Doctoral (PhD) dissertation

Genetic stability and epigenetic changes of *in vitro* cultivated apple (*Malus* × *domestica* Borkh.)

DOI: 10.54598/001410

Andrea Gulyás

Gödöllő

2021

The PhD School

Name: **Plant Science PhD School**

Discipline: **Plant- and Horticultural Sciences**

Head: **Dr. Lajos Helyes**
Professor, DSc.
MATE, Faculty of Agricultural and Environmental Sciences
Institute of Horticulture Technology Research

Supervisor: **Dr. Erzsébet Kiss**
Professor, CSc.
MATE, Faculty of Agricultural and Environmental Sciences Institute of Genetics, Microbiology and Biotechnology

Dr. Judit Dobránszki
Scientific advisor, DSc.
University of Debrecen, Faculty of the Agricultural and Food Science and Environmental Management, Centre for Agricultural Genomics and Biotechnology

……………………………… ………………………………

Dr. Erzsébet Kiss
Approval of the Supervisor

Dr. Judit Dobránszki
Approval of the Supervisor

……………………………….
Dr. Lajos Helyes
Approval of the Head of Doctoral School

Table of Contents

List of Abbreviations ... 5
1. Introduction ... 8
2. Objectives ... 9
3. Literature review ... 10
3.1. Apple ... 10
3.2. Apple domestication ... 11
3.3. Micropropagation ... 12
3.4. Epigenetic changes ... 14
3.4.1. Histone modification ... 15
3.4.2. MicroRNA ... 17
3.4.3. DNA methylation ... 17
3.5. Genetic marker ... 20
3.5.1. Microsatellite markers ... 21
3.6. Sequencing ... 22
3.6. 1. Next-generation sequencing ... 24
4. Materials and methods ... 27
4.1. Plant materials ... 27
4.2. DNA extraction ... 30
4.3. SSR analysis by ALF-Express II ... 30
4.4. Whole-genome Bisulfite Sequencing ... 32
4.4. 1. Sample preparation and Sequencing ... 32
4.4. 2. Bioinformatic analysis of Whole-genome Bisulfite Sequencing ... 33
5. Results and discussion ... 39
5.1. Microsatellite marker based selection ... 39
5.2. Analysis of Whole Genome Sequencing data ... 41
5.2.1. FastQC analysis ... 41
5.2.2. Trim Galore ... 44

5.2.3. Bismark ... 46

5.2.4. SeqMonk ... 48

5.2.5 Omicsbox ... 54

6. Conclusions and recommendations ... 94

7. New scientific results ... 98

8. Summary ... 99

9. Appendices ... 101

9.1. References ... 101

9.2. Supplementary data ... 124

10. ACKNOWLEDGEMENTS ... 126

List of Abbreviations

5-mC: 5-methylcytosine
5-hmC: 5-hydroxymethylcytosine
A: adenine
AAPT: aminoalcoholphosphotransferase
AAT: alcohol acyltransferase
ADH: alcohol dehydrogenase
AFLP: amplified fragment length polymorphism
ASPG: L-asparaginase
ATP: adenosine triphosphate
BA: benzyladenine
BAR: benzyladenine riboside
BIA: benzylisoquinoline alkaloid
C: cytosine
CDP: cytidine diphosphate
CML: calmodulin-like
CMT2: CHROMOMETHYLASE 2
CMT3: CHROMOMETHYLASE 3
dATP: deoxyadenosine triphosphate
dCTP: deoxycytidine triphosphate
ddNTP: dideoxynucleoside triphosphate
DEG: differentially expressed genes
dGTP: deoxyguanosine triphosphate
DMG: differentially methylated genes
DMR: differentially methylated regions
DNA: deoxyribonucleic acid
DNMT: DNA methyltransferase
DNMT1: DNA methyltransferase 1
DNMT3A: methyltransferase 3A
DNMT3B: methyltransferase 3B
DNMT3L: methyltransferase 3L
dNTP: deoxynucleoside triphosphate
DRM: domains rear-ranged methyltransferase
dTTP: deoxythymidine triphosphate
EB: ethidium bromide
EST: expressed Sequence Tag
ETP: ethanolaminephosphotransferase
FMO: flavin monooxygenases
G: guanine
GA_3: gibberellin acid 3
GDPD: glycerophosphodiester phosphodiesterase
GGT: glyoxylate aminotransferase
GO: gene ontology

GPI: glycosylphosphatidyl inositol
GSL: glucan synthase-like
GST: glutathione S-transferase
$HgCl_2$: mercury chloride
HTC: shikimate o-hydroxycinnamoyltransferase
IBA: indole-3-butyric acid
ISSR: inter simple sequence repeat
KEGG: Kyoto Encyclopedia of Genes and Genomes
LFC: log fold change
LOX: lipoxygenase
MA plot: Bland-Altman plot
$MgCl_2$: magnesium chloride
mRNA: messenger RNA
miRNA: microRNA
MPBQ MT: 2-methyl-6-phytyl-1,4-hydroquinone methyltransferase
MS: Murashige and Skoog basal medium
NaCl: Sodium Chloride
NAD-SDH: NAD^+-dependent sorbitol dehydrogenase
$NADP^+$: nicotinamide adenine dinucleotide phosphate oxicated
NADPH: nicotinamide adenine dinucleotide phosphate reduced
NAD-SDH: NAD^+-dependent sorbitol dehydrogenase
NAT: Arylamine N-acetyltransferase
NCBI: National Center for Biotechnology Information
NGS: Next-generation sequencing
NHGRI: National Human Genome Research Institute
NO: nitric oxide
non-TE: non-transposable element
NTPase: nucleoside triphosphatases
PAGE: polyacrylamide gel electrophoresis
PC: phosphatidylcholine
PCR: polymerase chain reaction
PE: paired-end
piRNA: piwi-interacting RNA
PL: pectate lyase
PME: pectin methylesterase
PTM: post-translational modification
Raf: raffinose
RAPD: randomly amplified polymorphic DNA
RNA: ribonucleic acid
Rubisco: Ribulose-l,5-bisphosphate carboxylase/oxygenase
SalR: salutaridine reductase
SDW: sterilized distilled water
SGS: second-generation sequencing
SiR: sulfite reductase
siRNA: small interfering RNA
snoRNA: small nucleolar RNA

SOLiD: Sequencing by Oligonucleotide Ligation and Detection
SSR: simple sequence repeat
SUMO: Small Ubiquitin−like Modifier
T: thymine
TBE: tris-borate-EDTA
TE: transposable element
Tween-20: polyoxyethylene sorbitan monolaurate
WGBS: whole-genome bisulfite sequencing
ZET: ZERZAUST

1. Introduction

Apple (*Malus* × *domestica* Borkh.) is one of the most popular and widely grown fruits around the world. The maturity of cultivated varieties, their shelf life, and the ability to transport their fruit makes it possible for apple to be consumed every month of the year. Apple production is estimated at 87.2 million tonnes (t) worldwide, with Asia (64.6%), Americas (11.2%) and Europe (19.6%) being the biggest contributors (FAOSTAT 2019), output that is 1.6% more than in 2018. Although there is an extremely high number (more than 7,500) of apple varieties (CORTÉS *et al*. 2019), only a small proportion of them is grown in commercial apple production. Apple has been grown for a long time, starting 4,000 years ago.

In Hungary, apple occupies a dominant position in the fruit production sector with 30,974 hectares (ha) in 2019, where the total harvested yield was 498,330 tonnes which are lower than the previous year (31,838 ha and 678,775 t) (KSH 2019). The most popular apple varieties in Hungary are Idared, Florina (Querina), Jonagold, Remo, Golden Delicious, Golden Delicious Reinders, Golden B, Jonathan M 40, Gala and Relinda. Plant biotechnology and micropropagation also provides an opportunity for breeders to produce new breeds or breeding lines (ALDWINCKLE and MALNOY 2009, MOYO *et al*. 2011). The *in vitro* micropropagation helps the production of uniform propagating material regardless of the season, place or climate and the rapid propagation of new varieties, breeding lines and variants gained from breeding.

Epigenetics involves hereditary changes in gene regulation which is not related to the changes of DNA sequence itself. The popularity of epigenetic research has been developing in the last decades and it has been found to play a key role in many genetic processes. However, the question may arise as whether, and if so, what changes in the genetic stock are caused by environmental effects, whether we should expect changes in gene expression due to epigenetic regulation (QUADRANA and COLOT 2016). Epigenetic changes, including methylation of DNA, histone proteins and microRNA may affect DNA function and RNA transcription. Epigenetic changes also play an important role in plant microproduction. Various methylation states can also occur during somatic embryogenesis and regeneration processes, which can result in gene expression changes during plant tissue culture but can also occur during *in vitro* micropropagation and in shoot cultures maintained for gene bank purposes (DUDITS and HESZKY 2000). Understanding methylation changes at the molecular level may contribute to the development of a better plant growth strategy for *in vitro* cultures (KARIM *et al*. 2016).

2. Objectives

Our aim was to investigate:

1. if there is an epigenetic change in *in vitro* tissue culture compared to the mother plant. Does *in vitro* micropropagation cause change in DNA sequence?
2. if there is any DNA-based change during *in vitro* micropropagation that can cause problems in the genetic stability of off-springs?
3. if there is any difference between the epigenomes of 'McIntosh' and 'Húsvéti rozmaring' apple scion cultivars as a function of environmental change?
4. if the epigenetic change is a reversible or irreversible process?

3. Literature review

3.1. Apple

Domesticated apple belongs to the *Rosaceae*, which is the most important crop family in temperate regions (DIRLEWANGER *et al.* 2002) and this family also includes pears (genus *Pyrus*), peaches (*Prunus persica* L.) and cherries (*Prunus avium* L.). However, the *Malus* genus consists of 30-50 species, the exact number cannot be determined as *Malus* species are very diverse and prone to hybridization, polyploidization and apomixis (JANICK *et al.* 1996, FORSLINE *et al.* 2003). Cultivated apples belong to the subfamily *Maloideae* (*Pomoideae*).

Taxonomic Tree:

- Domain: *Eukarya*
  - Kingdom: *Plantae*
    - Phylum: *Magnoliophyta*
      - Class: Magnoliopsida
        - Order: *Rosales*
          - Family: *Rosaceae*
            - Genus: *Malus*
              - Species: *Malus* × *domestica* Borkh.

The domesticated apple (*Malus* × *domestica* Borkh.) is one of the oldest grown temperate fruits. Apple trees are deciduous and tend to be 3 to 12 m in height with a dense, twig-filled crown. The leaves are 3 to 10 cm long, simple, with serrated margins. The flowers, which have five petals, maybe white, pink, and red in colour (IGNATOV and BODISHEVSKAYA 2011). The fruit is a globose pome, much smaller in size for wild species than cultivation. The fruits of cultivated apples contain one or three seeds with five star-like mats (NIKAM *et al.* 2013). The apple pome is a fleshy indehiscent fruit or also known as a false fruit because it comes from a part of the flower other than the ovary (DAR *et al.* 2015).

Apple, which is an extremely versatile crop, is one of the most important fruit-growing plants in the horticultural sector, while ornamental trees and shrubs of the subfamily are also of great value (TERPÓ 1987). Apples can be consumed raw, which is their most common form of consumption since fruits can be stored for a long time under appropriate conditions, but they are also used in significant quantities in the canning, confectionery, and spirit industries. The fruits are held in the dark at 1°C to 4°C in cold storage. But the storage life of the apple depends on the type and humidity (around 85%), too (LARRIGAUDIERE *et al.* 1997, DELONG *et al.* 2004).

Despite its exotic origin, apples are very well adapted to temperate climates and extreme weather conditions (-40°C to 50°C) (FORSLINE *et al*., 2003, IGNATOV and BODISHEVSKAYA, 2011).

The first apple genome sequence comes from a diploid cultivar, 'Golden Delicious', and its genome size was estimated at 742.3 Mbp with 57,386 genes (VELASCO *et al*. 2010). Most of the domesticated apple cultivars are diploid (2n=34) (WAY *et al*. 1991) with a haploid chromosome number of x=17, although, there are also some triploid (e.g., 'Jonagold') and tetraploid cultivas (e.g., 'Yellow Transparent'), as well (JANICK *et al*. 1996, KELLERHALS 2009, CONSIDINE *et al*. 2012).

3.2. Apple domestication

The apple tree may appear to be a native plant, but with exotic origin. Domesticated apple probably arose from a cross between the wild Central Asian apple, *Malus sieversii* (Ledeb.) M. Roem., and the European apple, *Malus sylvestris* (L.) Mill., having originated in Central Asia, nowadays in the region where Kazakhstan is found (HARRIS *et al*. 2002, KELLERHALS 2009, VELASCO *et al*. 2010, CORNILLE *et al*. 2014). The center of highest genetic diversity is in the Tian Shan region of Kazakhstan, near Almaty (JANICK *et al*. 1996). It is difficult to determine when apple was domesticated and whether this process was spontaneous or induced but may have occurred about 4000 years ago in the Near East (ZOHARY *et al*. 2000). Domesticated apples first spread to Europe and North Africa through human migration and trafficking, and then around the world (CORNILLE *et al*. 2012). The art of grafting originates in Mesopotamia (FORSLINE *et al*. 2003), which made it possible to breed apple trees. Grafting extended to Persians, then Greeks, then to the Romans, who perfected orchard management (GLADIEUX *et al*. 2010, HARRIS *et al*. 2002). During the Middle Ages, apples were grown in abbey gardens, throughout the eastern Mediterranean. Pruning, training, and skills of grafting have become very advanced (HANCOCK *et al*. 2008). In the 18th century, apple production became widespread in Germany, Northern, and Eastern Europe. At least 1,200 named varieties were known at this time (HANCOCK *et al*. 2008). At the beginning of the 17th century, the domesticated apple appeared in America after it was introduced by European colonists (JANICK *et al*. 1996, FAZIO *et al*. 2007). Nowadays, apple cultivars are grown in all temperate regions of the world.

3.3. Micropropagation

Breeding apples by the traditional (crossbreeding) method is a very slow and lengthy process, as the species is characterized by a long juvenile phase, so it takes 5-10 years before turning fruitful.

The genome of the apple is highly heterozygous and can therefore only be maintained by vegetative propagation and cloning, such as budding or grafting. Although these traditional propagation methods do not provide disease-free healthy plants, they depend on the season; in addition, they typically result in a low multiplication rate. This would be important in the production of genetically uniform shoots and rootstocks for commercial apple production.

Micropropagation has shown an opportunity for breeders. With the help of *in vitro* micropropagation, it is possible to produce large quantities of progeny plants with the same genetic stock in a shorter time. Unlike the traditional method, which is intermittent, *in vitro* micropropagated plants, can be produced all year round regardless of the weather. The micropropagation method provides an opportunity to produce pathogen-free propagating material. The production of various vegetative organs, tissues, or cells of a plant under sterile and regulated conditions is known as *in vitro* micropropagation (Figure 1). Clones are formed during microproduction from various parts of the plant, such as the bud or meristem, from which the whole plant can be regenerated. Significant results have also been achieved in plant breeding using biotechnological methods. Apple tissue culture began around the 1970s, when publications on *in vitro* apple cultivation first appeared (JONES 1967, ELLIOTT 1972). Tissue culture is the method of preserving or developing tissues *in vitro* in a way that enables differentiation and preservation of structure and/or function (TORRES 2012). The *in vitro* propagation is the aseptic culture of explants of tissues and organs in closed vessels using established culture media in a regulated environment for clonal propagation of plants using tissue, cell, and organ culture methods (SCHAFFER 1990).

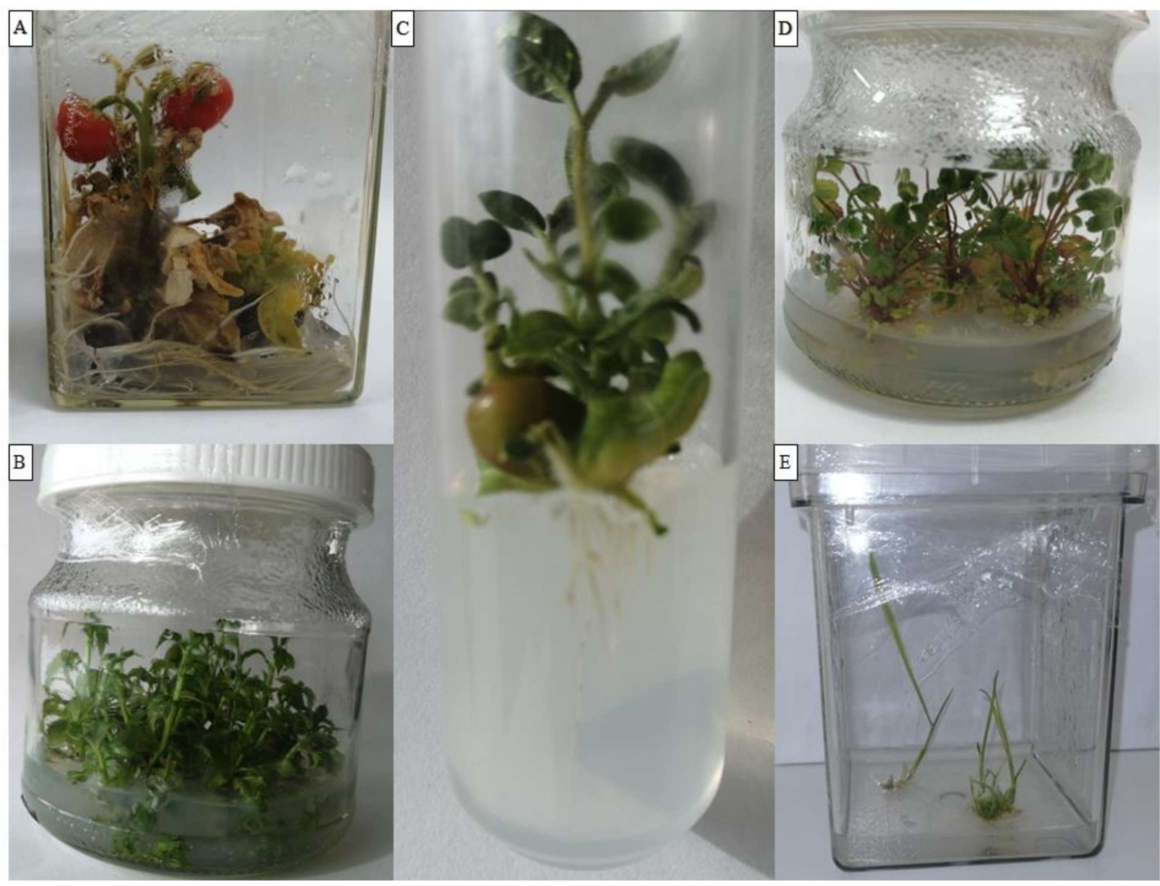


Figure 1. Micropropagated plants (A: tomato (*Solanum lycopersicum* L.), B: apple (*Malus* × *domestica* Borkh.), C: potato (*Solanum tuberosum* L.) D: wild strawberry (*Fragaria vesca* L.) E: common wheat (*Triticum aestivum* L.)) (University of Debrecen, photographer: Andrea Gulyás)

LÉVI and MARÓTI (1977) analysed the induction of callus development on explants isolated from different organs of the apple seedling (cotyledon, foliage, root, etc.) VIRŠCEK-MARN *et al*. (1996) analysed the genetic variation of regenerants from apple leaves *in vitro* using randomly amplified polymorphic DNA (RAPD). 'Golden Delicious Bovey' was studied with 25 primers. Using the RAPD method, 4 mutations were detected in 77 screened regenerants. Although this is only 0.05 mutation events per regenerant, but since RAPD markers are randomly distributed in the genome, so approximately 500 mutations are expected based on the genom size. Thus, the amount of somaclonal variations is large and may be agronomically important. OH *et al*. (2007) investigated the effects of stress on banana (*Musa* × *paradisiaca* L.) during *in vitro* tissue culture, which had an effect on the plant genome. The results of this study support the theory that the labile part of the genome changes during the generation of the somaclonal variation of bananas. Tissue culture thus continues to play an essential part of apple molecular breeding (PEIL *et al*. 2011) and biotechnology.

3.4. Epigenetic changes

Epigenetics was firstly characterized as a lot of interactions between qualities and the encompassing environment, which determines the phenotype or physical characteristics in a living being. WADDINGTON (1942) was the first to mention epigenetics, which interprets the process of evolution through which genotype creates phenotype. The 'epi' of epigenetics is Greek term ("over, upon, around") alludes to the variables happening above or around the more extensive occasions inside the genome. Since the introduction of the term epigenetics, number of studies in the scientific community suggest that gene functions might not be altered by more than just the changes in gene sequence. According to RUSSO *et al*. (1996) epigenetics is mitotically and / or meiotically inherited changes in gene function that cannot be explained by changes in DNA sequence. Epigenetics is defined as the study of changes in gene function that are stably inherited and do not involve changes in DNA sequence (BERGER *et al*. 2009). Epigenetics is explicitly connected with the molecular mechanisms that happen without changing the DNA sequences, and strongly affects regulation of gene expression and the structural dynamics of chromosomes (MURRELL *et al*. 2005). In fact, all cells in a multicellular organism contain the same genetic information, but during devlopment, every cell differentiates into a specific function (phenotype) without changing the DNA sequence. These cells originate after differentiation from a single pool of stem cells and contain the same genetic material but get functional variation from variations in their epigenetic patterns (ALLIS *et al*. 2007). The main types of epigenetic modifications may include:

- DNA methylation, i.e. chemical nucleotide modifications, including methylation of cytosine, the fifth base and thereby producing 5-methyl-cytosine (LISTER e*t al.* 2009a).
- histone modifications, i.e chemical modification of proteins closely related to genomic DNA, especially post-translational modification of histone proteins, as well as associations of mitotically heritable protein-DNA (SCHUETTENGRUBER *et al*. 2007).
- nucleosome positioning,
- small, non-coding RNAs (ribonucleic acid) (BRAIT and SIDRANSKY 2011). Based on the length of the molecule, we distinguish between small and long non-coding RNAs. Long non-coding RNAs are considered to be longer than 200 nucleotides in length (PARALKAR and WEISS 2013). Small non-coding RNAs comprise piwi-interacting RNAs (piRNAs), small interfering RNAs (siRNAs), microRNAs (miRNAs), and small nucleolar RNAs (snoRNAs) (YU and PESTELL 2012).

Of the epigenetic modifications, histone modification and DNA methylation have been most intensively studied by epigenetic researchers to explore their role in regulating gene expression (KAWANABE *et al*. 2016, SHEN *et al*. 2012) (Figure 2).

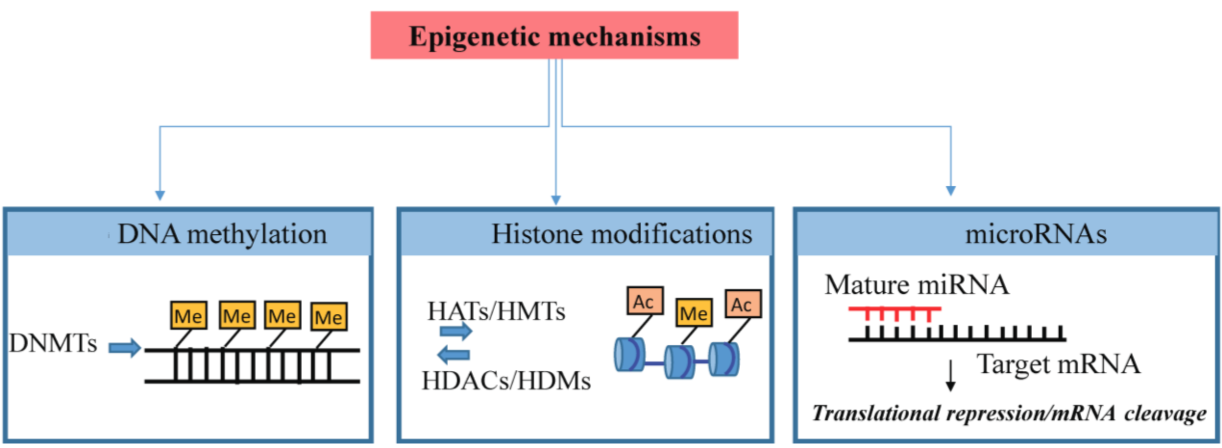


Figure 2. Epigenetic mechanisms. DNA methylation (Me) of cytosine residues results in transcriptional silencing, according to epigenetic processes. Methylation and acetylation (Ac) are the two main types of histone changes that can inhibit or activate gene transcription, respectively. MicroRNAs are small RNAs that can control gene expression by binding to specific mRNA regions and causes mRNA degradation or translation inhibition (CATUOGNO *et al*. 2018).

In 2016, BAI *et al*. performed anthocyanin biosynthesis with a triploid apple variety 'Mutsu' because this type does not contain transcribed *MdMYB1-2* and *MdMYB1-3* alleles. This variety has a green skin color, but red pigmentations appeared on it under a paper bag, suggesting that the *MdMYB1-2/-3* alleles were turned on, which triggered the expression of downstream anthocyanin biosynthesis genes in artificial ultraviolet-B light. Significant differences in DNA methylation and hystonine modifications were observed. The result shows that bag treatment causes epigenetic changes. DACCORD *et al*. (2017) sequenced a doubled-haploid (GDDH13) variety of 'Golden Delicious' and in the experiment, complete *de novo* sequencing was performed using Illumina, PacBio, and BioNano. Their research largely identified a new hyper-repeating retrotransposon sequence and observed DNA methylation patterns.

3.4.1. Histone modification

Histones are conservative proteins in a phylogenetic sense that undergo number of translational modifications while functioning. Large number of post-translational modification (PTMs) types have been identified and mapped, such as lysine acetylation and methylation, phosphorylation, ubiquitination, ADP-ribosylation arginine methylation, citrullination, and SUMOylation (ARNAUDO and GARCIA, 2013). Histones perform an important role in various nuclear processes such as replication, DNA repair, and transcription (KOUZARIDES 2007, BANNISTER and KOUZARIDES 2011). Histone molecules are positively charged proteins with a protruding tail of 15-38 amino acids at the N-terminus ("histone tail") that affects the structure

of the nucleosome. The main types of histones involved in DNA compression are H1, H2A, H2B, H3, and H4. However, there are some histone variants (e.g., H2A.Z) that have their own function. Histone modifications are dynamic and reversible; there are several epigenetic modifying enzymes responsible for the addition or removal of histone modifications (BANNISTER and KOUZARIDES 2011).

Plants contain a number of histone variants and enzymes that modify histones post-translationally and affect gene regulation. The use of chromatin immunoprecipitation and deep sequencing provide insight into the genomic distribution of histone variants and histones with specific post-translational modifications. (BERGER *et al*. 2009, ROUDIER *et al*. 2009). In plants, methylation, and deacetylation of H3K9 and H3K27 lead to gene repression, while acetylation and demethylation of H3K4 and H3K36 lead to gene activation and thus induce gene expression (LAURIA and ROSSI 2011) to change the structure of the package, which either activates the DNA for transcription or makes the structure more condensed so that the transcription machines cannot reach it.

Histone phosphorylation is also linked to the regulation of chromatin structure. Histone phosphorylation has been related to a number of processes including transcription, DNA repair, apoptosis, and chromosome condensation (CHEUNG *et al*. 2000).

Ubiquitin is a small regulatory protein (76 amino acids in length) found in almost all eukaryotic organisms' tissues. Its best known and most important function is to control the recycling of proteins. The carboxylic acid of the terminal glycine from the di-glycine motif in activated ubiquitin forms an amide bond to the epsilon amine of the lysine in the modified protein during ubiquitination, which is an enzymatic operation. Ubiquitination can alter the histones H3, H2B, and, in particular, H2A (WEST and BONNER 1980). Histone ubiquitination can be connected to transcriptionally active DNA regions and may inhibit chromatin condensation (MOORE *et al*. 2002).

Small Ubiquitin−like Modifier (SUMO) proteins are 100 amino acids in length. SUMOylation is a posttranslational modification that plays a role in a number of cellular processes, including nuclear cytosolic transport, transcriptional control, apoptosis, protein stability, stress response, and cell cycle progression (HAY 2005). SUMO proteins are linked to ubiquitin, and SUMOylation is regulated by an enzymatic cascade similar to ubiquitination. Unlike ubiquitin, SUMO does not act as a degradation tag for proteins.

3.4.2. MicroRNA

MicroRNAs were discovered in 1993 in *Caenorhabditis elegans* (a soil-dwelling nematode species) and were thought to be a specialty of nematodes (LEE *et al*. 1993, WIGHTMAN *et al*. 1993). However, they have since been detected in unicellular, multicellular organisms (algae, plants, inferior animals, mammals, etc.) representing different levels of phylogenesis (NAQVI *et al*. 2009). To date, a total of 38,589 mature miRNAs have been discovered thanks to their extensive study and developments in molecular biology techniques (www.mirbase.org). The miRNAs are single-stranded, non-coding small RNAs of about 20-24 nucleotides in length that, in most cases, act through inhibition of messenger RNAs (mRNAs) (translation block and/or mRNA degradation). The biological significance of miRNAs is multifaceted: they are involved in the regulation of almost all intracellular processes, including individual development, cell proliferation, cell differentiation, and apoptosis, and their dysfunctions may play a role in the pathomechanism of various diseases, including tumors and metabolic disorders in eucaryotes (BARTEL 2009, CARTHEW and SONTHEIMER 2009, KROL *et al*. 2010, POTESTÀ *et al*. 2020). Different organisms are characterized by a unique miRNA pattern.

At different stages of development, microRNAs are expressed on a cell-by-cell basis, thus contributing, among other things, to cell type-specific protein expression.

3.4.3. DNA methylation

DNA methylation is known to be an important factor in the regulation of gene expression (BERDASCO and ESTELLER 2010, JONES 2012). DNA methylation is the most widely studied epigenetic pathway (ESTELLER 2008), as it is one of the most important epigenetic mechanisms when the cytosine or adenine base becomes to be methylated in DNA (TURNER *et al*. 2013). FEINBERG and VOGELSTEIN (1983) were the first to describe DNA methylation. DNA methylation is a covalent modification of cytosine by the addition of a methyl group. Methylation of DNA occurs with 5'-carbon cytosine, usually when guanine (CpG or CG) is prevalent. These locations are known as CpG islands. Methylation is defined as either symmetrical or asymmetrical depending on the target sequence. Methylation of CpG and CHG is called symmetric, methylation of CHH is called asymmetric. In plants, DNA methylation occurs in the context of all three sequences; symmetrical CpG and CHG relationships and asymmetric CHH (H = A (adenine), C (cytosine) or T (thymine)) relationships (MIURA *et al*. 2001). Methylation in transposable elements and the promoter region of a gene contributes to silencing, on the other hand, methylation within the gene body causes gene expression (CHAN *et al*. 2005).

CpG site methylation is catalyzed by DNA methyltransferase (DNMT) (Figure 3). DNA methylation is performed by DNMT enzymes that are capable of generating or maintaining the required methylation pattern.

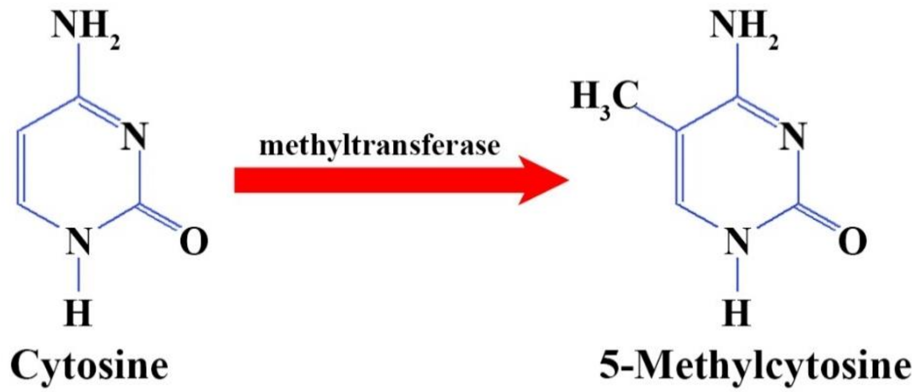


Figure 3. Methylation of cytosine. DNA methyltransferases are enzymes that methylate DNA (DNMTs). DNMTs transfer methyl groups at the 5 -position of cytosine residues (illustrated by Andrea Gulyás).

DNA methyltransferase 1 (DNMT1) methylates hemimethylated DNA, DNMT3A (methyltransferase 3A) and DNMT3B (methyltransferase 3B) generate new DNA methylation signals and are needed for DNA methylation samples during early development (JONES 2012). The DNMT1 enzyme is responsible for preserving the methylation of the newly formed DNA strand during replication (VILLA *et al*. 2004). The DNMT1 maintains methylation patterns and ideally binds hemi-methylated DNA. Methylation marks are thus preserved throughout the division of cells. On the other hand, the *de novo* DNA methyltransferases DNMT3A and DNMT3B establish „new" methylation patterns. DNMT3L (methyltransferase 3L) does not have enzymatic activity and is responsible for regulating the activity of other DNA methyltransferase enzymes (LI *et al*. 2012). DNMT2 is essentially a transfer RNA methyltransferase enzyme, but also has weak DNA methyltransferase activity (GOLL *et al*. 2006). Given the interrelationship between the methylation and siRNA pathways, it is not surprising that LISTER *et al*. (2008) discovered at least one methylated cytosine in 85% of genomic regions with small RNA sequence identity, accounting for 39% of all methylated sites (COKUS *et al*. 2008). The amount of methylated cytosine ($5^{m}C$) in plants varies by up to 30% when methylation occurs (COKUS *et al*. 2008). DNA methyltransferase1 (MET1) is a homologue of the conserved mammalian DNMT1, retains CG methylation in plants (CAO *et al.* 2000). A plant-specific CHROMOMETHYLASE 3 (CMT3) maintains a high level of CHG context in *Arabidopsis thaliana* (L.) Heynh. (thale cress), while CHH and CHG methylation are normally maintained by domains rear-ranged methyltransferases (DRMs) and CMT2 (CHROMOMETHYLASE 2) methyltransferase (ZEMACH *et al*. 2013). In a

broad wide range of plant species, cytosine methylation is commonly observed in transposons, TEs (transposable element), and other repeated sequences (ZEMACH *et al*. 2010). Methylation of DNA can deactivate TEs, preventing them from reactivating when methylated (TIRADO-MAGALLANES *et al*. 2017).

Gene expression is up-regulated or down-regulated by increased methylation of genomic DNA. Down-regulated gene expression allows plants to conserve energy for biotic or abiotic stress. In comparison, the reduction in methylation of resistance-related genes favours the activation of chromatin and the expression of novel genes, which provide long-term or permanent resistance to stress. DNA methylation acts as a marker, and this results in the modification of gene expression by either physically preventing transcription of the DNA or by using that marker as a binding point to recruit proteins which effect the higher order genetic structures (PHILLIPS 2008). DNA methylation is commonly considered to have a more long-term stable effect on the regulation of gene expression relative to other epigenetic mechanisms (JONES and TAKAI 2001).

Cell differentiation has been closely linked to DNA methylation (MEISSNER *et al*. 2008, HUANG and FAN 2010) and, in turn, has provided cells with a heritable identity. Methylation is most commonly observed in the CG sense in plants, as it is in most eukaryotes (LAW and JACOBSEN 2010). More than 99.98% of methylation is present in the CG context in differentiated human fetal fibroblasts (LISTER *et al*. 2009b), while only a small majority (55%) of methylcytosines are found in the CG context in *Arabidopsis* immature floral tissue (LISTER *et al*. 2008). Methylation in plant genomes has several common characteristics. Methylation within the CG context is more frequent than CHG or CHH within gene (COKUS *et al*. 2008, LISTER *et al*. 2008, SEYMOUR *et al*. 2014). When the MET1-3 methyltransferases were knocked out in *A. thaliana*, CHG methylation in euchromatic regions increased significantly, and CHG methylation within gene bodies was enriched, resulting on a similar profile to that of CG in wild-type plants (COKUS *et al*. 2008).

The methylation of CpG, CHG, and CHH in the model *A. thaliana* plant is 24%, 7.7%, and 1.7%, respectively (COKUS *et al*. 2008), while in maize (*Zea mays* L.) these values are 86.4%, 70.9%, and 1.2% respectively (WEST *et al*. 2014). Highly methylated levels (80–100%) in CpG regions, non-methylated and partially methylated levels (20–100%) in CHG regions, and non-methylated and weakly methylated levels (10%) in CHH regions have been observed in both plant species (COKUS *et al*. 2008, LISTER *et al*. 2008). *Chlamydomonas reinhardtii*, the green alga, has the lowest degree of DNA methylation among plants with CpG, CHG, and CHH contexts of 5.4%, 2.6%, and 2.5%, respectively (BARTELS *et al.* 2018). CpG, CHG, and CHH sites are methylated in 58.4%, 31%, and 5.1% of rice (*Oryza sativa* L.) leaves, respectively (NIEDERHUTH *et al.* 2016). The DNA in 19.9%, 30.5% and 49.6% of CpG, CHG and CHH

contexts, respectively is methylated in *Capsicum annuum* L. (pepper) (RAWOOF *et al.* 2019). These values vary significantly from those of other plants but are approximately comparable to those of tomato (*Solanum lycopersicum* L.) (ZHONG *et al.* 2013). In comparison to potato (*Solanum tuberosum* L.) and tomato (WANG *et al.* 2018), soybean (*Glycine max* (L.) Merr.) (AN *et al.* 2017), and field mustard (*Sinapis arvensis* L.). RAWOOF *et al.* (2019) found that *C. annuum* has the highest recorded global cytosine methylation level across different methylation contexts. At CG, CHG, and CHH sites in black cottonwood (*Populus trichocarpa* Torr. & A.Gray ex. Hook.), the average global DNA methylation level was 43.99%, 29.84%, and 11.57%, respectively (LIANG *et al.* 2019). Methylation of CHH occurred primarily in small, highly methylated regions in different *Poaceae* species, whereas it occurred primarily in long, low-methylated regions in soybean (NIEDERHUTH *et al.* 2016). The epigenome is influenced by DNA methylation as well as the silencing of repeated DNA sequences in the genome (BEWICK and SCHMITZ 2017). DU *et al.* (2020) examined 91 bud mutations of the 'Fuji' apple in the study. Using the genetic variation of 'Fuji' as a control, the characteristics of epigenetic variation in different varieties and mutant groups were examined. It was found that the global genomic DNA methylation level of the 91 bud index of 'Fuji' averaged 35.9%. The methylation pattern of the 'Fuji' mutants was largely identical to the methylation pattern of the original 'Fuji'. The methylation variation of CHG in the mutant group showed a large difference in its methylation levels. There was no association between genetic and epigenetic variation in the 'Fuji' mutant classes. DU *et al.* (2020) were detected 16 epigenetic outlier loci.

3.5. Genetic marker

A genetic marker is a phenotypic or genotypic characteristics that can be used as an experimental probe to label an individual, nucleus, chromosome, protein or DNA. A genetic marker can be linked to a genotype (molecular marker) or a phenotype (e.g., a morphological marker). There are two major groups of molecular markers, one is protein and the other is nucleic acid markers. Probably all allelic forms of the genome can be used as genetic markers. They can be used to explore the diversity of plant populations that plant breeders can use in varietal production. The first sense and scientifically sound plant genetic studies were performed by MENDEL (1866) with peas. Mendelian rules may be used to track the inheritance of molecular markers, which are unique fragments of DNA with recognizable chromosomal localization. A marker is usually a DNA segment of unknown function, but may be a gene or a locus on a chromosome that contains different sequences in different genotypes, alleles that are not involved in phenotyping. Their numbers are not infinite in mathematical terms alone, they generally have

no phenotypic or physiological effect (KISS 2005). The existance of DNA markers is not influenced by environmental factors. Properties of ideal DNA markers:

- high degree of polymorphism evenly distributed in the genome;
- codominant inheritance;
- easy availability, does not require prior information about the genome under study;
- easy and fast analysis;
- reproducibility;
- easy interchangeability of data between laboratories.

Most markers do not meet all requirements, only some. The choice of marker depends on the goal and task (KISS 2005).

