Molecular genetic analysis of somaclonal variations in *in vitro* cultured *Vanilla planifolia* Andrews

by

Chai Li Chin, BSc.

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

School of Biosciences, Faculty of Science

JUNE, 2019

Research Committee

Supervisor	**Dr. Chin Chiew Foan** (University of Nottingham, MY)
Co-supervisors	**Dr. Peter G. Alderson** (University of Nottingham, UK; retired) **Dr. Lai Kong Song** (University Putra Malaysia, MY) **Dr. Ho Wan Yong** (University of Nottingham, MY)

This study was financially supported by the Fundamental Research Grant Scheme (FRGS) by Ministry of Education (MOE) and MyBrain15 PhD Scholarship Scheme by Ministry of Higher Education (MOHE).

ABSTRACT

Somaclonal variations (SV) in micropropagated in the vanilla orchid cultivar *V. planifolia* Andrews are not well-studied so far. Understanding the occurrence of SV in vanilla is important for the development of an effective pre-screening tool for clonal fidelity in *in vitro* mass propagation of elite vanilla genotypes and for the selection process during *in vitro* crop improvements of vanilla, which has a narrow genetic basis due to being extensively cloned. In this research, two main regeneration systems, the direct nodal regeneration using optimal shoot induction medium (SIM, Murashige and Skoog (MS) medium + 1.0 mg/L 6-benzylaminopurine (BAP) and the indirect nodal regeneration using optimal callus induction medium (CIM, ½ MS + 2.5 mg/L α-naphthalene acetic acid (NAA) + 1.0 mg/L BAP) were established. Regenerants from both direct and indirect nodal regeneration pathways were studied at morphological, genetic, epigenetic and cytological levels. Morphological abnormalities, such as irregular leaf shape and sizes, and formation of multiple shoot apices, were observed in callus regenerants of vanilla. Further analyses at the genetic and epigenetic levels using inter-simple sequence repeats (ISSR) and methylation-sensitive amplified polymorphisms (MSAP) DNA markers also showed that the indirectly regenerated vanilla cultures were both genetically and epigenetically unstable, as opposed to the directly regenerated vanilla cultures. Using flow cytometry (FCM), relative genome size estimation of vanilla (2C = 5.5 ± 0.27 pg) was achieved with *P. sativum* as the internal control. Endopolyploidy was observed in different organ explants of vanilla, with irregular marginal replication ratios of 1.33 ± 0.02, 1.42 ± 0.02, 1.60 ± 0.01, 1.76 ± 0.01 and 1.69 ± 0.03 (instead of 2) between respective peaks, which is called "progressively partial endoreduplications, a unique phenomenon seen in some orchid species. The

marginal replication ratio was constant between different tissues, but ploidy population levels varied among different tissues. Axillary buds have a dominant 1.5C peak and relatively smaller populations of other endopolyploidy peaks, as opposed to leaf and root and callus tissues which have a relatively balanced population of different mixoploids. Prolonged culture period of callus cultures induced from nodal bud explants (with dominant 1.5C peak) showed a mixoploid population, which increased in the polyploidy population and overall nuclear content (pg) over prolonged culture duration up to 12 months. To obtain a global view of molecular stress response to prolonged culture duration of vanilla callus culture, we further studied by conducting *de novo* mRNA and microRNA-seq of the 3M, 6M and 12M vanilla calli. A series of genes involved in five key gene families, namely DNA methylation family (found in the KEGG pathway of cysteine and methionine metabolism), retrotransposable element family (in particular the *copia* and *gypsy*-like LTR retrotransposons), histone modification family (in particular the HDACs and HATs), DNA transcription factors (in particular the WRKY transcription factors) and microRNAs were found to be differentially expressed. Three miRNAs, namely the vpl_miR395g, vpl_miR166r and vpl_miR8155 expressed in prolonged vanilla callus cultures, were found to potentially regulate eight genes, which fell into the functional gene families of DNA methylation and retrotransposable elements; which results were validated by qRT-PCR. This study has shown that a direct nodal regeneration method is more suitable for clonal mass propagation of elite vanilla genotypes as it produced genetically and epigenetically stable regenerants; while the indirect nodal regeneration could be a potential tool for increasing the genetic diversity of vanilla through *in vitro* improvements. At the transcriptional level of study, we observed the potential cross-talk between retrotransposable elements, microRNA and DNA methylation, which could have led to the epigenetic regulation of SV in vanilla callus. For clonal fidelity screening process for SV in vanilla, it is recommended to combine morphological, genetic, epigenetic and cytological methods for more effective screening. The microRNAs and their putative target genes could be potential targets for *in vitro* crop improvements of vanilla as well as screening tools, however, further characterization work is required.

