmodel plants. American Journal of Botany, 99: 267–276.

Watson J.D., Crick F.H.C. (1953): Genetical implications of the structure of deoxyribonucleic acid. Nature, 171: 964–967.

Wullschleger S.D., Difazio S.P. (2003): Emerging use of gene expression microarrays in plant physiology. Comparative and Functional Genomics, 4: 216–224.

Xu Y., Feng S., Jiao Q., Liu C., Zhang W., Chen W., Chen X. (2012): Comparison of*MdMYB1* sequences and expression of anthocyanin biosynthetic and regulatory genes between *Malus domestica* Borkh. cultivar 'Ralls' and its blushed sport. Euphytica, 185: 157–170.

Yang S., Fresnedo-Ramirez J., Wang M., Cote L., Schweitzer P., Barba P., Takacs E.M., Clark M., Luby J., Manns D.C., Sacks G., Mansfield A.K., Londo J., Fennel A., Gadoury D., Reisch B., Cadle-Davidson L., Sun O. (2016): A next generation marker genotyping platform, (AmpSeq) in heterozygous crops: a case study for marker-assisted selection in grapevine. Horticulture Research, 3: 16002.

Zhang S., Chen W., Xin L., Gao Z., Hou Y., Yu X., Zhang Z., Qu S. (2014): Genomic variants of genes associated with three horticultural traits in apple revealed by genome re-sequencing. Horticulture Research, 1: 14045.

Zhao T., Liang D., Wang P., Liu J., Ma F. (2012): Genome wide analysis and expression profiling of the DREB transcription factor gene family in *Malus* under abiotic stress. Molecular Genetics and Genomics, 287: 423–436.

Received for publication April 21, 2017
Accepted after corrections March 23, 2018
Published online June 13, 2018