3.5.1. Microsatellite markers

All eukaryotic nuclear genomes contain microsatellite (LITT and LUTY 1989) or sequentially simple sequence repeat (SSR) (TAUTZ *et al.* 1986) motifs. Long before they were discovered it was known that satellite sequences included tandem repeat DNA regions, so the word "satellite sequence" was coined. Microsatellites range in size from 1 to 6 nucleotides and are mainly found in heterochromatic, intergenic regions of the genome as tandem repeats, but protein coding genes can also contains microsatellites. Before and after the tandem repeat sequences, there are conserved DNA segments for which polymerase chain reaction (PCR) primers can be designed and the microsatellite sequence between them can be amplified. A method for detecting polymorphic microsatellites was first described in human DNA in 1989 (WEBER and MAY 1989) and two years later microsatellite motifs were identified in plant genomes, as well (CONDIT and HUBBELL 1991). The analysis with SSR markers is excellently reproducible, the method is fast and easy to apply. The vast majority of SSR markers are codominantly inherited and, chromosome specific (RÖDER *et al.* 1998). SSR motifs consisting of short tandem repeat sequences are common, evenly distributed in the genome, unaffected by gene interactions, hypervariable, and show Mendelian inheritance. The results can be accurately interpreted between different laboratories and the marking process is easy to handle. They are especially well suited for polymorphism detection, variety recognition, parentage/pedigree analysis and (KISS 1999).

In 1997, GUILFORD and colleagues used first microsatellite markers in apples. In their experiments, they applied 21 different genotypes of apple DNA for their studies. In the course of the research, 14 SSR primers were developed, of which 4 primers resulted in complex patterns, which proved that microsatellite regions can multiply from several loci. Of the SSR primers used, 10 proved to be suitable for apple to identify different genotypes. Even 3 markers were sufficient for the molecular discrimination of the 21 apple cultivars studied (GUILFORD *et al.* 1997).

Subsequently, a number of scientific results were obtained from the analysis of apples with microsatellite markers, such as HOKANSON *et al.*, in 1998, examined a collection of 66 apple samples with 8 SSR markers. In this study the 8 primer pairs unambiguously differentiated all but seven pairs of the accessions in the collection of 66 apple genotypes. In the following years, more than a hundred SSR primers were developed. In 2002 LIEBHARD *et al*. tested 140 SSR primers on 8 apple cultivars. The markers they developed made it possible to identify 17 linkage groups in apples.

GALLI *et al*. (2005) studied 66 commercially available apple cultivars using 6 SSR markers. As a result, 55 polymorphic alleles were detected at the 6 SSR loci. Successful differentiation of all apple genotypes except for somatic mutants was achieved using only four (CH03g07, CH04e03, CH05d11, and CH05e03) SSR markers (GALLI *et al*. 2005). KIRÁLY (2013) set the goal of identifying apple varieties in the Carpathian Basin using SSR markers. In her experiments, 12 primers were used and a total of 72 apple varieties included. Beside genotyping microsatellite results revealed parent-progeny relationshps and triploidy of certain cultivars (KIRÁLY 2013).

The use of molecular marker data has proven to be extremely useful in organizing and managing genetic resource collections (VAN TREUREN *et al*. 2010). In particular, fingerprinting (HOKANSON *et al*. 1998, GALLI *et al*. 2005), genetic mapping (LIEBHARD *et al*. 2002, LIEBHARD *et al*. 2003), marker-assisted breeding (LONGHI *et al*. 2013), and the diversity of domestic and wild species (ZHANG *et al*. 2007, YAN *et al*. 2008).

A number of databases have been set up over the years to store and publish apple microsatellite marker data as HUNGARIAN APPLE MICROSATELLITE / SSR DATABASE (http://gbi.mkk.szie.hu/genetika/Alma%20mikroszatellit/Uj/Fooldal.html) or HiDRAS (https://sites.unimi.it/camelot/hidras/HiDRAS-SSRdb/pages/CompleteSRRtable.php)

3.6. Sequencing

Sequencing is the process of determining the order of the four different nucleotides of the deoxyribonucleic acid molecule, adenine, guanine, cytosine, and thymine. The first DNA revolution took place in the second half of the 1970s, where ALLAN MAXAM and WALTER GILBERT (1977) and FREDERICK SANGER *et al*. (1977) published the sequencing methods they developed. Both techniques greatly increased the availability of DNA sequences.

The essence of the Maxam and Gilbert method is that during the process, the single-stranded DNA sample with radiolabel (gamma-32P ATP (adenosine triphosphate)) is divided into four halves and four destructive reactions are performed. Four chemical treatments were performed on bases A + G, G, C + T, C, such as purification of purines (A + G) with formic acid,

methylation of guanine (G) with dimethyl sulfate, hydrolysis of pyrimidine (C + T) with hydrazine, and inhibition of the reaction of thymine hydrazine with sodium chloride. The modified DNA is removed by hot piperidine, and the resulting fragments are separated by electrophoresis (PAGE, polyacrylamide gel electrophoresis) and the labeled portions are identified by autoradiography (MAXAM and GILBERT 1977).

In Sanger's method, the DNA to be sequenced was denatured into single-stranded DNA and divided into four parts (tubes). DNA polymerase and four different deoxynucleoside triphosphates (dNTP) (dATP; deoxyadenosine triphosphate, dGTP; deoxyguanosine triphosphate, dCTP; deoxycytidine triphosphate, and dTTP; deoxythymidine triphosphate) were added to all four reactions. In all reactions dATP was radioactively labeled. One of the four dideoxynucleoside triphosphates (ddNTP) was added to each reaction. The first reaction for ddATP, ddGTP for the second, ddCTP for the third, and ddTTP for the fourth. DNA polymerase synthesized new DNA along the DNA template by incorporating the appropriate dNTPs and randomly specific ddNTPs. When a ddNTP was incorporated, strand elongation stopped because the ddNTP molecules lack a 3'-OH group inhibiting the binding of additional dNTPs. This results in DNA molecules of different lengths. Since only one form of ddNTP was present in each reaction, the last nucleotide of each molecule is known, allowing the DNA sequence to be reconstructed (SANGER *et al.* 1977).

Of the two techniques, Sanger sequencing became more prevalent because it used fewer toxic chemicals and radioisotopes and therefore became a defining tool in genomics. For 30 years, Sanger sequencing was the most used sequencing technology and was constantly evolving. It was parallelized and automated, resulting in better quality and low cost. Sanger sequencing was used to sequence the first complete human genome (INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM 2004). This form of sequencing continues to be widely used in diagnostics.

The Human Genome Project required a huge amount of material, time, and human effort, so new technology was needed to determine the DNA sequence. For this reason, the National Human Genome Research Institute (NHGRI) has launched a tender aimed at having a full human genome study under 1,000 USD. This marked the beginning of the development of next generation sequencing (NGS) technologies. The essence of NGS methods is that without gel electrophoresis, hundreds of thousands of base sequences are determined simultaneously with direct detection.

3.6. 1. Next-generation sequencing

Technologies for massively parallel sequencing of short DNA fragments are known as Second-generation sequencing (SGS) or Next-generation sequencing (METZKER 2010). NGS incorporates sequencing technologies such as Illumina, SOLiD, Roche 454, or Ion Torrent. All of these modern methods allow the rapid determination of the nucleotide sequence of DNA molecules much more efficiently than the older Sanger sequencing technique. With NGS, whole genomes or transcriptomes can be sequenced in parallel in a matter of hours at relatively low cost. Therefore, NGS methods are also called high-throughput sequencing technology (METZKER 2010, GOODWIN *et al*. 2016) (Table 1).

Table 1. Comparison of Next-generation sequencing platforms based on the detection chemistry, read length, run time, advantages and limitations

Platform	Chemistry	Read length	Run time	Advantages	Limitations
454 Life Sciences (Roche)	Pyro-sequencing	700 bp to 1 kb	23 hrs	Long reads	High homopolymer error rate
HiSeq (Illumina)	Reversible terminator	36– 150 bp	40 hrs-11 days	High-throughput; low error rate; paired-end reads	Short reads; long run time; decreasing read quality towards ends
SOLiD (Life Technologies)	Ligation	50–75 bp	8 days	Low error rate	Short reads; long run time
Ion Proton (Life Technologies)	Proton detection	200 bp	2–4 hrs	Desktop machine; short run time	Short reads; chimeras; homopolymer errors

The first next-generation device appeared in 2005, developed by 454 Life Sciences (MARGULIES *et al*. 2005). The 454 Life Sciences sequencer uses pyrosequencing technology based on detecting the activity of a DNA polymerase enzyme that synthesizes DNA using another light-emitting (luciferase) enzyme. Real-time DNA polymerase activity is detected using a chemiluminescent enzyme. When a nucleotide is incorporated into a DNA strand, pyrophosphate is formed. The amount of this is measured by the flash on the effect of the luciferase enzyme activity. In the process, one of the four possible nucleotides is introduced into the reaction space at a time; if it is complementary to the next nucleotide, it is bound into the strand by DNA polymerase. Unincorporated nucleotides are washed out and another nucleotide is added to the

system. This process is repeated cyclically. Roche 454 is capable of sequencing DNA of approximately 400-600 megabases (MARGULIES *et al*. 2005, MARDIS 2008).

In 2006, Applied Biosystems (now Thermo Fischer Scientific) developed a method based on oligonucleotide ligation and detection (Sequencing by Oligonucleotide Ligation and Detection, SOLiD). The method is based on short reading sequences (35-50 base fragments) and generates a fragment library (MARDIS 2008). Different adapters are ligated to the two ends of the fragments, and then the fibers are attached to magnetic beads to which primers are attached. and these primers hybridize to the adapters. Only one piece of DNA is attached to each bead. The fibers are then amplified by emulsion PCR on the surface of the beads, the template is denatured, and the beads are covalently bound to a sequencing plate. Ligation is performed with oligonucleotides that are octamers and are labeled with four types of fluorophores (PANDEY *et al*. 2008).

Semiconductor technology based sequencing (Ion Proton and Ion Torrent) can be considered as a variant of pyrosequencing that detects protons (hydrogen ions) released during dNTP incorporation (ROTHBERG *et al*. 2011). The templates to be sequenced are prepared on beads by emulsion PCR. All the reagents required for sequencing are present in this water-oil mixture. The template beads are loaded into the proton sensing wells, and then the sequencing reaction produces a proton release proportional to the incorporated nucleotides, which is detected by the built-in pH-sensor, that is by tiny pH meters that are present at high densitiy on the surface of a small semiconductor chip (CHOI *et al*. 2009).

Illumina is one of the most commonly used NGS methods for DNA sequencing (GOODWIN *et al*. 2016). In this process, the adapter of a free end of an already bound fragment interacts with the complementary primer fixed on the plate (PETTERSSON *et al*. 2009). Then, a double-stranded bridge is generated and after denaturation two single-stranded templates are produced, which are able to participate in a next amplification step (Figure 4).

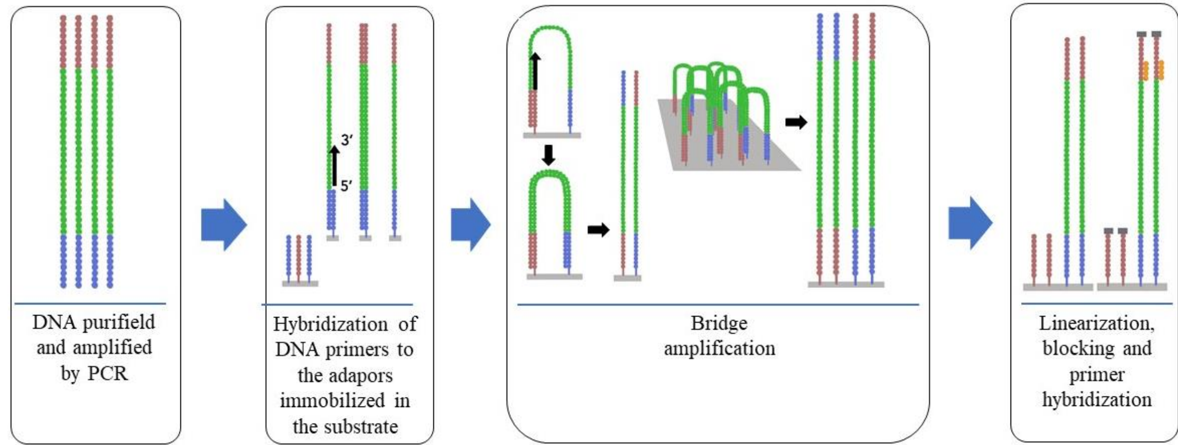


Figure 4. Immobilization and 3' expansion, bridge amplification, linearization, and hybridization are all used to create clusters from single-molecule DNA templates inside the sealed Illumina flow cell (based on https://www.illumina.com)

The sequencing itself is performed with four different fluorescently labeled nucleotides that are added to the reaction simultaneously and flooded simultaneously in the flow cell. If a nucleotide binds complementary to the template strand, a specific fluorescence colour is emitted, and the newly added base can be identified (MARDIS 2008). Before the next cycle, the fluorophores are cleaved and washed from the flow cell. At the end of the process, the machine saves the sequencing data as a simple text string on a FASTQ file (COCK *et al*. 2010). The Illumina platform has been used in many projects due to its high accuracy and throughput. Because genome sequencing projects require high coverage, and the Illumina platform (Figure 5) has the opportunity to do so in addition to short reading sequences. This technique can answer a number of biological questions for researchers in gene expression studies (FAN *et al*. 2013, LAHENS *et al*. 2017), as well as the schematic assembly of genomes and transcripts (ZHAO *et al*. 2017, ZHU *et al*. 2018) and the identification of polymorphic sites (GARRISON and MARTH 2012), and the detection of DNA methylation sites (FALCKENHAYN 2017).

Figure 5. Illumina platforms (NextSeq 500: University of Debrecen, HiSeq 2500: KU Leuven) (photographer: Andrea Gulyás)

4. Materials and methods

4.1. Plant materials

'McIntosh'

'McIntosh' (Figure 6) is the most popular apple cultivar in Eastern Canada and sells well in Europe (FERREE and WARRINGTON 2003). John McIntosh discovered the 'McIntosh' apple on his Dundela farm in Ontario (Upper Canada) in 1810 (FISHER and KITSON 1991) although it only became commercially available in 1870 (FERREE and WARRINGTON, 2003). 'McIntosh' apples are medium-sized. The tree grows between 2 to 8 m in height, its flowers are white and opens into clusters of five-petaled flowers. The fruit, which has a red and green skin, a tart flavour, and tender white flesh that is crisp, ripens in September (FERREE and WARRINGTON 2003). 'McIntosh' is very susceptible to apple scab, moderately resistant to fire blight (JACKSON 2003). The apple scab disease is caused by the fungal spores of *Venturia inaequalis* (HRAZDINA 2003). The spores infect the fruit, producing black spots. Even though infection does not affect the infected apple flesh, the fruit cannot be sold.

Figure 6. 'McIntosh' (A: mother plant, B: *in vitro* and C: acclimatized plants) (photographer: Andrea Gulyás)

'Húsvéti rozmaring'

'Húsvéti rozmaring' (Figure 7) appeared in Hungary in the first half of the 19th century. This apple has several other names like Honti apple, Toklyó apple or Entz rozmaring because some people thought that Ferenc Entz discovered this variety. It was produced in Hungary for a very

long time under commercial conditions, but by the 1990s it was only found in home gardens. The tree expands hemispherically, has a large crown, and it can reach 8 m in height. Every alternate year, the tree yields plenty of fruits, with low fruit setting in intermittent years. This may be a reason why this cultivar is no longer used in the large-scale farm production. The fruit is medium-sized, its skin is thick and waxy, its colour is greenish-yellow and pale red. The flesh is hard, white and the taste is sweet and sour (TOMCSÁNYI *et al*. 1982). It is disease-resistant to fire blight and powdery mildew, moderately resistant to apple scab (PAPP 2017), drought tolerant and undemanding to its environment. 'Húsvéti rozmaring' is easy to transport and store.

Figure 7. 'Húsvéti rozmaring' (A: mother plant, B: *in vitro* and C: acclimatized plants) (photographer: Andrea Gulyás)

In vivo mother plants

We call as mother plants apple samples that come from trees that have been exposed to the weather more than 20 years, grown in the field. The mother samples derived from the Research Institute of Újfehértó of Institutes for Agricultural Research and Educational Farm of University of Debrecen, where the apple trees used in the experiment can be found. 16-year-old *in vitro* shoot cultures sourced from these trees.

In vitro culture

The shoot samples ('McIntosh' and 'Húsvéti rozmaring') were taken from the collection of the University of Debrecen, Institutes for Agricultural Research and Educational Farm, Research Institute of Újfehértó. Leaves were removed from the shoot pieces and surface disinfected, as follows: initial washing in tap water containing 0.05% Tween-20 (polyoxyethylene sorbitan monolaurate) (Sigma-Aldrich, St. Louis, MO, USA). After removal of shoot tips, disinfection was

made with 70% ethanol (Sigma-Aldrich, Saint Louis, MO, USA) for 2 minutes, single wash in sterilized distilled water (SDW), a dip in 0.1% mercuric chloride ($HgCl_2$) (Sigma-Aldrich, St. Louis, MO, USA) solution for 3 minutes and three rinses in SDW. Surface-disinfected shoot tips were then placed on Murashige and Skoog (MS) basal medium (MURASHIGE and SKOOG 1962), which, in addition to MS macro- and microelements, vitamins, contained 3% sucrose (Sigma-Aldrich, St. Louis, MO, USA) and 0.7% agar (plant cell culture tested, SIGMA), 0.3 mg/L IBA (indole-3-butyric acid) (Sigma-Aldrich, St. Louis, MO, USA) and 0.2 mg/L GA_3 (gibberellin acid 3) (Sigma-Aldrich, St. Louis, MO, USA) growth regulators. The citokinin-content of the medium was 1.0 mg/L BA (benzyladenine) (Sigma-Aldrich, St. Louis, MO, USA) for 'McIntosh' and 0.5 mg/L BAR (benzyladenine riboside) for 'Húsvéti Rozmaring'. The pH of the medium was adjusted to 5.8 before autoclaving (Supplementary Table 1).

To maintain the shoots, 400 ml glass jars (OROSházaGLAS, Orosháza, Hungary) (4 shoots/jar; 50-ml MS medium/jar) were used, the shoots were grown in a culture room with a 16-h photoperiod (105 µmol m^{-2} s^{-1}) and 22±2 ºC. Shoots were transferred to fresh medium every 4 weeks.

For the experiments *in vitro* shoot cultures maintained in *in vitro* culture for 16 years were used.

Acclimatization

In vitro shoots with 45–50 mm long were placed vertically into root induction medium (R-72; DOBRÁNSZKI *et al*. 2000) (Supplementary Table 2) and put in a Memmert Din 40050-IP-20 incubator (Memmert, Germany) where cultures were incubated at 26°±2ºC in darkness for one week. After then plants were transferred to root induction medium (R6; DOBRÁNSZKI *et al*. 2000) (Supplementary Table 3) and put back then to the culture room for two weeks. Acclimatization was performed according to the method described by BOLAR *et al*. (1998). Plantlets were placed in Jiffy-7® pellets (Jiffy International, Norway), irrigated with 1.5 ml/L Previcur 607 SL (Bayer, Leverkusen, Germany) 1/10 MS, placed in a VEG-box (Sat-Elit Ltd, Budapest, Hungary) and completely sealed with plastic foil (folpack, Borsod Chemical Combinate, Hungary) to ensure humidity. Plants were grown in an acclimatization room with a 16-h photoperiod and 22±2 ºC at 70–80% relative humidity. After 3 weeks, plants were transferred to a plastic pots (8 cm in diameter) which contained a mixture of perlite and peat (1:3, v/v) and covered with plastic bags. When new leaves appeared, small holes were made in each bag. Acclimatized plants remained in the acclimatization room until they became one year old (Figure 8) (TEIXEIRA DA SILVA *et al*. 2019).

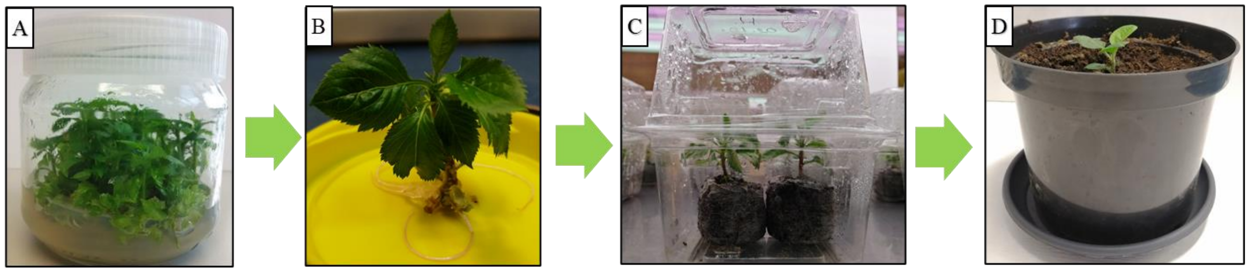


Figure 8. Acclimatization process: A: *in vitro* shoots, B: *in vitro* with roots, C: plantlet, D: acclimatized plant (photographer: Andrea Gulyás)

4.2. DNA extraction

Leaf samples were collected from three different trees from the two apple scions ('Húsvéti rozmaring' and 'McIntosh') to give a total of six plant samples. All samples (*in vitro*, acclimatized, mother plant) were lyophilized (Christ ALPHA 1-2 LD PLUS, Martin Christ Gefriertrocknungsanlagen GmbH, Osterode am Harz, Germany). Genomic DNA was isolated with the NucleoSpin Plant II kit (Macherey-Nagel, Germany) using the manufacturer's instructions. Three biological replicates and three technical replicates for the *in vitro* shoots, *in vivo* mother plants and one-year acclimatized plants were pooled (three replicates) for DNA extraction. The quality of the isolated DNA was checked by gel electrophoresis (Cleaver Scientific Ltd., Warwickshire, UK). Samples were also checked with a NanoDrop ND-1000 microcapillary spectrophotometer (NanoDrop Technologies/Thermo Scientific, Wilmington, DE, USA).

4.3. SSR analysis by ALF-Express II

Polymerase chain reactions with the extracted DNA were performed in an iCycler Thermal Cycler (Bio-Rad Laboratories, Hercules, CA, USA). Before the reaction, the DNA stock solutions were diluted to a concentration of 20 ng/µl in each case. The volume of the PCR mixture was 25 µl, which contained the following components: 2.5 µl of 10X DreamTaq Buffer (Fermentas, Thermo Fisher Scientific, CA, US) (containing 20 mM magnesium chloride ($MgCl_2$)), 0.5 µl of dNTP (10 mM) (Thermo Fisher Scientific, CA, US), 0.5 µl (10 µM) forward and reverse primer (Integrated DNA Technologies, Coralville, IA), 0.2 µl 5 U/µl DreamTaq DNA Polymerase (Fermentas, Thermo Fisher Scientific, CA, US), 2 µl template DNA (15 ng/µl) and 18.8 µl MQ water (high purity water Milli-Q water system, UK). The reaction mixture was prepared on ice. The primers used in the SSR (LIEBHARD *et al.* 2002) analysis are shown in Table 2. Table 3 lists the optimized PCR programs used for the SSR primer sets.

Table 2. Apple microsatellite primers (LIEBHARD *et al*. 2002)

Primer name	Forward (5'-3')	Reverse (5'-3')
CH01f02	ACCACATTAGAGCAGTTGAGG	CTGGTTTGTTTTCCTCCAGC
CH01h01	GAAAGACTTGCAGTGGGAGC	GGAGTGGGTTTGAGAAGGTT
CH01h02	AGAGCTTCGAGCTTCGTTTG	ATCTTTTGGTGCTCCCACAC
CH02c02	CTTCAAGTTCAGCATCAAGACAA	TAGGGCACACTTGCTGGTC
CH02c06	TGACGAAATCCACTAATGCA	GATTGCGCGCTTTAACAT
CH02c09	TTATGTACCAACTTTGCTAACCTC	AGAAGCAGCAGAGGAGGATG
CH02c11	TGAAGGCAATCACTCTGTGC	TTCCGAGAATCCTCTTCGAC
CH02d08	TCCAAAATGGCGTACCTCTC	GCAGACACTCACTATCTCTC
CH03a02	TTGTGGACGTTCTGTGTTGG	CAAGTTCAACAGCTCAAGATGA
CH03g07	AATAAGCATTCAAAGCAATCCG	TTTTTCCAAATCGAGTTTCGTT
CH04e03	TTGAAGATGTTTGGCTGTGC	TGCATGTCTGTCTCCTCCAT
CH04e05	AGGCTAACAGAAATGTGGTTTG	ATGGCTCCTATTGCCATCAT
CH04g10	CAAAGATGTGGTGTGAAGAGGA	GGAGGCAAAAAGAGTGAACCT
CH05c02	TTAAACTGTCACCAAATCCACA	GCGAAGCTTTAGAGAGACATCC
CH05c04	CCTTCGTTATCTTCCTTGCATT	GAGCTTAAGAATAAGAGAAGGGG
CH05d11	CACAACCTGATATCCGGGAC	GAGAAGGTCGTACATTCCTCAA
CH05e03	CGAATATTTTCACTCTGACTGGG	CAAGTTGTTGTACTGCTCCGAC

Table 3. Polymerase chain reactions (PCR)

	Temperature	Time
Initial denaturation	94°C	2 min
Stage 1: 10 cycles of		
DNA denaturation	94°C	30 sec
Annealing *	65°C	30 sec
Elongation	72°C	1 min
Stage 2: 24 cycles of		
DNA denaturation	94°C	30 sec
Annealing	56°C	30 sec
Elongation	72°C	1 min
Final elongation	72°C	5 min

* The annealing temperature is decreased by 1.0 °C every cycle until a specified annealing temperature is reached.

PCR products were separated on a 1.2% agarose (Serva Electrophoresis GmbH, Heidelberg, Germany) gel in 0.5x TBE buffer containing 5 µl (1 mg/ml) of ethidium bromide (EB, Biotium, Inc., Hayword, CA, USA). The gel was run in a Cleaver horizontal gel electrophoresis system (Cleaver Scientific Ltd, United Kingdom) in 0.5x TBE (Tris-borate-EDTA) buffer. The

results were documented using Alpha Innotech MultiImage^{TM} Light Cabinet (Alpha. Innotech Corp., San Leandro, CA, USA). Under UV light (313 nm), the sizes of DNA fragments were estimated with a 50-bp DNA ladder marker (Thermo Fisher Scientific, CA, US).

To determine the length of microsatellite alleles, an ALF-Express II (Amersham Pharmacia Biotech, Uppsala, Sweden) instrument was used, which automatically detects fluorescently labelled DNA fragments during polyacrylamide gel electrophoresis. Gel electrophoresis is vertical, fluorescently labelled fragments migrate downward toward the cathode due to their negative charge. The temperature of the gel is maintained by the flow of water (55°C) through the thermoplate. The thickness of the gel (0.3 mm) is determined by the thickness of the spacer between the two sheets of glass forming the cartridge. The 8% PAGE gel (ReproGel^{TM} High Resolution, GE Healthcare Bio-Sciences AB, Uppsal, Sweden) is a denaturing gel. Following the PCR described above, the PCR products were denatured) at 95°C was denatured for 5 minutes and placed on ice. Samples can be loaded in 40 pockets on top of the gel. 4 µl of external standard was pipetted in the first, twentieth, and fortieth pockets of the gel, and then 8 µl of the mixture of samples and internal standards was loaded in pockets 2-19 and 21-39. The standards consisted of 70 bp, 95 bp, 150 bp, 275 bp, 300 bp fragments. To determine the exact size of the alleles, forward primers were labelled with Cy5 fluorescent dye. During the electrophoresis, the laser beam passes through the gel and excites the fluorescently labelled DNA fragments. The emitted light is detected by a photodetector. Detector signals are collected and transmitted to the computer for processing. The signals in the form of vertices can be followed continuously during the run. By using the ALFwin Fragment Analyzer 1.00 software, the results can be evaluated, and the size of the fragments quantified.

4.4. Whole-genome Bisulfite Sequencing

4.4. 1. Sample preparation and Sequencing

Whole-genome Bisulfite Sequencing method was applied by using 100 ng of genomic DNA (*in vitro*, acclimatized and mother plant), to determine the status of cytosine methylation with the Pico Methyl-Seq Library Prep Kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's instructions. The libraries were quality checked (QC) on Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) to determine exact size and concentration. Whole-genome bisulfite sequencing (WGBS) was performed on Illumina HiSeq 2500 (Illumina, San Diego, CA, USA). WGBS sequencing provides a single base resolution of 5-mC (5-methylcytosine) in the genome, allowing profiling of the whole genome's DNA methylation. Library preparation, bisulfite conversion, paired-end sequencing on Illumina platforms, and primary data processing are all part of the study.

Sodium bisulfite conversion of genomic DNA to detect unmethylated and methylated cytosines is a standard for DNA methylation analysis. Sodium bisulfite conversion provides single nucleotide resolution of 5-mC of the genome. It involves the deamination of unmodified cytosines to uracil, which leaving the modified bases 5-mC and 5-hmC (5-hydroxymethylcytosine). The uracils are amplified in PCR reaction as thymines, whereas 5-mC or 5-hmC residues get amplified as cytosines. The comparison of sequence information between the reference genome and bisulfite-treated DNA can provide information about cytosine methylation patterns.

Conventional methods for NGS library preparation for WGBS typically involve fragmenting genomic DNA, end repairing, adapterization, bisulfite conversion, followed by limited amplification with PCR.

4.4. 2. Bioinformatic analysis of Whole-genome Bisulfite Sequencing

A number of bioinformatics programs provide opportunities to evaluate whole genome sequencing data. It is up to the analyst to decide which bioinformatic programs are best for him or her to evaluate his or her experiment. There are several options to choose from for a process, such as Trim Galore or Trimmomatic. Both scripts are used to cut Illumina (FASTQ) data as well as remove adapters (for Trim Galor this is done by Cutadapt). The following bioinformatics programs were used to evaluate the whole genome sequencing of apples (Figure 9).

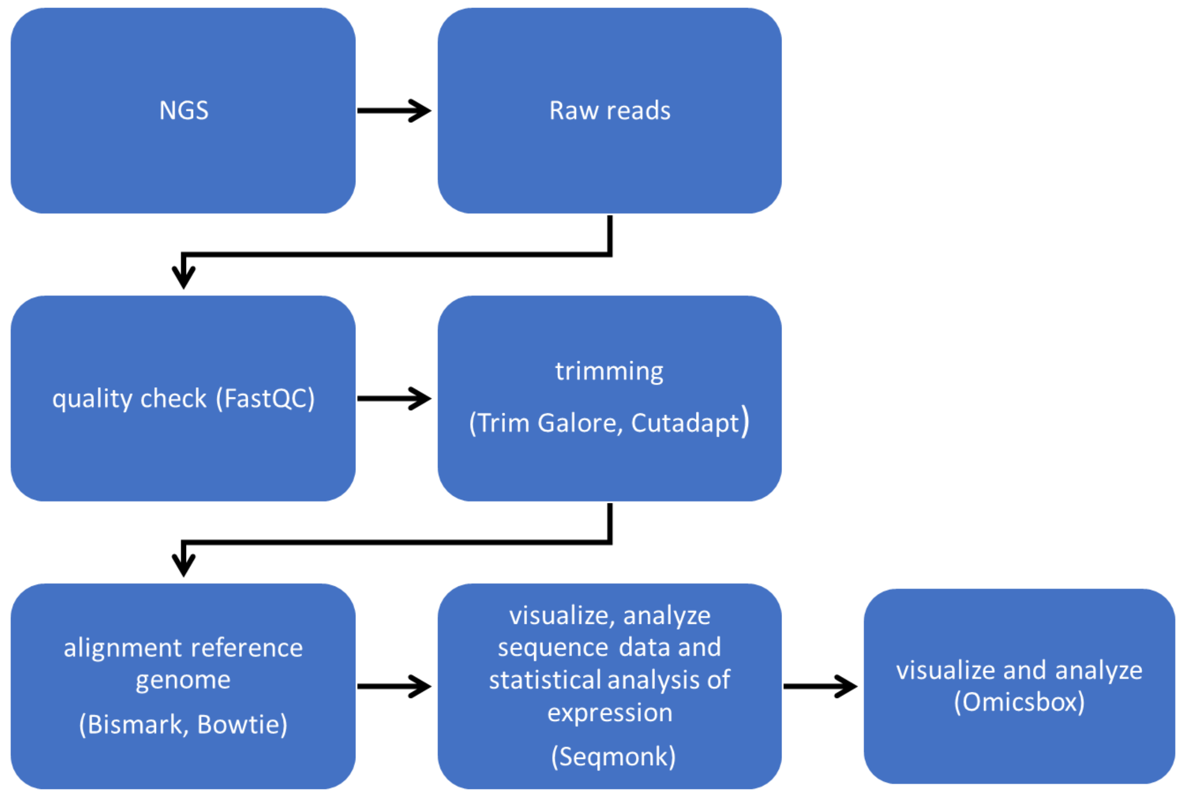


Figure 9. Bioinformatic pipeline of Whole-genome Bisulfite Sequencing (illustrated by Andrea Gulyás)

Sequence data (nucleotide sequences) generated by Illumina HiSeq 2500 are transformed into ASCII text format as a FASTQ file. Raw reads were analysed using the FastQC program (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The raw read files and comparison files are used to launch the data analysis process. FastQC v0.11.5 (https://github.com/s-andrews/FastQC) provides the ability to perform quality control testing on raw high-performance sequence data. The analysis performed by the tool ensures that the data is qualitatively correct and free of problems or distortions. With the help of the program, we can check the:

- Basic Statistics;
- Per Base Sequence Quality;
- Per Sequence Quality Scores;
- Per Base Sequence Content;
- Per Sequence GC content;
- Per Base N Content;
- Sequence Length Distribution;
- Sequence Duplication Levels;
- Overrepresented Sequences;
- Adapter Content;
- Kmer Content.

For basic statistics, the table contains important information such as sequence size, how long the sequence is, and %GC (Figure 10). This percentage may refer to a particular segment of DNA or the percentage of GC content in the entire genome.

Basic Statistics

Measure	Value
Filename	C_ACAGTG_L001_R1_001_val_1.fq.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	91867470
Sequences flagged as poor quality	0
Sequence length	20-114
%GC	27

Figure 10. Basic statistics table of the FastQC program

One of the most important statistics is the Per base sequence quality report (Figure 11). The mean and distribution of content scores (Y-axis) for each location in a population of thousands of reads are depicted in this graph (X-asis). This graph shows the distribution of quality scores across all reads at each point in the read. This plot will tell us if there were any issues during the sequencing process and if we need to contact the sequencing facility.

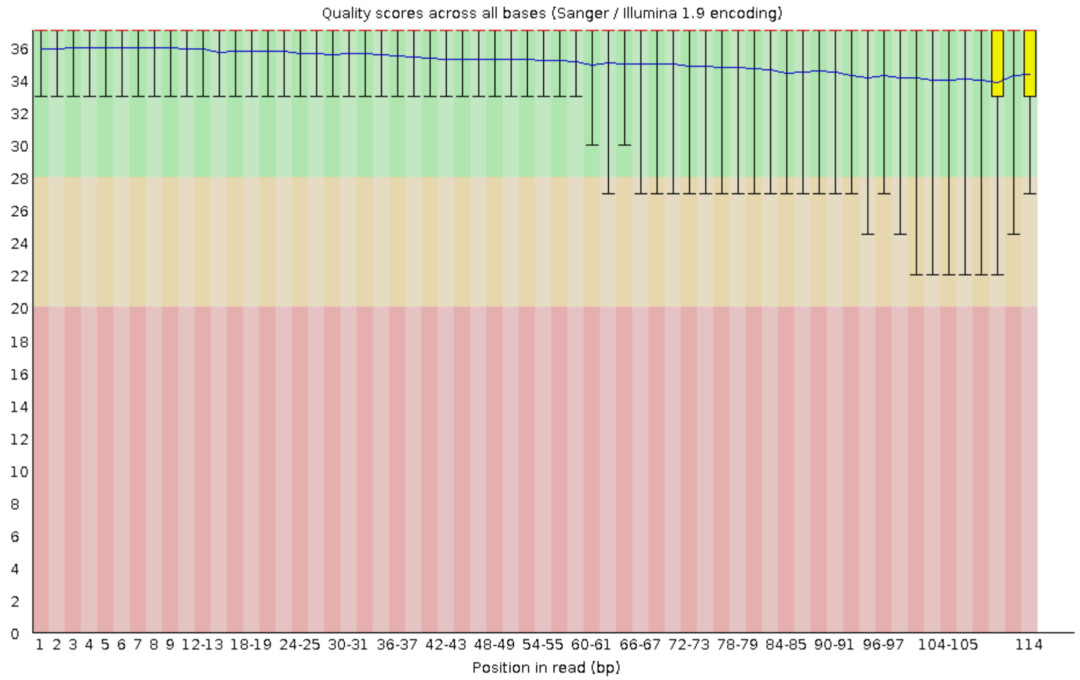


Figure 11. Per base sequence quality of reverse reads before trimming (very good quality calls (green), fair quality calls (orange), and low quality calls (red))

The quality scores are shown on the graph's y-axis. The better the base call, the higher the score. The graph's context divides the y axis into three categories: very good quality calls (green), fair quality calls (orange), and low quality calls (red). Since the quality of calls degrades as the read progresses on most platforms, it's normal to see base calls dropping into the orange region near the end of a read.

A QC analysis can reveal issues with the sequencer or with the starting library content. Trimming base pairs from the raw read will help to overcome certain anomalies. This is supported by the pipeline, which includes a script that can trim a specified number of base pairs. The pipeline is set not to trim any bases from the raw nucleotide sequence at the start. Based on the QC outcome, this option can be updated in the Makefile.

Trim Galore v0.5.0 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) is a script to automate quality and adapter trimming. Trim Galore determines the location of the sequence cleavage where the partial amount is minimal. Trim Galore uses Cutadapt (MARTIN 2011) and FastQC for quality and adapter cutting for FASTQ files, used to trim the raw reads. At

the 3' end, it trims 1 bp from each read, trims low-quality bases (Phred score 20) at the 3' end, removes all adapter sequences at the 3' end, and finally removes read-pairs that are less than 20 bp long. Phred 20 indicates that 1/100 can be defective. (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). The indication of Illumina adapters wasn't needed for Trim Galore as it recognized them automatically. Trim Galore v0.5.0 removed the adapters and 12 bp from the 5' end of the reads (BOLGER *et al*. 2014) with Cutadapt v1.15.

After sequence quality control, we used Bismark v0.17.0 (https://www.bioinformatics.babraham.ac.uk/projects/bismark/), which contains codes and functions to map bisulfite-treated readings to a reference genome. It makes methylation calls and specifies the methylation states of cytosines. Bismark maps bisulfite converted gene reads to the genome using the short read aligner Bowtie 2. Using Bowtie v.2.3.4 (LANGMEAD *et al*. 2009, LANGMEAD and SALZBERG 2012), we collected an average of 9.7 Gb/Illumina PE read sample (approximate sequence depth=17x) with *Malus* × *domestica* GDDH13 Whole Genome v1.1 reference apple genome (DACCORD *et al*. 2017). Bowtie v2.3.4 is an open-source alignment program that maps the short DNA sequence to the reference genomes.

Results from Bismark (BAM files) can be displayed and evaluated using SeqMonk v.1.41.0 (https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/) (Figure 12) (ANDREWS 2007). The program provides the ability to visualize and analyse sequence data such as chromosomes, introns, and exons. It also enables statistical analysis of the data broken down into different regions. The program can be used to scroll through the evaluation of Next-generation sequencing, but it can also be used for any other data set. It has the advantage of using a public data set (https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/).

Figure 12. SeqMonk (List Panel: A list of all the imported and created files, Chromosome Panel: A quick bird's eye view of data signal on the chromosomes, Track Panel: A detailed view of annotation and data tracks)

The assembled apple genes were annotated using the Blast2GO feature in OmicsBox (Biobam, Valencia, Spain) (Figure 13). Genes were blasted for the highest sequence homology against the National Center for Biotechnology Information (NCBI) Taxonomy *Viridiplantae* database (http://www.ncbi.nlm.nih.gov/taxonomy) and subsequently assigned to their corresponding gene ontology (GO) terms (GÖTZ *et al*. 2008). OmicsBox GO for functional annotation and analysis was used, which consists of three parts: the biological process, the molecular function and the cellular component.