Dedicated to my beloved family and late Ms. Wong Quin Nee

ACKNOWLEDGEMENTS

First of all, my deepest gratitude to my principal supervisor, Dr Chin Chiew Foan, for her continuous support, guidance and trust throughout this project. Besides my mentor, I'd also like to thank the rest of my thesis committee: Prof. Peter G. Alderson, Dr. Ho Wan Yong and Dr. Lai Kok Song (UPM), for their insightful comments, support and encouragements. Deep appreciation to the assessor of this work, especially Dr. Winnie Yap, for her critical comments which have incented me to widen my research perspectives.

My biggest love goes to my wonderful family and loving relatives, as their warmth and love have been the biggest pillar of support through this journey. Special thanks to my beloved friends and colleagues, particularly, late Ms. Wong Quin Nee, CK, Xiao Ying, Marcus, Shiau Ying, Hui Hui, Wai Kuan, Boon Chin, Nengi, Johnathan, Bobby, Zheng Yew, De Yan, Ee Leen and Jimmy, for their stimulating discussions, for the sleepless nights working together and all the mental and physical support through the challenges faced throughout this PhD duration. I'm also grateful to the BRC lab committee and our lab technicians, especially Ms Siti Lin, Ms Norasyikin, Ms Asma and Ms Shankari, who provided assistance during the conduct of my laboratory experiments.

My acknowledgement to the Ministry of Science, Technology and Innovation (MOSTI) Malaysia for providing funding to support this research through Fundamental Research Grant Schemes (FRGS) and the Ministry of Higher Education (MOHE) for supporting my studies through the MyBrain15 PhD Scholarship programme. Heartfelt appreciation to Prof. Sean May and Dr. Marcos Castellano from Nottingham Arabidopsis Stock Centre (NASC) in the UK campus for their kind guidance and support on the microarray wet-lab and

data analysis. My sincere thanks also to Dr. Yap Swee Keong, Ms. Nancy Liew and Dr. Dilan from the Institute of Biosciences (IBS) in UPM for their advices and technical support on RNA extraction, QC and experimental design for RNA-seq. Thanks to Dr. Sunir Malla and. Ms. Victoria Wright from the Deep Seq Department from the UK campus for advising and carrying out the RNA-seq wet lab. Also thanks to Dr. Ajit Singh for his expertise advices on experimental design and statistics along the way. Not forgetting Prof. Ing. Jaroslav Doležel from The Centre of Plant Structural and Functional Genomics (CIRAD) in France, for his kind donation of internal standard seeds and his kind sharing of work and advices on plant ploidy and flow cytometry work. Additionally, I'd like to also thank Dr Ho Wai Kuan from Crops for the future Research Centre (CFFRC) and Mr Jaeyres Jani (Bioeasy Sdn Bhd) for their advices and support on bioinformatics of mRNA and microRNAs respectively. Last but not least, thanks to Mr. K.K. Yong from KK vanilla plantation in Jerantut, Pahang, for his kind sharing on his experiences with vanilla growing, and also his kind donation of plant materials used for this research study.

TABLE OF CONTENTS

ABSTRACT III

ACKNOWLEDGEMENTS VI

TABLE OF CONTENTS VIII

LIST OF FIGURES XV

LIST OF TABLES XXIII

LIST OF APPENDICES XXVII

LIST OF ABBREVIATIONS XXVIII

CHAPTER I - INTRODUCTION 1

1.1. BACKGROUND AND OBJECTIVES OF STUDY 2

1.2. THESIS OUTLINE AND PROJECT OVERVIEW 4

1.3. LITERATURE REVIEW 7

1.3.1. INTRODUCTION TO VANILLA ORCHIDS 7

1.3.1.1. Origin, uses and nutraceutical properties of vanilla 7

1.3.1.2. Vanilla history and cultivation 8

1.3.1.3. Vanilla ethymology and taxonomy 9

1.3.1.4. Vanilla botany and cultivation 12

1.3.1.5. Vanilla flowering and pollination 13

1.3.1.6. Botany of green vanilla pod 16

1.3.1.7. When the greens become black – vanilla pods curing 18

1.3.1.8. Diseases of vanilla orchids 20

1.3.1.9. Vanilla macro- and micropropagation 22

1.3.1.9.1. In vitro conservation through shoot tips ***22***

1.3.1.9.2. In vitro germination and hybridization using seeds ***22***

1.3.1.9.3. Micropropagation ***23***

1.3.1.10. Global market and demands of vanilla 23

1.3.2. INTRODUCTION TO PLANT TISSUE CULTURE (PTC) AND SOMACLONAL VARIATION (SV) 26

1.3.2.1. Introduction to plant tissue culture (PTC) 26

1.3.2.2. Definitions and implications of SV in agriculture 29

1.3.2.3. Possible mechanisms associated to SV 37

1.3.2.3.1. Genetic mechanisms associated to SV ***39***

1.3.2.3.2. Epigenetic mechanisms associated to SV ***42***

1.3.2.4. Possible factors influencing rate of SV 45

1.3.2.4.1. Explant source ***45***

1.3.2.4.2. Genotype ***45***

1.3.2.4.3. Plant growth regulators (PGRs) ***46***

1.3.2.4.4. Culture time and long-term cultures ***46***

1.3.2.4.5. Type of TC-induced response ***47***

1.3.2.5. Tools used for SV detection 47
1.3.3. MOLECULAR TOOLS APPLIED IN THIS STUDY 51
1.3.3.1. Inter-simple sequence repeats (ISSR) markers for genetic studies 51
1.3.3.1.1. DNA fingerprinting and ISSR marker 51
1.3.3.1.2. Principles of ISSR marker 53
1.3.3.1.3. Applications of ISSR marker 55
1.3.3.2. Methylation-sensitive Amplified Polymorphisms (MSAPs) markers for epigenetic studies 55
1.3.3.2.1. Cytosine methylations in plants 55
1.3.3.2.2. Principles of MSAP markers 56
1.3.3.2.3 Applications of MSAP for methylation analysis in PTC... 59
1.3.3.3. Flow cytometry (FCM) for cytological studies 59
1.3.3.3.1. Principles of Flow cytometry (FCM) 59
1.3.3.3.2. Applications of FCM in plants 61
1.3.3.3.3. Endopolyploidy in plants 62
1.3.3.4. Transcriptome profiling 64
1.3.3.4.1. Next-generation transcriptome sequencing (NGS) 64
1.3.3.4.2. Reverse transcription quantitative real-time PCR (RT-qPCR) 68
1.3.3.4.1.1. Nomenclature and standardization of terminologies 68
1.3.3.4.1.2. Principles of RT-qPCR: mRNA and miRNA 68
1.3.3.4.1.3. Applications of RT-qPCR 72
1.3.3.4.3. Plant microRNAs: Principles and bioinformatics 73
1.3.3.4.3.1. Definitions, databases and nomenclature 73
1.3.3.4.3.2. Biogenesis and functions of miRNAs in plants 73
1.4.3.4.3.2. Principles of bioinformatics for plant microRNAs 75

CHAPTER II - *IN VITRO* REGENERATION OF VANILLA PLANIFOLIA ANDREWS 81

2.1. INTRODUCTION 82

2.2. MATERIALS AND METHODS 84

2.2.1. INSTRUMENT AND APPARATUS PREPARATION 84
2.2.2. CHEMICALS, MEDIA AND STOCK PLANT GROWTH REGULATORS PREPARATION 84
2.2.3. CULTURE CONDITION 85
2.2.4. SOURCE OF PLANT MATERIALS 85
2.2.5. ESTABLISHMENT OF *V. PLANIFOLIA* ASEPTIC CULTURES 87
2.2.6. VANILLA NODAL, LEAF AND ROOT TIPS CULTURE 88
2.2.7. SHOOT PROLIFERATION AND MAINTENANCE, AND ROOTING 89
2.2.8. VANILLA SEED CULTURE 89
2.2.9. MORPHOLOGICAL OBSERVATION AND HISTOLOGICAL ANALYSIS 90
2.2.10. STATISTICAL ANALYSIS 90

2.3. RESULTS 91

2.3.1. ESTABLISHMENT OF VANILLA ASEPTIC CULTURES 91

2.3.2. NODAL CULTURE (NT) 94
2.3.2.1. Direct shoot regeneration 95
2.3.2.2. Indirect regeneration 95
2.3.3. ROOT TIPS CULTURE 97
2.3.4. LEAF CULTURE 102
2.3.5. SEEDS CULTURE 102
2.4. DISCUSSION 104
2.4.1. ESTABLISHMENT OF ASEPTIC CULTURE OF VANILLA 104
2.4.2. NODAL CULTURE – THE DIRECT AND INDIRECT REGENERATION 106
2.4.2.1. Direct shoot regeneration 107
2.4.2.2. Indirect shoot regeneration 107
2.4.3. ROOT TIPS CULTURE AS THE POTENTIAL ALTERNATIVE REGENERATION 108
2.4.4. LOW GERMINATION RATE IN SEEDS CULTURE 111
2.5. CONCLUSIONS AND FUTURE PROSPECTS 112
CHAPTER III - MORPHOLOGIC, GENETIC AND EPIGENETIC STABILITIES OF VANILLA PLANTS DERIVED FROM DIRECT AND INDIRECT NODAL REGENERATIONS AS REVEALED BY ISSR AND MSAP 114
3.1. INTRODUCTION 115
3.2. MATERIALS AND METHODS 117
3.2.1. CHEMICAL REAGENTS AND APPARATUS 117
3.2.2. OLIGONUCLEOTIDES SYNTHESIS AND PRIMERS DESIGN 117
3.2.3. PLANT MATERIALS 118
3.2.4. OVERALL METHODOLOGY 118
3.2.5. GENOMIC DNA EXTRACTION OF *V. PLANIFOLIA* 120
3.2.6. NUCLEIC ACID QUANTIFICATION AND QUALITY CHECK 120
3.2.7. INTER-SIMPLE SEQUENCE REPEATS - POLYMERASE CHAIN REACTION (ISSR-PCR) FINGERPRINTING 120
3.2.7.1. Optimization of PCR conditions 121
3.2.7.2. Statistical analysis 121
3.2.7.3. Analysis of genetic diversity 122
3.2.8. METHYLATION-SENSITIVE AMPLIFIED POLYMORPHISMS - POLYMERASE CHAIN REACTION (MSAP-PCR) FINGERPRINTING OF *V. PLANIFOLIA* 123
3.2.8.1. Construction of MSAP adapters 123
3.2.8.2. DNA double digestion with methylation-sensitive and – insensitive REs 123
3.2.8.3. Adapter Ligation 124
3.2.8.4. Pre-selective amplification (+1) 124
3.2.8.5. Selective amplification (+2) 125