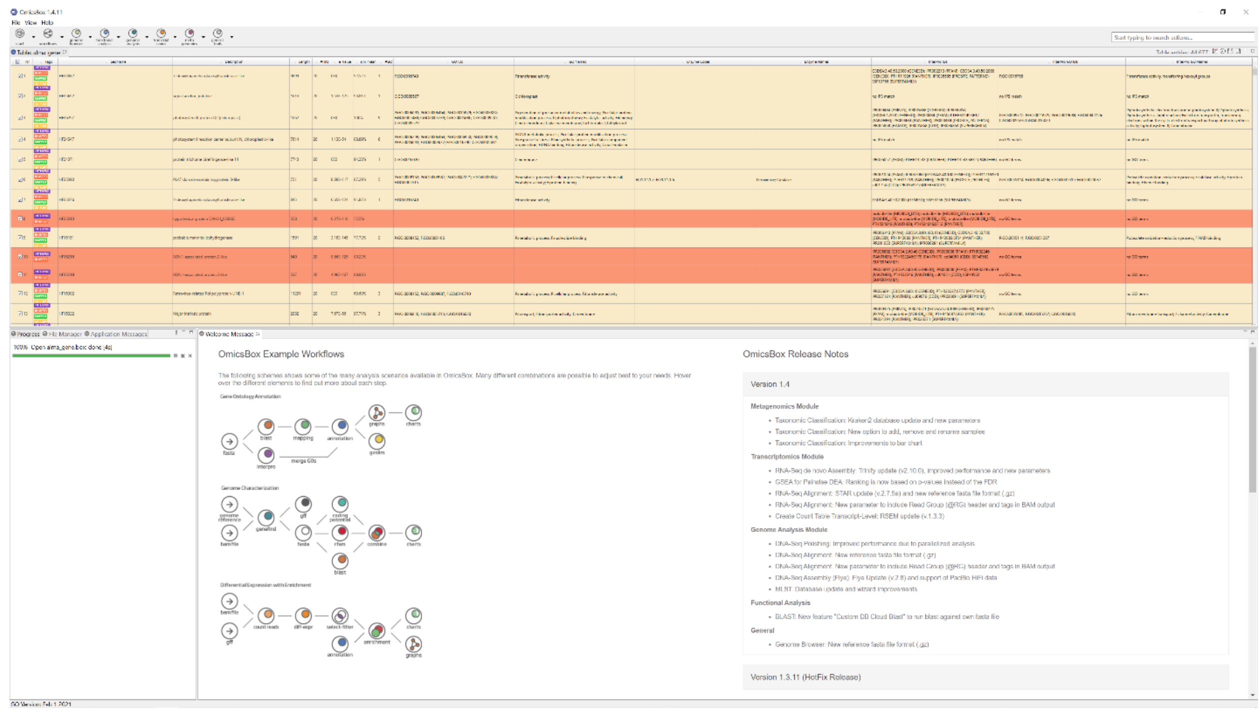


Figure 13. Omicsbox.The upper part of the image shows the apple genome (SeqName, Description, Length, GO IDs, InterproNames, etc.). At the bottom of the Omicsbox is an example of how to build a workflow.

The Kyoto Encyclopedia of Genes and Genomes (KEGG) database integrates genes from fully sequenced genomes into higher-level functions of the cell and body. The KEGG database is commonly used in high-throughput sequencing technologies including comparative proteomic analysis, which identifies proteins involved in the accumulation of soluble sugars and organic acids in the ripe fruit of wild *Malus* species (MA *et al.* 2019).

5. Results and discussion

5.1. Microsatellite marker based selection

The allele sizes determined by Alf-Express II device at 17 microsatellite loci are compiled in Table 4. No length polymorphism could be observed between the mother plant, *in vitro*, and acclimatized apple cultivars (GULYÁS *et al*. 2019a).

Table 4. Comparison of apple microsatellite allele lengthes (bp) (GULYÁS *et al*. 2019a)

SSR marker name	'Húsvéti rozmaring' mother plant	'Húsvéti rozmaring' *in vitro*	'Húsvéti rozmaring' acclimatized	'McIntosh' mother plant	'McIntosh' *in vitro*	'McIntosh' acclimatized
CH01f02	170:184	170:184	170:184	174:206	174:206	174:206
CH01h01	112:112	112:112	112:112	114:116	114:116	114:116
CH01h02	203:206	203:206	203:206	249:249	249:249	249:249
CH02c02	171:185:191	171:185:191	171:185:191	179:183	179:183	179:183
CH02c06	216:220:252	216:220:252	216:220:252	230:254	230:254	230:254
CH02c09	241:247	241:247	241:247	231:255	231:255	231:255
CH02c11	222:232	222:232	222:232	226:226	226:226	226:226
CH02d08	211:217	211:217	211:217	211:229	211:229	211:229
CH03a02	135:145	135:145	135:145	119:157	119:157	119:157
CH03g07	119:129	119:129	119:129	126:166	126:166	126:166
CH04e03	193:197:203	193:197:203	193:197:203	185:199	185:199	185:199
CH04e05	174:182	174:182	174:182	184:204	184:204	184:204
CH04g10	135:135	135:135	135:135	139:143	139:143	139:143
CH05c02	168:172	168:172	168:172	168:168	168:168	168:168
CH05c04	185:207	185:207	185:207	207:207	207:207	207:207
CH05d11	187:195:205	187:195:205	187:195:205	173:175	173:175	173:175
CH05e03	162:172:190	162:172:190	162:172:190	162:162	162:162	162:162

GUILFORD *et al*. (1997) were the first to develop and apply SSR markers for identifying apple cultivars and study polymorphism. Di- [(GA)n and (GT)n)] and trinucleotide [(CAG)n and (AAG)n] repeats were identified. LIEBHARD *et al.* (2002) developed 140 new SSR markers in the apple genome, of which we also used the markers for our own research. These markers were

chosen because several studies have shown that they are suitable for polymorphism and relatedness studies between apple cultivars (GALLI *et al.* 2005, SIKORSKAITE *et al.* 2012, URRESTARAZU *et al.* 2012). In some loci three alleles appeared in 'Húsvéti rozmaring', which is not expected if the tested genotype is diploid. However already GUILFORD and colleagues (1997) observed that from the 14 primers they had developed 4 resulted in complex pattern, proving that microsatellite regions can multiply from several loci. Similar results were obtained by WICHMANN *et al.* (2010), when they analysed Hungarian apple cultivars with a European standard set of microsatellite markers.

RAHMAN and RAJORA (2001) examined *Populus tremuloides* Michx. (Trembling aspen) donor trees under *in vitro* micropropagation conditions. One somaclonal variant was observed from the 13 donors in the PTR5 locus and one somaclonal variant from the 4 donors in the PTR2 locus. At the SSR locus, this meant that a new larger bp allele appeared in the PTR2 locus, while a new third allele in the PTR5 locus, which was explained by duplication of a new chromosome or chromosome segment. PALOMBI and DAMIANO (2002) studied genetic stability in kiwi (*Actinidia chinensis var. deliciosa* (A.Chev.) A.Chev.) plants *in vitro*. Differences between 'Hayward' clones were detected with SSR markers in the presence of a 100 bp and 190 bp fragment, which were missing from the other genotypes examined. FEUSER *et al.* (2003) identified the pineapple (*Ananas comosus* (L.) Merr.) plant as genetically stable with RAPD primers *in vitro*. KHAI and LANG (2005) distinguished between phenotypic and genotypic mutants of the indica rice plant *in vivo*. The phenotypic mutants differed externally, however, they could not show a significant difference with SSR markers. However, the genotypic mutants did not differ phenotypically, but showed significant differences with SSR markers. MODGIL *et al.* (2005) identified the apple plant as genetically stable *in vitro* by RAPD. JIN *et al.* (2008) studied tree cotton (*Gossypium arboreum* L.) plant *in vitro* under the influence of 2,4-D and kinetin hormone. Significant difference with SSR and RAPD primers was revealed due to the combination effects of hormones. PRADO *et al.* (2010) studied grape (*Vitis vinifera* L.) plants *in vitro* regenerated by somatic embryogenesis. SSR markers could identify a mutant allele (231 bp) compared to the normal one (237 bp) in the 'Torrontés' variety only in one case. Their research could not prove that this had anything to do with micropropagation, the medium used, or possible hormone combinations.

5.2. Analysis of Whole Genome Sequencing data

ADUSUMALLI *et al*. (2015) perfectly summarized how many opportunities a researcher may have to process his/her data. Data quality can be checked with a number program, and the following steps list several programs to choose from and leave it to the researcher to see which programs are the best for whole-genome bisulfite sequencing analysis. KRUEGER *et al.* (2012) wrote a review about investigating DNA methylome analysis using short bisulfite sequencing data. It was also a challenge for them to analyse the BS-seq data to see which programs were appropriate and which were the ones that gave the best results. Based on the articles available, it can be said that there is no program combination that always brings the best results, so there are different programs that researchers can choose from and decide which are the most suitable for their research/data. We did the same when we tried to analyse our own data. Based on the literature and our own experience, we selected the programs that were used to analyse the data obtained after bisulfite sequencing.

5.2.1. FastQC analysis

After sequencing the Illumina HiSeq 2500 2x150 bp, 12 files were recovered. The average file size was 7.23 GB. Sequencing results were recovered in fastq.gz format, which was analysed using bioinformatic softwares. First, we used the FastQC program to check the quality of the obtained raw reads. After illumina sequencing, many row readers are checked by researchers using the FastQC program (DETTMAN *et al.* 2012, TYLER *et al.* 2018, KOROSTIN *et al.* 2020). The program had 11 quality reports. As mentioned in the Material and Method, one of the important tables for FastQC was the base statistics. Table 5 shows the total sequence size, sequence length, and %GC of the ‘McIntosh’ and ‘Húsvéti rozmaring’ apple scions we sequenced.

Table 5. Basic statistics by FastQC on the two apple scions (‘McIntosh’ and ‘Húsvéti rozmaring’)

Samples	Total Sequences	Sequence length	%GC
‘McIntosh’ acclimatized R1	107,618,474	20-114	26
‘McIntosh’ acclimatized R2	107,618,474	20-114	26
‘McIntosh’ *in vitro* R1	94,768,039	20-114	28
‘McIntosh’ *in vitro* R2	94,768,039	20-114	28
‘McIntosh’ mother plant R1	91,867,470	20-114	27

'McIntosh' mother plant R2	91,867,470	20-114	27
'Húsvéti rozmaring' acclimatized R1	98,435,581	20-114	27
'Húsvéti rozmaring' acclimatized R2	98,435,581	20-114	27
'Húsvéti rozmaring' *in vitro* R1	95,521,186	20-114	28
'Húsvéti rozmaring' *in vitro* R2	95,521,186	20-114	28
'Húsvéti rozmaring' mother plant R1	93,708,990	20-114	27
'Húsvéti rozmaring' mother plant R2	93,708,990	20-114	27

The designations R1 and R2 derive from the sequencing of both strands of DNA. This was important because if we had sequenced only one of the strands, we would not have obtained a complete picture (only the CpG) of the methylation of the entire genome (CpG, CHG and CHH).

Based on the Table 5, it can be observed that the sequence sizes do not differ much from each other, except for 'McIntosh' acclimatized R1 and 'McIntosh' acclimatized R2, whose total sequence length stands out from the others. 'McIntosh' mother plant R1 and 'McIntosh' mother plant R2 had the lowest values (91867470) of the plant samples tested. The %GC is also nearly the same among all tested apples.

In addition to Per base sequence quality (Figure 14, 15), the Per base sequence content report is important for evaluating the sequencing. The proportion of each base location in a file for which each of the four regular DNA bases (G (gray), A (green), T (red), C (blue)) has been named is plotted by content. In the libraries there is little difference between the different bases of the sequence (G, T and A, T), we can say that the lines run almost parallel to each other. The lines reflect the relative number of bases in the entire genome.

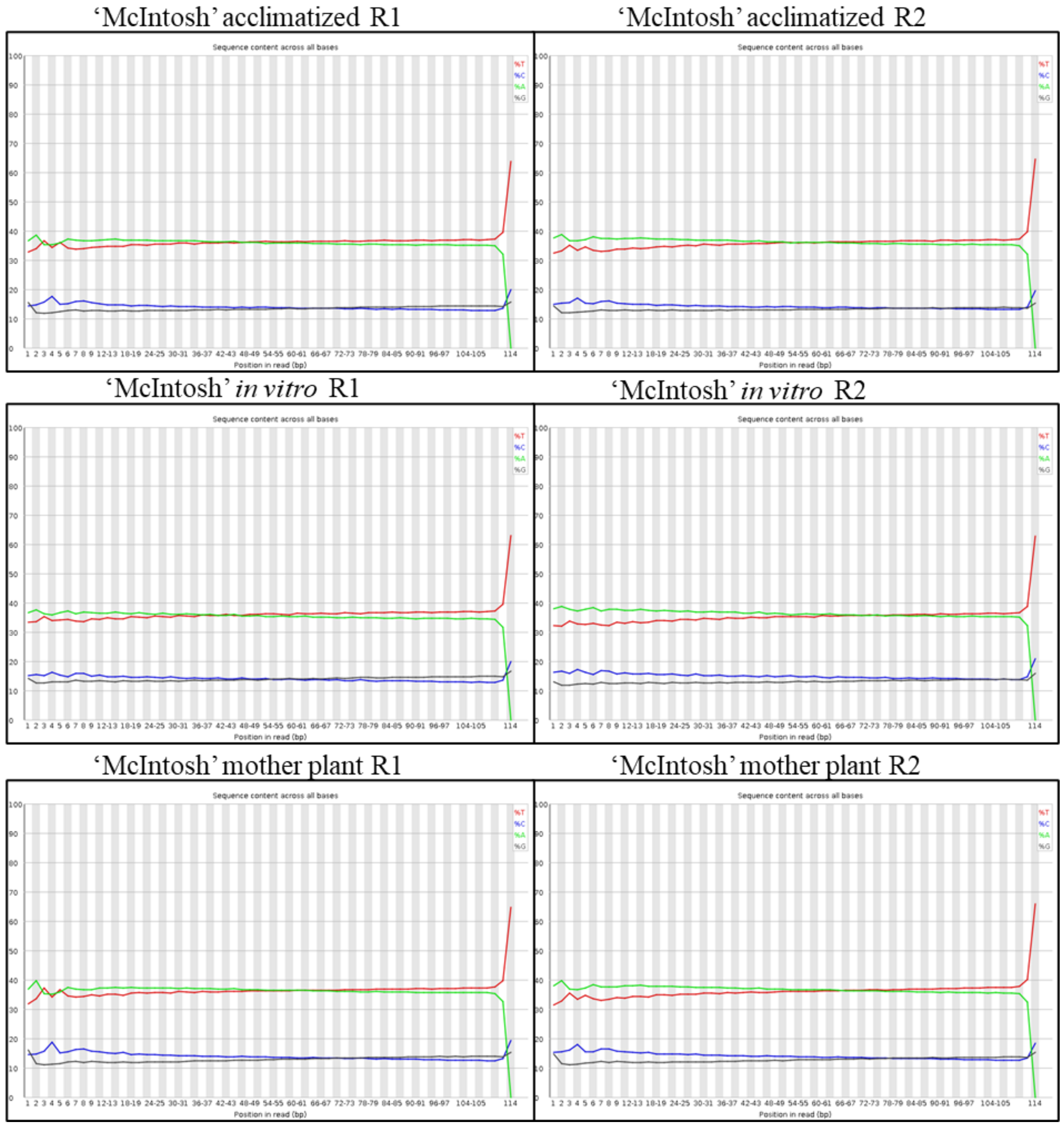


Figure 14. 'McIntosh' (acclimatized, *in vitro* and mother plant) per base sequence content report (G (gray), A (green), T (red), C (blue))

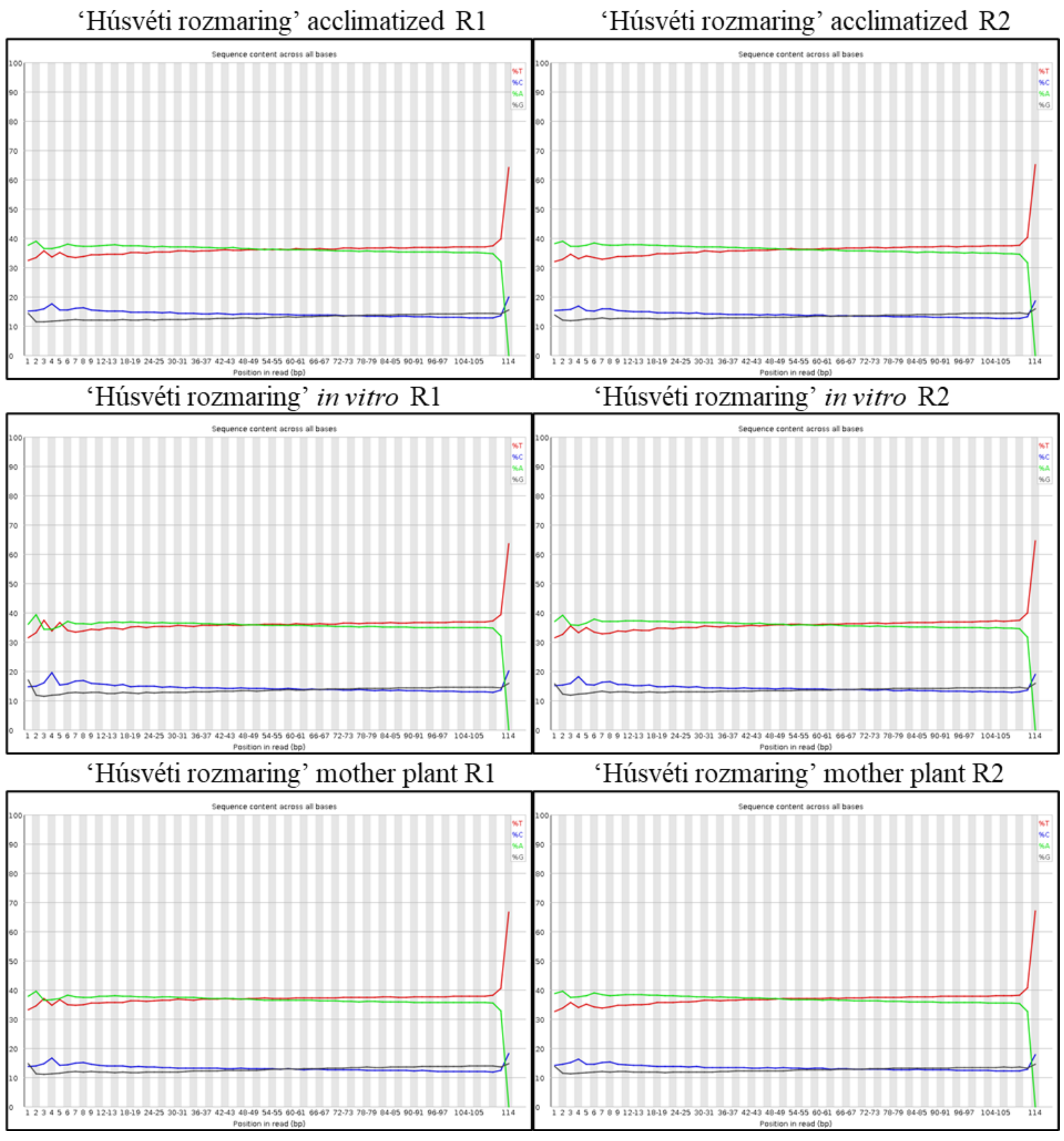


Figure 15. ‘Húsvéti rozmaring’ (acclimatized, *in vitro* and mother plant) per base sequence content report (G (gray), A (green), T (red), C (blue))

5.2.2. Trim Galore

There are also several programs for trimming data such as Btrim (KONG 2011), Cutadapt (MARTIN 2011), Kraken (DAVIS *et al.* 2013), Sickle (JOSHI and FASS 2011), Trim Galore (ANDREWS 2012), Trimmomatic (BOLGER *et al.* 2014) etc. We chose Trim Galore because it was also used in complete genome sequencing published in several scientific papers (CARMONA *et al.* 2013, LUND *et al*. 2017, MERKEL *et al.* 2019)

Trim Galore is a program that has no graphical interface and can only run files controlled by FastQC by command line. At the end of the program, we get a lot of data about the raw read. The total basepairs processed length is the amount of Surviving Reads, which is the amount you removed from the program (Table 6).

Table 6. Trim Galore report on the two apple scions

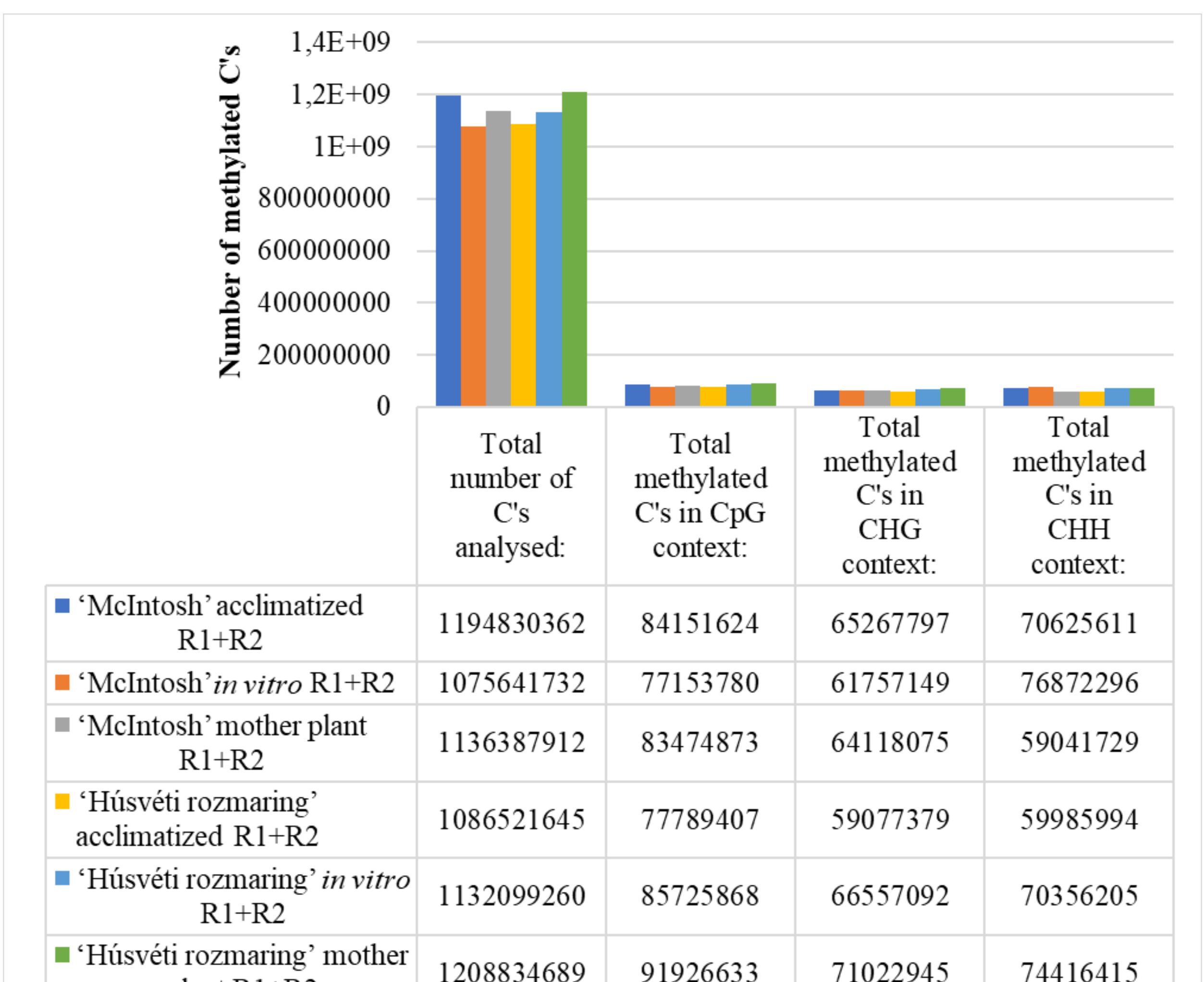
Sample	Raw paired-reads (Total basepairs processed)	Surviving Reads	Removed Reads (Quality-trimmed)	Removed Reads Percentage (%)
'McIntosh' acclimatized R1	13,646,776,500	13,538,096,741	108,679,759	0.8
'McIntosh' acclimatized R2	13,646,776,500	13,423,348,097	223,428,403	1.6
'McIntosh' *in vitro* R1	12,046,214,754	11,954,723,079	91,491,675	0.8
'McIntosh' *in vitro* R2	12,046,214,754	11,867,936,123	178,278,631	1.5
'McIntosh' mother plant R1	11,626,545,042	11,541,963,744	84,581,298	0.7
'McIntosh' mother plant R2	11,626,545,042	11,477,424,281	149,120,761	1.3
'Húsvéti rozmaring' acclimatized R1	12,496,611,204	12,396,149,077	100,462,127	0.8
'Húsvéti rozmaring' acclimatized R2	12,496,611,204	12,324,778,044	171,833,160	1.4
'Húsvéti rozmaring' *in vitro* R1	12,089,468,538	11,997,185,927	92,282,611	0.8
'Húsvéti rozmaring' *in vitro* R2	12,089,468,538	11,924,564,245	164,904,293	1.4
'Húsvéti rozmaring' mother plant R1	11,870,656,056	11,779,779,199	90,876,857	0.8
'Húsvéti rozmaring' mother plant R2	11,870,656,056	11,712,987,790	157,668,266	1.3

Based on the Table 6, it can be seen that most of the raw paired was given by 'McIntosh' acclimatized R1 and 'McIntosh' acclimatized R2 and the least one was given by 'Húsvéti rozmaring' mother plant R1 and 'Húsvéti rozmaring' mother plant R2. At the percentage of reads removed, it was observed that the highest percentage of reads was removed (1.6%) in the case of the acclimatized R2 'McIntosh'. 'McIntosh'R1 was the mother plant from which the lowest read value was removed. Based on the Table 6, it can be said that on average 0.9% of the reads was removed ads during trimming.

5.2.3. Bismark

GREHL *et al.* (2020) research suggests that although there are pipelines that are already optimized, unfortunately many important plants are still missing this process. Here, eight read mappers were compared: Bismark, BismarkBwt2, BSMAP, BS-Seeker2, Bwameth, GEM3, Segemehl, and GSNAP to assess the effect of read mapping results on DMR (differentially methylated regions) prediction. The user need the most accuracy, and Bismark, which requires the least memory and results in accuracy and a large number of individually mapped readings. That is why we used the Bismark program. Using Bismark, we were able to determine the total number of C’s analysed, the amount of total methylated CpG, CHG, and CHH, and the amount of total unmenthylated CpG, CHG and CHH.

Table 7. Total number of methylated C’s in CpG, CHG and CHH contexts

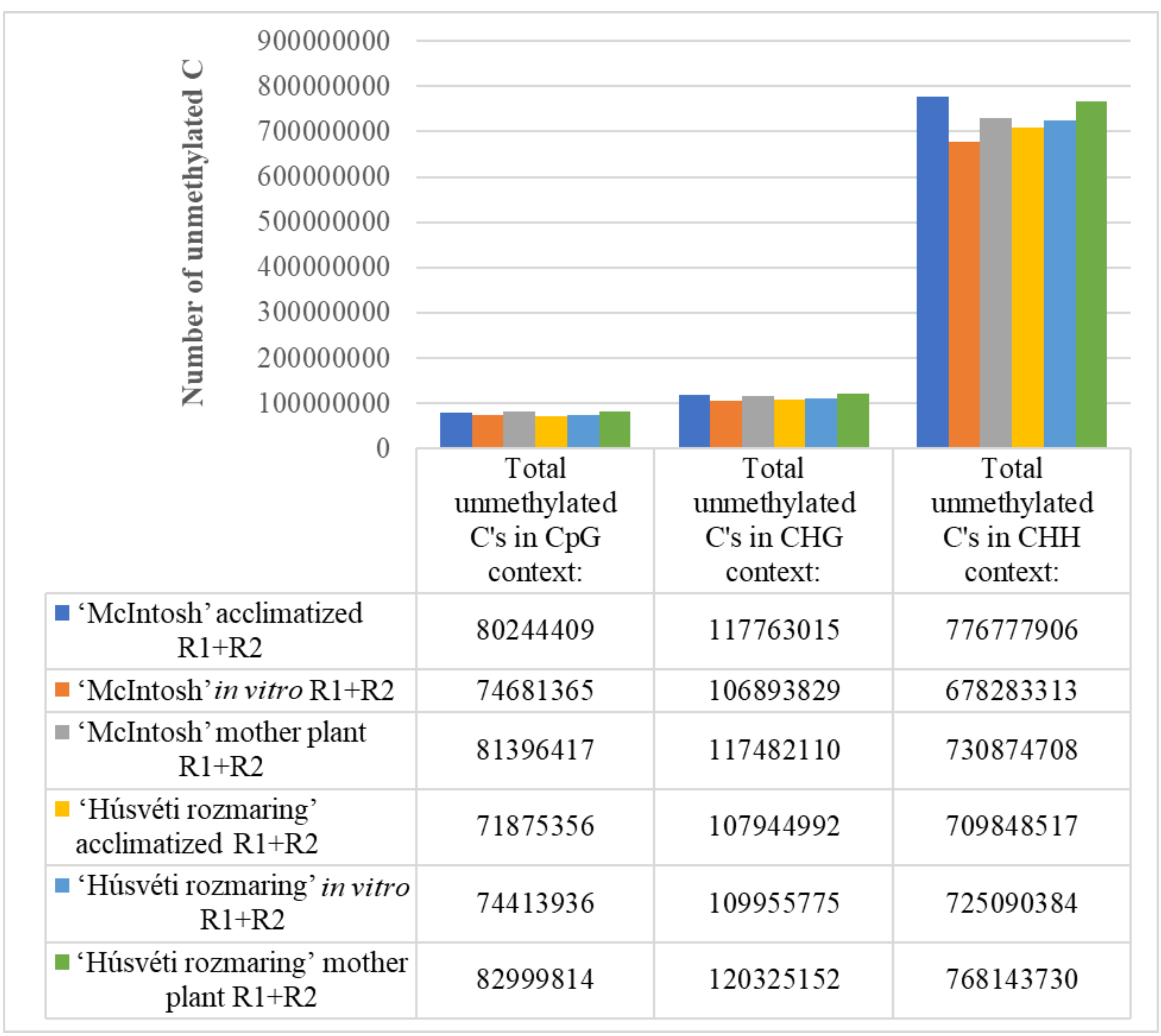


	Total number of C's analysed:	Total methylated C's in CpG context:	Total methylated C's in CHG context:	Total methylated C's in CHH context:
‘McIntosh’ acclimatized R1+R2	1194830362	84151624	65267797	70625611
‘McIntosh’ *in vitro* R1+R2	1075641732	77153780	61757149	76872296
‘McIntosh’ mother plant R1+R2	1136387912	83474873	64118075	59041729
‘Húsvéti rozmaring’ acclimatized R1+R2	1086521645	77789407	59077379	59985994
‘Húsvéti rozmaring’ *in vitro* R1+R2	1132099260	85725868	66557092	70356205
‘Húsvéti rozmaring’ mother plant R1+R2	1208834689	91926633	71022945	74416415

The highest value in the total number of C’s analysed (Table 7) was obtained with the ‘Húsvéti rozmaring’ mother plant R1+R2 (1,208,834,689). This value was highly distinguished from the other samples examined. On average, the total number of C’s analysed 1,139,052,600

was 6,12 % lower than this value for 'Húsvéti rozmaring' mother plant R1+R2. 'Húsvéti rozmaring' mother plant R1+R2 also outperformed all samples in the total number of CpG. For CpG, the average was 83,370,364.2. In the case of total CHG, although 'Húsvéti rozmaring' mother plant R1+R2 stands out. In this context the average of methylated cytosines were 68,549,708.3. It can be observed that of the 6 plants samples we examined, 'McIntosh' mother plant R1 + R2 'Húsvéti rozmaring' acclimatized R1 + R2 had a much lower methylated cytosine number.

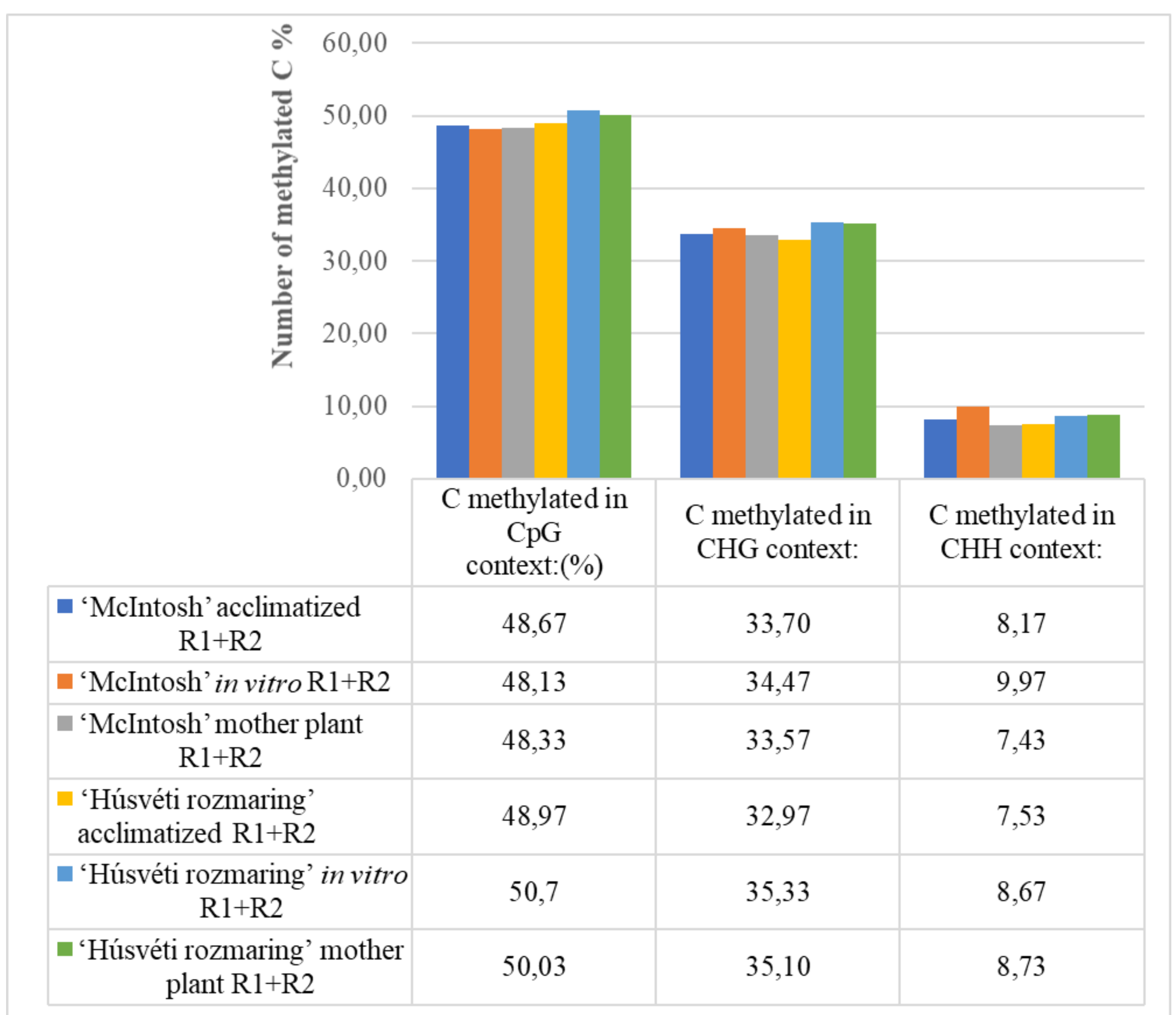
Table 8. Total unmethylated C's in CpG, CHG and CHH contexts



	Total unmethylated C's in CpG context:	Total unmethylated C's in CHG context:	Total unmethylated C's in CHH context:
'McIntosh' acclimatized R1+R2	80244409	117763015	776777906
'McIntosh' *in vitro* R1+R2	74681365	106893829	678283313
'McIntosh' mother plant R1+R2	81396417	117482110	730874708
'Húsvéti rozmaring' acclimatized R1+R2	71875356	107944992	709848517
'Húsvéti rozmaring' *in vitro* R1+R2	74413936	109955775	725090384
'Húsvéti rozmaring' mother plant R1+R2	82999814	120325152	768143730

Table 8 shows how many unmethylated contexts were in all the cases examined, in total, CpG, CHG, and CHH respectively. Even for unmethylated contexts, 'Húsvéti rozmaring' mother plant R1 + R2 shows the highest values. The same was observed for CHG as for CpG. For CHH, nearly identical results were obtained except for 'McIntosh' *in vitro* R1 + R2 (678,283,313) where the number of unmethylated cytosines was much lower.

Table 9. C methylated in CpG, CHG and CHH contexts



	C methylated in CpG context:(%)	C methylated in CHG context:	C methylated in CHH context:
‘McIntosh’ acclimatized R1+R2	48,67	33,70	8,17
‘McIntosh’ *in vitro* R1+R2	48,13	34,47	9,97
‘McIntosh’ mother plant R1+R2	48,33	33,57	7,43
‘Húsvéti rozmaring’ acclimatized R1+R2	48,97	32,97	7,53
‘Húsvéti rozmaring’ *in vitro* R1+R2	50,7	35,33	8,67
‘Húsvéti rozmaring’ mother plant R1+R2	50,03	35,10	8,73

The methylated cytosin for CpG is shown in Table 9, it can be observed that the values are almost the same, only the values of ‘Húsvéti rozmaring’ *in vitro* R1+R2 and ‘Húsvéti rozmaring’ mother plant R1+R2 are higher than the others (Table 9). In the CHH context, however, the highest value was found ‘McIntosh’ *in vitro* for R1 + R2.

5.2.4. SeqMonk

After Bismark, the resulting CpG, CHG, and CHH methylation levels were visualized and analysed using SeqMonk. We used SeqMonk to compare our samples based on the GDDH13 Version 1.1 reference genom (https://iris.angers.inra.fr/gddh13/). Each context, CpG, CHG, and CHH had to be imported into the program to get a complete picture of the ‘McIntosh’ and ‘Húsvéti rozmaring’ apple scions, what we studied. The settings required for the analysis were aided by "Exercises: Differential Methylation" (https://www.bioinformatics.babraham.ac.uk/training/Methylation_Course/Differential%20meth

ylation%20Exercises.pdf) published by SeqMonk, which contained all the necessary information. With the help of the program, we were able to import the CpG, CHG CHH tables of two apples into a txt file (GULYÁS *et al*. 2019b) and we also looked at the IDs that differ significantly in the CpG, CHG and CHH contexts, if we do not break them down into varieties. Respectively, we also created a table that shows the IDs that are not significant in the three contexts. Also, with the help of SeqMonk, we compared both *in vitro* and acclimatized plants of the mother plant regardless of cultivars. After multiple-test correction, methylation dynamics is determined by analyzing data sets using the SeqMonk chi-square (χ^2) filter with a p value requirement of < 0.05.

The SeqMonk help to visualize our results, which we present through the scatter plots, based on the generated statistical analysis in logarithmic fold change (LFC) values. Differentially methylated genes (DMG) were described as genes with significant differences in DNA methylation (DMGs) in either their promoter or coding regions, as determined by the χ^2 test. The χ^2 test was visualized using a SeqMonk-generated MA plot (Bland-Altman plot), in which the differences in measurements between two of the three samples in all permutations were measured by transforming the data into M (log ratio) and A (mean average) scales using SeqMonk, and then plotting these values (BLAND and ALTMAN 1999). The DNA methylation patterns of samples from various environments in both scion cultivars were determined and compared using four MA plot tests. In ‘McIntosh’ (Figure 16) and ‘Húsvéti rozmaring’ (Figure 17), paired similarities in DNA methylation patterns in CpG, CHG, and CHH contexts were determined between acclimatized and *in vitro* plants, acclimatized and mother plants, and *in vitro* and mother plants of each cultivar separately, and by considering the average of all cultivars (Figure 18). *In vitro*, acclimatized, and mother plants of ‘McIntosh’ and ‘Húsvéti rozmaring’ were also compared for DNA methylation patterns between cultivars (Figure 19).

DMGs were downregulated in *in vitro* plants relative to acclimatized plants when DNA methylation patterns of ‘McIntosh’ were compared (Figure 16A, 16D, 16G). The level of methylation in *in vitro* plants was higher than in mother plants, indicating that DMGs were upregulated, resulting in lower methylation in the mother plant (Figure 16C, 16F, 16I). When comparing the MA plots of acclimatized and mother plants (Figure 16B, 16E, 16H), no significant directional differences were seen, suggesting that after acclimatization, the DNA methylation pattern became identical to that of the mother plant.

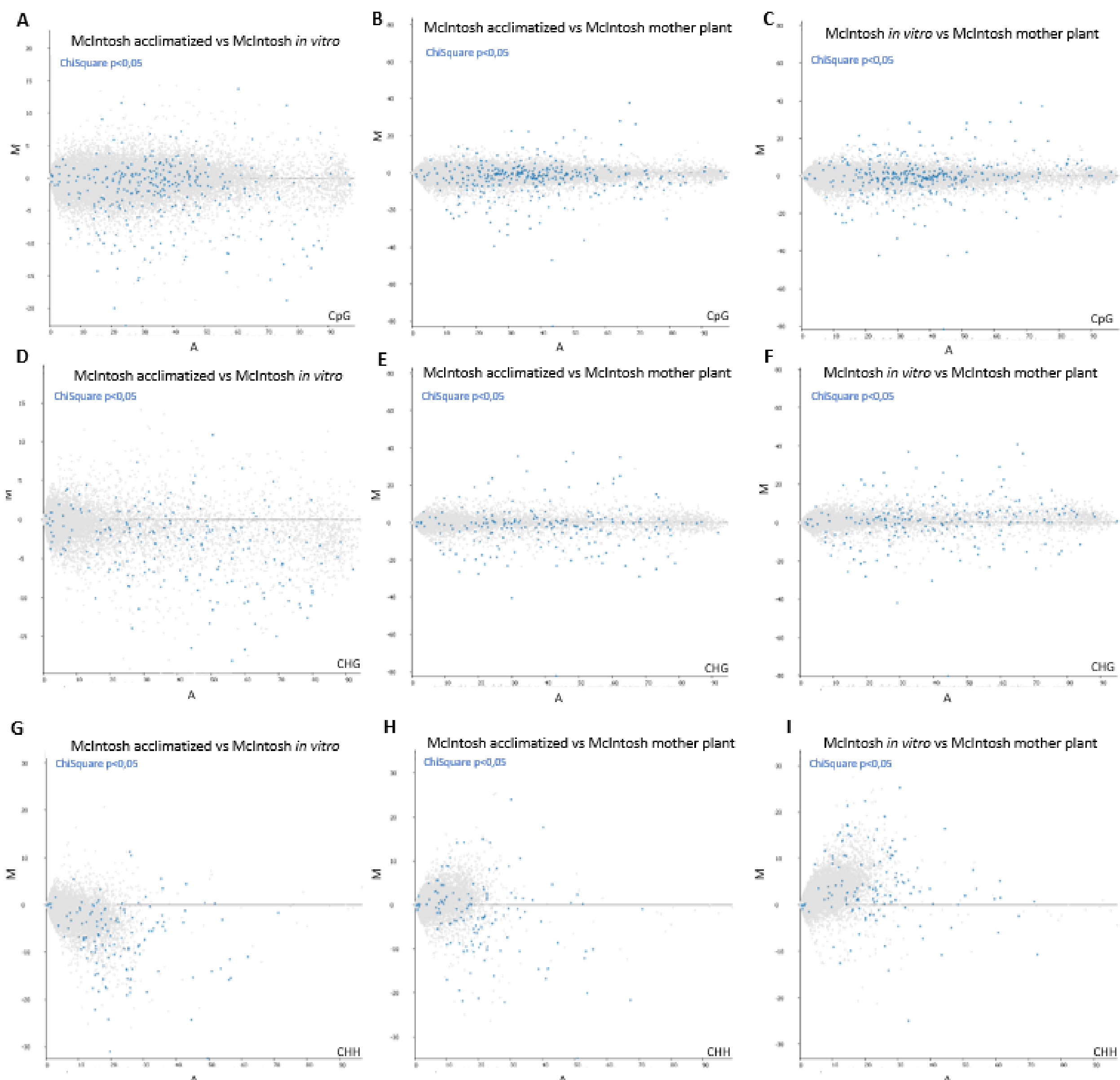


Figure 16. Differences in DNA methylation levels in CpG (A, B, C), CHG (D, E, F), and CHH (G, H, I) contexts between acclimatized and *in vitro* plants (A, D, G), acclimatized and mother plants (B, E, H), and *in vitro* and mother plants (C, F, I) of scion cultivar 'McIntosh' (grey dots: non-significantly differently methylated genes, blue dots: significantly differently methylated genes) (GULYÁS *et al.* 2019b)

DMG was reduced in *in vitro* plants compared to acclimatized plants (Figures 17A, 17D, 17G) in all contexts. In the mother plant, DMGs were upregulated, thus methylation levels were lower compared to *in vitro* (Figures 17C, 17F, 17I). No significant differences were found between acclimatized and mother plants (Figures 17B, 17E, 17H) here either. The same was true for 'Húsvéti rozmaring' as it was for 'McIntosh'.

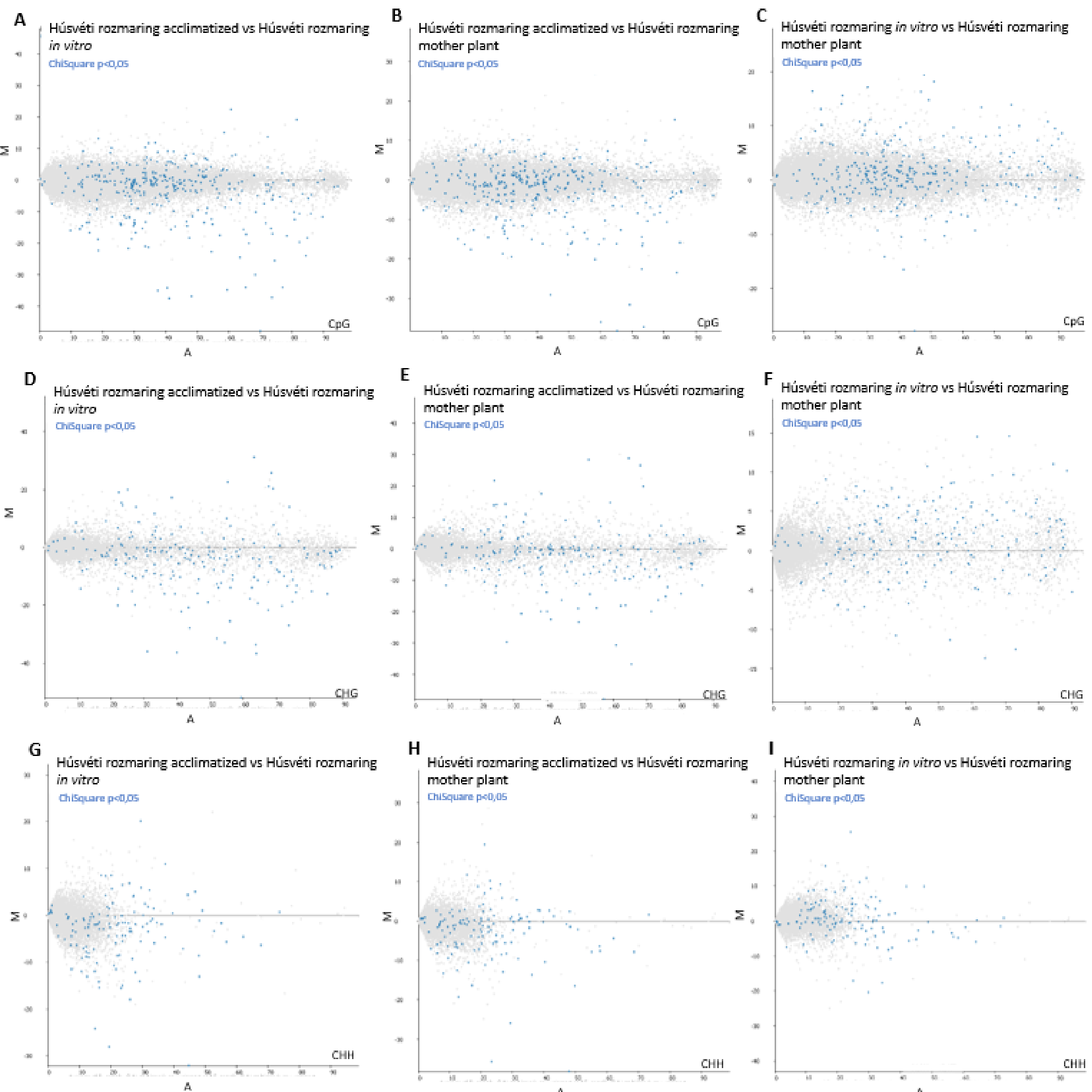


Figure 17. Differences in DNA methylation levels in CpG (A, B, C), CHG (D, E, F), and CHH (G, H, I) contexts between acclimatized and *in vitro* plants (A, D, G), acclimatized and mother plants (B, E, H), and *in vitro* and mother plants (C, F, I) of scion cultivar 'Húsvéti rozmaring' (grey dots: non-significantly differently methylated genes, blue dots: significantly differently methylated genes) (GULYÁS *et al*. 2019b)

Both up- and downregulation of DMGs were found in all three contexts when the two scion cultivars, 'McIntosh' and 'Húsvéti rozmaring', were compared (CpG, Figure 18A, 18B, 18C; CHG, Figure 18D, 18E, 18F; CHH, Figure 18G, 18H, 18F). The two scion cultivars had different methylation patterns, showing that the epigenome was differently regulated and in a cultivar-specific method during apple adaptation to various environments. Our results show that cultivar-specific DNA methylation patterns may play a regulatory role in sensing and respond by regulating stress-responsive gene expression independently of the environment.

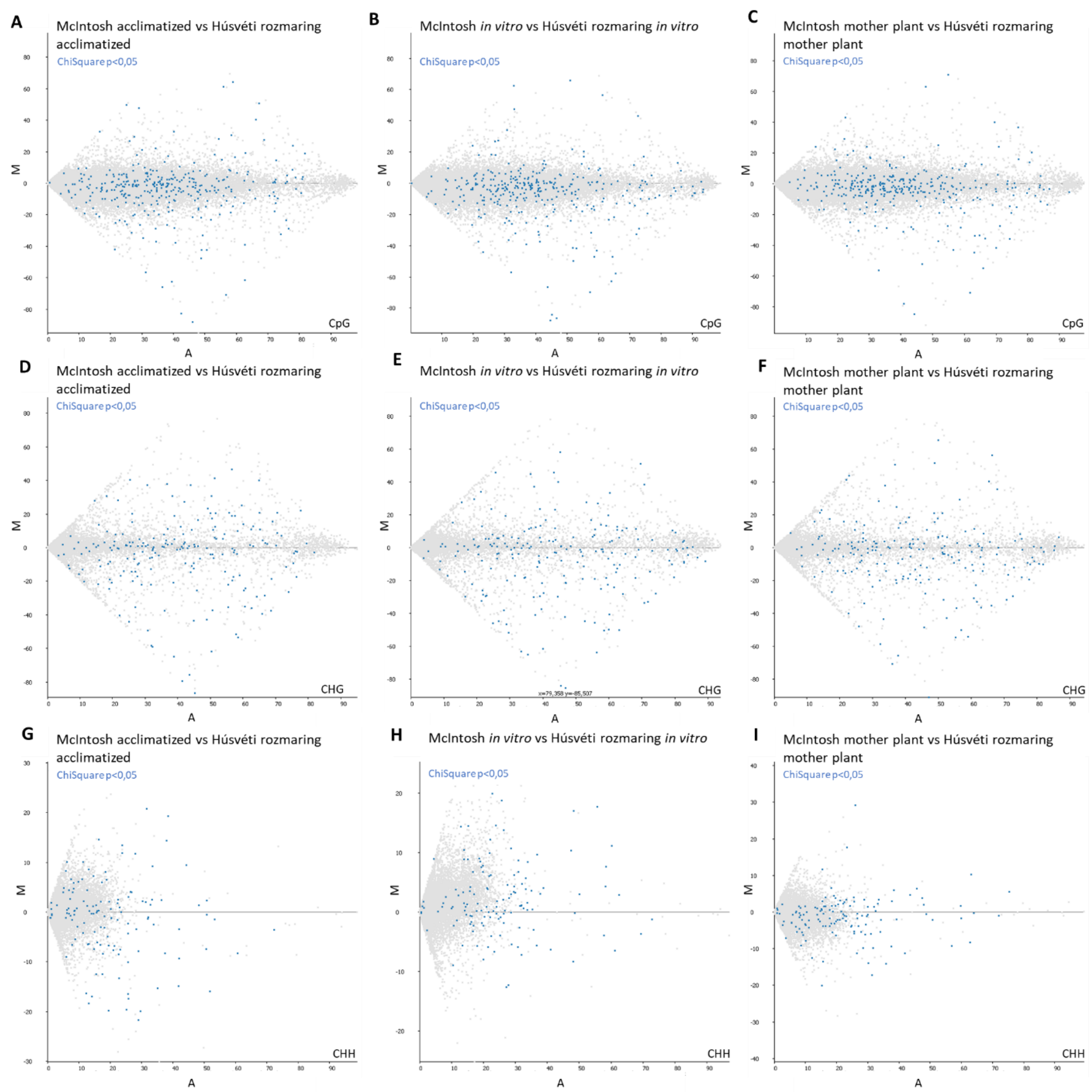


Figure 18. Differences in DNA methylation levels in CpG (A, B, C), CHG (D, E, F), and CHH (G, H, I) contexts between apple scion cultivars 'McIntosh' and 'Húsvéti rozmaring' in acclimatized (A, D, G), *in vitro* (B, E, H), and mother plants (C, F, I) (grey dots: non-significantly differently methylated genes, blue dots: significantly differently methylated genes) (GULYÁS *et al*. 2019b)

In both cultivars and all three contexts, significant differences in DNA methylation patterns were found. When apples were examined in general, we compared the *in vitro* to the acclimatized, the mother plant to the acclimatized, and the *in vitro* to the mother plant. We found that acclimatized plants were much less methylated than *in vitro* (Figure 19A, 19D, 19G). When compared to the acclimatized with the mother plant, the methylation pattern tended to be similar (Figure 19B, 19E, 19H).

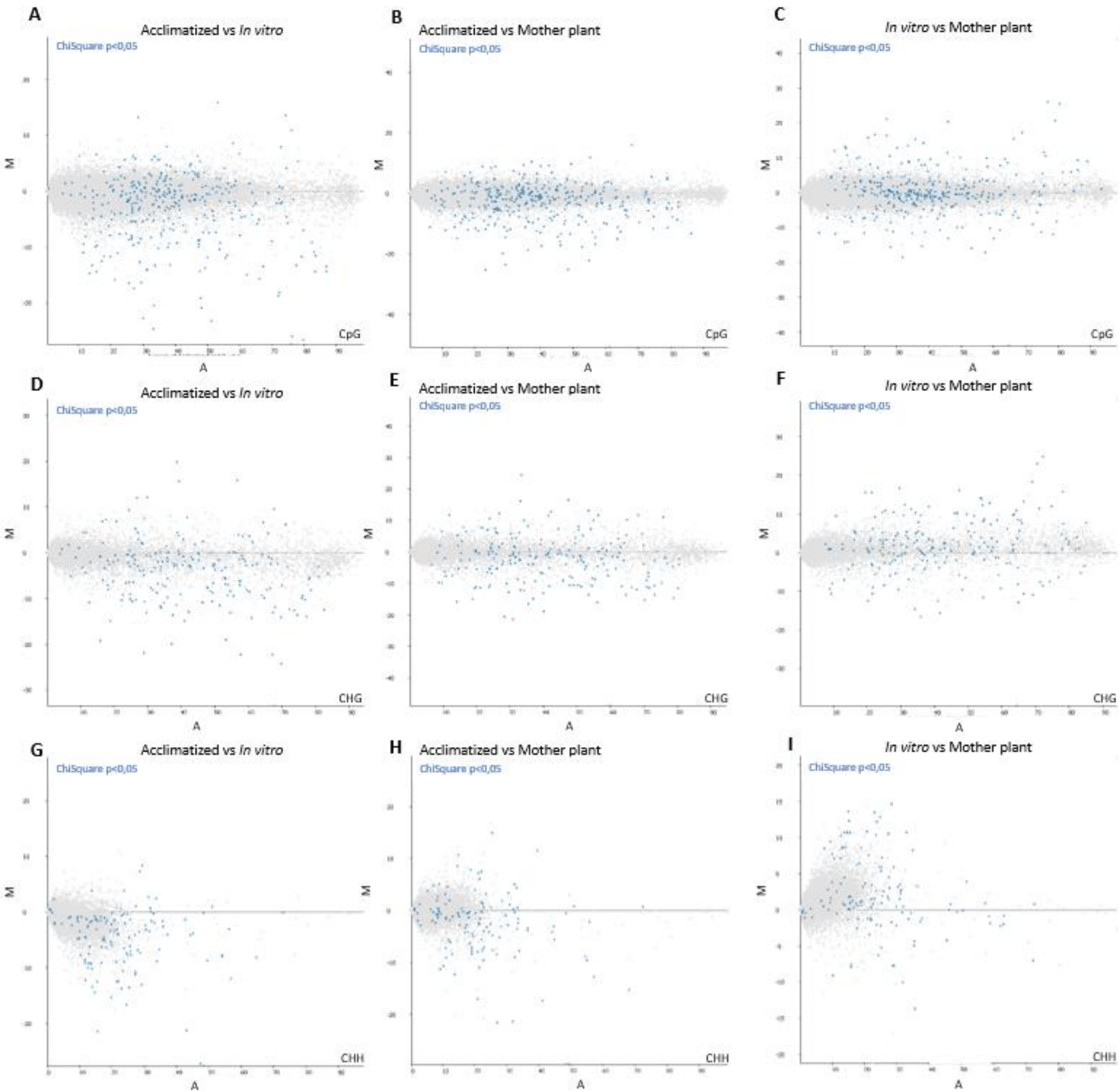


Figure 19. Differences in DNA methylation levels in CpG (A, B, C), CHG (D, E, F), and CHH (G, H, I) contexts between acclimatized and *in vitro* plants (A, D, G), acclimatized and mother plants (B, E, H), and *in vitro* and mother plants (C, F, I) based on the average of both scion cultivars (grey dots: non-significantly differently methylated genes, blue dots: significantly differently methylated genes) (GULYÁS *et al*. 2019b)

Using SeqMonk, we were able to illustrate the differences between the different apple samples in all three contexts. ROTHKEGEL *et al.* (2020) studied different cultivars of cherries (*Prunus avium* (L). L.). For low-cool 'Royal Dawn' and high-temperature 'Kordia' cultivars, flower bud methylome was examined using MethylC-seq. Their results were visualized using SeqMonk, among others, as described by BUTLER *et al.* (2013) also when *Pseudomonas syringae* pv. *actinidiae* were studied.

5.2.5 Omicsbox

The tables obtained by SeqMonk were further examined using Omicsbox (GULYÁS *et al*. 2019b). GO annotation was performed with Blast2 GO v5.1.12 (CONESA and GÖTZ 2008) based on the *Malus* × *domestica* GDDH13 Whole Genome v1.1 (DACCORD *et al*. 2017). Based on the χ^2 results for variations in DNA methylation patterns and their similarities, DMGs were chosen for GO annotation when differences were larger than ±10.00% in CpG and CHG, or larger than the average change in methylation (±1.35%) in CHH. For functional mapping and GO annotation, several databases were used [NCBI Nr: non-redundant protein database (ftp://ftp.ncbi.nlm.nih.gov/blast/db/); Swissprot-Uniprot database (https://www.uniprot.org/); Kyoto Encyclopedia of Genes and Genomes (KEGG; https://www.genome.jp/kegg/);GO; InterproScan; https://www.ebi.ac.uk/interpro/] by BlastX-fast with the E-value cut-off set to 10^{-3} (GULYÁS *et al*. 2019b). For functional mapping of biological processes, molecular functions, and cellular compounds, all assembled DMGs, based on the χ^2 test were considered. Gene Ontology (GO) annotation was used to specify the functions of all DMGs in the three environments, with focusing on the biological processes, the molecular functions, and the cellular components of green plants (*Viridiplantae*), as these may play important roles throughout *in vitro* culture. A total of 45,116 genes were examined using the program. Significant differences were identified in the differentially methylated genes in a total of 586 genes (334 in CpG, 201 in CHG, and 131 in CHH), of which only 446 were found in the DNA methylation Blast2GO database. This shows a 1.3% difference in gene DNA methylation pattern in response to a change in environment, such as *in vivo* versus *in vitro*, or when the *in vitro* conditions were replaced by an *in vivo* environment. In all three contexts, 72 DMGs were similar in CHG and CpG, five DMGs were identical in CHG and CHH, and seven DMGs were identical in CpG and CHH. Furthermore, four DMGs were found to be similar in all three cases, while 259, 128 and 123 DMGs were found to be unique to the CpG, CHG, and CHH contexts, respectively (Figure 20).

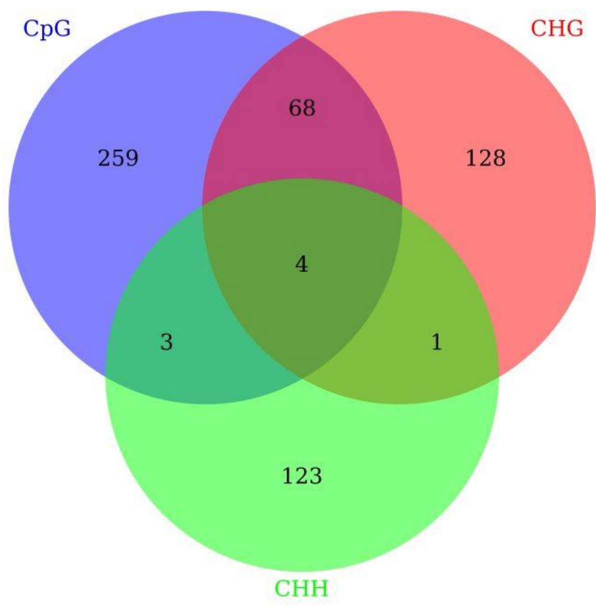


Figure 20. The CpG, CHG, and CHH contexts, the sequence distribution of significant differences in DNA methylation in genes (GULYÁS *et al*. 2019b)

DMGs identified as MD07G1113000 (protein transport), MD08G1041600 (extracellular space), MD09G1054800 (phosphatidic acid binding), and MD10G1265800 (not annotated) were methylated in all three contexts in *in vitro* shoots. The highest level of methylated DMGs was found in *in vitro* tissue culture. Some DMGs that participate in oxidation-reduction processes, metabolism, and biosynthesis (KAWAKATSU *et al.* 2017, BOUYER *et al.* 2017) but are not needed during *in vitro* culture were primarily methylated. In 'McIntosh' the amount of DNA methylation of DMGs was lower than in 'Húsvéti rozmaring'. After acclimatization, the amount of methylation in DMGs decreased in both cultivars.

Table 10. Biological proccesses in 'McIntosh' affected by methylation (yellow: present and number of DMGs, blue: absent of DMGs)

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
cell wall organization	2								
response to stress	2	2						2	
regulation of transcription, DNA-templated	2		2			2		2	2
serine family amino acid metabolic process	2								
metal ion transport	2			1					
lignin biosynthetic process	3	3		1	1		3		
coumarin biosynthetic process	3	3		1	1		3	3	
pigment biosynthetic process	3	3		1	1		3	3	
protein phosphorylation	3	2		4	4			4	

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
oxidation-reduction process	7	5	3	5	2	3	8	6	2
carbohydrate metabolic process	3								
stilbene biosynthetic process	3	3		1	1		3	3	
alkaloid biosynthetic process	3	3		1	1		3	3	
tyrosine metabolic process	3	3		1	1		3	3	
riboflavin metabolic process	3	3		1	1		3	3	
proteolysis		2	2			2		2	2
organelle organization		2				2			2
regulation of transcription by RNA polymerase II		2							
peptide metabolic process		2						2	
regulation of cellular protein metabolic process		2						2	
developmental process		2						2	
response to chemical		2						2	
methylation		2			1			2	
transmembrane transport		3					2	3	
peptide transport		2						2	
ribosome biogenesis			2						2
cellular response to stimulus			2			3			3
sucrose metabolic process			2			2			2
regulation of protein complex assembly			2			2			2
translation			2						2
coenzyme metabolic process			2						2
innate immune response			2			2			2
nucleotide metabolic process			2						
defense response to other organism			2			2			2
ribose phosphate metabolic process			2						
response to abiotic stimulus			2			2			2
starch metabolic process			2			2			2
aspartate metabolic process				1					
cytokinin biosynthetic process				1					
alanine metabolic process				1					
fructose metabolic process				1					
ion transmembrane transport				1		2			
response to oxidative stress				1					
cellular transition metal ion homeostasis				1					
mannose metabolic process				1					
aspargine metabolic process				1					
protein repair				1					
cellular protein metabolic process						2			
positive regulation of cellular process						2			2
small molecule metabolic process						2			
generation of precursor metabolites and energy						2			2
phosphorylation						3	2		2
regulation of response to stimulus						2			2
regulation of biological quality						2			2

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
cellular protein modification process							2		
carbohydrate derivative process							2		
cellular component organization or biogenesis							2		
carboxylic acid metabolic process									2
purine ribonucleotide metabolic process									2
transport			2						2
cellular catabolic process		2						2	

In the 'McIntosh' cultivar, we were able to identify 62 biological processes affected by methylation. Oxidation and reduction processes showed significant changes in all conditions and contexts based on Omicsbox analysis. A total of 126 DMGs were identified between the mother plant and *in vitro* plant, of which 44, 51, and 31 were methylated in the contexts of CpG, CHG, and CHH, respectively. The DMGs between the mother plant and *in vitro* plants affected mainly the biosynthetic and metabolic processes. Examples of such processes are lignin, coumarin stilbene, carbohydrate and alkali biosynthesis. The regulation of transcription, DNA-templated; serine family amino acid metabolic process; metal ion transport and carbohydrate metabolic process were always methylated in CpG context. In CHG context DMGs were related to proteolysis; organelle organization; regulation of transcription by RNA polymerase II; peptide metabolic process; regulation of cellular protein metabolic process; developmental process; response to chemical; methylation; transmembrane transport; peptide transport and cellular catabolic process. The ribosome biogenesis; cellular response to stimulus; sucrose metabolic process; regulation of protein complex assembly; translation; coenzyme metabolic process; innate immune response; nucleotide, metabolic process; defense response to other organism; ribose phosphate metabolic process; response to abiotic stimulus; starch metabolic process and transport were always methylated in CHH context. A total of 84 DMGs were identified between acclimatized and *in vitro* plants, of which 27, 16, and 41 were methylated in the contexts of CpG, CHG, and CHH, respectively The differences between acclimatized and *in vitro* plants affected metal ion transport; aspartate metabolic process; cytokinin biosynthetic process; alanine metabolic process; fructose metabolic process; ion transmembrane transport; response to oxidative stress; cellular transition metal ion homeostasis; mannose metabolic process; asparagine metabolic process and protein repair in CpG context. In CHG context DMGs were related to the methylation. It can be observed that regulation of transcription, DNA-templated; proteolysis; organelle organization; cellular response to stimulus; sucrose metabolic process; regulation of protein complex assembly; innate immune response; defense response to other organism; response to abiotic stimulus; starch metabolic process; cellular protein metabolic process; positive regulation

of cellular process; small molecule metabolic process; generation of precursor metabolites and energy; phosphorylation; regulation of response to stimulus; regulation of biological quality is related to the exclusive methylation of CHH context. A total of 140 DMGs were identified between the mother and acclimatized plants, of which 41, 54 and 45 were methylated in CpG, CHG, and CHH contexts, respectively. The cellular protein modification process; carbohydrate derivative process; cellular component organization or biogenesis were always methylated in CpG context. In the pattern of biological processes, in response to stress; peptide metabolic process; regulation of cellular protein metabolic process; developmental process; response to chemical; methylation; peptide transport; cellular catabolic process appeared in CHG context. The organelle organization; ribosome biogenesis; cellular response to stimulus; sucrose metabolic process; regulation of protein complex assembly; translation; coenzyme metabolic process; innate immune response; defense response to other organism; response to abiotic stimulus; starch metabolic process; positive regulation of cellular process; generation of precursor metabolites and energy; regulation of response to stimulus; regulation of biological quality; carboxylic acid metabolic process; purine ribonucleotide metabolic process; transport were always methylated in CHH context. The regulation of transcription, DNA-templated and proteolysis were always methylated in CHG and CHH context. The coumarin biosynthetic process; pigment biosynthetic process; stilbene biosynthetic process; alkaloid biosynthetic process; tyrosine metabolic process and riboflavin metabolic process were always methylated in CpG and CHG context.

Table 11. Biological proccesses in 'Húsvéti rozmaring' affected by methylation (yellow: present and number of DMGs, blue: absent of DMGs)

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
serine family amino acid metabolic process	1								
lignin biosynthetic process					2		2	2	
coumarin biosynthetic process					2		2	2	
pigment biosynthetic process					2		2	2	
protein phosphorylation	1	1		8	8		4	5	
oxidation-reduction process		2		5	5	2	5	3	4
carbohydrate metabolic process	1			4			3	1	
stilbene biosynthetic process					2		2	2	
alkaloid biosynthetic process		1			3		2	2	
tyrosine metabolic process					2		2	2	
riboflavin metabolic process					2		2	2	
proteolysis			2			2			
organelle organization						2			
transmembrane transport				3	2				
ribosome biogenesis			2			2			2

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
cellular response to stimulus			2			2			3
sucrose metabolic process						2			2
regulation of protein complex assembly			2						
translation			2			2			2
coenzyme metabolic process			2						2
nucleotide metabolic process			2						
response to abiotic stimulus			2						
starch metabolic process						2			2
aspartate metabolic process	1								
cytokinin biosynthetic process	1								
alanine metabolic process	1								
ion transmembrane transport						2			2
asparagine metabolic process	1								
positive regulation of cellular process						2			
regulation of biological quality				3		2			
cellular component organization or biogenesis					2		2		2
purine ribonucleotide metabolic process						2			2
regulation of transcription, DNA-templated	1		2						2
cellular catabolic process					3		2		
glutathione metabolic process		1						1	
toxin catabolic process		1						1	
cellular carbohydrate metabolic process			2		2				
response to biotic stimulus			5			5			3
defense response			5			5			3
ribosome phosphate metabolic process			2						
organic substance transport			2		2	2			
organic cyclic compound biosynthetic process				3					
ion transport				3					
cellular nitrogen compound biosynthetic process				3					
aromatic compound biosynthetic process				3					
heterocycle biosynthetic process				3					
alpha-amino acid metabolic process				5					
nucleobase-containing small molecule metabolic process					2				
regulation of macromolecule metabolic process					2				
regulation of nitrogen compound metabolic process					2				
transcription, DNA-template					2	2			
regulation of cellular metabolic process					2				
regulation of primary metabolic process					2				
drug transport					2				
signal transduction					2			1	
nitrogen compound transport					2				
generation of precursor metabolites and energy						2			2
regulation of RNA metabolic process							2		
methylation								1	

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
transcription, DNA-template								1	
purine nucleobase transmembrane transport								1	
flavonoid biosynthetic process								1	
pyrimidine nucleobase metabolic process								1	
ATP transport								1	
purine nucleobase metabolic								1	
ADP transport								1	
rRNA processing								1	
ribosomal small subunit biogenesis								1	

In ‘Húsvéti rozmaring’, 8 DMGs were identified in the CpG context, 6 in CHG, and 34 in CHH in mother plant vs. *in vitro* plant comparison. The genes affecting the serine family amino acid metabolic process; carbohydrate metabolic process; aspartate metabolic process; cytokinin biosynthetic process; alanine metabolic process; asparagine metabolic process; regulation of transcription, DNA-templated processes only in CpG context were DMGs. While the genes affecting oxidation-reduction process; alkaloid biosynthetic process; glutathione metabolic process; toxin catabolic process processes, only in CHG contexts were significantly differently methylated. The genes affecting, proteolysis; ribosome biogenesis; cellular response to stimulus, regulation of protein complex assembly; translation; coenzyme metabolic process; nucleotide metabolic process; response to abiotic stimulus; cellular carbohydrate metabolic process; response to biotic stimulus; defense response; ribosome phosphate metabolic process; organic substance transport processes only in CHH context were significantly differently methylated. In the comparison of acclimatized vs. *in vitro* plants of ‘Húsvéti rozmaring’, these numbers are 46 in CpG, 57 in CHG and 40 in CHH, which are much higher than mother plant vs. *in vitro* plant or mother plant and acclimatized plant comparison. So, it can be said that there were many more methylated genes in acclimatized vs. *in vitro*, suggesting that they were down-regulated resulting in high methylation levels. In the CpG context, DMGs were related to carbohydrate metabolic process; organic cyclic compound biosynthetic process; ion transport; cellular nitrogen compound biosynthetic process; aromatic compound biosynthetic process; heterocycle biosynthetic process; alpha-amino acid metabolic process. In CHG context DMGs were related to the lignin biosynthetic process; coumarin biosynthetic process; pigment biosynthetic process; stilbene biosynthetic process; alkaloid biosynthetic process; tyrosine metabolic process; riboflavin metabolic process; cellular component organization or biogenesis; cellular catabolic process; cellular carbohydrate metabolic process; nucleobase-containing small molecule metabolic process; regulation of macromolecule metabolic process; regulation of nitrogen compound metabolic process; regulation of cellular metabolic process; regulation of primary metabolic process; drug transport; signal

transduction, participating in nitrogen compound transport processes. In the CpG and CHG contexts were DMGs in protein phosphorylation; transmembrane transport processes, while in the CpG, CHG and CHH contexts had DMGs in oxidation-reduction process. The DMGs were present in only CHH context at proteolysis; organelle organization; ribosome biogenesis; cellular response to stimulus; sucrose metabolic process; translation; starch metabolic process; ion transmembrane transport; positive regulation of cellular process; purine ribonucleotide metabolic process; response to biotic stimuli; defense response; generation of in precursor metabolites and energy processes, in CpG and CHH contexts at regulation of biological quality processes, in CHG and CHH contexts at organic substance transport; transcription, DNA-template process. In the comparison between the mother plant and acclimatized plant of 'Húsvéti rozmaring', we found 32 DMGs in CpG, 35 in CHG, and 33 in CHH contexts. This comparison contained less DMGs than the previous comparison (acclimatized vs. *in vitro* plants). DMGs were present in CpG and CHG contexts at lignin biosynthetic process; coumarin biosynthetic process; pigment biosynthetic process; protein phosphorylation; carbohydrate metabolic process; stilbene biosynthetic proces; alkaloid biosynthetic process; tyrosine metabolic process; riboflavin metabolic process. In CpG, CHG and CHH contexts, the DMGs were participating in oxidation-reduction process. In CHH context, the DMGs were affecting in ribosome biogenesis; cellular response to stimulus; sucrose metabolic process; translation; coenzyme metabolic process; starch metabolic process; ion transmembrane transport; purine ribonucleotide metabolic process; regulation of transcription, DNA-templated; response to biotic stimulus; defense response; generation of precursor metabolites and energy processes. The cellular component organization or biogenesis in CpG and CHH contexts, and cellular catabolic process; regulation of RNA in metabolic process processes in CpG context were differentially methylated. In the CHG context, the DMGs were present in glutathione metabolic process; toxin catabolic process; signal transduction processes; methylation; transcription, DNA-template; purine nucleobase transmembrane transport; flavonoid biosynthetic process; pyrimidine nucleobase metabolic process; ATP transport; purine nucleobase metabolic; ADP transport; rRNA processing; ribosomal small subunit biogenesis.

Table 12. Molecular functions in 'McIntosh' affected by methylation (yellow: present and number of DMGs, blue: absent of DMGs)

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
oxidoreductase activity, acting on paired donors with incorporation or reduction of molecular oxygen	4			2					
carbohydrate derivative binding	4						3		

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
small molecule binding	4								
monooxygenase activity	4			2					
drug binding	4						3		
anion binding	5						5		
catalytic activity, acting on a protein	5	6				3	3		3
transferase activity	6						5		7
protein binding	11	12	4	11	6	5	13	13	6
hydrolase activity	9	6		2			5	6	6
metal ion binding	8	4				4	6	5	3
heterocyclic compound binding	8								
organic cyclic compound binding	8								
DNA binding		3	2			3		3	3
transferase activity, acyl groups		3							
monophenol monooxygenase activity		3			1		3	3	
transcription regulator activity		3							
catechol oxidase activity		3			1		3	3	
transferase activity, phosphorus-containing groups		3							
ATP binding		5	2	4	4	4		7	3
transferase activity, transferring hexosyl groups			2						
structural constituent of ribosome			2						
endopeptidase activity			2						
transition metal ion binding			2	4					
ATPase activity			2			3			
transferase activity, transferring acyl groups other than amino-acyl groups			2						
oxidoreductase activity			2						
protein kinase activity				4	4			4	
S-adenosylmethionine-dependent methyltransferase activity					1				
oxidoreductase activity, oxidizing metal ions					1				
copper ion binding					1				
kinase activity						3			
nucleotide binding							3		
nucleic acid binding							3		
transferase activity, transferring acyl groups								3	

A total of 153 DMGs were identified between the 'McIntosh' mother plant and *in vitro* plant (Table 12), of which 80 were in CpG, 51 in CHG, and 22 in CHH contexts. DMGs were present in CHG context at DNA binding; transferase activity, acyl groups; monophenol monooxygenase activity; transcription regulator activity; catechol oxidase activity; transferase activity, phosphorus-containing groups. We found genes which were differently methylated in CHH context at transferase activity, transferring hexosyl groups; structural constituent of ribosome; endopeptidase activity; transition metal ion binding; ATPase activity; transferase activity, transferring acyl groups other than amino-acyl groups; oxidoreductase activity processes,

while in CpG context had DMGs in oxidoreductase activity, acting on paired donors with incorporation or reduction of molecular oxygen; carbohydrate derivative binding; small molecule binding; monooxygenase activity; drug binding; anion binding; transferase activity; heterocyclic compound binding; organic cyclic compound binding processes. In the CpG and CHG contexts were DMGs in catalytic activity, acting on a protein; metal ion binding, while in CHG and CHH contexts had DMGs in DNA binding; ATP binding. In the comparison of acclimatized vs. *in vitro* plants of ‘McIntosh’, these numbers are 29 in CpG, 19 in CHG and 25 in CHH. In CpG context DMGs were related to oxidoreductase activity, acting on paired donors with incorporation or reduction of molecular oxygen; monoxygenase activity; transition metal ion binding. We found genes which were differently methylated in CHG context at S-adenosylmethionine-dependent methyltransferase activity; oxidoreductase activity; oxidizing metal ion; copper ion binding. The ATPase activity and kinase activity in CpG context were differentially methylated. In the CpG, CHG and CHG contexts were DMGs in protein binding; ATP binding. DMGs were present in CpG and CHG context at protein kinase activity. In the comparison between the mother plant and acclimatized plant, we found 55 DMGs in CpG, 47 in CHG, and 25 in CHH contexts. In CpG context DMGs were related to carbohydrate derivative binding; drug binding; anion binding; nucleotide binding; nucleic acid binding. The protein kinase activity and transferase activity, transferring acyl groups were always methylated in CHG context. We found genes which were differently methylated in CpG and CHG context at monophenol monooxygenase activity and catechol oxidase activity. The DNA binding and ATP binding in CHG and CHH context were differentially methylated. In CpG and CHH contexts DMGs were related to catalytic activity, acting on a protein; transferase activity. The protein binding hydrolase activity; metal ion binding functions were the only one that appeared in all comparisons and in all contexts.