3.2.8.6. Optimization of RE digestion and PCR conditions.....125
3.2.8.7. Screening of MSAP pre-selective (+1) and selective (+2) primers 125
3.2.8.8. Gel Electrophoresis 126
3.2.8.9. MSAP data scoring and analysis 126
3.2.8.10. MSAP data interpretation.....126

3.3. RESULTS 128

3.3.1. MORPHOLOGICAL ABNORMALITY OF *IN VITRO* CULTURED *V. PLANIFOLIA*.....128
3.3.2. DNA EXTRACTION OF *V. PLANIFOLIA* 130
3.3.3. ISSR-PCR FINGERPRINTING 130
3.3.3.1. Optimization of ISSR-PCR Parameters.....130
3.3.3.2. SV screening with ISSR-PCR 132
3.3.4. MSAP-PCR FINGERPRINTING.....137
3.3.4.1. Optimization of MSAP-RE digestion conditions.....137
3.3.4.2. Optimization of MSAP-PCR condition.....138
3.3.4.3. Selection between selective +1 or +2 primers 139
3.3.4.4. SV screening with MSAP-PCR.....142

3.4. DISCUSSIONS 145

3.5. CONCLUSIONS AND FUTURE PROSPECTS..... 147

CHAPTER IV - SYSTEMIC VARIATIONS OF CELLULAR ENDOPOLYPLOIDY IN DIFFERENT TISSUES OF *V. PLANIFOLIA* AND ITS IMPLICATIONS IN MICROPROPAGATION 148

4.1. INTRODUCTION 149

4.2. MATERIALS AND METHODS..... 151

4.2.1. CHEMICALS REAGENTS AND APPARATUS 151
4.2.2. SOURCE OF PLANT MATERIALS.....151
4.2.3. PLANT NUCLEI EXTRACTION AND SAMPLE PREPARATIONS 154
4.2.4. FLOW CYTOMETRIC ANALYSES 155
4.2.5. FORMULAS FOR GENOME SIZE CALCULATION 156
4.2.6. STATISTICAL ANALYSES 157

4.3. RESULTS 157

4.3.1. GENOME SIZE ESTIMATION OF VANILLA MOTHER PLANT (MP) USING INTERNAL STANDARDS 157
4.3.2. SYSTEMIC ENDOPOLYPLOIDY VS EXPLANT TYPE IN VANILLA 159
4.3.3. FLOW CYTOMETRIC ANALYSES IN *IN VITRO* CULTURED VANILLA CALLI 162

4.4. DISCUSSION 165

4.4.1. SELECTION OF SUITABLE INTERNAL STANDARDS.....165

4.4.2. SYSTEMIC DIFFERENTIAL PROGRESSIVELY PARTIAL ENDOPOLYPLOIDY (PPE) AND ITS IMPLICATIONS IN PTC OF VANILLA....166
4.4.3. PROLONGED CALLUS CULTURE DURATION AND SV.....169