Table 13. Molecular functions in ‘Húsvéti rozmaring’ affected by methylation (yellow: present and number of DMGs, blue: absent of DMGs)

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
oxidoreductase activity, acting on paired donors with incorporation or reduction of molecular oxygen							3		
anion binding									2
protein binding	1	2	3	17	13	5	11	10	5
hydrolase activity				8	4		5		
metal ion binding			3		5				
DNA binding			3			3			3
monophenol monooxygenase activity								2	
catechol oxidase activity								2	

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
ATP binding	1	1	2	10	10	2	6	6	
transferase activity, transferring hexosyl groups						2			2
structural constituent of ribosome			2			2			2
endopeptidase activity			2			2			
transition metal ion binding				5			3		2
ATPase activity			2						
transferase activity, transferring acyl groups other than amino-acyl groups	1		2			2			2
oxidoreductase activity				5					3
protein kinase activity		1		8	8		4	5	
nucleotide binding									2
nucleic acid binding					3			3	
asparaginase activity	1								
alpha-L-fucosidase activity	1								
sequence-specific-DNA binding	1								
quercetin 3-O-glucosyltransferase activity	1								
intramolecular lyase activity	1								
flavonol 7-O-beta-glucosyltransferase activity	1								
hydrolase activity, acting on ester bonds	1								
quercetin 7-O-glucosyltransferase activity	1								
glutathione transferase activity		1							
salutaridine reductase (NADPH) activity		1							
lyase activity			2			2			
oxidoreductase activity, acting on diphenols and related substances as donors, oxygen as acceptor					3				
transmembrane transporter activity					3			2	
UDP-glucosyltransferase activity					3		3	2	
calcium ion binding						2			2
ATPase activity, coupled to transmembrane movement of substance						2			2
ion transmembrane transporter activity						2			2
monooxygenase activity							3		
zinc ion binding								2	

In the case of ‘Húsvéti rozmaring’, a total of 270 genes were methylated in the molecular functions (Table 13). A total of 38 DMGs were identified in the ‘Húsvéti rozmaring’ mother plant vs. *in vitro* plant comparison, of which 11 were in CpG, 6 in CHG, and 21 in CHH contexts. In the asparaginase activity; alpha-L-fucosidase activity; sequence-specific-DNA binding; quercetin 3-O-glucosyltransferase activity; intramolecular lyase activity; flavonol 7-O-beta-glucosyltransferase activity; hydrolase activity, acting on ester bonds; quercetin 7-O-glucosyltransferase activity DMGs were in CpG context. The protein kinase activity; glutathione transferase activity; salutaridne reductase (NADPH) activity in CHG context were differentially methylated. We found genes which were differently methylated in CHH at metal ion binding; DNA binding; structural constituent of ribosome; endopeptidase activity; ATPase activity; lyase

activity. In CpG, CHG and CHG context DMGs were related to protein binding; ATP binding. In the CpG and CHH contexts were DMGs in transferase activity, transferring acyl groups other than amino-acyl groups. In the acclimatized vs. *in vitro* comparison, 131 DMGs were identified, of which 53, 52, and 26 were found in the contexts of CpG, CHG, and CHH, respectively. In the transition metal ion binding; oxidoreductase activity DMGs were in CpG context. We found genes which were differently methylated in CHG at nucleic acid binding; oxidoreductase activity, acting on diphenols and related substances as donors, oxygen as acceptor; transmembrane transporter activity; UDP-glucosyltransferase activity. The DNA binding; transferase activity, transferring hexosyl groups; structural constituent of ribosome; transferase activity, transferring acyl groups other than amino-acyl groups; lyase activity; calcium ion binding; ATPase activity, coupled to transmembrane movement of substance; endopeptidase activity; ion transmembrane transporter activity in CHH context were differentially methylated. In the hydrolase activity, protein kinase activity DMGs were in CpG and CHG context. We found genes which were differently methylated in CpG, CHG and CHH at protein binding; ATP binding. In the mother plant vs. acclimatized plant comparison, 101 DMGs were found, of which 38, 34, and 29 were in the CpG, CHG, or CHH contexts, respectively. We found genes which were differently methylated in CpG at oxidoreductase activity, acting on paired donors with incorporation or reduction of molecular oxygen; hydrolase activity; monooxygenase activity. The monophenol monooxygenase activity; catechol oxidase activity; nucleic acid binding; transmembrane transporter activity; zinc ion binding in CHG context were differentially methylated. In the CHH contexts were DMGs in anion binding; DNA binding; transferase activity, transferring hexosyl groups; structural constituent of ribosome; transferase activity, transferring acyl groups other than amino-acyl groups; oxidoreductase activity; nucleotide binding; calcium ion binding; ATPase activity, coupled to transmembrane movement of substance; ion transmembrane transporter activity, while in CpG and CHG contexts had DMGs in ATP binding; protein kinase activity; UDP-glucosyltransferase activity. The transition metal ion binding in CpG and CHH contexts were differentially methylated. The protein binding function was the only one that appeared in all comparisons and in all contexts. ATP binding was found to be in almost all contexts, with the exception of the mother plant vs. acclimatized plant in CHH context.

Table 14. Cellular components in 'McIntosh' affected by methylation (yellow: present and number of DMGs, blue: absent of DMGs)

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
integral component of plasma membrane	1						1		
extracellular region	1						1		
protein serine/threonine phosphatase complex	1						1		
cytoplasm	1			1					
cell wall	2						1		
nucleus	3		2	1		2	2	4	2
transcription factor complex	2						1		
intrinsic component of plasma membrane		2						2	
catalytic complex		2							
nuclear lumen		2							
intracellular non-membrane-bounded organelle		2				3		2	
integral component of membrane		7	5	2	2	8		7	6
cytoplasmic part		4						4	
endoplasmic reticulum			2			2			2
intracellular organelle part			2			3		2	3
ribosome			2						2
plasmodesma				1	1		1		
plasma membrane				1					
cytosol					1				
membrane protein complex						2			2
intracellular organelle lumen								2	
protein-containing complex								3	

A total of 22 processes were identified in the cellular processes (Table 14), which were affected by different methylation. Between the mother plant and *in vitro* plant, 43 DMGs were found, of which 11, 19, and 13 were differently methylated in the CpG, CHG, and CHH contexts, respectively. In the CpG context was DMGs in integral component of plasma membrane; extracellular region; protein serine/threonine phosphatase complex; cytoplasm; cell wall; nucleus; transcription factor complex, while in CHG context had DMGs in intrinsic component of plasma membrane; catalytic complex; nuclear lumen; intracellular non-membrane-bounded organelle; cytoplasmic part. The endoplasmic reticulum; intracellular organelle part; ribosome in CHH context were differentially methylated. In the CHG and CHH contexts were DMGs in integral component of membrane. A total of 30 DMGs were identified between acclimatized plant and *in vitro* plant, of which 6 were methylated in the context of CpG, 4 in CHG, and 20 in CHH contexts. We found genes which were differently methylated in CpG at cytoplasm; nucleus; plasma

membrane, while in CHG context had DMGs in cytosol. The intracellular non-membrane-bounded organelle; endoplasmic reticulum; intracellular organelle part; membrane protein complex in CHH context were differentially methylated. We found genes which were differently methylated in CpG and CHG at plasmodesma. The integral component of membrane was found in all contexts. A total of 51 DMGs were found between the mother plant and the acclimatized plant, of which 8 in CpG, 26 in CHG, and 17 in CHH were differently methylated. The DMGs were in a similar motif at integral component of plasma membrane; extracellular region; protein serine/threonine phosphatase complex; cell wall; nucleus; transcription factor complex in CpG context, while in CHG context had DMGs in intrinsic component of plasma membrane; intracellular non-membrane-bounded organelle; integral component of membrane; cytoplasmic part and endoplasmic reticulum; intracellular organelle part; ribosome in CHH context in the mother plant and *in vitro* plant as well as, in the mother plant and acclimatized plant comparisons.

Table 15. Cellular components in 'Húsvéti rozmaring' affected by methylation (yellow: present and number of DMGs, blue: absent of DMGs)

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
cytoplasm		1		4	2		4		
nucleus	2			3	2	2	2	2	2
transcription factor complex	1								
integral component of membrane	1		6	10	9	6	7	6	6
endoplasmic reticulum			3						2
intracellular organelle part									2
ribosome			2			2			2
plasma membrane				2		1			
membrane protein complex			2						2
extrinsic component of membrane						1			
lysosome						1			
vacuolar proton-transporting V-type ATPase, V0 domain						1			
nucleosome						1			
Sec61 translocon complex						1			
anchored component of membrane						1			
Golgi apparatus						1			
Arp2/3 protein complex						1			
plastid						1			
vacoular proton-transporting V-type ATPase complex						1			
photosystem II oxygen evolving complex						1			
small-subunit processome								1	
chloroplast membrane								1	
membrane		1							

In the ‘Húsvéti rozmaring’ mother plant vs. *in vitro* plant comparison (Table 15), 20 DMGs were identified, of which 5, 2, and 13 genes were differently methylated in CpG, CHG and CHH contexts, respectively. The nucleus; transcription factor complex; integral component of membrane had CpG, cytoplasm; membrane had DMGs in CHG context. The endoplasmic reticulum; ribosome; membrane protein complex had DMGs only in CHH context while in CpG and CHH contexts had DMGs in integral component of membrane. In the acclimatized vs. *in vitro* plant comparison, 54 DMGs were found, of which 19 were in CpG, 13 in CHH and 22 in CHH contexts. In CHH context, DMGs were present in the ribosome; extrinsic component of membrane; lysosome; vacoular proton-transporting V-type ATPase, V0 domain; nucleosome; Sec61 translocon complex; anchored component of membrane; Golgi apparatus; Arp2/3 protein complex; plastid; vacuolar proton-transporting V-type ATPase complex; photosystem II oxygen evolving complex. The cytoplasm at DMGs in CpG and CHG context while in CpG and CHH contexts had DMGs in plasma membrane. The nucleus; integral component of membrane found to be in all contexts in the acclimatized vs. *in vitro* plant comparison. For the mother plant vs. acclimatized plant, 39 DMGs were found, of which 13 were in CpG, 10 in CHG, and 16 in CHH contexts. The endoplasmic reticulum; intracellular organelle part; ribosome; membrane protein complex only was affected if the methylation occurred in CHH context, while small-subunit process; chloroplast membrane process was affected by different methylation only in CHG context. We found genes which were differently methylated in CpG, CHG and CHH at nucleus; integral component of membrane, while in CpG context had DMGs in cytoplasm.

Table 16. Biological processes in ‘McIntosh’ vs. ‘Húsvéti rozmaring’ affected by methylation (yellow: present and number of DMGs, blue: absent of DMGs)

	Acclimatized vs. acclimatized			*In vitro* vs. *in vitro*			Mother plant vs. mother plant		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
gene expression	4						6		
cellular macromolecule biosynthetic process	4			6			6		
response to stimulus	4			5				4	
cellular component biogenesis	4			4			4		
transmembrane transport	4	5		4	4			4	
organonitrogen compound biosynthetic process	4				3		4		
DNA metabolic process	4			4			5		
regulation of macromolecule metabolic process	5								
regulation of macromolecule process	5								
regulation of cellular metabolic process	5			6					
regulation of primary metabolic process	5	4		6					
protein phosphorylation	5	7		6	8		5	7	
oxidation- reduction process	12	7	3	9	5	3	9	7	3

	Acclimatized vs. acclimatized			*In vitro* vs. *in vitro*			Mother plant vs. mother plant		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
alpha-amino acid metabolic process	7			6			5	4	
cellular nitrogen compound biosynthetic process	6			6			6		
organic cyclic compound biosynthetic process	6			5			6		
heterocycle biosynthetic process	5			4			5		
aromatic compound biosynthetic process	5			4			5		
cellular component organization	5	5		5				4	
drug metabolic process	5						4		
response to chemical		4			3				
regulation of nitrogen compound metabolic process	5	4		6					
regulation of cellular macromolecule biosynthetic process		4							
secondary metabolic process		4							
peptide transport		4						4	
regulation of gene expression		4		4					
alkaloid biosynthetic process		4						4	
small molecule metabolic process		6							
transcription, DNA-templated		5							2
ribosome biogenesis			2						2
cellular response to stimulus			2			2			
phosphorylation			2			2			
sucrose metabolic process			2			2			2
organelle organization			2		3				
translation			2						2
coenzyme metabolic process			2			2			2
positive regulation of cellular process			2						
innate immune response			2			2			2
generation of precursor metabolites and energy			2						
carboxylic acid metabolic process			2			2			2
regulation of response to stimulus			2						
nucleobase-containing compound biosynthetic process			2						
defense response to other organism			2			2			2
purine ribonucleotide metabolic process			2						
regulation of biological quality			2						
ion transmembrane transport			2						
starch metabolic process			2			2			2
RNA metabolic process				4					
cellular amide metabolic process				4		2			
intracellular protein transport					3				
regulation of transcription, DNA-templated					3	2		4	
nucleobase-containing small molecule metabolic process					3				
proteolysis						2			2
cellular protein metabolic process						2			
regulation of protein complex assembly						2			
nucleotide metabolic process						2			2

	Acclimatized vs. acclimatized			*In vitro* vs. *in vitro*			Mother plant vs. mother plant		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
ribose phosphate metabolic process						2			2
organic substance transport						2			2
response to abiotic stimulus						2			
response to stress							4		
regulation of RNA metabolic process							4		
cellular catabolic process							4		
lipid biosynthetic process							4		
cellular lipid metabolic process							4		
regulation of cellular process									2

In this comparison, we examined what changes occurred in biological processes, molecular functions, and cellular composition when compared 'McIntosh' with 'Húsvéti rozmaring' in the same environmental origin. In the comparison of acclimatized 'McIntosh' vs. acclimatized 'Húsvéti rozmaring' (Table 16), 210 DMGs were identified, of which 104 were in CpG, 67 in CHG, and 39 in CHH contexts. We identified DMGs in CpG context at gene expression; cellular macromolecule biosynthetic process; response to stimulus; cellular component biogenesis; organonitrogen compound biosynthetic process; DNA metabolic process; regulation of macromolecule metabolic process; regulation of macromolecule process; regulation of cellular metabolic process; regulation of primary metabolic process; protein phosphorylation; alpha-amino acid metabolic process; cellular nitrogen compound biosynthetic process; organic cyclic compound biosynthetic process; heterocycle biosynthetic process; aromatic compound biosynthetic process; drug metabolic process. In CHG context, DMGs were present in the response to chemical; regulation of cellular, macromolecule biosynthetic process; secondary metabolic process; peptide transport; regulation of gene expression; alkaloid biosynthetic process; small molecule metabolic process; transcription and DNA-templated. In CHH context, we observed genes in the ribosome biogenesis; cellular response to stimulus; phosphorylation; sucrose metabolic process; organelle organization; translation; coenzyme metabolic process; positive regulation of cellular process; innate immune response; generation of precursor metabolites and energy; carboxylic acid metabolic process; regulation of response to stimulus; nucleobase-containing compound biosynthetic process; defense response to other organism; purine ribonucleotide metabolic process; regulation of biological quality; ion transmembrane transport and starch metabolic process were differently methylated. The transmembrane transport; regulation of primary metabolic process; protein phosphorylation; cellular component organization and regulation of nitrogen compound metabolic process had DMGs in CpG and CHG contexts. The oxidation- reduction process found to be in all contexts in the comparison of acclimatized 'McIntosh' vs. acclimatized 'Húsvéti rozmaring'. In *in vitro* 'McIntosh' vs. *in vitro*

‘Húsvéti rozmaring’, 170 DMGs were identified, of which 98 were in CpG, 35 in CHG, and 37 in CHH contexts. In CpG context, we identified differently methylated genes in cellular macromolecule biosynthetic process; response to stimulus; cellular component biogenesis; DNA metabolic process; regulation of cellular metabolic process; regulation of primary metabolic process; alpha-amino acid metabolic process; cellular nitrogen compound biosynthetic process; organic cyclic compound biosynthetic process; heterocycle biosynthetic process; aromatic compound biosynthetic process; cellular component organization; regulation of nitrogen compound metabolic process; regulation of gene expression; RNA metabolic process and cellular amide metabolic process. We specified genes in CHG context at organonitrogen compound biosynthetic process; response to chemical; organelle organization; intracellular protein transport and nucleobase-containing small molecule metabolic process only process, which were differently methylated. DMGs were in the CHH context at cellular response to stimulus; phosphorylation; sucrose metabolic process; coenzyme metabolic process; innate immune response; carboxylic acid metabolic process; defense response to other organism; starch metabolic process; proteolysis; cellular protein metabolic process; regulation of protein complex assembly; nucleotide metabolic process; ribose phosphate metabolic process; organic substance transport and response to abiotic stimulus process. We specified genes in CpG and CHG contexts at protein phosphorylation and transmembrane transport. In the comparison of mother plants of the two cultivars, 163 DMGs were found, of which 90 were in CpG, 42 in CHG, and 31 in CHH contexts. In CpG context, we found differently methylated genes in gene expression process; cellular macromolecule biosynthetic process; cellular component biogenesis; organonitrogen compound biosynthetic process; DNA metabolic process; cellular nitrogen compound biosynthetic process; organic cyclic compound biosynthetic process; heterocycle biosynthetic process; aromatic compound biosynthetic process; drug metabolic process; response to stress; regulation of RNA metabolic process; cellular catabolic process; lipid biosynthetic process and cellular lipid metabolic process. We identified DMGs in CHG context at response to stimulus; transmembrane transport; cellular component organization; peptide transport; alkaloid biosynthetic process and regulation of transcription, DNA-templated. We observed genes in CHH context at transcription, DNA-templated; ribosome biogenesis; sucrose metabolic process; translation; coenzyme metabolic process; innate immune response; carboxylic acid metabolic process; defense response to other organism; starch metabolic process; proteolysis; nucleotide metabolic process; ribose phosphate metabolic process; organic substance transport process and regulation of cellular process, which were differently methylated. We specified genes in CpG and CHG context at alpha-amino acid metabolic process. It can also be read from the table that only the oxidation-reduction process was the only one where we found DMGs in all comparisons and in all contexts.

Table 17. Molecular functions in 'McIntosh' vs. 'Húsvéti rozmaring' affected by methylation (yellow: present and number of DMGs, blue: absent of DMGs)

	Acclimatized vs. acclimatized			*In vitro* vs. *in vitro*			Mother plant vs. mother plant		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
monooxygenase activity	5								
ligase activity	5						5		
oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	5								
protein kinase activity	5	7		6	8		5	7	
transition metal binding	5								
hydrolase activity	6	11		10	8		7	7	
protein binding	18	16	4	20	14	4	19	14	4
ATP binding	9	11	3	9	12	4	8	11	3
transmembrane transporter activity		5							
oxidoreductase activity		7	2	9	5	2	9	7	2
metal ion binding		7		8	6			6	
nucleic acid binding		9		7	9				
transferase activity, transferring hexosyl groups			2			2			2
structural constituent of ribosome			2						2
catalytic activity, acting on protein			2						
calcium ion binding			2			2			
ATPase activity, coupled to transmembrane movement of substances			2						
lyase activity			2						
ion transmembrane transporter activity			2						
transferase activity, transferring acyl groups other than amino-acyl groups			2			2			2
kinase activity			2			2			2
endopeptidase activity						2			2
DNA binding						3	5	5	3
ATPase activity						2			2
transition metal ion binding							5		2

In the acclimatized 'McIntosh' plant vs. acclimatized 'Húsvéti rozmaring' plant comparison (Table 17), 158 DMGs were found, of which 58 were in CpG, 73 in CHG and 27 in CHH contexts. In CpG context, we identified DMGs at monooxygenase activity; ligase activity; oxidoreductase activity, acting on paired donors with incorporation or reduction of molecular oxygen and transition metal binding. In CHG context, genes at transmembrane transporter activity; metal ion binding and nucleic acid binding process were differently methylated, while in CHH context at transferase activity, transferring hexosyl groups; structural constituent of ribosome; catalytic activity, acting on protein; calcium ion binding; ATPase activity, coupled to transmembrane movement of substances; lyase activity; ion transmembrane transporter activity; transferase activity, transferring acyl groups other than amino-acyl groups and kinase activity

process were differently methylated. The protein kinase activity was differently methylated in CpG and CHG contexts. In *in vitro* 'McIntosh' plant vs. *in vitro* 'Húsvéti rozmaring' plant, 156 DMGs, of which were 69 in CpG, 62 in CHG, and 25 in CHH contexts. In CpG and CHG contexts, DMGs were present in protein kinase activity; hydrolase activity; metal ion binding and nucleic acid binding. In CpG, CHG and CHH contexts, we observed DMGs at protein binding; ATP binding and oxidoreductase activity processes. We identified genes only in CHH context at transferase activity, transferring hexosyl groups; calcium ion binding; transferase activity, transferring acyl groups other than amino-acyl groups; kinase activity; endopeptidase activity and ATPase activity, which were differently methylated. In the comparison of mother plants of the two cultivars, 146 DMGs were found, of which were 63 in CpG, 57 in CHG and 26 in CHH contexts. We determined DMGs in CpG context at ligase activity. In CpG and CHG context, DMGs were affecting in protein kinase activity and hydrolase activity. In CpG, CHG and CHH context, we observed genes at protein binding; ATP binding; oxidoreductase activity and DNA binding process, which were differently methylated. We identified DMGs in metal ion binding processes which were present only in CHG context, while DMGs were present in transferring hexosyl groups; structural constituent of ribosome; transferase activity, transferring acyl groups other than amino-acyl groups; kinase activity; ATPase activity and endopeptidase activity process in CHH context. In CpG and CHH contexts, DMGs were identified in transition metal ion binding process. It can also be seen from the table that we found DMGs for both protein binding and ATP binding in all comparisons and in all contexts. For oxidoreductase activity, we found DMGs in all comparisons except in the CpG context at comparison of acclimatized plants of the two cultivars.

Table 18. Cellular components in 'McIntosh' vs. 'Húsvéti rozmaring' affected by methylation (yellow: present and number of DMGs, blue: absent of DMGs)

	Acclimatized vs. acclimatized			*In vitro* vs. *in vitro*			Mother plant vs. mother plant		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
catalytic complex	2								
plasma membrane	2	3		3					
ribonucleoprotein complex	2			3					
intracellular non-membrane-bounded organelle	2			3		3			
cytoplasmic part	2			4			4		
intracellular organelle part	3		2	3		3	3		3
nucleus	5	3		6	3		7	3	2
integral component of membrane	16	16	7	17	13	8	11	13	7
organelle membrane		3			3			3	
organelle subcompartment		3							
cytoplasmic vesicle		3			3			3	
endomembrane system		4			3			3	

	Acclimatized vs. acclimatized			*In vitro* vs. *in vitro*			Mother plant vs. mother plant		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
protein-containing complex		8			8		6	8	
membrane protein complex			2			2			2
endoplasmic reticulum			2			3			3
ribosome			2						2

In the case of acclimatized plants of the two cultivars in cellular components (Table 18), we identified 92 DMGs, of which 34 were in CpG, 43 in CHG and 15 in CHH contexts. We observed DMGs only in CpG context at catalytic complex; ribonucleoprotein complex; intracellular non-membrane-bounded organelle; cytoplasmic part process and intracellular organelle part, while DMGs were in plasma membrane and nucleus process in CpG and CHG contexts. In CpG, CHG and CHH contexts, DMGs were present in integral component of membrane process, while only in CHG context were genes with different methylation in organelle membrane; organelle subcompartment; cytoplasmic vesicle; endomembrane system and protein-containing complex process. In CHH context, DMGs were in membrane protein complex; endoplasmic reticulum; intercellular organelle part and ribosome process. In the comparison of *in vitro* plants of the two cultivars in cellular components, 91 DMGs were found, of which 39 were in CpG, 33 in CHG and 19 in CHH contexts. We determined DMGs only in CpG context at plasma membrane; ribonucleoprotein complex and cytoplasmic part process, while genes at nucleus process were differently methylated in CpG and CHG contexts. DMGs were present in CpG and CHH contexts at intracellular non-membrane-bounded organelle and intercellular organelle part process, while genes in integral component of membrane process, were differently methylated in the CpG, CHG and CHH contexts. We found genes in CHG context at organelle membrane; cytoplasmic vesicle; endomembrane system and protein-containing complex process, while DMGs were present in membrane protein complex and endoplasmic reticulum process in the CHH context. In the comparison of mother plants of the two cultivars in cellular components, 83 DMGs were identified, of which 31 were in CpG, 33 in CHG, and 19 in CHH contexts. The cytoplasmic part and intercellular organelle part were differently methylated only in CpG content. The intracellular organelle part; membrane protein complex; endoplasmic reticulum and ribosome were differently methylated in CHH contexts. In CpG, CHG and CHH contexts, DMGs were present at the nucleus and integral component of membrane intercellular organelle part, while genes in CHG context was differently methylated at the organelle membrane; cytoplasmic vesicle and endomembrane system process. We determined DMGs in CpG and CHG contexts at protein-containing complex. It can also be read from the table that only the integral component of membrane was the only one where we found DMGs in all comparisons and in all contexts.

Table 19. Biological processes in average of both scion affected by methylation (yellow: present and number of DMGs, blue: absent of DMGs)

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
transmembrane transport								4	
organonitrogen compound biosynthetic process								2	
regulation of macromolecule metabolic process				2					
regulation of cellular metabolic process				2					
regulation of primary metabolic process				2					
protein phosphorylation	2	1		5	4		3	5	
oxidation- reduction process	5	6	2	4	3	3	2	4	3
alpha-amino acid metabolic process								2	
regulation of nitrogen compound metabolic process				2					
alkaloid biosynthetic process	3	4		2	3		1		
transcription, DNA-templated			2						
ribosome biogenesis	1					2			2
cellular response to stimulus						3			3
phosphorylation									2
sucrose metabolic process			2			2			2
organelle organization						2			2
translation						2			2
coenzyme metabolic process			2			2			2
positive regulation of cellular process						2			2
innate immune response			2			2			
generation of precursor metabolites and energy						2			2
carboxylic acid metabolic process									2
regulation of response to stimulus						2			
defense response to other organism			2			2			
purine ribonucleotide metabolic process						2			2
regulation of biological quality				2		2			2
ion transmembrane transport						2	1		2
starch metabolic process			2			2			2
cellular amide metabolic process			2						
regulation of transcription, DNA-templated	1					2		2	2
proteolysis	1		2			2			
regulation of protein complex assembly						2			2
nucleotide metabolic process			2						
organic substance transport								3	
response to abiotic stimulus						2			
response to stress	2	2					2	2	
lignin biosynthetic process	3	3		2	2		1		
coumarin biosynthetic process	3	3		2	2		1		
pigment biosynthetic progress	3	3		2	2		1		
ion transport				2				2	
riboflavin metabolic process	3	3		2	2		1		

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
carbohydrate metabolic process				2	1				
stilbene biosynthetic process	3	3		2	2		1		
tyrosine metabolic process	3	3		2	2		1		
RNA-dependent DNA biosynthetic process					1				
flavonoid biosynthetic process					1		1		
xenobiotic transport					1				
ATP transport					1				
drug transmembrane transport					1				
regulation of transcription by RNA polymerase II					1				
glutathione metabolic process		1			1		1		
ribosomal small subunit biogenesis					1				
mitotic chromosome condensation					1				
ADP transport					1				
defense response					1				5
rRNA processing					1				
toxin catabolic process		1			1		1		
cell differentiation		1			1				
response to biotic stimulus					1				5
response to oxygen-containing compound	1	1					1		
protein dephosphorylation		1					1		
oligopeptide transmembrane transport		1					1		
response to acid chemical	1	1					1		
purine nucleobase transmembrane transport							1		
regulation of protein ubiquitination		1					1		
serine family amino acid metabolic process	1						2		
RNA processing								2	
ribonucleoprotein complex biogenesis								2	
drug transport								2	
nitrogen compound transport								3	
cytokinin biosynthetic process	1								
fructose metabolic process	1								
ALTERNATE of GO:0006355	1								
mannose metabolic process	1								
aspartate metabolic process	2								
alanine metabolic process	2								
asparagine metabolic process	2								
methylation		1							
pyrimidine nucleobase metabolic process		1							
protein SUMOylation		1							
recruitment of 3'-end processing factors to RNA polymerase II holoenzyme complex		1							
purine nucleobase metabolic process		1							
histone modification		1							
positive regulation of transcription elongation from RNA polymerase II promoter		1							
ribose phosphate metabolic process			2						

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
carbohydrate derivate metabolic process	1						1		
regulation of biological process			2						

In the mother plant vs. *in vitro* plant comparison, 117 DMGs were identified, of which 47, 46 and 24 were in the CpG, CHG, and CHH contexts, respectively. In the CpG and CHG contexts, DMGs were present in protein phosphorylation; alkaloid biosynthetic process; response to stress; lignin biosynthetic process; coumarin biosynthetic process; pigment biosynthetic progress; riboflavin metabolic process; stilbene biosynthetic process; tyrosine metabolic process; response to oxygen-containing compound and response to acid chemical processes. DMGs were in the CpG, CHG and CHH contexts at the oxidation-reduction process, while genes were differently methylated at transcription, DNA-templated; sucrose metabolic process; coenzyme metabolic process; innate immune response; defense response to other organism; starch metabolic process; cellular amide metabolic process; nucleotide metabolic process; ribose phosphate metabolic process and regulation of biological process only in CHH context. We identified DMGs in CpG context at ribosome biogenesis; regulation of transcription DNA-templated; serine family amino acid metabolic process; cytokinin biosynthetic process; fructose metabolic process; ALTERNATE of GO: 0006355; mannose metabolic process; aspartate metabolic process; alanine metabolic process; asparagine metabolic process and carbohydrate derivate metabolic process, while in CHG context, DMGs were present in glutathione metabolic process; toxin catabolic process; cell differentiation; protein dephosphorylation; oligopeptide transmembrane transport; regulation of protein ubiquitination; methylation; pyrimidine nucleobase metabolic process; protein SUMOylation; recruitment of 3'-end processing factors to RNA, polymerase II holoenzyme complex; purine nucleobase metabolic process; histone modification and positive regulation of transcription elongation from RNA polymerase II promoter processes. In the acclimatized plant vs. *in vitro* plant comparison (Table 19), 117 DMGs were identified, of which 37, 38, 42 were in the CpG, CHG, and CHH contexts, respectively. In the CpG context, DMGs were present at regulation of macromolecule metabolic process; regulation of cellular metabolic process; regulation of primary metabolic process; regulation of nitrogen compound metabolic process and ion transport processes. We observed DMGs in CpG and CHG contexts at protein phosphorylation; alkaloid biosynthetic process; lignin biosynthetic process; coumarin biosynthetic process; pigment biosynthetic progress; riboflavin metabolic process; carbohydrate metabolic process; stilbene biosynthetic process and tyrosine metabolic process, while genes were differently methylated at oxidation-reduction process in CpG, CHG and CHH contexts. DMGs were present only in CHH context at ribosome biogenesis; cellular response to stimulus; sucrose metabolic process; organelle

organization; translation; coenzyme metabolic process; positive regulation of cellular process; innate immune response; generation of precursor metabolites and energy; regulation of response to stimulus; defense response to other organism; purine ribonucleotide metabolic process; ion transmembrane transport; starch metabolic process; regulation of transcription, DNA-templated; proteolysis; regulation of protein complex assembly and response to abiotic stimulus. In the CpG and CHH contexts, we defined genes which were differently methylated at regulation of biological quality, while DMGs were present only in CHG context at RNA-dependent DNA in biosynthetic process; flavonoid biosynthetic process; xenobiotic transport; ATP transport; drug transmembrane transport; regulation of transcription by RNA polymerase II; glutathione metabolic process; ribosomal small subunit biogenesis; mitotic chromosome condensation; ADP transport; defense response; rRNA processing; toxin catabolic process; cell differentiation and response to biotic stimulus. In the mother plant vs. acclimatized plant comparison, 108 DMGs were identified, of which 27, 35 and 46 were in the CpG, CHG and CHH contexts, respectively. We observed DMGs only in CHG context at transmembrane transport; organonitrogen compound biosynthetic process; alpha-amino acid metabolic process; organic substance transport; ion transport; RNA processing; ribonucleoprotein complex biogenesis; drug transport and nitrogen compound transport. DMGs were present in CpG and CHG contexts at protein phosphorylation and response to stress processes, while genes were differently methylated at oxidation-reduction process in CpG, CHG and CHH contexts. In the CpG context, we analysed DMGs at alkaloid biosynthetic process; lignin biosynthetic process; coumarin biosynthetic process; pigment biosynthetic progress; riboflavin metabolic process; stilbene biosynthetic process; tyrosine metabolic process; flavonoid biosynthetic process; glutathione metabolic process; toxin catabolic process; response to oxygen-containing compound; protein dephosphorylation; oligopeptide transmembrane transport; response to acid chemical; purine nucleobase transmembrane transport; regulation of protein ubiquitination; serine family amino acid metabolic process and carbohydrate derivate metabolic process. DMGs were present in CHH context at ribosome biogenesis; cellular response to stimulus; phosphorylation; sucrose metabolic process; organelle organization; translation; coenzyme metabolic process; positive regulation of cellular process; generation of precursor metabolites and energy; carboxylic acid metabolic process; purine ribonucleotide metabolic process; regulation of biological quality; starch metabolic process; regulation of protein complex assembly; defense response and response to biotic stimulus processes. DMGs were present in CpG and CHH contexts at ion transmembrane transport while genes were differently methylated at regulation of transcription, DNA-templated in CHG and CHH contexts.

Table 20. Molecular functions in average of both scion affected by methylation (yellow: present and number of DMGs, blue: absent of DMGs)

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
monooxygenase activity				3					
ligase activity							3		
oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen				3					
protein kinase activity	2			5	4			5	
hydrolase activity				5	2		2	4	
protein binding	5		3	16	12	5	9	12	5
ATP binding	2	2		6	6	2	4	8	2
transmembrane transporter activity							3		
oxidoreductase activity						2	2	4	2
metal ion binding	4	3	3		3		3		
nucleic acid binding					2			5	
transferase activity, transferring hexosyl groups			2	3		2			2
structural constituent of ribosome						2			2
catalytic activity, acting on protein		3							2
calcium ion binding						2			2
ATPase activity, coupled to transmembrane movement of substances						2			2
ion transmembrane transporter activity						2			2
transferase activity, transferring acyl groups other than amino-acyl groups	2	2	2			2			2
endopeptidase activity			2			2			
DNA binding	2	2	3			2			3
transition metal ion binding				3					
drug transmembrane transporter activity					2				
active transmembrane transporter activity					2			3	
monophenol monooxygenase activity	3	3			2				
catechol oxidase activity	3	3			2				
UDP-glucosyltransferase activity					2				
protein serine/threonine kinase activity							2		
isomerase activity	2								
asparaginase activity	2								
transcription factor activity, transcription factor binding		2							
enzyme binding		2							
transferase activity, transferring phosphorus-containing groups		2							
adenyl ribonucleotide binding			2						
anion binding			3						

In the mother plant vs. *in vitro* plant comparison, 71 DMGs were identified, of which 27, 24, and 20 were in the CpG, CHG, and CHH contexts, respectively. In the CpG context, DMGs were present at protein kinase activity; isomerase activity and asparaginase activity, while genes

were differently methylated in CpG and CHH contexts at protein binding process. We identified genes which were differently methylated in CpG and CHG contexts at the ATP binding; monophenol monooxygenase activity and catechol oxidase activity processes, while DMGs were present at metal ion binding; transferase activity, transferring acyl groups other than aminoacyl groups and DNA binding processes in CpG, CHG and CHH contexts. We analysed DMGs only in CHH context at transferase activity, transferring hexosyl groups; endopeptidase activity; adenyl ribonucleotide binding and anion binding processes, while genes were differently methylated in CHG context at catalytic activity, acting on protein; transcription factor activity, transcription factor binding; enzyme binding and transferase activity, transferring phosphorus-containing groups. In the acclimatized vs. *in vitro* plant comparison (Table 20), 108 DMGs were identified, of which 44, 39, 25 were in the CpG, CHG and CHH contexts, respectively. We identified DMGs only in CpG context at monooxygenase activity; oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen and transition metal ion binding processes, while genes were differently methylated at protein kinase activity and hydrolase activity process in CpG and CHG contexts. In CpG, CHG and CHH contexts, DMGs were present at protein binding and ATP binding processes, while genes were differently methylated only in CHH context at oxidoreductase activity; structural constituent of ribosome; calcium ion binding; ATPase activity, coupled to transmembrane movement of substances; ion transmembrane transporter activity; transferase activity, transferring acyl groups other than amino-acyl groups; endopeptidase activity and DNA binding. In CHG context, we observed DMGs at metal ion binding; nucleic acid binding; drug transmembrane transporter activity; active transmembrane transporter activity; monophenol monooxygenase activity; catechol oxidase activity and UDP-glucosyltransferase activity.

In the mother plant vs. acclimatized comparison, 95 DMGs were identified, of which 28, 41, and 26 were in the CpG, CHG, and CHH contexts, respectively. In CpG context, DMGs were present at ligase activity; transmembrane transporter activity, metal ion binding and protein serine / threonine kinase activity processes, while genes were differently methylated only in CHG context at protein kinase activity; nucleic acid binding and active transmembrane transporter activity. In the CpG and CHG contexts, we identified DMGs at hydrolase activity processes, while genes were differently methylated in CpG, CHG and CHH contexts at protein binding; ATP binding processes and oxidoreductase activity. DMGs were present in CHH context at transferase activity, transferring hexosyl groups; structural constituent of ribosome; catalytic activity, acting on protein; calcium ion binding; ATPase activity, coupled to transmembrane movement of substances; ion transmembrane transporter activity; transferase activity, transferring acyl groups other than amino-acyl groups and DNA binding.

Table 21. Cellular components in average of both scion affected by methylation (yellow: present and number of DMGs, blue: absent of DMGs)

	Mother plant vs. *in vitro*			Acclimatized vs. *in vitro*			Mother plant vs. acclimatized		
	CpG	CHG	CHH	CpG	CHG	CHH	CpG	CHG	CHH
catalytic complex								2	
plasma membrane			1	1				2	
ribonucleoprotein complex								2	
cytoplasmic part								4	
intracellular organelle part						3			
nucleus	2	1	2	3	1	2	2	2	2
integral component of membrane	1		4	6	4	7		8	7
membrane protein complex						2			2
endoplasmic reticulum			2			2			2
intercellular organelle part									3
ribosome			1			2			2
transcription factor complex	1			1					
plasmodesma				1			1		
cytoplasm		2		2			2		
chromatin					1				
nuclear condensin complex					1				
small-subunit processome					1				
chloroplast membrane					1				
integral component of plasma membrane		1					1		
protein serine/threonine phosphatase complex		1					1		
Cdc73/Paf1 complex		1							
extrinsic component of membrane			1						
nucleosome			1						
anchored component of membrane			1						
plastid			1						
photosystem II oxygen evolving complex			1						

In the mother plant vs. *in vitro* plant comparison (Table 21), 25 DMGs were identified, of which 4, 6, and 15 were in the CpG, CHG, and CHH contexts, respectively. We observed DMGs only in CHH context at plasma membrane; endoplasmic reticulum; ribosome; extrinsic component of membrane; nucleosome; anchored component of membrane; plastid and photosystem II oxygen evolving complex, while genes were differently methylated in CpG, CHG and CHH contexts at nucleus. In CpG and CHH contexts, DMGs were present at integral component of plasma membrane, while genes were differently methylated in CpG context at transcription factor complex process. In CHG context, DMGs were present at cytoplasm; integral component of plasma membrane; protein serine/threonine phosphatase complex and Cdc73/Paf1 complex processes. In the acclimatized plant vs. *in vitro* plant comparison (Table 21), 41 DMGs were identified, of which 14, 9, and 18 were in the CpG, CHG, and CHH contexts, respectively. We

observed DMGs only in the CpG context at plasma membrane; transcription factor complex; plasmodesma and cytoplasm, while genes were differently methylated in CHH context at membrane protein complex; endoplasmic reticulum and ribosome. In CpG, CHG and CHH contexts, DMGs were present at nucleus and integral component of membrane processes, while genes were differently methylated only in CHG context at chromatin; nuclear condensin complex; small-subunit processome and chloroplast membrane. In the mother plant vs. acclimatized plant comparison, 45 DMGs were identified, of which 7, 20, and 18were in the CpG, CHG, and CHH contexts, respectively. In CHG and CHH contexts, DMGs were present at nucleus, integral component of membrane, while genes were differently methylated only in CHG context at catalytic complex; plasma membrane; ribonucleoprotein complex and cytoplasmic part. We identified DMGs in CpG, CHG and CHH contexts at nucleus, while genes were differently methylated only in CHH context at membrane protein complex; endoplasmic reticulum; intercellular organelle part and ribosome. In CpG context, DMGs were present at plasmodesma; cytoplasm; integral component of plasma membrane in protein serine / threonine phosphatase complex processes.

LIN *et al*. (2019) also analysed the whole-genome bisulfite sequencing reveals of the role for DNA methylation in variants from callus culture of pineapple using Omicsbox (before Blast2GO). FENG *et al*. (2019) studied the molecular mechanism of 'Fuji' and 'Shandingzi''s resistance to *Marssonina coronaria* infection with transcriptome analysis. They discovered 4898 DEGs (differentially expressed gene) in 'Fuji' and 2897 DEGs in 'Shandingzi', respectively. The specific DEGs in 'Shandingzi' were mainly correlated with the biotic stimulus response, defensive response, stress response, and oxidoreductase function, as revealed by GO enrichment analysis of the co-expressed DEGs by Omicsbox. Although in the case of apple whole genome sequencing, the results have not yet been visualized using Omicsbox before, but LIU *et al*. (2017) continuously collected rice ovules from the beginning of megaspore mother cell meiosis until the mature female gametophyte forming time. Global DNA methylation patterns in the ovules of a high-frequency female-sterile line and a wild-type rice line were compared using whole-genome bisulfite sequencing. They discovered more DMGs enriched in cellular portion, reproduction control, metabolic pathway, and other pathways based on functional annotation and KEGG pathway analysis of DMGs. In specific, several ovule formation genes and plant hormone-related genes in the two rice lines had substantially different methylation patterns.

In the case of *Diuraphis noxia* (Russian wheat aphid), DNA methylation and demethylation are regulated by functional DNA methyltransferases and DnTET enzymes, which were studied in 2020 by DU PREEZ *et al.* using Omicsbox, among others. The study examined epigenomas of different Russian wheat aphid biotypes (i.e., SA1 and SAM) in two related Russian virulences to

elucidate its role in virulence in this species. A total of 40 differentially methylated genes were identified, and less and more virulent biotypes were found to contain significantly different differentially methylated genes.

KEGG

Table 22. Dynamic changes of DMGs relatide to key enzymes involved in 'McIntosh' and 'Húsvéti rozmaring' apple cultivars plotted with KEGG map

Pathway	Sequences number	Enzyme
Alanine, aspartate and glutamate metabolism	3	EC 2.6.1.44, EC 3.5.1.1
alpha-Linolenic acid metabolism	2	EC 1.1.1.1, EC 2.3.1.16
Aminobenzoate degradation	3	EC 3.1.3.41
Arginine and proline metabolism	1	EC 1.14.13.39
Arginine biosynthesis	1	EC 1.14.13.39
Benzoate degradation	1	EC 2.3.1.16
beta-Alanine metabolism	1	EC 1.4.3.21
Betalain biosynthesis	4	EC 1.14.18.1
Biosynthesis of antibiotics	7	EC 2.6.1.57, EC 1.14.13.39, EC 2.5.1.47, EC 1.1.1.2, EC1.1.1.1, EC 2.3.1.16, EC 1.1.1.44
Biosynthesis of unsaturated fatty acids	1	EC 2.3.1.16
Caffeine metabolism	1	EC 2.3.1.5
Caprolactam degradation	1	EC 1.1.1.2
Chloroalkane and chloroalkene degradation	1	EC 1.1.1.1
Cyanoamino acid metabolism	4	EC 3.2.1.21, EC 1.14.13.39, EC 3.5.1.1
Cysteine and methionine metabolism	3	EC 2.6.1.57, EC 2.5.1.47, EC 2.6.1.44
Drug metabolism - cytochrome P450	5	EC 1.14.13.8, EC 2.5.1.18, EC 1.1.1.1
Drug metabolism - other enzymes	5	EC 2.3.1.5, EC 3.1.1.1, EC 2.5.1.18
Ether lipid metabolism	1	EC 2.7.8.1
Ethylbenzene degradation	1	EC 2.3.1.16
Fatty acid degradation	2	EC 1.1.1.1, EC 2.3.1.16
Fatty acid elongation	1	EC2.3.1.16
Flavonoid biosynthesis	2	EC 1.14.11.9, EC 2.3.1.133
Fructose and mannose metabolism	1	EC 1.1.1.14
Galactose metabolism	1	EC 3.2.1.22
Geraniol degradation	1	EC 2.3.1.16
Glutathione metabolism	3	EC 2.5.1.18, EC 1.1.1.44
Glycerolipid metabolism	2	EC 3.2.1.22, EC 1.1.1.2
Glycerophospholipid metabolism	2	EC 3.1.4.46, EC 2.7.8.1
Glycine, serine and threonine metabolism	3	EC 1.1.1.1, EC 2.6.1.44, EC 1.4.3.21
Glycolysis / Gluconeogenesis	2	EC 1.1.1.2, EC 1.1.1.1
Glycosphingolipid biosynthesis - globo and isoglobo series	1	EC 3.2.1.22
Isoquinoline alkaloid biosynthesis	7	EC 2.6.1.57, EC 1.10.3.1, EC 1.14.18.1, EC 1.1.1.248, EC 1.4.3.21
Lysine biosynthesis	1	EC 2.6.1.57
Lysine degradation	3	EC 2.1.1.43

Metabolism of xenobiotics by cytochrome P450	3	EC 2.5.1.18, EC 1.1.1.1
Naphthalene degradation	1	EC 1.1.1.1
Nitrotoluene degradation	1	EC 2.3.1.5
Novobiocin biosynthesis	1	EC 2.6.1.57
Other glycan degradation	1	EC 3.2.1.51
Oxidative phosphorylation	1	EC 1.6.99.3, EC 1.6.5.3
Pentose and glucuronate interconversions	5	EC 1.1.1.14, EC 3.1.1.11, EC 4.2.2.2, EC 1.1.1.2
Pentose phosphate pathway	1	EC 1.1.1.44
Phenylalanine metabolism	2	EC 2.6.1.57, EC 1.4.3.21
Phenylalanine, tyrosine and tryptophan biosynthesis	1	EC 2.6.1.57
Phenylpropanoid biosynthesis	3	EC 3.2.1.21, EC 2.3.1.133, EC 1.11.1.7
Phosphonate and phosphinate metabolism	1	EC 2.7.8.1 - EPT
Purine metabolism	20	EC 3.6.1.15, EC 3.6.1.3, EC 2.7.7.6
Pyrimidine metabolism	3	EC 2.7.7.6
Retinol metabolism	1	EC 1.1.1.1
Sphingolipid metabolism	1	EC 3.2.1.22
Starch and sucrose metabolism	4	EC 3.2.1.21, EC 2.4.1.34, EC 3.2.1.39, EC 2.4.1.12, EC 2.4.1.25
Stilbenoid, diarylheptanoid and gingerol biosynthesis	1	EC 2.3.1.133
Sulfur metabolism	2	EC 2.5.1.47, EC 3.6.3.25
T cell receptor signaling pathway	3	EC 3.1.3.16
Th1 and Th2 cell differentiation	3	EC 3.1.3.16
Thiamine metabolism	17	EC 3.6.1.15
Tropane, piperidine and pyridine alkaloid biosynthesis	2	EC 2.6.1.57, EC 1.4.3.21
Tyrosine metabolism	8	EC 2.6.1.57, EC 1.10.3.1, EC 1.1.1.1, EC 1.14.18.1, EC 4.1.1.68, EC 1.4.3.21
Ubiquinone and other terpenoid-quinone biosynthesis	1	EC 2.1.1.295
Valine, leucine and isoleucine degradation	1	EC 2.3.1.16

KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis was performed to identify differentially regulated biological pathways by apple cultivars from different environments (Table 22). In the analysis, we found a total of 60 pathways that can be affected due to different methylation in the various environments regarding the two apple cultivars examined. This means a total of 166 DMGs and 43 enzymes.

The ethanolaminephosphotransferase (ETP) (EC 2.7.8.1) plays role in degradation pathways, such as phosphonate and phosphinate metabolism, glycerophospholipid metabolism, ether lipid metabolism. The ETP (also known as aminoalcoholphosphotransferases; AAPTs; EC 2.7.8.1 and EC 2.7.8.2) catalyze the 1,2-diacylglycerols with cytidine diphosphate (CDP)-aminoalcohols to form phosphatidylaminoalcohols. GOODE and DEWEY (1999) characterized GmAAPT1 (AAPT of *Glycine max*) as a hybridization probe to designated AtAAPT1 and AtAAPT2 (AAPT of *A. thaliana*). GmAAPT1 and AtAAPT1 showed similar Ca^{2+} and cytidine

5'-monophosphate (CMP) inhibition. AAPTs are able to catalyze the formation of CDP-choline and diacyglycerol from phosphatidylcholine (PC) in the presence of the CMP.

The gentiobiase (or β-glucosidases EC 3.2.1.21) plays role in phenylpropanoid biosynthesis and degradation pathways, such as cyanoamino acid metabolism, starch and sucrose metabolism. The β-glucosidases play a key role in several important biological processes in all living organisms. Toxic β-glucosidic substrates and β-glucosidases are stored in subcellular and tissue parts in plants (POULTON 1990). Damage to cell walls by pests brings the enzyme and substrate together, go to the hydrolysis of substrate and release of toxic aglycones (SADDER and SUWWAN 2006).

The asparaginase II (EC 3.5.1.1) plays role in degradation pathways, such as cyanoamino acid metabolism, alanine, aspartate and glutamate metabolism. The L-asparaginases (ASPGs) catalyze the hydrolysis of L-asparagine and cause asparagine and ammonia. ASPGs are expressed in sink tissues of higher plants, where they should metabolize transported asparagine (LEA *et al.* 2007).

The RNA polymerase (EC 2.7.7.6) plays role in degradation pathways, such as pyrimidine metabolism, purine metabolism. DNA-dependent RNA polymerase thus plays a role in the first step of the process of expressing genetic information encoded in cellular DNA. Other enzymes that synthesize a heteroribonucleic acid polymer have occurred in some plants, however only RNA is used as template for the synthesis of a complementary product (DUDA 1976).

Flavanone 3-beta-hydroxylase (EC 1.14.11.9) plays role in flavonoid biosynthesis. This enzyme is a 2-oxoglutarate dependent dioxygenase, which catalyzes the synthesis of dihydrokaempferol, what common precursor for of the flavons, anthocyans and proanthocyanids (OWENS *et al*. 2008).

Alanine-glyoxylate transaminase (EC 2.6.1.44) plays role in glycine, serine and threonine metabolism; alanine, aspartate and glutamate metabolism; cysteine and methionine metabolism pathways. Plant peroxisomal glyoxylate aminotransferases play a role in photorespiratory pathway. Glyoxylate aminotransferase (GGT) activity has been observed in several locations in different plant and algae (LIEPMAN *et al*. 2003).

Endo-1,3-beta-D-glucosidase (EC 3.2.1.39) plays role in starch and sucrose metabolism. Endo-1,3-beta-gluconase degrades the cell wall polymer, callose. Glycosylphosphatidyl inositol (GPI)-anchored proteins are usually located at the cell surface and regulate a high range of biological processes. ZERZAUST (ZET) encodes a GPI-anchored beta-1,3-glucanase, which localizes to the cell wall (VADDEPALLI *et al*. 2017).

Lactoperoxidase (EC 1.11.1.7) plays role in phenylpropanoid biosynthesis. Lactoperoxidase is belong to class III peroxidases, which are generally secreted into the cell wall

or the vacuole. Class III plant peroxydases catalyse the reduction of H_2O_2 by taking electrons to various donor molecules such as auxin, phenolic compounds or secondary metabolites (COSIO and DUNAND 2009).

Adenylpyrophosphatase (EC 3.6.1.3; now Ca^{2+} transporter, EC 7.2.2.10) plays role in purine metabolism. Calmodulin-like (CML) proteins, as part of the second messender system, play a major role in plant development and in coordinating plant stress tolerance (ASTEGNO *et al*. 2017).

Amine oxidase (EC 1.4.3.21) plays role in glycine, serine and threonine metabolism; isoquinoline alkaloid biosynthesis; tyrosine metabolism; beta-alanine metabolism; phenylalanine metabolism; tropane, piperidine and pyridine alkaloid biosynthesis pathways. Nitric oxide is a key signaling molecule in plants, which regulate a wide range of physiological processes, such as development process, germination or flowering, as well as the adaptive response to biotic and abiotic stresses (GROß *et al*. 2017).

Pectate lyase (EC 4.2.2.2) plays role in pentose and glucuronate interconversions. Pectate lyases (PLs), a family of depolymerizing enzymes, such as α-1,4-glycosidic linkages in homogalacturonan of pectate (demethylated pectin) by β-elimination and produce unsaturated oligogalacturonides. These enzymes have been identified in several plant pathogenic microorganisms by releasing plant cell wall oligogalacturonides, which act as defense elicitors (PALUSA *et al*. 2007).

Salutaridine reductase (NADPH; nicotinamide adenine dinucleotide phosphate) (EC 1.1.1.248) plays role in isoquinoline alkaloid biosynthesis. Benzylisoquinoline alkaloids (BIAs) are a group of nitrogen-containing plant secondary metabolites. Salutaridine reductase (SalR) is a BIA biosynthetic gene found in *Papaver somniferum* L. (opium poppy) that is responsible for the synthesis of morphinan alkaloids. SalR was isolated and overexpressed in *P. somniferum*, but also from *Papaver* species that do not produce morphinan alkaloids (ZIEGLER *et al.* 2009).

L-iditol 2-dehydrogenase (EC 1.1.1.14) plays role in fructose and mannose metabolism; pentose and glucuronate interconversions. NAD^+-dependent sorbitol dehydrogenase (NAD-SDH) is a key enzyme in sorbitol metabolism, which plays a role in regulating sink strength and determining the quality of apple fruit. NAD-SDH is localized mostly in the cytoplasm and chloroplast of the fruit and leaves (WANG *et al*. 2009).

The α-galactosidase (EC 3.2.1.22) plays role in glycerolipid metabolism; galactose metabolism; glycosphingolipid biosynthesis and sphingolipid metabolism. The α-galactosidase plays a role in biosynthetic and hydrolytic pathways in raffinose (Raf) metabolism, especially under certain abiotic stress conditions (drought, high salinity or high temperature). A α-

galactosidase reporter construct using the ATSIP2 promoter, which show that ATSIP2 is strongly expressed in sink tissues of *A. thaliana* (PETERS *et al*. 2010).

Nucleoside-triphosphate phosphatase (EC 3.6.1.15) plays role in purine metabolism; thiamine metabolism. The nucleoside triphosphatases (NTPases) are activated by arginine or lysine. NTPases represent most common protein fold that can comprise up to 18% of all gene products in a cell (SHALAEVA *et al*. 2018).

The pectinesterase (EC 3.1.1.11) plays role in pentose and plucuronate interconversions. The endogenous pectin methylesterase (PME) yields acidic pectin with lower degree of esterification, which can cross-link with polyvalent cations such as Ca^{2+} to form insoluble pectate precipitates to pectin-degrading polygalacturonases (ZHI *et al*. 2008).

The 4-nitrophenylphosphatase (EC 3.1.3.41) plays role in aminobenzoate degradation. The 4-nitrophenylphosphatases comprise a superfamily of enzymes involved in the transport of charged substrates across biological membranes. Members of the P-type ATPases are found in all kingdoms of life. This family of enzymes have got different groups based on their substrate specificity, example heavy metal-transporting P-type ATPases belongs to type I class; mono- and divalent cation-transporting P-type ATPases belongs to type II class (BRAMKAMP *et al.* 2004).

The nitric-oxide synthase (NADPH) (EC 1.14.13.39) plays role in arginine biosynthesis; arginine and proline metabolism; cyanoamino acid metabolism and biosynthesis of antibiotics. The nitric oxide (NO) plays a role in activating resistance responses. NO potentiates pathogen- and reactive oxygen species (ROS)-induced cell death and NO synthase inhibitors block hypersasitive resistance response in *Arabidopsis* (CHANDOK *et al*. 2003).

The α-L-fucosidase (EC 3.2.1.51) plays role in glycan degradation. The α-L-fucosidases catalyze the removal of terminal L-fucose residues linked via α-1,2, α-1,3, α-1,4 or α-1,6 bonds to reducing end of N-acetyl glucosamine of oligosaccharide chains. Mammalian α-L-fucosidases are lysosomal enzymes involved in the hydrolytic degradation of fucose-containing molecules (INTRA *et al.* 2007).

The 2-methyl-6-phytyl-1,4-hydroquinone methyltransferase (MPBQ MT; EC 2.1.1.295) was present in ubiquinone and other terpenoid-quinone biosynthesis. The MPBQ MT were constitutively expressed in corn. The MPBQ and methylated MPBQ are also a substrate for tocopherols in corn (NAQVI *et al*. 2011).

Shikimate o-hydroxycinnamoyltransferase (HTC; EC 2.3.1.133) plays role phenylpropanoid biosynthesis, stilbenoid, diarylheptanoid and gingerol biosynthesis, flavonoid biosynthesis pathways. Transcriptome data from roots and nodules of *Datisca glomerata* (C.Presl) Baill. (durango root) and *Medicago truncatula* Gaertn. (barrelclover) were analysed by differential expression between tissues and relative expression levels of transcripts within each transcriptome

were compared to better understand the expression of phenylpropanoid and flavonoid biosynthetic genes in roots and nodules. The transcriptome of *D. glomerata* showed low expression of genes encoding the major enzymes in the lignin-monolignol branch of the phenylpropanoid pathway. The transcript of shikimate O-hydroxycinnamoyltransferase (HCT), the first enzyme in this branch, was expressed in the 48th percentile (GIFFORD *et al*. 2018).

O-acetylserine sulfhydralase (also named as cysteine synthase) (EC 2.5.1.47) plays role in sulfur metabolism, biosynthesis of antibiotics, cysteine and methionine metabolism in the pathways. It is a pyridoxal-phosphate protein. Some alkyl thiols, cyanide, pyrazole and some other heterocyclic compounds can act as acceptors. A kinetic assay can be used to detect *in vitro* sulfite reductase (SiR) activity in plants based on a coupled reaction in which the sulfide produced is converted to cysteine in the assay medium by O-acetylserine sulfhydralase and its substrate, O-acetylserine. In crude desalted protein extracts, a new kinetic assay for SiR activity has been created. In contrast to current protocols, pre-treatment of the protein with tungstate increased SiR activity by 29 and 12% in *Arabidopsis* and tomato leaf, respectively, and the addition of NADPH to the reaction medium increased SiR activity by 1.6 and 2.8-fold in *Arabidopsis* and tomato, respectively (BRYCHKOVA *et al*. 2012).

Acetyl-CoA C-acyltransferase (EC 2.3.1.16) participate in degradation pathways, such as the valine, leucine and isoleucine degradation, geraniol degradation, biosynthesis of unsaturated fatty acids, fatty acid elongation, biosynthesis of antibiotics, fatty acid degradation, alpha-Linolenic acid metabolism, benzoate degradation, ethylbenzene. This enzyme is also found in eukaryotes and prokaryotes and is involved in degradation processes such as fatty acid beta-oxidation, which was identified by LIU *et al*. (2016) by analyzing a transcriptome at four different stages of fruit development in *Symplocos paniculata* Miq. (Asiatic sweetleaf) as one of the most important enzymes involved in fatty acid biosynthesis and metabolism.

Monophenol monooxygenase (also known as tyrosinase) (EC 1.14.18.1) plays role in the isoquinoline alkaloid biosynthesis, tyrosine metabolism, betaline biosynthesis pathways. It is found in many living things, such as plants, fungi, bacteria, insects. This enzyme is involved in the synthesis of betalains and melanin. Plant polyphenol oxidase is normally a constitutive enzyme, although it can be caused in certain cases, biotic stresses, such as the rupture of plant tissues (caused by pathogens and insects), can bring the enzyme and its natural substrates into contact, resulting in browning. They can be found in a variety of locations inside the plant cell, including vacuoles, protective cells, and epithelial cells (SCHNABL *et al*. 1986, HUTZLER *et al.* 1998, FERRERES *et al*. 2011).

NADH dehydrogenase (EC 1.6.99.3) plays role in Oxidative phosphorylation pathways. Iron-sulfur centers are found in this flavoprotein. NADH dehydrogenase can serve as an acceptor

after preparations have been subjected to certain cytochrome treatments. Two protons are extruded from the cytoplasm or the intramitochondrial or stromal compartment under normal conditions. NADH: ubiquinone reductase (H^+-translocating) is found in a mitochondrial complex as EC 7.1.1.2 (2020). WANG *et al.* (2017) examined genes encoding protein-coding in grapes using an SSR marker. A total of 4,337 genes containing SSRs were found among 29,971 genes encoding protein-coding and 5,384 SSRs. These genes were annotated with GO and KEGG. Some of the genes have been shown to be involved in the synthesis and metabolism of secondary metabolites, the synthesis of flavones or anthocyanins, the development and morphology of plant organs, and tolerance to biotic or abiotic stresses. Some genes functioned in the oxidative phosphorylation pathway as NADH dehydrogenase (VIT_00s2376g00010 gene).

Protein 2-phosphatase (EC 3.1.3.16) plays role in the T cell receptor signaling pathway, Th1 and Th2 cell differentiation pathways. Protein phosphatase is a phosphatase enzyme that extracts a phosphate group from its substrate protein's phosphorylated amino acid residue. BAI *et al.* (2017) New Zealand flax (*Phormium tenax* J.R.Forst. & G.Forst.) were studied for drought resistance by molecular methods. During the experiment, the plants were examined under normal and drought stress. A total of 4380 DMGs were found, of which 2698 can be related to drought stress. They observed that two genes encoding protein 2-phosphatase were upregulated during drought stress.

Dehydrogenase ($NADP^+$-dependent, decarboxylating) (EC 1.1.1.44) plays role in antibiotic biosynthesis, glutathione metabolism, pentose-phosphate pathways. KAHL (1974) studied potato tuber. The two dehydrogenases of the oxidative pentose phosphate cycle, glucose-6-phosphate- and 6-phosphogluconate dehydrogenase, are only weakly active in resting potato tuber tissue. Slicing the skin, on the other hand, significantly increases the activity of both enzymes. The increased activity caused by slicing is the result of more distinct activation/inactivation of seven 6-phosphogluconate dehydrogenase isozymes and improved action of at least five glucose-6-phosphate dehydrogenase isozymes. He concluded that most of increased activity of the two enzymes are the result of *de novo* synthesis rather than activation of pre-existing proenzymes.

Aldehyde reductase ($NADP^+$) (EC 1.1.1.2) plays role in glycerolipid metabolism, biosynthesis of antibiotics, glycolysis/gluconeogenesis, pentose and glucuronate interconversions, caprolactam degradation pathways. It is a zinc protein. Alcohol dehydrogenase ($NADP^+$) group, some of its members oxidize only primary alcohols while the other part also acts on secondary alcohols. MORITA *et al.* (2007) cloned and functionally analysed a new aldo-keto reductase from *Aloe arborescens* Mill. (candelabra aloe). Aldo-keto reductase found in *A. arborescens* showed similarity to plant polyketide reductases, so they looked at the activity of the recombinant enzyme

expressed in *Escherichia coli* and showed no detectable activity, instead Aldo-Keto reductase catalyzed the reduction of various carbonyl compounds by NADPH, benzaldehyde and DL-glycerol aldehyde in *A. arborescen*.

Arylamine N-acetyltransferase (NAT; EC2.3.1.5) plays role in drug metabolism - other enzymes, caffeine metabolism and nitrotoluene degradation pathways. A wide specificity for aromatic amines, like serotonin, allows it to catalyze acetyl transfer between arylamines without the need for CoA. NAT catalyze the transfer of an acetyl group from the acetyl-CoA to the nitrogen or oxygen atom of primary arylamines, hydrazines and their N-hydroxylated metabolites. NAT play a role in detoxification and metabolic activation of xenobiotics in *Salmonella enterica* (DELOMÉNIE *et al*. 2001).

ATP-binding protein CysA (EC 3.6.3.25) plays role in sulfur metabolism pathways. The existence of two identical ATP-binding domains distinguishes ABC-type ATPases. During the transport process, it is not phosphorylated. Sulfate and thiosulfate anions are imported by a bacterial enzyme. Unfortunately, there was no such research in plants at the moment, but LIU *et al*. (2020) examined the genus *Ewingella* using comparative genomic analysis. Genes associated with stress response were also examined in the experiment in five Ewingella americana strains. In addition. All five *E. americana* strains had five genes coding for sulfate and thiosulfate import ATP-binding protein CysA, DedA protein, S-formylglutathione hydrolase, TsgA protein, and S-(hydroxymethyl) glutathione dehydrogenase for detoxification, and a cold shock protein CspA for cold stress tolerance.

Polyphenol oxidase (EC 1.10.3.1) is identified in isoquinoline alkaloid biosynthesis and tyrosine metabolism pathways. Catalyzes the oxidation of catechol to the corresponding o-quinone as a type 3 copper protein. A number of substituted catechols are also affected by the enzyme. MURUMKAR and CHAVAN (1985) studied the changes in the levels of individual inorganic constituents and the activity of some enzyme systems in the aging leaves of chickpeas (*Cicer arietinum* L.). The activity of acid phosphatase, alkaline phosphatase, ATPase, inorganic pyrophosphatase, and 3-phosphoglycerate phosphatase all increased as leaf aging progressed. Leaf senescence was also linked to an increase in peroxidase and polyphenol oxidase activity, a significant decrease in pyruvate kinase activity, and a small increase in aldolase activity.

The disproportionating enzyme or 4-α-glucanotransferase (EC 2.4.1.25) is involved in the starch and sucrose metabolism pathway. D-enzyme is the name given to the plant enzyme. TAKAHA (1996) investigated D-enzyme extracted from potato tubers with molecular instruments. As a results, it was discovered that D-enzyme's mRNA accumulates when starch biosynthesis is most active, but decreases when starch is degraded. These findings seem to

contradict the popular belief that D-enzyme is involved in starch degradation and may also play a role in starch synthesis.

The flavin-containing monooxygenase (EC 1.14.13.8) is involved in the drug metabolism-cytochrome P450 pathway. Hydrazines, phosphines, boron-containing compounds, sulfides, selenides, iodide, as well as basic, secondary, and tertiary amines are all substrates for this monooxygenase. This monooxygenase differs from others in that it produces a relatively stable hydroperoxy flavin intermediate. According to HOU *et al.* (2010), *Arabidopsis* flavin monooxygenases (FMO) play an important role in plant growth and development by synthesizing auxin and other signaling molecules.

The cellulose synthase (UDP-forming) (EC 2.4.1.12) plays role in the starch and sucrose metabolism pathway. Proteins of *Arabidopsis* have been investigated by bioinformatics methods to better understand the salt-response mechanism (MEILI *et al.* 2012). In their analysis, 292 salt-response proteins were found in the plant. Such was, among others, Cellulose synthase A catalytic subunit 1 (UDP-forming) (EC 2.4.1.12), which we also identified in our research.

The α-carboxylesterase (EC 3.1.1.1) plays role in the drug metabolism-other enzymatic pathway. Carboxylesterase has a broad specificity, is catalyzed by enzymes derived from microsomes, and hydrolyzes vitamin A esters. Long-term callus culture was used to study differential gene expression at different concentrations and the development of alkaloids from *Cereus peruvianus* C.F.Först callus grown on medium supplemented with different concentrations of tyrosine, 2,4-D, kinetin, and NaCl. Tyrosine at 200 mg/l produced the most alkaloids and induced two α-carboxylesterases (EC 3.1.1.1; EST-6 and EST-7 isozymes), one α/β-arylesterase (EC 3.1.1.2) in the callus, one α/β-acetylesterase (EC 3.1.1.6) and two carboxylesterases (EC 3.1.1.1; EST -3 and EST-12 isozymes) was found to be more intensely stained (KLEBER DA ROCHA *et al.* 2005).

The reductase (H^+-translocating) (EC 1.6.5.3) plays role in the oxidative phosphorylation pathway. The largest complex in the mitochondrial respiratory chain is NADH-ubiquinone oxidoreductase (EC 1.6.5.3). In eukaryotes, it has more than 40 subunits, as MEYER *et al.* discovered in *Arabidopsis* mutants in 2011. Four subcomplexes of the membrane arm of NADH-ubiquinone oxidoreductase with apparent molecular weights of 200, 400, 450, and 650 kDa were observed in their experiment.

The transaminase (EC 2.6.1.57) plays role in the lysine biosynthesis, isoquinoline alkaloid biosynthesis, novobiocin biosynthesis, tyrosine metabolism, phenylalanine, tyrosine and tryptophan biosynthesis, biosynthesis of antibiotics, cysteine and methionine metabolism, phenylalanine metabolism, tropane, piperidine and pyridine alkaloid biosynthesis pathway. A pyridoxal phosphate protein. SAHU *et al.* (2015) attempted to develop simple sequence repeats

markers to transaminase for *Centella asiatica* (L.) Urb. (Gotu Kola), but the developed primers were also tested in different genera of the *Apiaceae* family. Thus, routes to SSRs containing Expressed Sequence Tag (EST) were mapped using KEGG. Several enzymes have been identified, such as: UDP-arabinopyranose mutase (EC 5.4.99.30), transaminase (EC 2.6.1.57) 2-alkenal reductase [NAD (P) $^+$] (EC1.3.1.74), etc.

The 2-Hydroxyhepta-2,4-diene-1,7-dioate isomerase (EC 4.1.1.68) plays role in the tyrosine metabolism pathway. In *Escherichia coli*, it is a part of the 4-hydroxyphenylacetate degradation pathway. Unfortunately, there was no such research in plants at the moment. The phylogenetics of *Methanococcus jannaschii* proteins have been studied (KYRPIDES *et al*. 1999). 301 universal proteins have been identified that can be grouped into 246 biochemical functions, such as amino acid biosynthesis, cell envelope, cellular processes, Central intermediary metabolism, where the enzyme (2-Hydroxyhepta-2,4-diene-1,7-dioate isomerase; EC 4.1.1.68) which we also identified during our experiments.

The N-methyltransferase (EC 2.1.1.43) plays role in the lysine degradation pathway. N-methyltransferase established 1976, amended 1982, amended it a year later, and deleted it in 2019. Now described by EC 2.1.1.354-362. In 1995, KLEIN and HOUTZ studied the cloning and developmental expression of the large subunit of pea (*Pisum sativum* L.) ribulose-1,5-bisphosphate carboxylase/oxygenase N-methyltransferase. The broad subunit N-methyltransferase (Protein methylase Ill, Rubisco LSMT, EC 2.1.1.43) of Ribulose-l,5-bisphosphate carboxylase/oxygenase (Rubisco) catalyzes methylation of the e-amino group of Lys-14 in the Rubisco LS. A derived N-terminal amino acid sequence with chloroplast trans- sit peptide-like features was discovered.

The callose synthase (also named as 1,3-beta-glucan synthase) (EC 2.4.1.34) plays role in the starch and sucrose metabolism pathway. VOIGT *et al*. in 2006, the expression of eight glucan synthase-like (GSL) genes in the stem, leaf blade and thorn of wheat (*Triticum aestivum* L.) was examined. Organ-specific expression of six GSL genes and significant differences in expression levels were detected. There were also differences in callose synthase function and overall callose levels in the organs. Their findings show that a group of GSL genes involved in general or organ-specific developmental processes are tightly regulating callose synthesis.

Glycerophosphodiester phosphodiesterase (GDPD) (AT5G55480) (EC 3.1.4.46) plays role in the glycerophospholipid metabolism pathway. The GDPD is a broad specificity for glycerophosphodiesters. In study of ROOMI *et al.* (2018) many enzymes involved in cell wall metabolism, such as the GDPD (AT5G55480), were found to be up-regulated. During glycerol metabolism, this enzyme converts glycerophosphodiester to glycerol-3-phosphate and alcohols, which is important for many physiological processes in living organisms.

Alcohol dehydrogenase (ADH; EC 1.1.1.1) plays role in the glycine, serine and threonine metabolism, chloroalkane and chloroalkene degradation, tyrosine metabolism, retinol metabolism, naphthalene degradation, biosynthesis of antibiotics, metabolism of xenobiotics by cytochrome P450, drug metabolism-cytochrome P450, glycolysis/gluconeogenesis, fatty acid degradation, alpha-linolenic acid metabolism pathway. It is a zink protein. The enzyme has a broad specificity for primary or secondary alcohols or hemi-acetals; however, it oxidizes methanol much more poorly than ethanol. Greensleeves apples from a transgenic line with a high suppression of ethylene biosynthesis were used by DEFILIPPI *et al*. (2005). Aroma volatile-related enzymes such as alcohol acyltransferase (AAT), ADH, and lipoxygenase (LOX) were studied in peel and flesh tissues, as well as amino acids and fatty acids as aroma volatile precursors. AAT enzyme activity showed a pattern in parallel with ethylene regulation but ADH and LOX are independent of ethylene modulation.

Glutathione S-transferases (GST; EC 2.5.1.18) plays role in the drug metabolism-other enzymes, glutathione metabolism, metabolism of xenobiotics by cytochrome P450, drug metabolism-cytochrome P450 pathway. A glutathione transferase group of enzymes of broad specificity. The addition of aliphatic epoxides and arene oxides to glutathione, the reduction of polyol nitrate to polyol and nitrile by glutathione, some isomerization reactions, and disulfide interchange are all catalyzed by GSTs. The response to abiotic stress is mediated by GSTs (EC 2.5.1.18). The *LeGSTU2* GST gene from tomato was cloned and functionally characterized in the study of XU *et al.* (2015).

6. Conclusions and recommendations

We started our experiments with checking the genetic homogeneity of the *in vitro* plantlets maintained in sterile shoot cultures for 16 years by using microsatellite markers compared to their mother plants ('McIntosh' and 'Húsvéti rozmaring'). We chose this marker type because there are a lot of studies available in the literature proving the suitability and reliability of SSR analyses to objectively genotype the varieties, to determine parentage, to check the true-to-typness of cultivars in various species including the genus Malus (Liebhard *et al*. 2002). Based on our results, we found no difference with the 17 SSR markers we selected on the mother plants, *in vitro*, and between the acclimatized plants, either the 'McIntosh' or in 'Húsvéti rozmaring'. JIN *et al*. studied the genetic fidelity of regenerated plants and spontaneous shoot regeneration from the leaves of the apple rootstock 'Pingyitiancha' (*Malus hupehensis* var. *pinyiensis*) in 2014. All banded profiles of the regenerated plants are monomorphic and similar to those of the mother plant, as determined by SSR markers. This means that the SSR markers they used to fail to prove an effective duration polymorphism between the *in vitro* and the mother plant. The study uses RAPD and SSR markers to determine genetic fidelity in sugarcane plants that were regenerated by direct organogenesis (PANDEY *et al*. 2012). There was no evidence of somaclonal variation in RAPD banding patterns produced by PCR amplification using 20 random primers. The plantlets' RAPD patterns were similar to the mother plant's pattern, which means that the direct adventitious organogenesis did not result in somaclonal variation detectable by RAPD. Similarly, SSR banding pattern analysis using 15 primers revealed no signs of somaclonal variation. RAI *et al*. (2012) analysed the genetic fidelity of 1-year-old guava (*Psidium guajava* L.) plants developed from *in vitro* somatic embryogenesis. 6 SSR primer pairs generated reproducible and transparent bands with sizes varying from 100 to 300 bp. SSR primers produced monomorphic amplicons in all regenerated plants. Microtubers were produced *in vitro* from potato shoot cultures and then stored on three different media for a year (TIWARI *et al*. 2013). A total of 38 primers were applied in the study: 10 RAPD, 11 ISSR (inter simple sequence repeat), 12 SSR, and 5 AFLP (amplified fragment length polymorphism) primers. SSR, ISSR, and RAPD analyses revealed 100% genetic identity between the mother plant and its shoot cultures in clusters, while AFLP analysis revealed 85 to 100% genetic similarity. Based on their findings, they concluded that AFLP is the best tool for demonstrating the genetic stability of *in vitro* preserved potato microtubules, against SSR, ISSR, and RAPD.

While other researchers have found differences, such as MODGIL *et al*. (2005) who used RAPD primers in an *in vitro* and mother plant study of MM106 apples. A total of 129 evaluable fragments were found in their study, including 18 polymorphs among all individuals studied. In i*n vitro* regenerated pineapple, KOHPAII *et al*. (2017) searched for somaclonal variation. All morphological, cellular, and biochemical observations provided evidence for the existence of somaclonal variation in pineapple plants. Molecularly, this was demonstrated using ISSR markers. ISSR locus analysis showed genetic differences between the mother plant and the regenerated progenies. Based on the literature, it can be said that there have been studies that have been able to determine genetic differences among *in vitro*, acclimatized plant and mother plant. Thus, it would be worthwhile to examine our samples not only with SSR markers, but also with other DNA fingerprinting techniques (RAPD, ISSR, AFLP etc.) that have been used by other researchers.

MSAP profiling was performed on the leaf tissues of six plants (*Salvadora persica* L. (arak)), *Commiphora wightii* (Arn.) Bhandari (Indian bdellium-tree), male and female *Simmondsia chinensis* (Link) C.K.Schneid. (Jojoba), *Jatropha curcas* L. (physic nut), and female genotype of *Withania coagulans* (Stocks) Dunal (Indian cheese maker) developing *in vivo* and *in vitro*. *In vivo*-grown tissues of *S. persica*, *S. chinensis* (male), and *W. coagulans* and *in vitro*-grown tissues of *C. wightii*, *S. chinensis* (female), and *J. curcas* had higher percentage methylation in genome. The plasticity of the genomes of plants developing in two distinct habitats is reflected in the differences in DNA methylation and polymorphism in genomes. Under *in vitro* and *in vivo* conditions, different patterns of DNA methylation of homologous nucleotide sequences and polymorphism in the methylated DNA in tissues indicate that these fragments could be involved in complex processes controlling plant growth and development (RATHORE *et al*. 2015). The DNA methylation genome of the rice embryo and endosperm has been extensively profiled, according to XING *et al*. (2015) in comparison to the endosperm, the embryos were primarily hypermethylated with non-transposable element (non-TE) genes. Hypermethylated genes are substantially more common in low-expressed genes. Methylation was significantly reduced in the early developmental stages (2-3 days after pollination), suggesting that demethylation plays a dominant role in early endosperm growth and that DNA methylation may be directly controlled by stable negatively associated genes for DNA methylation shift and expression change. During seed growth, both rice subspecies had varying DNA methylation profiles in the embryo and endosperm, with the highest methylation level 6 days after pollination. The systematic characterization of complex DNA methyloma in rice seed production aids in the understanding of epigenetic regulation's effects and mechanisms in seed development.

PERRIN *et al*. (2020) observed DNA methylation patterns of mature apple trees to juvenile seedlings resulting from selfing. They did not observe global genome-wide change in DNA

methylation patterns and levels with WGBS. They identified variations in DNA methylation patterns which localized in gene regions involved in photosynthesis. Epigenomic and transcriptomic analysis indicate that grafted plants are at an intermediate phase between an adult tree and seedling, which inherit part of the DNA methylation patterns and levels of their donor tree.

When comparing the DMG patterns of ‘McIntosh’, DMGs were downregulated in *in vitro* plants compared to acclimatized plants. The DNA methylation level was higher in *in vitro* plants than in mother plants. There was no considerable directional changes in MA plots when acclimatized and mother plants were compared, indicating that after acclimatization, the DNA methylation pattern became similar to mother plant. According to GO annotation function of the DMGs were present in:

i) metabolic processes, cellular, phosphorylation, and biosynthetic processes among their biological functions;
ii) regulatory sequences for catalytic, kinase and transferase activity among their molecular functions;
iii) sequences characteristic of membrane components, intracellular parts and cell parts among their cellular components.

The DNA methylation patterns of ‘Húsvéti rozmaring’ showed a similar tendency in all three contexts (CpG, CHG and CHH), as for ‘McIntosh’. DMGs were present in:

i) metabolic, phosphorylation, and biosynthetic processes among their biological functions;
ii) regulatory sequences for catalytic, kinase and transferase activity among their molecular functions;
iii) sequences characteristic of membrane components and cell parts among their cellular components.

We observed several DMGs in the three different environmental conditions, which confirm the findings of a series of earlier studies that have also implicated DNA methylation in the regulation of genes controlling pathways in plant development or tissue differentiation, during embryogenesis under *in vitro* propagation (US-CAMAS *et al.* 2014). DNA methylation is a key regulatory system in gene expression and plant development under stress conditions (CRISP *et al.* 2016, LÄMKE and BÄURLE 2017). We identified DMGs between ‘McIntosh’ and ‘Húsvéti rozmaring’ scion cultivars, which were both up- and down-regulated in all three contexts, independent of the environment. DMGs were present in:

i) metabolic cellular, biosynthetic and transfer processes, and phosphorylation among their biological functions;
ii) catalytic, transferase, kinase and hydrolase activities among their molecular functions;

iii) membrane compounds, cell parts, intracellular parts and cytoplasmic parts among their cellular components.

Our results suggest that the potential role of cultivar-specific DNA methylation constitute an important regulatory mechanism for sensing and responding to stress, such as drought and salinity (RAJKUMAR *et al*. 2019), as well as adaptation to other biotic and abiotic stresses (VIGGIANO and DE PINTO 2017). In general, DMGs were detected in both cultivars and in all three contexts:

i) *in vitro* plants were highly methylated compared to acclimatized plants (*in vitro* plants were downregulated);

ii) *in vitro* plants were downregulated relative to mother plants;

iii) DNA methylation patterns became similar to mother plants after acclimatization in both scion cultivars.

The genotype of the donor plant influences the level of DNA methylation and pattern of regenerants and the ability of the genotype how to adopt to the *in vitro* environment (MACHCZYŃSKA *et al*. 2014).