4.5. CONCLUSION AND FUTURE PROSPECTS..... 170

CHAPTER V – *DE NOVO* TRANSCRIPTOME AND MICRORNA PROFILING OF LONG-TERM CULTURED VANILLA CALLUS 171

5.1. INTRODUCTION 172

5.2. MATERIALS AND METHODS..... 175

5.2.1. CHEMICALS REAGENTS, KITS AND APPARATUS.....175
5.2.2. OLIGONUCLEOTIDE NUCLEOTIDE SYNTHESIS AND PRIMER SOURCE 175
5.2.3. SOURCE OF PLANT MATERIALS.....175
5.2.4. OVERVIEW OF RESEARCH METHODOLOGY.....176
5.2.5. TOTAL RNA EXTRACTION, RNA QUANTITY AND QUALITY CHECK 178
5.2.6. MRNA-SEQUENCING AND DATA ANALYSIS.....179
5.2.6.1. mRNA poly-A enrichment, bar-coded sequencing libraries construction and sequencing 179
5.2.6.2. Raw sequence pre-processing and *de novo* transcriptome assembly 179
5.2.6.3. Construction of a Digital Gene Expression Atlas and Annotation Pipeline 180
5.2.7. SMALL RNA SEQUENCING – MICRORNA ANALYSIS.....181
5.2.7.1. Constructing and sequencing small RNA libraries.....181
5.2.7.2. Sequence analysis of miRNA.....182
5.2.8. INTEGRATION OF MRNA-SEQ AND MIRNA-SEQ DATA.....184
5.2.9. RT-QPCR VALIDATIONS OF DEGS AND DEMS IN PROLONGED VANILLA CALLUS CULTURE.....185
5.2.9.1. RT-qPCR Primer Design.....185
5.2.9.1.1. mRNA Primer Design.....185
5.2.9.1.2. miRNA Primer Design.....186
5.2.9.2. cDNA library construction.....186
5.2.9.2.1. mRNA.....187
5.2.9.2.2. miRNA.....187
5.2.9.3. Standard qPCR Conditions.....188
5.2.9.3.1. mRNA.....188
5.2.9.3.2. miRNA.....189
5.2.9.4. Annealing Temperature Optimizations 190
5.2.9.4.1. mRNA.....190
5.2.9.4.2. miRNA.....190
5.2.9.5. Validation of qPCR Amplification Efficiency 191
5.2.9.6. Reference Stability Test.....191
5.2.9.7. Data Analysis for Relative Quantification of Gene Expression 192

5.3. RESULTS 194

5.3.1. TOTAL RNA QUALITY AND QUANTITY 194
5.3.2. MRNA SEQUENCE PROFILES 195
5.3.2.1. Sequence processing and *de novo* assembly 195
5.3.2.2. Functional annotation and classifications of differentially expressed genes (DEGs) 197
5.3.2.2.1. BLAST, GO-mapping and Annotation Project Statistics 197
5.3.2.2.2 Global transcriptome dynamics 199
5.3.2.2.3. DGE and time course expression analysis 201
5.3.2.2.4. Gene annotations and classifications 202
5.3.2.3. DGE analysis of vanilla calli over time 216
5.3.2.3.1. Expression of methylation and histone modification-related genes 216
5.3.2.3.2. Expression of TFs, phytohormone- and stress-related genes 218
5.3.2.3.3. Expression of TE-related genes 219
5.3.3. SMALL RNA SEQUENCE PROFILES 221
5.3.3.1. Sequence processing and filtering 221
5.3.3.2. Small RNA Annotation 222
5.3.3.3. Global micronome dynamics 223
5.3.3.4. Differential miRNA expressions (DME) 224
5.3.4. MIRNA-MRNA TARGET CORRELATIONS 226
5.3.4.1. Target prediction and correlation analysis 226
5.3.4.2. Conserved Protein Domains and miRNAs Target Sites .. 227
5.3.4.2.1. DNA methyltransferase and miRNA 228
5.3.4.2.2. Co-factor-binding and miRNA 229
5.3.4.2.3. DNA integration and miRNA 230
5.3.4.2.4. DNA binding and miRNA 232
5.3.5. PRIMER DESIGN FOR RT-QPCR 234
5.3.6. ANNEALING TEMPERATURE OPTIMIZATIONS 235
5.3.6.1. mRNA 235
5.3.6.2. miRNA 238
5.3.7. ESTIMATION OF REFERENCE GENES STABILITY 240
5.3.8. QPCR AMPLIFICATION EFFICIENCY OF GENES 242
5.3.8.1. mRNA 242
5.3.8.2. miRNA 243
5.3.9. MELT CURVE ANALYSIS 244
5.3.10. VALIDATION OF RNA-SEQ DATA BY RT-QPCR 246

5.4. DISCUSSION 252

5.4.1. BLAST AND FUNCTIONAL GENE ANNOTATION 252
5.4.2. DNA METHYLATION PROFILE 253
5.4.3. RETROTRANSPOSON PROFILE 254
5.4.4. HISTONE PROFILE 255
5.4.4.1. HDACs: HDAC2, HDAC19, SAP18 255

5.4.1.2. Histone Cores: H2A, H3, H4 and H3K4 256
5.4.1.3. HATs: ELP and GCN5 256
5.4.5. STRESS REGULATING TFS MIGHT BE KEY REGULATORS OF THE RTS EXPRESSION 257
5.4.6. MIR166R, MIR395G AND MIR8155 AS POTENTIAL SUPPRESSOR FOR DNA METHYLATION AND RT GENES 257
5.4.7. SELECTION OF SUITABLE REFERENCE CANDIDATES FOR RT-QPCR 258
5.4.8. RT-QPCR VALIDATION OF RNA-SEQ DATA 262
5.5. CONCLUSION AND FUTURE PROSPECTS 262
CHAPTER VI – GENERAL DISCUSSION AND PERSPECTIVES 264
6.1. SUMMARY OF FINDINGS 265
6.2. IMPLICATION OF FINDINGS 266
6.3. FUTURE RECOMMENDATIONS 269
APPENDICES 271
APPENDIX A. THE LIST OF STANDARD SOLUTIONS, CHEMICALS, ENZYMES AND KITS USED IN THE STUDY AND THEIR RESPECTIVE MANUFACTURERS. 272
APPENDIX B. EQUIPMENT, APPARATUS AND SOFTWARE USED IN THIS STUDY. 275
APPENDIX C. CHEMICAL BUFFERS USED IN THIS STUDY AND THEIR BUFFER COMPONENTS. 277
APPENDIX D. FORMULATION OF MEDIA USED IN THIS STUDY. 279
APPENDIX E. LIST OF ADAPTER AND PRIMER SEQUENCES USED IN ISSR AND MSAP ANALYSIS (**CHAPTER 3**). 280
APPENDIX F. LIST OF PLANT (*VIRIDIPLANTAE*) WITH AVAILABLE MIRNA SEQUENCES DEPOSITED IN MIRBASE VERSION 21.0, WHICH WAS USED FOR MIRNA ANNOTATION IN THIS STUDY. 281
APPENDIX G. IDENTIFICATION OF CONSERVED MIRNA FAMILIES IN *VANILLA PLANIFOLIA* WITH CORRESPONDING ISOFORMS. 284
APPENDIX H. CONTIG SEQUENCES OF MRNA USED IN QRT-PCR ANALYSIS, IN FASTA FORMAT. 298
APPENDIX I. LIST OF PRIMER SEQUENCES USED IN QRT-PCR ANALYSIS (**CHAPTER 6**). 305
APPENDIX J. REFERENCE STABILITY PRE-SCREENING RESULTS FOR 86 POTENTIAL MIRNA CANDIDATES FROM RNA-SEQ DATA, COMPUTED BY FIVE ALGORITHMS (REFFINDER, DELTA CT, BESTKEEPER, NORMFINDER AND GENORM) 306
REFERENCES 310
LIST OF PUBLICATIONS 361

LIST OF FIGURES

Figure 1.1.	Overview of overall research project outline.	6
Figure 1.2.	Schematic diagram showing plant forms of a vanilla orchid.	13
Figure 1.3.	*Vanilla planifolia* vines, flowers and pods in KK vanilla plantation in Jerantut, Pahang, Malaysia.	14
Figure 1.4.	Illustration of (a) longitudinal section and (b) hand pollination process of vanilla flower	16
Figure 1.5.	Vanilla seed pod.	17
Figure 1.6.	Main stages of the vanilla curing process, also defined as controlled ripening process.	19
Figure 1.7.	Vanilla diseases.	21
Figure 1.8.	Producer prices of world top three vanilla producers (in USD per tonnes) from 1991-2014.	24
Figure 1.9.	World vanilla beans production from 1994 to 2014.	25
Figure 1.10.	World area harvested, yield and production of vanilla from 2004 to 2014.	26
Figure 1.11.	Routes of plant regeneration in tissue cultures *in vitro* from different explants.	27
Figure 1.12.	A schematic illustration of different types of callus that can be induced *in vitro*: friable, compact, shooty, embryonic and rooty calli.	29
Figure 1.13.	Somaclonal variants of banana showing (a) abnormal, spindle-shaped leaf lamina, (b-c) leaf texture/colour variants and (d) abnormal split-fingers at banana brunch.	35
Figure 1.14.	Somaclonal variants of oil palms with mantled fruits forms.	36

Figure 1.15.	*In-vitro* culture and molecular changes caused in the process.	38
Figure 1.16.	Schematic illustration of changes in chromatin structure via DNA methylation, histone modification, and small RNA-directed DNA methylation (RdDM).	42
Figure 1.17.	Lineology of DNA-based markers (Time Clock).	52
Figure 1.18	Schematic representation of the inter-simple sequence repeat (Inter-SSR-PCR or ISSR-PCR).	54
Figure 1.19.	Schematic diagram of MSAP process.	57-58
Figure 1.20.	A schematic of the primary components that comprise a flow cytometer.	60
Figure 1.21.	Endopolyploidy in plants.	63
Figure 1.22.	Timeline of DNA sequencing technology evolution, from Sanger to NGS, and the cost per raw megabase of DNA sequenced.	65
Figure 1.23.	Outline of Illumina sequencing by synthesis (SBS) process.	67
Figure 1.24.	Two-step reverse transcription polymerase chain reaction (RT-PCR).	69
Figure 1.25.	Real-time PCR output and calculation of $C_{t.}$	70
Figure 1.26.	Melt Curve Analysis (MCA) of qPCR.	71
Figure 1.27.	The miRNA biogenesis and general mechanism of miRNA-mediated regulation of stress-responsive gene expression in plants to attain stress tolerance.	74
Figure 1.28.	Nucleotide pairing between plant miRNAs and target mRNAs	76
Figure 2.1.	Source of vanilla tissue culture mother plants for nodal cuttings and seed pods.	86
Figure 2.2.	Vanilla post-surface sterilization.	92