7. New scientific results

1. As firstly in the international scientific community, we published the DNA methylome of diploid apple and identified DMGs in the whole genome of apple with whole genome bisulfite sequencing technique between the mother, acclimatized and *in vitro* plants.
2. We observed similar global DNA methylation levels for both ‘McIntosh’ and ‘Húsvéti rozmaring’ under different environmental conditions.
3. We found DMGs in the *in vitro* plants at all of the CpG, CHG and CHH contexts, which were caused by the micropropagation process in laboratory environmental conditions.
4. We could justify that there are very low but significant differences between acclimatized and mother plants, that means the DNA methylation pattern of acclimatized plant became very similar to that of the mother plant.
5. We published first, by investigating the full genome, that the methylome of apple that has been sustained *in vitro* for more than 16 years began to regress during one year of acclimatization and the 1-year-old acclimatized plant showed a very close DNA methylation pattern like the mother plant. This demonstrates that *in vitro* environment as a stressor can be quenched by adaptation and thereby can be restored by acclimatization.
6. We observed DMGs between the two scion cultivars (‘McIntosh’ and ‘Húsvéti rozmaring’). The different scion cultivars were acclimatized to the three environmental conditions (mother plant, acclimatized plant and *in vitro* plant) in different ways. There are more easily adaptable scion cultivars to micropropagation, such as ‘Húsvéti rozmaring’, which regulate only few genes to adapt to stress.

8. Summary

Plant micropropagation plays an important role in the process of variety maintenance, gene conservation and pathogen removal, which is an integral part of the production of plant propagation material. *In vitro* laboratory conditions can be characterized as a constant stress situation in the life cycle of a plant, to which individual plant species, but even plant varieties, adapt differently. This adaptation may have multiple phenotypic or genotypic manifestations. DNA methylation changes, levels, or patterns play a key role in the regulation of gene expression. Depending on the DNA methylation pattern, plants can regulate their gene expression processes, which are required for rapid adaptation to changing environmental factors.

In our research, we monitored the changes of the DNA methylation in 'McIntosh' and 'Húsvéti rozmaring' apple cultivars in the mother plant, acclimatized plant, and *in vitro* plants of long term culture. For this, we chose the most modern sequencing method currently available, which provides the most accurate information about whole genome-wide changes in DNA methylation within the the plant, in all CpG, CHG, and CHH contexts. We were the first to determine the methylation pattern and global levels of diploid apple DNA using the WGBS method. We could not detect a difference between global DNA methylation levels between each environment. The two apple cultivars globally seek to balance DNA methylation levels in all three contexts. However, significant differences in DNA methylation pattern were already observed at gene level (DMG). In both apple cultivars, the same trend was observed in all three contexts and in all three environmental circumstances. In *in vitro* tissue culture, the DNA methylation pattern exhibits a state of hypermethylation (the methyl group in the 5-methylcytosine nucleotide is more methylated in DNA) in several DMGs, demonstrating that the *in vitro* plant methylates its genes that are not necessary under artificial laboratory conditions compared to an acclimatized or mother plant. GO annotation was used to determine gene functions. These genes are primarily involved in metabolic, phosphorylation, biosynthetic processes, biological regulation of biological function; regulatory sequences for catalytic, kinase and transferase activity of molecular functions; sequences characteristic for membrane components and cell parts of cellular components play a role.

Regarding the comparison of acclimatized and mother plant, the same trend was observed for both 'McIntosh' and 'Húsvéti rozmaring' varieties. During acclimatization, the methylation level of DMGs showed a decrease and the DNA methylation pattern of the acclimatized plant based on MA plots showed almost the same state as the methylation pattern of the mother plant DNA. Based on these results, it can be said that metabolic processes, cellular, phosphorylation, and biosynthetic processes in biological functions; regulatory sequences for catalytic, kinase and

transferase activity in molecular functions; genes involved in the sequences characteristic of membrane components, intracellular parts and cell parts in cellular components are required for the acclimatization process and for acclimation to the mother plant's environment. We also identified significant differences in the adaptation of 'McIntosh' and 'Húsvéti rozmaring' scions to micropropagation. The 'Húsvéti rozmaring' methylated fewer genes, so fewer DMGs were identified compared to the 'McIntosh' variety. Thus, it can be stated that 'Húsvéti rozmaring' was more easily adapted to the conditions of *in vitro* micropropagation in metabolic cellular, biosynthetic and transfer processes, and phosphorylation in biological functions; catalytic, transferase, kinase and hydrolase activities in molecular functions; membrane compounds, cell parts, intracellular parts and cytoplasmic parts in a cellular component with DMGs playing a role.

DMGs were categorized into biological, molecular, and cellular processes using the KEGG database to obtain a more accurate picture of the status of each DMG. In our studies, the identified DMGs were mainly involved in phosphonate and phosphinate metabolism, thiamine metabolism, tyrosine metabolism, biosynthesis of antibiotics, isoquinoline alkaloid biosynthesis processes, etc..

Overall, *in vitro* tissue culture causes a change in DNA methylation level and a change in DNA methylation pattern of the genes in the plant, which can vary from environment to environment and from plant variety to plant variety. But it can also be stated that the maintenance of *in vitro* micropropagation for decades does not cause an irreversible change in the plant compared to the parent plant since the investigated *in vitro* plantlets were kept in *in vitro* culture for 16 years.

9. Appendices

9.1. References

ADUSUMALLI, S., MOHD OMAR, M. F., SOONG, R., BENOUKRAF, T. (2015): Methodological aspects of whole-genome bisulfite sequencing analysis. *Briefings in Bioinformatics*, 16(3) pp. 369-379. https://doi.org/10.1093/bib/bbu016

ALDWINCKLE, H., MALNOY, M. (2009): Plant regeneration and transformation in the *Rosaceae*. Transgen. *Plant Journal,* 3 1–39. p.

ALLIS, C.D., JENUWEIN, T., REINBERG, D., CAPARROS M-L. (2007): Epigenetics. Cold Spring Harbor Laboratory Press. *Cold Spring Harbor*, NY

AN, Y.Q.C., GOETTEL, W., HAN, Q., BARTELS, A., LIU, Z., XIAO, W. (2017): Dynamic changes of genome-wide DNA methylation during soybean seed development. *Scientific Reports,* 7(1) pp. 1-14. https://doi.org/10.1038/s41598-017-12510-4

ANDREWS, S. (2007): SeqMonk: A tool to visualise and analyse high throughput mapped sequence data.

ANDREWS, S. (2012): Trim Galore, a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files.

ARNAUDO, A.M., GARCIA, B.A. (2013): Proteomic characterization of novel histone post-translational modifications. *Epigenetics & Chromatin,* 6(1) pp. 1-7. https://doi.org/10.1186/1756-8935-6-24

ASTEGNO, A., BONZA, M.C., VALLONE, R., LA VERDE, V., D'ONOFRIO, M., LUONI, L., MOLESINI, B., DOMINICI, P. (2017): *Arabidopsis* calmodulin-like protein CML36 is a calcium (Ca^{2+}) sensor that interacts with the plasma membrane Ca^{2+}-ATPase isoform ACA8 and stimulates its activity. *Journal of Biological Chemistry*, 292(36) 15049-15061. https://doi.org/10.1074/jbc.M117.787796

BAI, S., TUAN, P.A., SAITO, T., HONDA, C., HATSUYAMA, Y., ITO, A., MORIGUCHI, T. (2016): Epigenetic regulation of MdMYB1 is associated with paper bagging-induced red pigmentation of apples. *Planta,* 244(3) pp. 573-586. https://doi.org/10.1007/s00425-016-2524-4

BAI, Z-Y., WANG, T., WU, Y-H., WANG, K., LIANG, Q.-Y., PAN, Y-Z., JIANG, B.-B., ZHANG, L., LIU, G.-L., JIA, Y., LIU, Q.L. (2017): Whole-transcriptome sequence analysis of differentially expressed genes in *Phormium tenax* under drought stress. *Scientific Reports*, 7(1) pp. 1-9. https://doi.org/10.1038/srep41700

BANNISTER, A.J., KOUZARIDES, T. (2011): Regulation of chromatin by histone modifications. *Cell Research,* 21(3) pp. 381-395. https://doi.org/10.1038/cr.2011.22

BARTEL, D.P. (2009): MicroRNAs: target recognition and regulatory functions. *Cell,* 136(2) pp.

215-233. https://doi.org/10.1016/j.cell.2009.01.002

BARTELS, A., HAN, Q., NAIR, P., STACEY, L., GAYNIER, H., MOSLEY, M., HUANG, Q., PEARSON, J., HSIEH, T.F., AN, Y.Q., XIAO, W. (2018): Dynamic DNA methylation in plant growth and development. *International Journal of Molecular Sciences,* 19(7) pp. 2144. https://doi.org/10.3390/ijms19072144

BERDASCO, M., ESTELLER, M. (2010): Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Developmental Cell,* 19(5) pp. 698-711. https://doi.org/10.1016/j.devcel.2010.10.005

BERGER, S.L., KOUZARIDES, T., SHIEKHATTAR, R., SHILATIFARD, A. (2009): An operational definition of epigenetics. *Genes & Development,* 23(7) pp. 781-783. https://doi.org/10.1101/gad.1787609

BEWICK, A.J., SCHMITZ, R.J. (2017): Gene body DNA methylation in plants. *Current Opinion in Plant Biology,* 36 pp. 103-110. https://doi.org/10.1016/j.pbi.2016.12.007

BLAND, J.M., ALTMAN, D.G. (1999): Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8(2) pp. 135-160. https://doi.org/10.1177/096228029900800204

BOLAR, J.P., NORELLI, J.L., ALDWINCKLE, H.S., HANKE, V. (1998): An efficient method for rooting and acclimation of micropropagated apple cultivars. *HortScience*, 33(7) pp.1251-1252. https://doi.org/10.21273/HORTSCI.33.7.1251

BOLGER, A.M., LOHSE, M., USADEL, B. (2014): Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics,* 30(15) pp. 2114-2120. https://doi.org/10.1093/bioinformatics/btu170

BOUYER, D., KRAMDI, A., KASSAM, M., HEESE, M., SCHNITTGER, A., ROUDIER, F., COLOT, V. (2017): DNA methylation dynamics during early plant life. *Genome Biology,* 18(1) pp. 1-12. https://doi.org/10.1186/s13059-017-1313-0

BRAIT, M., SIDRANSKY, D. (2011): Cancer epigenetics: above and beyond. *Toxicology Mechanisms and Methods,* 21(4) pp. 275-288. https://doi.org/10.3109/15376516.2011.562671

BRAMKAMP, M., GASSEL, M., ALTENDORF, K. (2004): FITC binding site and p-nitrophenyl phosphatase activity of the Kdp-ATPase of *Escherichia coli*. *Biochemistry,* 43(15) pp. 4559-4567. https://doi.org/10.1021/bi030198a

BRYCHKOVA, G., YARMOLINSKY, D., VENTURA, Y., SAGI, M. (2012): A novel in-gel assay and an improved kinetic assay for determining *in vitro* sulfite reductase activity in plants. *Plant and Cell Physiology*, 53(8) pp. 1507-1516. https://doi.org/10.1093/pcp/pcs084

BUTLER, M.I., STOCKWELL, P.A., BLACK, M.A., DAY, R.C., LAMONT, I.L., POULTER, R.T. (2013): *Pseudomonas syringae* pv. *actinidiae* from recent outbreaks of kiwifruit bacterial

canker belong to different clones that originated in China. *PloS One,* 8(2) p.e57464. https://doi.org/10.1371/journal.pone.0057464

CAO, X., SPRINGER, N.M., MUSZYNSKI, M.G., PHILLIPS, R.L., KAEPPLER, S., JACOBSEN, S.E. (2000): Conserved plant genes with similarity to mammalian *de novo* DNA methyltransferases. *Proceedings of the National Academy of Sciences*, 97(9) pp. 4979-4984. https://doi.org/10.1073/pnas.97.9.4979

CARMONA, J.J., IZZI, B., JUST, A.C., BARUPAL, J., BINDER, A.M., HUTCHINSON, J., HOFMANN, O., SCHWARTZ, J., BACCARELLI A., MICHELS, K.B. (2013): Comparison of multiplexed reduced representation bisulfite sequencing (mRRBS) with the 450K Illumina Human BeadChip: from concordance to practical applications for methylomic profiling in epigenetic epidemiologic studies. *Epigenetics & Chromatin,* 6(1) pp. 1-2. https://doi.org/10.1186/1756-8935-6-S1-P36

CARTHEW, R.W., SONTHEIMER, E.J. (2009): Origins and mechanisms of miRNAs and siRNAs. *Cell,* 136(4) pp. 642-655. https://doi.org/10.1016/j.cell.2009.01.035

CATUOGNO, S., ESPOSITO, C.L., UNGARO, P. AND DE FRANCISCIS, V. (2018): Nucleic acid aptamers targeting epigenetic regulators: an innovative therapeutic option. *Pharmaceuticals,* 11(3) pp. 79. https://doi.org/10.3390/ph11030079

CHAN, S.W.L., HENDERSON, I.R., JACOBSEN, S.E. (2005): Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nature Reviews Genetics,* 6(5) pp. 351-360. https://doi.org/10.1038/nrg1601

CHANDOK, M.R., YTTERBERG, A.J., VAN WIJK, K.J., KLESSIG, D.F. (2003): The pathogen-inducible nitric oxide synthase (iNOS) in plants is a variant of the P protein of the glycine decarboxylase complex. *Cell*, 113(4) pp. 469-482. https://doi.org/10.1016/S0092-8674(03)00350-7

CHEUNG, P., TANNER, K.G., CHEUNG, W.L., SASSONE-CORSI, P., DENU, J.M., ALLIS, C.D. (2000): Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. *Molecular Cell,* 5(6) pp. 905-915. https://doi.org/10.1016/S1097-2765(00)80256-7

CHOI, M., SCHOLL, U.I., JI, W., LIU, T., TIKHONOVA, I.R., ZUMBO, P., NAYIR, A., BAKKALOĞLU, A., ÖZEN, S., SANJAD, S., NELSON-WILLIAMS, C. (2009): Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proceedings of the National Academy of Sciences,* 106(45) pp. 19096-19101. https://doi.org/10.1073/pnas.0910672106

COCK, P.J., FIELDS, C.J., GOTO, N., HEUER, M.L. RICE, P.M. (2010): The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research,* 38(6) pp. 1767-1771. https://doi.org/10.1093/nar/gkp1137

COKUS, S.J., FENG, S., ZHANG, X., CHEN, Z., MERRIMAN, B., HAUDENSCHILD, C.D., PRADHAN, S., NELSON, S.F., PELLEGRINI, M., JACOBSEN, S.E. (2008): Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature,* 452(7184) pp. 215-219. https://doi.org/10.1038/nature06745

CONDIT, R., HUBBELL, S. (1991): Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. *Genome,* 34 pp. 66-71.

CONESA, A., GÖTZ, S. (2008): Blast2GO: a comprehensive suite for functional analysis in plant genomics. *International Journal of Plant Genomics,* 2008. https://doi.org/10.1155/2008/619832

CONSIDINE, M.J., WAN, Y., D'ANTUONO, M.F., ZHOU, Q., HAN, M., GAO, H., WANG, M. (2012): Molecular genetic features of polyploidization and aneuploidization reveal unique patterns for genome duplication in diploid *Malus*. *PloS One,* 7(1) p.e29449. https://doi.org/10.1371/journal.pone.0029449

CORNILLE, A., GIRAUD, T., SMULDERS, M.J.M., ROLDAN-RUIZ, I., GLADIEUX, P. (2014): The domestication and evolutionary ecology of apples. *Trends in Genetics,* 30(2) pp. 57-65. https://doi.org/10.1016/j.tig.2013.10.002

CORNILLE, A., GLADIEUX, P., SMULDERS, M.J.M., ROLDÁN-RUIZ, I., LAURENS, F., CAM, B.L., NERSESYAN, A., CLAVEL, J., OLONOVA, M., FEUGEY, L., GABRIELYAN, I., ZHANG, X., TENAILLON, M.I., GIRAUD, T. (2012): New insight into the history of domesticated apple: secondary contribution of the European wild apple to the genome of cultivated varieties. *PloS Genetics,* 8(5) p.e1002703. https://doi.org/10.1371/journal.pgen.1002703

CORTÉS, V., CUBERO, S., BLASCO, J., ALEIXOS, N., TALENS, P. (2019): In-line application of visible and near-infrared diffuse reflectance spectroscopy to identify apple varieties. *Food and Bioprocess Technology,* 12(6) pp. 1021-1030. https://doi.org/10.1007/s11947-019-02268-0

COSIO, C., DUNAND, C. (2009): Specific functions of individual class III peroxidase genes. *Journal of Experimental Botany*, 60(2) pp. 391-408. https://doi.org/10.1093/jxb/ern318

CRISP, P.A., GANGULY, D., EICHTEN, S.R., BOREVITZ, J.O., POGSON, B.J. (2016): Reconsidering plant memory: Intersections between stress recovery, RNA turnover, and epigenetics. *Science Advances*, 2(2) p.e1501340. https://doi.org/10.1126/sciadv.1501340

DACCORD, N., CELTON, J.M., LINSMITH, G., BECKER, C., CHOISNE, N., SCHIJLEN, E., VAN DE GEEST, H., BIANCO, L., MICHELETTI, D., VELASCO, R., DI PIERRO, E.A., GOUZY, J., REES, D.J.G., GUÉRIF, P., MURANTY, H., DUREL, C.E., LAURENS, F., LESPINASSE, Y., GAILLARD, S., AUBOURG, S., QUESNEVILLE, H., WEIGEL, D., VAN DE WEG, E., TROGGIO, M., BUCHER, E. (2017): High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nature Genetics,* 49(7) pp.1099–1106. https://doi.org/10.1038/ng.3886

DAR, J.A., WANI, A.A., DHAR, M.K. (2015): Morphological, biochemical and male-meiotic characterization of apple (*Malus* × *domestica* Borkh.) germplasm of Kashmir Valley. *Chromosome Botany,* 10(2) pp. 39-49. https://doi.org/10.3199/iscb.10

DAVIS, M.P., VAN DONGEN, S., ABREU-GOODGER, C., BARTONICEK, N., ENRIGHT, A.J. (2013): Kraken: a set of tools for quality control and analysis of high-throughput sequence data. *Methods,* 63(1) pp. 41-49. https://doi.org/10.1016/j.ymeth.2013.06.027

DEFILIPPI, B.G., DANDEKAR, A.M., KADER, A.A. (2005): Relationship of ethylene biosynthesis to volatile production, related enzymes, and precursor availability in apple peel and flesh tissues. *Journal of Agricultural and Food Chemistry*, *53*(8) pp. 3133-3141. https://doi.org/10.1021/jf047892x

DELOMÉNIE, C., FOUIX, S., LONGUEMAUX, S., BRAHIMI, N., BIZET, C., PICARD, B., DENAMUR, E., DUPRET, J.M. (2001): Identification and functional characterization of arylamine N-acetyltransferases in eubacteria: evidence for highly selective acetylation of 5-aminosalicylic acid. *Journal of Bacteriology*, 183(11) pp. 3417-3427. https://doi.org/10.1128/JB.183.11.3417-3427.2001

DELONG, J.M., PRANGE, R.K., HARRISON, P.A. (2004): The influence of 1-methylcyclopropene on 'Cortland' and 'McIntosh' apple quality following long-term storage. *HortScience,* 39(5) pp. 1062-1065. https://doi.org/10.21273/HORTSCI.39.5.1062

DETTMAN, J.R., RODRIGUE, N., MELNYK, A.H., WONG, A., BAILEY, S.F., KASSEN, R. (2012): Evolutionary insight from whole-genome sequencing of experimentally evolved microbes. *Molecular Ecology,* 21(9) pp. 2058-2077. https://doi.org/10.1111/j.1365-294X.2012.05484.x

DIRLEWANGER, E., COSSON, P., TAVAUD, M., ARANZANA, M.J., POIZAT, C., ZANETTO, A., ARÚS, P., LAIGRET, F. (2002): Development of microsatellite markers in peach *(Prunus persica* (L.) Batsch) and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). *Theoretical and Applied Genetics,* 105 pp. 127–138. https://doi.org/10.1007/s00122-002-0867-7

DOBRÁNSZKI, J., MAGYAR-TÁBORI, K., JÁMBOR-BENCZÚR, E., LAZÁNYI, J., BUBÁN, T., SZALAI, J. (2000): Influence of aromatic cytokinins on shoot multiplication and their after-effects on rooting of apple cv. Húsvéti rozmaring. *International Journal of Horticultural Science*, 6(4) pp. 84-87. https://doi.org/10.31421/IJHS/6/4/231

DU, X., WANG, Y., LIU, M., LIU, X., JIANG, Z., ZHAO, L., TANG, Y., SUN, Y., ZHANG, X., LIU, D., SONG, L. (2020): The assessment of epigenetic diversity, differentiation, and structure in the 'Fuji'mutation line implicates roles of epigenetic modification in the occurrence of different mutant groups as well as spontaneous mutants. *PloS One*, 15(6) p.e0235073.

https://doi.org/10.1371/journal.pone.0235073

DU PREEZ, P.H., BREEDS, K., BURGER, N.F.V., SWIEGERS, H.W., TRUTER, J.C., BOTHA, A.M. (2020): DNA Methylation and demethylation are regulated by functional DNA methyltransferases and DnTET enzymes in *Diuraphis noxia*. *Frontiers in Genetics*, 11, p. 452. https://doi.org/10.3389/fgene.2020.00452

DUDA, C.T. (1976): Plant RNA polymerases. *Annual Review of Plant Physiology*, 27(1), pp. 119-132.

DUDITS, D., HESZKY, L. (2000): Növényi biotechnológia és géntechnológia. *Agroinform Kiadóház Rt.* Budapest

ELLIOTT, R.F. (1972): Axenic culture of shoot apices of apple. *New Zealand Journal of Botany,* 10(2) pp. 254-258. https://doi.org/10.1080/0028825X.1972.10429153

ESTELLER, M. (2008): Epigenetics in cancer. *New England Journal of Medicine,* 358(11) pp. 1148-1159. https://doi.org/10.1056/NEJMra072067

FALCKENHAYN, C. (2017): The methylome of the marbled crayfish *Procambarus virginalis*. *Doctoral dissertatio*n. Ruperto-Carola University of Heidelberg, Germany

FAN, X.D., WANG, J.Q., YANG, N., DONG, Y.Y., LIU, L., WANG, F.W., WANG, N., CHEN, H., LIU, W.C., SUN, Y.P., WU, J.Y. (2013): Gene expression profiling of soybean leaves and roots under salt, saline–alkali and drought stress by high-throughput Illumina sequencing. *Gene,* 512(2) pp. 392-402. https://doi.org/10.1016/j.gene.2012.09.100

FAOSTAT (2019): http://www.fao.org/faostat/en/#data/QC

FAZIO, G., ALDWINCKLE, H.S., VOLK, G.M., RICHARDS, C.M., JANISIEWICZ, W.J., FORSLINE, P.L. (2007): Progress in evaluating *Malus sieversii* for disease resistance and horticultural traits. In *XII EUCARPIA Symposium on Fruit Breeding and Genetics,* 814 pp. 59-66. https://doi.org/10.17660/ActaHortic.2009.814.2

FEINBERG, A.P., VOGELSTEIN, B. (1983): A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry,* 132(1) pp. 6-13. https://doi.org/10.1016/0003-2697(83)90418-9

FENG, H., LI, H., ZHANG, M., SONG, Y., YUAN, G., HAN, Q., HUANG, L. (2019): Responses of Fuji (*Malus domestica*) and Shandingzi (*Malus baccata*) apples to Marssonina coronaria infection revealed by comparative transcriptome analysis. *Physiological and Molecular Plant Pathology*, 106 pp.87-95. https://doi.org/10.1016/j.pmpp.2018.12.007

FERREE, D.C., WARRINGTON, I.J. (eds) (2003): Apples: botany, production, and uses. *CABI*, Wallingford, UK pp. 74-75. https://doi.org/10.1079/9780851995922.0000

FERRERES, F., FIGUEIREDO, R., BETTENCOURT, S., CARQUEIJEIRO, I., OLIVEIRA, J., GIL-IZQUIERDO, A., PEREIRA, D.M., VALENTÃO, P., ANDRADE, P.B., DUARTE, P.,

ROS BARCELÓ A., SOTTOMAYOR, M. (2011): Identification of phenolic compounds in isolated vacuoles of the medicinal plant *Catharanthus roseus* and their interaction with vacuolar class III peroxidase: an H_2O_2 affair?. *Journal of Experimental Botany*, 62(8) pp. 2841-2854. https://doi.org/10.1093/jxb/erq458

FEUSER, S., MELER, K., DAQUINTA, M., GUERRA, M.P., NODARI, R.O. (2003): Genotypic fidelity of micropropagated pineapple (*Ananas comosus*) plantlets assessed by isozyme and RAPD markers. *Plant Cell, Tissue and Organ Culture*, *72*(3) pp. 221-227. https://doi.org/10.1023/A:1022321405683

FISHER, D.V., KITSON, J.A. (1991): The apple. Quality and Preservation of Fruits. *CRC* Press Inc., Florida In: ESKIN, N.A. M. (ed) pp. 45-66. https://doi.org/10.1201/9781351076159-4

FORSLINE, P.L., ALDWINCKLE, H.S., DICKSON, E.E., LUBY, J.J., HOKANSON, S.C. (2003): Collection, maintenance, characterization, and utilization of wild apples of Central Asia. In: JANICK, J. (ed), *Horticultural Reviews* John Wiley & Sons, 29 pp. 1–62. https://doi.org/10.1002/9780470650868.ch1

GALLI, Z., HALÁSZ, G., KISS, E., HESZKY, L., DOBRÁNSZKI, J. (2005): Molecular identification of commercial apple cultivars with microsatellite markers. *HortScience,* 40(7) pp. 1974-1977. https://doi.org/10.21273/HORTSCI.40.7.1974

GARRISON, E., MARTH, G. (2012): Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv*:1207.3907.

GIFFORD, I., BATTENBERG, K., VANIYA, A., WILSON, A., TIAN, L., FIEHN, O., BERRY, A.M. (2018): Distinctive patterns of flavonoid biosynthesis in roots and nodules of *Datisca Glomerata* and *Medicago* spp. Revealed by metabolomic and gene expression profiles. *Frontiers in Plant Science*, 9 pp. 1463. https://doi.org/10.3389/fpls.2018.01463

GLADIEUX, P., ZHANG, X.-G., RÓLDAN-RUIZ, I., CAFFIER, V., LEROY, T., DEVAUX, M., VAN GLABEKE, S., COART, E., LE CAM, B. (2010): Evolution of the population structure of *Venturia inaequalis*, the apple scab fungus, associated with the domestication of its host. *Molecular Ecology,* 19 pp. 658– 674. https://doi.org/10.1111/j.1365-294X.2009.04498.x

GOODE, J.H., DEWEY, R.E. (1999): Characterization of aminoalcoholphosphotransferases from *Arabidopsis thaliana* and soybean. *Plant Physiology and Biochemistry*, 37(6) pp. 445-457. https://doi.org/10.1016/S0981-9428(99)80049-7

GOODWIN, S., MCPHERSON, J.D., MCCOMBIE, W.R. (2016): Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics,* 17(6) pp. 333. https://doi.org/10.1038/nrg.2016.49

GOLL, M.G., KIRPEKAR, F., MAGGERT, K.A., YODER, J.A., HSIEH, C.L., ZHANG, X., GOLIC, K.G., JACOBSEN, S.E., BESTOR, T.H. (2006): Methylation of tRNAAsp by the

DNA methyltransferase homolog Dnmt2. *Science,* 311 pp. 395–398. https://doi.org/10.1126/science.1120976

GÖTZ, S., GARCÍA-GÓMEZ, J.M., TEROL, J., WILLIAMS, T.D., NAGARAJ, S.H., NUEDA, M.J., ROBLES, M., TALÓN, M., DOPAZO, J., CONESA, A. (2008): High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research,* 36(10) pp. 3420-3435. https://doi.org/10.1093/nar/gkn176

GREHL, C., WAGNER, M., LEMNIAN, I., GLASER, B., GROSSE, I. (2020): Performance of mapping approaches for whole-genome bisulfite sequencing data in crop plants. *Frontiers in Plant Science,* 11 pp. 176. https://doi.org/10.3389/fpls.2020.00176

GROß, F., RUDOLF, E.E., THIELE, B., DURNER, J., ASTIER, J. (2017): Copper amine oxidase 8 regulates arginine-dependent nitric oxide production in *Arabidopsis thalian*a. *Journal of Experimental Botan*y, 68(9) pp. 2149-2162. https://doi.org/10.1093/jxb/erx10

GUILFORD, P., PRAKASH, S., ZHU, J.M., RIKKERINK, E., GARDINER, S., BASSETT, H., FORSTER, R. (1997): Microsatellites in *Malus* × *domestica* (apple): abundance, polymorphism and cultivar identification. *Theoretical and Applied Genetics,* 94(2) pp. 249-254. https://doi.org/10.1007/s001220050407

GULYÁS, A., DOBRÁNSZKI, J., KISS, E., HIDVÉGI, N. (2019a): Genetikai stabilitás és génexpressziós vizsgálatok alma hajtástenyészetekben. *Kertgazdaság,* 51(4) pp. 17-29.

GULYÁS, A., DOBRÁNSZKI, J., KISS, E., TEIXEIRA DA SILVA, J.A., POSTA, K., HIDVÉGI, N. (2019b): Changes in DNA methylation pattern of apple long-term *in vitro* shoot culture and acclimatized plants. *Journal of Plant Physiology*, 239 pp. 18-27. https://doi.org/10.1016/j.jplph.2019.05.007

HANCOCK, J.F., LUBY, J.J., BROWN, S.K., LOBOS, G.A. (2008): Apples. *Temperate Fruit Crop Breeding Springer*, Dordrecht pp. 1-38. https://doi.org/10.1007/978-1-4020-6907-9_1

HARRIS, S.A., ROBINSON, J.P., JUNIPER, B.E. (2002): Genetic clues to the origin of the apple. *TRENDS in Genetics,* 18(8) pp. 426-430. https://doi.org/10.1016/S0168-9525(02)02689-6

HAY, R.T. (2005): SUMO: a history of modification. *Molecular Cell,* 18(1) pp. 1-12. https://doi.org/10.1016/j.molcel.2005.03.012

HOKANSON, S. C., SZEWC-MCFADDEN, A. K., LAMBOY, W. F., MCFERSON, J. R. (1998): Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a *Malus* × *domestica* Borkh. core subset collection. *Theoretical and Applied Genetics*, 97(5-6) pp. 671-683.

HOU, X., LIU, S., PIERRI, F., DAI, X., QU, L.J., ZHAO, Y. (2010): Allelic analyses of the *Arabidopsis* YUC1 locus reveal residues and domains essential for the functions of YUC family of flavin monooxygenases. *Journal of Integrative Plant Biology*, 53(1) pp. 54-62.

https://doi.org/10.1111/j.1744-7909.2010.01007.x

HRAZDINA, G. (2003): Response of scab-susceptible (McIntosh) and scab-resistant (Liberty) apple tissues to treatment with yeast extract and *Venturia inaequalis*. *Phytochemistry,* 64(2) pp. 485-492. https://doi.org/10.1016/S0031-9422(03)00150-X

HUANG, K., FAN, G. (2010): DNA methylation in cell differentiation and reprogramming: an emerging systematic view. *Regenerative Medicine,* 5(4) pp. 531-544. https://doi.org/10.2217/rme.10.35

HUTZLER, P., FISCHBACH, R., HELLER, W., JUNGBLUT, T.P., REUBER, S., SCHMITZ, R., VEIT, M., WEISSENBÖCK, G., SCHNITZLER, J.P. (1998): Tissue localization of phenolic compounds in plants by confocal laser scanning microscopy. *Journal of Experimental Botany*, 49(323) pp. 953-965. https://doi.org/10.1093/jxb/49.323.953

IGNATOV, A., BODISHEVSKAYA, A. (2011): *Malus*. *Wild Crop Relatives: Genomic and Breeding Resources Springer*, Berlin, Heidelberg, pp. 45-64. https://doi.org/10.1007/978-3-642-16057-8_3

INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM (2004): Finishing the euchromatic sequence of the human genome. *Nature,* 431(7011) pp. 931. https://doi.org/10.1038/nature03001

INTRA, J., PEROTTI, M.E., PAVESI, G., HORNER, D. (2007): Comparative and phylogenetic analysis of α-l-fucosidase genes. *Gene*, 392(1-2) pp. 34-46. https://doi.org/10.1016/j.gene.2006.11.002

JACKSON, J.E. (2003): The biology of apples and pears. *Cambridge University Press* UK pp. 42 https://doi.org/10.1017/CBO9780511542657

JANICK, J., CUMMINS, J.N., BROWN, S.K., HEMMAT, M. (1996): Apples. *Fruit Breeding,* pp. 1–77.

JIN, S., MUSHKE, R., ZHU, H., TU, L., LIN, Z., ZHANG, Y., ZHANG, X. (2008): Detection of somaclonal variation of cotton (*Gossypium hirsutum*) using cytogenetics, flow cytometry and molecular markers. *Plant Cell Reports,* 27(8) pp. 1303-1316. https://doi.org/10.1007/s00299-008-0557-2

JIN, W., WANG, Y., WANG, H. (2014): Adventitious shoot regeneration from leaves of apple rootstock ‘Pingyitiancha’ (*Malus hupehensis* var. *pinyiensis*) and genetic fidelity of regenerated plantlets using SSR markers. *Canadian Journal of Plant Science*, 94(8), pp. 1345-1354. https://doi.org/10.4141/cjps2013-357

JONES, P.A. (2012): Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics,* 13(7) pp. 484-492. https://doi.org/10.1038/nrg3230

JONES, P.A., TAKAI, D. (2001): The role of DNA methylation in mammalian epigenetics. *Science, 293*(5532) pp. 1068-1070. https://doi.org/10.1126/science.1063852

JONES, O.P. (1967): Effect of benzyl adenine on isolated apple shoots. *Nature,* 215(5109) pp. 1514-1515. https://doi.org/10.1038/2151514a0

JOSHI, N.A., FASS, J. (2011): Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software].

KAHL, G. (1974): *De novo* synthesis of glucose-6-phosphate-(EC 1.1.1.49) and 6-phosphogluconate dehydrogenase (EC 1.1.1.44) in plant storage tissue slices. *Zeitschrift für Naturforschung C*, 29(11-12) pp.700-704.

KARIM, R., NURUZZAMAN, M., KHALID, N., HARIKRISHNA, J.A. (2016): Importance of DNA and histone methylation in *in vitro* plant propagation for crop improvement: a review. *Annals of Applied Biology,* 169(1) pp. 1-16. https://doi.org/10.1111/aab.12280

KAWAKATSU, T., NERY, J.R., CASTANON, R., ECKER, J.R. (2017): Dynamic DNA methylation reconfiguration during seed development and germination. *Genome Biology*, 18(1) pp.1-12. https://doi.org/10.1186/s13059-017-1251-x

KAWANABE, T., ISHIKURA, S., MIYAJI, N., SASAKI, T., WU, L.M., ITABASHI, E., TAKADA, S., SHIMIZU, M., TAKASAKI-YASUDA, T., OSABE, K., PEACOCK, W.J., DENNIS, E.S., FUJIMOTO, R. (2016): Role of DNA methylation in hybrid vigor in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences,* 113(43) pp. E6704–E6711. https://doi.org/10.1073/pnas.1613372113

KELLERHALS, M. (2009): Introduction to apple (*Malus* × *domestica*). In Genetics and genomics of *Rosaceae Springer*, New York, NY. pp. 73-84. https://doi.org/10.1007/978-0-387-77491-6_4

KHAI, T.H., LANG, T.N. (2005): Using SSR markers to identify allele variation of somaclonal mutants in indica rice. *Omonrice*, 13 pp.121-125.

KIRÁLY, I. (2013): Kárpát-medencei almafajták jellemzése pomológiai vizsgálatokkal és mikroszatellit alapú molekuláris markerezéssel. *Doctoral dissertation*, Budapesti Corvinus Egyetem

KISS, E. (2005): Molekuláris növénynemesítés. In: HESZKY, L., FÉSŰS, L., HORNOK, L. *Mezőgazdasági Biotechnológia Agrofórum Kiadó, Budapest* pp. 194-210.

KISS, E. (1999): Növényi molekuláris genetika I. *Egyetemi jegyzet*. SZIE Genetika és Növénynemesítés Tanszék, Gödöllő

KLEBER DA ROCHA, L., JOSÉ BRAZ DE OLIVEIRA, A., APARECIDA MANGOLIN, C. (2005): Effect of different culture medium components on production of alkaloid in callus tissues of *Cereus peruvianus* (*Cactaceae*). *Acta Scientiarum Biological Sciences*, pp. 37-41. https://doi.org/10.4025/actascibiolsci.v27i1.1355

KLEIN, R.R., HOUTZ, R.L. (1995): Cloning and developmental expression of pea ribulose-1, 5-bisphosphate carboxylase/oxygenase large subunit N-methyltransferase. *Plant Molecular*

Biology, 27(2) pp.249-261.

KOHPAII, F.N., FARAHANI, F., NOORMOHAMMADI, Z. (2017): Somaclonal variation in the *in vitro* regenerated pineapple (*Ananas comosus*): Investigation of the cellular characteristics biochemical specificities and ISSR markers. *Phytologia Balcanica*, 23(1) pp.73-83.

KONG, Y. (2011): Btrim: a fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. *Genomics,* 98(2) pp. 152-153. https://doi.org/10.1016/j.ygeno.2011.05.009

KOROSTIN, D., KULEMIN, N., NAUMOV, V., BELOVA, V., KWON, D., GORBACHEV, A. (2020): Comparative analysis of novel MGISEQ-2000 sequencing platform vs Illumina HiSeq 2500 for whole-genome sequencing. *Plos One*, 15(3) p.e0230301. https://doi.org/10.1371/journal.pone.0230301

KOUZARIDES, T. (2007): Chromatin modifications and their function. *Cell,* 128(4) pp. 693-705. https://doi.org/10.1016/j.cell.2007.02.005

KROL, J., LOEDIGE, I., FILIPOWICZ, W. (2010): The widespread regulation of microRNA biogenesis, function and decay. *Nature Reviews Genetics,* 11(9) pp. 597-610. https://doi.org/10.1038/nrg2843

KRUEGER, F., KRECK, B., FRANKE, A., ANDREWS, S.R. (2012): DNA methylome analysis using short bisulfite sequencing data. *Nature Methods,* 9(2): pp. 145. https://doi.org/110.1038/Nmeth.1828

KSH (2019): https://www.ksh.hu/stadat_files/mez/hu/mez0025.html

KYRPIDES, N., OVERBEEK, R., OUZOUNIS, C. (1999): Universal protein families and the functional content of the last universal common ancestor. *Journal of Molecular Evolution*, 49(4) pp. 413-423. https://doi.org/10.1007/PL00006564

LAHENS, N.F., RICCIOTTI, E., SMIRNOVA, O., TOORENS, E., KIM, E.J., BARUZZO, G., HAYER, K.E., GANGULY, T., SCHUG, J. AND GRANT, G.R. (2017): A comparison of Illumina and Ion Torrent sequencing platforms in the context of differential gene expression. *BMC Genomics, 18*(1) pp. 1-13. https://doi.org/10.1186/s12864-017-4011-0

LÄMKE, J., BÄURLE, I. (2017): Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. *Genome Biology*, 18(1) pp.1-11. https://doi.org/10.1186/s13059-017-1263-6

LANGMEAD, B., SALZBERG, S.L. (2012): Fast gapped-read alignment with Bowtie 2. *Nature Methods,* 9(4) pp. 357. https://doi.org/10.1038/nmeth.1923

LANGMEAD, B., TRAPNELL, C., POP, M., SALZBERG, S.L. (2009): Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology,* 10(3) pp. 1-10. https://doi.org/10.1186/gb-2009-10-3-r25

LARRIGAUDIERE, C., GRAELL, J., SALAS, J., VENDRELL, M. (1997): Cultivar differences

in the influence of a short period of cold storage on ethylene biosynthesis in apples. *Postharvest Biology and Technology,* 10(1) pp. 21-27. https://doi.org/10.1016/S0925-5214(97)87274-2

LAURIA, M., ROSSI, V. (2011): Epigenetic control of gene regulation in plants. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms,* 1809(8) pp. 369-378. https://doi.org/10.1016/j.bbagrm.2011.03.002

LAW, J.A., JACOBSEN, S.E. (2010): Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics,* 11(3) pp. 204-220. https://doi.org/10.1038/nrg2719

LEA, P.J., SODEK, L., PARRY, M.A., SHEWRY, P.R., HALFORD, N.G. (2007) Asparagine in plants. *Annals of Applied Biology*, *150*(1) pp.1-26. https://doi.org/10.1111/j.1744-7348.2006.00104.x

LEE, R.C., FEINBAUM, R.L., AMBROS, V. (1993): The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell,* 75(5) pp. 843-854. https://doi.org/10.1016/0092-8674(93)90529-Y

LÉVI, É., MARÓTI, M. (1977): Almaembrió-kultúra és szövettenyészet. *Botanikai Közlemények,* 64 (2) pp. 75-85.