Figure 2.3.	(a) Liquid culture test of vanilla explants after one-month post-surface sterilization, and (b) Gram stain test of bacterial isolates isolated from vanilla cultures.	93
Figure 2.4.	Direct single, multiple and abnormal shoot regenerations of vanilla nodal cultures.	95
Figure 2.5.	Different shooting or callusing responses in BAP treatment with various NAA levels.	96
Figure 2.6.	Different callus morphologies in NAA treatments with various BAP levels.	96
Figure 2.7.	Callus formation from 2,4-D treatments.	97
Figure 2.8.	Response of *in vitro* root tips culture to different hormones.	98
Figure 2.9.	The progressive development of root-derived callus to PLB and multiple shoots mass of root tips of *V. planifolia.*	100
Figure 2.10.	Light micrograph showing the sequence of development of callus from cross section of vanilla root tip.	101
Figure 2.11.	Vanilla seeds germination and PLB formation.	103
Figure 2.12.	Shoot tip of a vanilla plant bearing *Bacillus amyloliquefaciens*.	106
Figure 2.13.	Sporadical callus formation (a-b) in aerial root tips of *V. planifolia* root tips partially submerged in MS medium supplemented with 1.0 mg/L BAP.	108
Figure 3.1.	Overview of genetic and epigenetic molecular markers analysis workflow of vanilla.	119
Figure 3.2.	Somaclones with distinct morphological appearances.	129
Figure 3.3.	Gel image of genomic DNA extracted from somaclones samples, in comparison with λ-DNA in the dilution of 25, 50, 75, 100 and 200 ng, on 0.8% agarose gel in 1x TAE buffer.	130

Figure 3.4.	Optimization of ISSR-PCR parameters.	132
Figure 3.5.	Screening of 20 selected somaclones regenerated from callus using the ISSR primer ISSR2.	133
Figure 3.6.	Screening of 20 selected somaclones regenerated from callus using the ISSR primer ISSR4.	134
Figure 3.7.	Screening of 20 selected somaclones regenerated from callus using the ISSR primer ISSR6.	135
Figure 3.8.	Dendrogram for ISSR analysis of somaclones.	136
Figure 3.9.	Optimization of RE digestion paramters for MSAP analysis.	137
Figure 3.10.	Optimization of MSAP-PCR parameters.	138
Figure 3.11.	Primers screening for *EcoRI*/Hpa*II* (E/H) selective amplification (+1) primer pairs.	139
Figure 3.12.	Primers screening for *EcoRI*/Hpa*II* (E/H) selective amplification (+2) primer pairs.	140
Figure 3.13.	Dendrogram of genetic distances between tissue culture regenerants.	144
Figure 4.1.	Seeds germination of internal standard plants.	152
Figure 4.2.	Diagrammatic representation of the direct and indirect plant regeneration pathways for vanilla.	154
Figure 4.3.	Bi-parametric dot plot of FL2-A versus FL2-H (a) used for gating plant nuclei populations.	156
Figure 4.4.	(a) *V. planifolia* leaf tip PI-stained nuclei (a) with and (b) without *R. sativus* as an internal control.	159
Figure 4.5.	Nuclear DNA content and distributions of endopolyploid nuclei in vegetative tisssues of *V. planifolia.*	160
Figure 4.6.	Variation in ploidy levels of nuclei populations from different tissues of vanilla.	161
Figure 4.7.	Ploidy changes in prolonged callus culture of vanilla.	164

Figure 5.1.	Overview of (a) RNA-Seq and (b) RT-qPCR workflow.	177
Figure 5.2.	Schematization of the miRNA bioinformatics pipeline.	182
Figure 5.3.	Integration pipeline for miRNA-mRNA target relations and expression correlations.	184
Figure 5.4.	The RNA profiles from Agilent 2100 Bioanalyzer analysis.	195
Figure 5.5.	Top-hit species distribution of BLAST results for vanilla to *nr* (*viridiplantae*) databases.	198
Figure 5.6.	Gene ontology classification of assembled transcripts.	200
Figure 5.7.	KEGG pathway for cysteine and methionine metabolism of genes expressed in vanilla callus culture.	203
Figure 5.8.	Conserved protein domain architectures for putative histone lysine N-methyltransferase (HMT) family members expressed in vanilla callus cultures.	206
Figure 5.9.	Conserved protein domain architectures for putative histone deacetylase (HDAC) family members expressed in vanilla callus cultures.	207
Figure 5.10.	Conserved protein domain architectures for putative histone acetyltransferase (HAT) family members expressed in vanilla callus cultures.	208
Figure 5.11.	Conserved protein domains and multiple sequence alignment (MSA) of (a) vanilla putative H2A variant 3 (H2A3) proteins.	210
Figure 5.12.	Conserved protein family domain of vpl_contig_60898 encoding for COMPASS-like H3K4 histone methylase component WDR5A protein (XP_020701228.1).	211
Figure 5.13.	Conserved protein domain (PLN00035 superfamily, provisional histone H4) and multiple	212

sequence alignment (MSA) of vanilla sequences encoding conserved histone H4.