LI, X.Q., GUO, Y.Y., DE, W. (2012): DNA methylation and microRNAs in cancer. *World Journal of Gastroenterology: WJG*, 18(9) pp. 882. https://doi.org/10.3748/wjg.v18.i9.882

LIANG, L., CHANG, Y., LU, J., WU, X., LIU, Q., ZHANG, W., SU, X., ZHANG, B. (2019): Global methylomic and transcriptomic analyses reveal the broad participation of DNA methylation in daily gene expression regulation of *Populus trichocarpa*. *Frontiers in Plant Science,* 10 pp. 243. https://doi.org/10.3389/fpls.2019.00243

LIEBHARD, R., GIANFRANCESCHI, L., KOLLER, B., RYDER, C. D., TARCHINI, R., VAN DE WEG, E., GESSLER, C. (2002): Development and characterisation of 140 new microsatellites in apple (*Malus* × *domestica* Borkh.). *Molecular Breeding,* 10(4) pp. 217-241. https://doi.org/10.1023/A:1020525906332

LIEBHARD, R., KELLERHALS, M., PFAMMATTER, W., JERTMINI, M., GESSLER, C. (2003): Mapping quantitative physiological traits in apple (*Malus*× *domestica* Borkh.). *Plant Molecular Biology,* 52(3) pp. 511-526. https://doi.org/10.1023/A:1024886500979

LIEPMAN, A.H., OLSEN, L.J. (2003): Alanine aminotransferase homologs catalyze the glutamate: glyoxylate aminotransferase reaction in peroxisomes of *Arabidopsis*. *Plant Physiology*, 131(1) pp. 215-227. https://doi.org/10.1104/pp.011460

LIN, W., XIAO, X.O., ZHANG, H., LI, Y., LIU, S., SUN, W., ZHANG, X., WU, Q. (2019): Whole-genome bisulfite sequencing reveals a role for dna methylation in variants from callus culture of pineapple (*Ananas comosus* L.). *Genes*, 10(11) p.877. https://doi.org/10.3390/genes10110877

LISTER, R., O'MALLEY, R.C., TONTI-FILIPPINI, J., GREGORY, B.D., BERRY, C.C., MILLAR, A.H., ECKER, J.R., (2008): Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell,* 133(3) pp. 523-536. https://doi.org/10.1016/j.cell.2008.03.029

LISTER, R., ECKER, J.R. (2009a): Finding the fifth base: Genome-wide sequencing of cytosine methylation. *Genome Research,* 19 (6) pp. 959–966. https://doi.org/10.1101/gr.083451.108

LISTER, R., PELIZZOLA, M., DOWEN, R.H., HAWKINS, R.D., HON, G., TONTI-FILIPPINI, J., NERY, J.R., LEE, L., YE, Z., NGO, Q.M., EDSALL, L. (2009b): Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature,* 462(7271) pp. 315-322. https://doi.org/10.1038/nature08514

LITT, M., LUTY, J.A. (1989): A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat with in the cardiac muscle actin gene. *American Journal of Human Genetics,* 44 pp.397 401

LIU, H., WU, Y., CAO, A., MAO, B., ZHAO, B., WANG, J. (2017): Genome-wide analysis of DNA methylation during ovule development of female-sterile rice fsv1. *G3: Genes, Genomes, Genetics*, 7(11) pp.3621-3635. https://doi.org/10.1534/g3.117.300243

LIU, Q., SUN, Y., CHEN, J., LI, P., LI, C., NIU, G., JIANG, L. (2016): Transcriptome analysis revealed the dynamic oil accumulation in *Symplocos paniculata* fruit. *BMC Genomics,* 17(1) pp. 1-17. https://doi.org/10.1186/s12864-016-3275-0

LIU, Z., SHENG, H., OKORLEY, B. A., LI, Y., & SOSSAH, F. L. (2020): Comparative Genomic Analysis Provides Insights into the Phylogeny, Resistome, Virulome, and Host Adaptation in the Genus Ewingella. *Pathogens*, 9(5) pp. 330. https://doi.org/10.3390/pathogens9050330

LONGHI, S., CAPPELLIN, L., GUERRA, W., COSTA, F. (2013): Validation of a functional molecular marker suitable for marker-assisted breeding for fruit texture in apple (*Malus*× *domestica* Borkh.). *Molecular Breeding,* 32(4) pp. 841-852. https://doi.org/10.1007/s11032-013-9912-2

LUND, R.J., HUHTINEN, K., SALMI, J., RANTALA, J., NGUYEN, E. V., MOULDER, R., GOODLETT, D.R., LAHESMAA R., CARPÉN, O. (2017): DNA methylation and transcriptome changes associated with cisplatin resistance in ovarian cancer. *Scientific Reports,* 7(1) pp. 1-11. https://doi.org/10.1038/s41598-017-01624-4

MA, B., DING, Y., LI, C., LI, M., MA, F., YUAN, Y. (2019): Comparative proteomic analysis reveals key proteins linked to the accumulation of soluble sugars and organic acids in the mature fruits of the wild *Malus* species. *Plants*, 8(11) p. 488. https://doi.org/10.3390/plants8110488

MACHCZYŃSKA, J., ORŁOWSKA, R., MAŃKOWSKI, D.R., ZIMNY, J., BEDNAREK, P.T., (2014): DNA methylation changes in triticale due to *in vitro* culture plant regeneration and

consecutive reproduction. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 119(2), pp. 289-299. https://doi.org/10.1007/s11240-014-0533-1

MARDIS, E.R. (2008): Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics,* 9 pp. 387-402. https://doi.org/10.1146/annurev.genom.9.081307.164359

MARGULIES, M., EGHOLM, M., ALTMAN, W.E., ATTIYA, S., BADER, J.S., BEMBEN, L.A., BERKA, J., BRAVERMAN, M.S., CHEN, Y.J., CHEN, Z., DEWELL, S.B. (2005): Genome sequencing in microfabricated high-density picolitre reactors. *Nature,* 437(7057) pp. 376-380. https://doi.org/10.1038/nature03959

MARTIN, M. (2011): Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal,* 17(1) pp. 10-12. https://doi.org/10.14806/ej.17.1.200

MAXAM, A.M., GILBERT, W. (1977): A new method for sequencing DNA. *Proceedings of the National Academy of Sciences,* 74(2) pp. 560-564. https://doi.org/10.1073/pnas.74.2.560

MENDEL, G. (1866): Versuche über Pflanzenhybriden. Verhandlungen des naturforschenden Vereines in Brünn, 4 pp. 3-47

MEILI, G., WEIXI, G., XUEJUAN, Y., CHUNXI, Z., FUJUN, L., XIN, L. (2012): Data mining of *Arabidopsis thaliana* salt-response proteins based on bioinformatics analysis. *Plant Omics Journal*, 5(2) pp. 75-78.

MEISSNER, A., MIKKELSEN, T.S., GU, H., WERNIG, M., HANNA, J., SIVACHENKO, A., ZHANG, X., BERNSTEIN, B.E., NUSBAUM, C., JAFFE, D.B., GNIRKE, A., JAENISCH, R., LANDER, E.S. (2008): Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature,* 454 pp.766–770. https://doi.org/10.1038/nature07107

MERKEL, A., FERNÁNDEZ-CALLEJO, M., CASALS, E., MARCO-SOLA, S., SCHUYLER, R., GUT, I.G., HEATH, S.C. (2019): gemBS: high throughput processing for DNA methylation data from bisulfite sequencing. *Bioinformatics,* 35(5) pp. 737-742. https://doi.org/10.1093/bioinformatics/bty690

METZKER, M.L. (2010): Sequencing technologies—the next generation. *Nature Reviews Genetics*, 11(1) pp. 31-46. https://doi.org/10.1038/nrg2626

MEYER, E.H., SOLHEIM, C., TANZ, S.K., BONNARD, G., MILLAR, A.H. (2011): Insights into the composition and assembly of the membrane arm of plant complex I through analysis of subcomplexes in *Arabidopsis* mutant lines. *Journal of Biological Chemistry*, 286(29) pp. 26081-26092. https://doi.org/10.1074/jbc.M110.209601

MIURA, A., YONEBAYASHI, S., WATANABE, K., TOYAMA, T., SHIMADA, H., KAKUTANI, T. (2001): Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis. Nature,* 411(6834) pp. 212-214. https://doi.org/10.1038/35075612

MODGIL, M., MAHAJAN, K., CHAKRABARTI, S.K., SHARMA, D.R., SOBTI, R.C. (2005):

Molecular analysis of genetic stability in micropropagated apple rootstock MM106. *Scientia Horticulturae*, 104(2) pp. 151-160. https://doi.org/10.1016/j.scienta.2004.07.009

MOORE, S.C., JASON, L., AUSIÓ, J. (2002): The elusive structural role of ubiquitinated histones. *Biochemistry and Cell Biology,* 80(3) pp. 311-319. https://doi.org/10.1139/o02-081

MORITA, H., MIZUUCHI, Y., ABE, T., KOHNO, T., NOGUCHI, H., ABE, I. (2007): Cloning and functional analysis of a novel aldo-keto reductase from *Aloe arborescens*. *Biological and Pharmaceutical Bulletin*, 30(12) pp.2262-2267. https://doi.org/10.1248/bpb.30.2262

MOYO, M., BAIRU, M.W., AMOO, S.O., VAN STADEN, J. (2011): Plant biotechnology in South Africa: Micropropagation research endeavours, prospects and challenges. *South African Journal of Botany,* 77(4) pp. 996-1011. https://doi.org/10.1016/j.sajb.2011.06.003

MURASHIGE, T., SKOOG, F. (1962): A revised medium for rapid growth and bio agsays with tohaoco tissue cultures. *Physiologia Plantarum*, 15(3) pp. 473-497. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x

MURRELL, A., RAKYAN, V.K., BECK, S. (2005): From genometo epigenome. *Human Molecular Genetics,* 14(1) pp. 3–10. https://doi.org/10.1093/hmg/ddi110

MURUMKAR, C.V., CHAVAN, P.D. (1985): Changes in the inorganic status and enzyme activities in senescent leaves of chickpea, *Cicer arietinum* L. *Acta Societatis Botanicorum Poloniae*, 54(4) pp.391-401. https://doi.org/10.5586/asbp.1985.034

NAQVI, A.R., ISLAM, M.N., CHOUDHURY, N.R., HAQ, Q.M.R. (2009): The fascinating world of RNA interference. *International Journal of Biological Sciences, 5*(2) pp. 97. https://doi.org/10.7150/ijbs.5.97

NAQVI, S., FARRÉ, G., ZHU, C., SANDMANN, G., CAPELL, T., CHRISTOU, P. (2011): Simultaneous expression of *Arabidopsis* ρ-hydroxyphenylpyruvate dioxygenase and MPBQ methyltransferase in transgenic corn kernels triples the tocopherol content. *Transgenic research*, 20(1) pp. 177- 181. https://doi.org/10.1007/s11248-010-9393-6

NIEDERHUTH, C.E., BEWICK, A.J., JI, L., ALABADY, M.S., DO KIM, K., LI, Q., ROHR, N.A., RAMBANI, A., BURKE, J.M., UDALL, J.A., EGESI, C. (2016): Widespread natural variation of DNA methylation within angiosperms. *Genome biology,* 17(1) pp. 1-19. https://doi.org/10.1186/s13059-016-1059-0

NIKAM, P.H., KAREPARAMBAN, J.A., JADHAV, A.P., KADAM V.J. (2013): Development and validation of high-performance thin layer chromatographic method for ursolic acid in *Malus domestica* peel. *Indian Journal of Pharmaceutical Sciences,* 75(4) pp. 486-489. https://doi.org/10.4103/0250-474X.119815

OH, T.J., CULLIS, M.A., KUNERT, K., ENGELBORGHS, I., SWENNEN, R., CULLIS, C.A. (2007): Genomic changes associated with somaclonal variation in banana (*Musa* spp.).

Physiologia Plantarum, 129(4) pp. 766-774. https://doi.org/10.1111/j.1399-3054.2007.00858.x

OWENS, D.K., CROSBY, K.C., RUNAC, J., HOWARD, B.A., WINKEL, B.S. (2008): Biochemical and genetic characterization of *Arabidopsis* flavanone 3β-hydroxylase. *Plant Physiology and Biochemistry*, 46(10) pp. 833-843. https://doi.org/10.1016/j.plaphy.2008.06.004

PALOMBI, M., DAMIANO, C. (2002): Comparison between RAPD and SSR molecular markers in detecting genetic variation in kiwifruit (*Actinidia deliciosa* A. Chev). *Plant Cell Reports*, 20(11) pp. 1061-1066. https://doi.org/10.1007/s00299-001-0430-z

PALUSA, S.G., GOLOVKIN, M., SHIN, S.B., RICHARDSON, D.N., REDDY, A.S. (2007): Organ-specific, developmental, hormonal and stress regulation of expression of putative pectate lyase genes in *Arabidopsis*. *New phytologist*, 174(3) pp. 537-550. https://doi.org/10.1111/j.1469-8137.2007.02033.x

PANDEY, R.N., SINGH, S.P., RASTOGI, J., SHARMA, M.L., SINGH, R.K. (2012): Early assessment of genetic fidelity in sugarcane (*'Saccharum officinarum'*) plantlets regenerated through direct organogenesis with RAPD and SSR markers. *Australian Journal of Crop Science*, 6(4) pp. 618-624.

PANDEY, V., NUTTER, R.C. PREDIGER, E. (2008): Applied biosystems solid™ system: ligation-based sequencing. Next Generation Genome Sequencing: Towards Personalized. *Medicine,* pp. 29-42. https://doi.org/10.1002/9783527625130.ch3

PAPP, D. (2017): Almafajták és nemesítési génforrások betegség-ellenállóságának karakterizálása fenotipizálással és molekuláris markerezéssel. *PhD thesis* Gödöllő pp. 40-96.

PARALKAR, V.R., WEISS, M.J. (2013): Long noncoding RNAs in biology and hematopoiesis. *Blood, The Journal of the American Society of Hematology,* 121(24) pp. 4842-4846. https://doi.org/10.1182/blood-2013-03-456111

PEIL, A., KELLERHALS, M., HÖFER, M., FLACHOWSKY, H. (2011): Apple breeding—from the origin to genetic engineering. *Fruit Vegetable and Cereal Science and Biotechnology,* 5 pp. 118-138.

PERRIN, A., DACCORD, N., ROQUIS, D., CELTON, J.M., VERGNE, E., BUCHER, E. (2020): Divergent DNA methylation signatures of juvenile seedlings, grafts and adult apple trees. *Epigenomes*, 4(1) pp. 4. https://doi.org/10.3390/epigenomes4010004

PETERS, S., EGERT, A., STIEGER, B., KELLER, F. (2010): Functional identification of *Arabidopsis* ATSIP2 (At3g57520) as an alkaline α-galactosidase with a substrate specificity for raffinose and an apparent sink-specific expression pattern. *Plant and Cell Physiology*, 51(10) pp. 1815-1819.https://doi.org/10.1093/pcp/pcq127

PETTERSSON, E., LUNDEBERG, J., AHMADIAN, A. (2009): Generations of sequencing technologies. *Genomics,* 93(2) pp. 105-111.

PHILLIPS, T. (2008): The role of methylation in gene expression. *Nature Education,* 1(1) pp. 116.

POTESTÀ, M., ROGLIA, V., FANELLI, M., PIETROBONO, E., GISMONDI, A., VUMBACA, S., TSANGUEU, R.G.N., CANINI, A., COLIZZI, V., GRELLI, S. AND MINUTOLO, A. (2020): Effect of microvesicles from Moringa oleifera containing miRNA on proliferation and apoptosis in tumor cell lines. *Cell Death Discovery,* 6(1) pp. 1-17. https://doi.org/10.1038/s41420-020-0271-6

POULTON, J.E. (1990): Cyanogenesis in plants. *Plant Physiology*, *94*(2) pp.401-405. https://doi.org/10.1104/pp.94.2.401

PRADO, M.J., RODRIGUEZ, E., REY, L., GONZÁLEZ, M.V., SANTOS, C., REY, M. (2010): Detection of somaclonal variants in somatic embryogenesis-regenerated plants of *Vitis vinifera* by flow cytometry and microsatellite markers. *Plant Cell, Tissue and Organ Culture* (PCTOC), 103(1) pp. 49-59. https://doi.org/10.1007/s11240-010-9753-1

QUADRANA, L., COLOT, V. (2016): Plant transgenerational epigenetics. *Annual Review of Genetics,* 50 pp. 467-491. https://doi.org/10.1146/annurev-genet-120215-035254

RAHMAN, M., RAJORA, O. (2001): Microsatellite DNA somaclonal variation in micropropagated trembling aspen (*Populus tremuloides*). *Plant Cell Reports*, *20*(6) pp. 531-536. https://doi.org/10.1007/s002990100365

RAI, M.K., PHULWARIA, M., GUPTA, A.K., SHEKHAWAT, N.S., JAISWAL, U., 2012. Genetic homogeneity of guava plants derived from somatic embryogenesis using SSR and ISSR markers. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 111(2), pp.259-264. https://doi.org/10.1007/s11240-012-0190-1

RAJKUMAR, M.S., SHANKAR, R., GARG, R., JAIN, M. (2019): Role of DNA methylation dynamics in desiccation and salinity stress responses in rice cultivars. *BioRxiv*, p.558064. https://doi.org/10.1101/558064

RATHORE, M.S., MASTAN, S.G., AGARWAL, P.K. (2015): Evaluation of DNA methylation using methylation-sensitive amplification polymorphism in plant tissues grown *in vivo* and *in vitro*. *Plant Growth Regulation*, 75(1) pp.11-19. https://doi.org/10.1007/s10725-014-9926-8

RAWOOF, A., AHMAD, I., RAMCHIARY, N. (2019): Epigenome landscape in *Capsicum* genome. *In The Capsicum Genome Springer*, Cham pp. 187-199. https://doi.org/10.1007/978-3-319-97217-6_11

RÖDER, M.S., KORZUN, V., WENDEHAKE, K., PLASCHKE, J., TIXIER, M.H., LEROY, P., GANAL, M.W. (1998): A microsatellite map of wheat. *Genetics,* 149(4) pp. 2007-2023. https://doi.org/10.1093/genetics/149.4.2007

ROOMI, S., MASI, A., CONSELVAN, G.B., TREVISAN, S., QUAGGIOTTI, S., PIVATO, M., ARRIGONI, G., YASMIN, T., CARLETTI, P. (2018): Protein profiling of *Arabidopsis* roots

treated with humic substances: insights into the metabolic and interactome networks. *Frontiers in Plant Science*, 9 1812. https://doi.org/10.3389/fpls.2018.01812

ROTHBERG, J.M., HINZ, W., REARICK, T.M., SCHULTZ, J., MILESKI, W., DAVEY, M., LEAMON, J.H., JOHNSON, K., MILGREW, M.J., EDWARDS, M., HOON, J. (2011): An integrated semiconductor device enabling non-optical genome sequencing. *Nature,* 475(7356) pp. 348-352. https://doi.org/10.1038/nature10242

ROTHKEGEL, K., SANDOVAL, P., SOTO, E., ULLOA, L., RIVEROS, A., LILLO-CARMONA, V., JAVIER CÁCERES-MOLINA, J., ALMEIDA A.M., MENESES, C. (2020): Dormant but active: chilling accumulation modulates the epigenome and transcriptome of *Prunus avium* during bud dormancy. *Frontiers in Plant Science,* 11 pp. 1115. https://doi.org/10.3389/fpls.2020.01115

ROUDIER, F., TEIXEIRA, F.K., COLOT, V. (2009): Chromatin indexing in *Arabidopsis*: an epigenomic tale of tails and more. *Trends in Genetics,* 25(11) pp. 511-517. https://doi.org/10.1016/j.tig.2009.09.013

RUSSO, V.E.A., RIGGS, A.D., MARTIENSSEN, R.A. (1996): Epigenetic mechanisms of gene regulation. *Cold Spring Harbor Laboratory Press*

SADDER, M.T., SUWWAN, M.A. (2006): Detection of *de novo* β-glucosidase (EC 3.2.1.21) Enzyme Activity in Tomato (*Solanum lycopersicum* L.) in 4-methylumbelliferyl ß-D-glucoside Zymograms. *Jordan Journal of Agricultural Sciences*, *2*(2). pp. 75-82.

SAHU, J., DAS TALUKDAR, A., DEVI, K., CHOUDHURY, M.D., BAROOAH, M., MODI, M. K., SEN, P. (2015): E-microsatellite markers for *Centella asiatica* (Gotu Kola) genome: validation and cross-transferability in *Apiaceae* family for plant omics research and development. *Omics: a Journal of Integrative Biology*, 19(1) pp. 52-65. https://doi.org/10.1089/omi.2014.0113

SANGER, F., NICKLEN, S., COULSON, A. R. (1977): DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences,* 74(12) pp. 5463-5467. https://doi.org/10.1073/pnas.74.12.5463

SCHAFFER, W.I. (1990): Terminology associated with cell, tissue and organ culture, molecular biology and molecular genetics. *In vitro Cellular & Developmental Biology,* 26(1) pp. 97-101. https://doi.org/10.1007/BF02624162

SCHNABL, H., WEISSENBÖCK, G., SCHARF, H. (1986): *In vivo*-microspectrophotometric characterization of flavonol glycosides in *Vicia faba* guard and epidermal cells. J*ournal of Experimental Botany*, 37(1) pp. 61-72. https://doi.org/10.1093/jxb/37.1.61

SCHUETTENGRUBER, B., CHOURROUT, D., VERVOORT, M., LEBLANC, B., CAVALLI, G. (2007): Genome regulation by polycomb and trithorax proteins. *Cell,* 128 pp. 735–745. https://doi.org/10.1016/j.cell.2007.02.009

SEYMOUR, D.K., KOENIG, D., HAGMANN, J., BECKER, C., WEIGEL, D. (2014): Evolution of DNA methylation patterns in the *Brassicaceae* is driven by differences in genome organization. *PLoS Genet,* 10(11) p.e1004785. https://doi.org/10.1371/journal.pgen.1004785

SHALAEVA, D.N., CHEREPANOV, D.A., GALPERIN, M.Y., GOLOVIN, A.V., MULKIDJANIAN, A.Y. (2018): Evolution of cation binding in the active sites of P-loop nucleoside triphosphatases in relation to the basic catalytic mechanism. *Elife*, 7, p.e37373. https://doi.org/10.7554/eLife.37373

SHEN, H., HE, H., LI, J., CHEN, W., WANG, X., GUO, L., PENG, Z., HE, G., ZHONG, S., QI, Y., TERZAGHI, W., DENG, X.W. (2012): Genome-wide analysisof DNA methylation and gene expression changes in two *Arabidopsis* ecotypes and their reciprocal hybrids. *Plant Cell,* 24 pp. 875–892. https://doi.org/10.1105/tpc.111.094870

SIKORSKAITE, S., GELVONAUSKIENE, D., STANYS, V., BANIULIS, D. (2012): Characterization of microsatellite loci in apple (*Malus* × d*omestica* Borkh.) cultivars. *Žemdirbystė Agriculture*, 99(2) pp. 131-138.

TAKAHA, T. (1996): Isolation of disproportionating enzyme (EC 2.4. 1.25) from potato and investigation of its role in starch metabolism. *Doctoral dissertation*, University of Edinburgh

TAUTZ, D., TRICK, M., DOVER, G.A. (1986): Cryptic simplicity in DNA is a major source of genetic variation. *Nature,* 322 pp. 652-656.

TEIXEIRA DA SILVA, J.A., GULYÁS, A., MAGYAR-TÁBORI, K., WANG, M.R., WANG, Q.C., DOBRÁNSZKI, J. (2019): *In vitro* tissue culture of apple and other *Malus* species: recent advances and applications. *Planta,* 249(4) pp. 975-1006. https://doi.org/10.1007/s00425-019-03100-x

TERPÓ, A. (1987): Növényrendszertan az ökonómbotanika alapjaival 1-2. *Mezőgazdasái Kiadó*, Budapest.

TIRADO-MAGALLANES, R., REBBANI, K., LIM, R., PRADHAN, S., BENOUKRAF, T. (2017): Whole genome DNA methylation: beyond genes silencing. *Oncotarget,* 8(3) pp. 5629. https://doi.org/10.18632/oncotarget.13562

TIWARI, J.K., CHANDEL, P., GUPTA, S., GOPAL, J., SINGH, B.P., BHARDWAJ, V. (2013): Analysis of genetic stability of *in vitro* propagated potato microtubers using DNA markers. *Physiology and Molecular Biology of Plants*, 19(4) pp.587-595. https://doi.org/10.1007/s12298-013-0190-6

TOMCSÁNYI, P., BÖDECS, L., MAJOROS, L. (1982): A gyümölcsfajtákról: Almagyümölcsűek és bogyósok In: LELKES, L., APJOK F. (ed) *Mezőgazdasági Könyvkiadó Vállalat*, Budapest pp.18-19.

TORRES, K.C. (2012): Tissue culture techniques for horticultural crops. *Springer Science &*

Business Media, pp.285.

TURNER, A.P., LONES, M.A., FUENTE, L.A., STEPNEY, S., CAVES, L.S., TYRRELL, A.M. (2013): The incorporation of epigenetics in artificial gene regulatory networks. *BioSystems,* 112(2) pp. 56-62. https://doi.org/10.1016/j.biosystems.2013.03.013

TYLER, A.D., MATASEJE, L., URFANO, C. J., SCHMIDT, L., ANTONATION, K.S., MULVEY, M.R., CORBETT, C.R. (2018): Evaluation of Oxford Nanopore's MinION sequencing device for microbial whole genome sequencing applications. *Scientific Reports*, 8(1) pp. 1-12. https://doi.org/10.1038/s41598-018-29334-5

URRESTARAZU, J., MIRANDA, C., SANTESTEBAN, L.G., ROYO, J.B. (2012): Genetic diversity and structure of local apple cultivars from Northeastern Spain assessed by microsatellite markers. *Tree Genetics & Genomes*, 8(6) pp. 1163-1180. https://doi.org/10.1007/s11295-012-0502-y

US-CAMAS, R., RIVERA-SOLÍS, G., DUARTE-AKÉ, F., DE-LA-PENA, C. (2014): *In vitro* culture: an epigenetic challenge for plants. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 118(2) pp.187-201. https://doi.org/10.1007/s11240-014-0482-8

VADDEPALLI, P., FULTON, L., WIELAND, J., WASSMER, K., SCHAEFFER, M., RANF, S., SCHNEITZ, K. (2017): The cell wall-localized atypical β-1, 3 glucanase ZERZAUST controls tissue morphogenesis in *Arabidopsis thaliana*. *Development*, 144(12) pp. 2259-2269. https://doi.org/10.1242/dev.152231

VAN TREUREN, R., KEMP, H., ERNSTING, G., JONGEJANS, B., HOUTMAN, H., VISSER, L. (2010): Microsatellite genotyping of apple (*Malus* × *domestica* Borkh.) genetic resources in the Netherlands: application in collection management and variety identification. *Genetic Resources and Crop Evolution,* 57(6) pp. 853-865. https://doi.org/10.1007/s10722-009-9525-0

VELASCO, R., ZHARKIKH, A., AFFOURTIT, J., DHINGRA, A., CESTARO, A., KALYANARAMAN, A., FONTANA, P., BHATNAGAR, S. K., TROGGIO, M., PRUSS, D., SALVI, S., PINDO, M., BALDI, P., CASTELLETTI, S., CAVAIUOLO, M., COPPOLA, G., COSTA, F., COVA, V., DAL RI, A., GOREMYKIN, V., KOMJANC, M., LONGHI, S., MAGNAGO, P., MALACARNE, G., MALNOY, M., MICHELETTI, D., MORETTO, M., PERAZZOLLI, M., SI-AMMOUR, A., VEZZULLI, S., ZINI, E., ELDREDGE, G., FITZGERALD, L.M., GUTIN, N., LANCHBURY, J., MACALMA, T., MITCHELL, J. T., REID, J., WARDELL, B., KODIRA, C., CHEN, Z., DESANY, B., NIAZI, F., PALMER, M., KOEPKE, T., JIWAN, D., SCHAEFFER, S., KRISHNAN, V., WU, C., CHU, V. T., KING, S. T., VICK, J., TAO, Q., MRAZ, A., STORMO, A., STORMO, K., BOGDEN, R., EDERLE, D., STELLA, A., VECCHIETTI, A., KATER, M. M., MASIERO, S., LASSERRE, P., LESPINASSE, Y., ALLAN, A.C., BUS, V., CHAGNE, D., CROWHURST, R. N., GLEAVE,

A. P., LAVEZZO, E., FAWCETT, J. A., PROOST, S., ROUZE, P., STERCK, L., TOPPO, S., LAZZARI, B., HELLENS, R. P., DUREL, C. E., GUTIN, A., BUMGARNER, R. E., GARDINER, S. E., SKOLNICK, M., EGHOLM, M., VAN DE PEER, Y., SALAMINI, F., VIOLA, R. (2010): The genome of the domesticated apple (*Malus* × *domestica* Borkh.). *Nature Genetics,* 42(10) pp. 833-839. https://doi.org/10.1038/ng.654

VILLA, R., DE SANTIS, F., GUTIERREZ, A., MINUCCI, S., PELICCI, P. G., DI CROCE, L. (2004): Epigenetic gene silencing in acute promyelocytic leukemia. *Biochemical Pharmacology,* 68(6) pp. 1247-1254. https://doi.org/10.1016/j.bcp.2004.05.041

VIGGIANO, L., DE PINTO, M.C. (2017): Dynamic DNA Methylation Patterns in Stress Response. In: RAJEWSKY N., JURGA S., BARCISZEWSKI J. (eds) *Plant Epigenetics. RNA Technologies. Springer, Cham* pp. 281-302. https://doi.org/10.1007/978-3-319-55520-1_15

VIRŠCEK-MARN, M., JAVORNIK, B., ŠTAMPAR, F., BOHANEC, B. (1996): Assessment of genetic variation among regenerants from *in vitro* apple leaves using molecular markers. *In Eucarpia Symposium on Fruit Breeding and Genetics,* 484 pp. 299-304.

VOIGT, C.A., SCHÄFER, W., SALOMON, S. (2006): A comprehensive view on organ-specific callose synthesis in wheat (*Triticum aestivum* L.): glucan synthase-like gene expression, callose synthase activity, callose quantification and deposition. *Plant Physiology and Biochemistry*, 44(4) pp.242-247. https://doi.org/10.1016/j.plaphy.2006.05.001

WADDINGTON, C.H. (1942): “The Epigenotype”. *Endeavour,* 1 pp. 18–20. https://doi.org/10.17660/ActaHortic.1998.484.52

WANG, P., SU, L., JIANG, X., CHEN, Y., REN, F., WANG, Y. (2017): Identification and Analysis of SSRs Derived from Protein-coding Genes in Grape. *Agricultural Science & Technology*, 18(9) pp. 1579-1584.

WANG, L., XIE, J., HU, J., LAN, B., YOU, C., LI, F., WANG, Z., WANG, H. (2018). Comparative epigenomics reveals evolution of duplicated genes in potato and tomato. *The Plant Journal*, 93(3) pp.460-471. https://doi.org/10.1111/tpj.13790

WANG, X.L., XU, Y.H., PENG, C.C., FAN, R.C., GAO, X.Q. (2009): Ubiquitous distribution and different subcellular localization of sorbitol dehydrogenase in fruit and leaf of apple. *Journal of Experimental Botany*, 60(3) pp. 1025-1034. https://doi.org/10.1093/jxb/ern347

WAY, R., ALDWINCKLE, H., LAMB, R., REJMAN, A., SANSAVINI, S., SHEN, T., WATKINS, R., WESTWOOD, M. YOSHIDA, Y. (1991): Apples (*Malus*). *Genetic Resources of Temperate Fruit and Nut Crops,* 290 pp. 3-46. https://doi.org/10.17660/ActaHortic.1991.290.1

WEBER, J.L., MAY, P.E. (1989): Abundant class of human DNA polymorphism which can be typed using the polymerase chain reaction. *American Journal of Human Genetics,* 44 pp. 388-396.

WEST, M.H., BONNER, W.M. (1980): Histone 2B can be modified by the attachment of ubiquitin. *Nucleic Acids Research,* 8(20) pp. 4671-4680. https://doi.org/10.1093/nar/8.20.4671

WEST, P.T., LI, Q., JI, L., EICHTEN, S.R., SONG, J., VAUGHN, M.W., SCHMITZ, R.J., SPRINGER, N.M. (2014): Genomic distribution of H3K9me2 and DNA methylation in a maize genome. *PLoS One,* 9(8): p.e105267. https://doi.org/10.1371/journal.pone.0105267

WICHMANN, B., GALLI, ZS., SZABÓ, T., KOVÁCS, L., HESZKY, L., KISS, E. (2010): Alma kereskedelmi és tájfajták elkülönítése európai standard SSR markerekkel. *Kertgazdaság* 42:(1) pp. 68-75.

WIGHTMAN, B., HA, I., RUVKUN, G. (1993): Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell,* 75(5) PP. 855-862. https://doi.org/10.1016/0092-8674(93)90530-4

XING, M.Q., ZHANG, Y.J., ZHOU, S.R., HU, W.Y., WU, X.T., YE, Y.J., WU, X.X., XIAO, Y.P., LI, X., Xue, H.W. (2015): Global analysis reveals the crucial roles of DNA methylation during rice seed development. *Plant Physiology*, *168*(4), pp.1417-1432. https://doi.org/10.1104/pp.15.00414

XU, J., XING, X.J., TIAN, Y.S., PENG, R. H., XUE, Y., ZHAO, W., YAO, Q.H. (2015): Transgenic *Arabidopsis* plants expressing tomato glutathione S-transferase showed enhanced resistance to salt and drought stress. *PloS One*, 10(9) p.e0136960. https://doi.org/10.1371/journal.pone.0136960

YAN, G., LONG, H., SONG, W., CHEN, R. (2008): Genetic polymorphism of *Malus sieversii* populations in Xinjiang. *China. Genetic Resources and Crop Evolution,* 55(1) pp. 171-181. https://doi.org/10.1007/s10722-007-9226-5

YU, Z., PESTELL, R.G. (2012): Small non-coding RNAs govern mammary gland tumorigenesis. *Journal of Mammary Gland Biology and Neoplasia,* 17(1) pp. 59-64. https://doi.org/10.1007/s10911-012-9246-4

ZEMACH, A., KIM, M.Y., HSIEH, P.H., COLEMAN-DERR, D., ESHED-WILLIAMS, L., THAO, K., HARMER, S.L., ZILBERMAN, D. (2013): The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell,* 153(1) pp. 193-205. https://doi.org/10.1016/j.cell.2013.02.033

ZEMACH, A., MCDANIEL, I. E., SILVA, P., ZILBERMAN, D. (2010): Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science,* 328(5980) pp. 916-919. https://doi.org/10.1126/science.1186366

ZHANG, C., CHEN, X., HE, T., LIU, X., FENG, T., YUAN, Z. (2007): Genetic structure of *Malus sieversii* population from Xinjiang, China, revealed by SSR markers. *Journal of Genetics and Genomi,* 34(10) pp. 947-955. https://doi.org/10.1016/S1673-8527(07)60106-4

ZHAO, F., DURNER, J., WINKLER, J.B., TRAIDL-HOFFMANN, C., STROM, T.M., ERNST, D., FRANK, U. (2017): Pollen of common ragweed (*Ambrosia artemisiifolia* L.): Illumina-based de novo sequencing and differential transcript expression upon elevated NO2/O3. *Environmental Pollution,* 224 pp. 503-514. https://doi.org/10.1016/j.envpol.2017.02.032

ZHI, X., ZHANG, Y., HU, X., WU, J., LIAO, X. (2008): Inactivation of apple pectin methylesterase induced by dense phase carbon dioxide. *Journal of Agricultural and Food Chemistry*, 56(13) pp. 5394-5400. https://doi.org/10.1021/jf800260c

ZHONG, S., FEI, Z., CHEN, Y.R., ZHENG, Y., HUANG, M., VREBALOV, J., MCQUINN, R., GAPPER, N., LIU, B., XIANG, J., SHAO, Y. (2013): Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nature Biotechnology,* 31(2) pp. 154-159. https://doi.org/10.1038/nbt.2462

ZHU, C., LI, X., ZHENG, J. (2018): Transcriptome profiling using Illumina-and SMRT-based RNA-seq of hot pepper for in-depth understanding of genes involved in CMV infection. *Gene,* 666 pp.123-133. https://doi.org/10.1016/j.gene.2018.05.004

ZIEGLER, J., FACCHINI, P. J., GEIßLER, R., SCHMIDT, J., AMMER, C., KRAMELL, R., VOIGTLÄNDER, S., GESELL, A., PIENKNY, S., BRANDT, W. (2009): Evolution of morphine biosynthesis in opium poppy. *Phytochemistry*, 70(15-16) pp. 1696-1707. https://doi.org/10.1016/j.phytochem.2009.07.006

ZOHARY, D., HOPF, M. (2000): Domestication of plants in the old world: the origin and spread of cultivated plants in West Asia, Europe, and the Nile Valley. *Oxford University Press Oxon,* pp. 316. https://doi.org/10.1016/B0-12-226865-2/00079-1

9.2. Supplementary data

Supplementary Table 1. MURASHIGE and SKOOG (MS) basal medium (MURASHIGE and SKOOG 1962)

MS (L)	‘McIntosh’	‘Húsvéti rozmaring’
MS macro		
KNO_3	1.9 g/L	1.9 g/L
NH_4NO_3	1.65 g/L	1.65 g/L
$CaCl_2x2H_2O$	0.44 g/L	0.44 g/L
$MgSO_4x7H_2O$	0.37 g/L	0.37 g/L
KH_2PO_4	0.17 g/L	0.17 g/L
MS micro		
$MnSO_4xH_2O$	0.0169 g/L	0.0169 g/L
$ZnSO_4x7\ H_2O$	0.0086 g/L	0.0086 g/L
H_3BO_3	0.0062 g/L	0.0062 g/L
KJ	0.00083 g/L	0.00083 g/L
$CuSO_4x5H_2O$	0.000025 g/L	0.000025 g/L
$Na_2MoO_4x2H_2O$	0.00025 g/L	0.00025 g/L
$CoCl_2x6H_2O$	0.000025 g/L	0.000025 g/L
Inozitol (1ml/L)	100 mg/L	100 mg/L
Apple iron		
NaFeEDTA	0.04059 g/L	0.04059 g/L
$FeSO_4x7H_2O$	0.0278 g/L	0.0278 g/L
Vitamins		
Thiamine (B_1)	0.8mg/L	0.8mg/L
Pyridoxine (B_6)	1.0 mg/L	1.0 mg/L
Nicotinic acid (B_3)	1.0 mg/L	1.0 mg/L
Glicin	4.0 mg/L	4.0 mg/L
Biotin	0.08 mg/L	0.08 mg/L
BA	1.0 ml/L	
BAR		0.5 ml/L
IBA	0.3 ml/L	0.3 ml/L
GA_3	0.2 ml/L	0.2 ml/L
Humus	1.0 ml/L	

Sucrose	30g/L	
Agar	5.5 g/L	
pH	5.8	

Supplementary Table 2. Root induction medium (DOBRÁNSZKI *et al*. 2000)

R72	1l
½ MS macro	50.0 ml/L
½ MS micro	0.5 ml/L
½ apple iron	5.0 ml/L
B_1 vitamin	0.5 mg/L
Inositol	100.0 mg/L
IBA	2.0 mg/L
Agar	7.5 g/L
Sucrose	20.0g/L
pH	5.5

Supplementary Table 3. Root induction medium (DOBRÁNSZKI *et al*. 2000)

R6	1l
½ MS macro	50.0 ml/L
½ MS micro	0.5 ml/L
½ apple iron	5.0 ml/L
Wuxal®	2.0 ml/L
Inositol	50.0 mg/L
Agar	6.5 g/L
Sucrose	30.0g/L
pH	5.5

10. ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Ezsébet Kiss for giving me the opportunity to complete my PhD. Dr. Judit Dobránszkin for supporting my research and providing me with the opportunity to complete my PhD and for her support and assistance in preparing publications, dissertations, and theses. I would like to thank Norbert Hidvégi for his help and advice in bioinformatics. To Jonathan to depict the evaluation of the data, and to Huba and Micy for their selfless help and advice. I would also like to thank my family and my friends for selflessly supporting me during my PhD years. Respectively to all my present and old colleagues who contributed to the birth of this dissertation.