Figure 5.14. Bar chart showing classifications of vanilla callus transcriptome sequences with BLAST match with the transposable element sequences in MIPS Repeat Element Database. 214

Figure 5.15. Heat maps of different methylation genes expressed in vanilla calli over time period 3M, 6M and 12M. 217

Figure 5.16. Heat maps of different histone protein-encoding genes expressed in vanilla calli over time period 3M, 6M and 12M. 218

Figure 5.17. Heat maps of different transcription factors expressed in vanilla calli over time period 3M, 6M and 12M. 219

Figure 5.18. Heat maps of different retrotransposon genes expressed in vanilla calli over time period 3M, 6M and 12M. 220

Figure 5.19. Length distribution of sequenced small RNAs of sample 6M-1 after pre-processing. 222

Figure 5.20. Heat map graphs of significantly expressed miRNAs during pairwise comparisons between (a) 3M versus 6M, and (b) 6M vs 12M. 225

Figure 5.21. Sequence alignment during target prediction of vpl-miR395g and one of its potential mRNA target, vpl_contig_14640 at a ComiR score of 4.0. 226

Figure 5.22. Venn diagram showing significant (FDR-corrected *p*-value $\leq$ 0.05) correlations between miRNA and mRNA expressions. 227

Figure 5.23. BLAST Hit Alignment score of top ten hits to vpl_contig_9944 to *nr* database. 228

Figure 5.24. BLAST Hit Alignment score of top ten hits to vpl_contig_14640 to *nr* database. 229

Figure 5.25.	Illustration of BLAST sequence alignment and positions of conserved regions of vpl_contig_24683.	231
Figure 5.26.	Illustration of BLAST sequence alignment and positions of conserved regions of vpl_contig_14640.	233
Figure 5.27.	Preliminary results of Primer-BLAST report using Vpl_contig_24863 and its forward and reverse primer pair.	234
Figure 5.28.	Agarose gel electrophoresis (2.5% w/v) of *V. planifolia* cDNA amplified by PCR to determine the optimum annealing temperature of each reference genes primer pair.	236
Figure 5.29.	Gradient Annealing Temperature PCR Profiles of *V. planifolia* cDNA amplified by mRNA target genes primer pairs on 2.5% (w/v) agarose gel.	237
Figure 5.30	Melt peak in gradient PCR of reference and target miRNAs.	239
Figure 5.31	Standard curve of *vpl_contig_9944*. Standard curve was constructed from seven series of two-fold dilution.	242
Figure 5.32.	Standard curve of *vpl-miR395g-3p*. Standard curve was constructed from seven series of two-fold dilution.	243
Figure 5.33.	Log amplification plot (a), melt curve (b) and melt peak (c) of one of the qPCR amplification experiments.	245
Figure 5.34.	The expression levels of predicted miRNAs and mRNAs.	247
Figure 5.35.	RT-qPCR results of differentially expressed miRNAs and mRNAs of vanilla callus cultures under prolonged culture durations of 3M, 6M and 12M.	248

Figure 5.36. RT-qPCR results of differentially expressed miRNAs and mRNAs of vanilla callus cultures under prolonged culture durations of 3M, 6M and 12M. 249

Figure 5.37. RT-qPCR results of differentially expressed miRNAs and mRNAs of vanilla callus cultures under prolonged culture durations of 3M, 6M and 12M. 250

LIST OF TABLES

Table 1.1.	Taxonomic classification of *Vanilla planifolia.*	10
Table 1.2.	A list of vanilla species and their geographical locations	10-11
Table 1.3.	A list of commercially exploited horticultural crop varieties which originated from SV.	31-34
Table 1.4.	Tissue culture-induced transposable elements in plants.	40-41
Table 1.5.	Modifications in histones and small RNA levels detected in plant cells/tissues cultured *in vitro.*	44
Table 1.6.	Techniques used for detecting cell culture-induced genomic changes.	48-49
Table 1.7.	Strengths and weaknesses of different marker systems for the assessment of clonal fidelity.	50
Table 1.8.	Functions of conserved plant microRNAs (miRNAs) for which computational prediction and experimental evidence have been published.	77-78
Table 1.9.	Plant microRNA (miRNA) distribution.	82
Table 2.1.	Hormonal treatments for the culture of different vanilla explants.	88
Table 2.2.	Results of surface sterilization using carbendazim, after one month.	93
Table 2.3.	The influence of BAP, TDZ, NAA or 2,4-D hormonal treatments on nodal explants when cultured on semi-solid MS medium, after 90 days of culture.	94
Table 2.4.	The influence of sole applications of BAP, TDZ, NAA or 2,4-D hormonal treatments on callus response of root tips explants when cultured on semi-solid MS medium, after 90 days of culture.	98

Table 2.5.	The influence of BAP or TDZ in combination with 2,4-D and NAA hormonal treatments on callus response of root tips explant after 90 days of culture.	99
Table 2.6.	Seed culture of vanilla.	102
Table 3.1.	MSAP-profiled methylation types and the scoring scheme	127
Table 3.2	Morphological description of tissue culture vanilla somaclones.	128
Table 3.3.	Vanilla ISSR primers amplification properties: primer sequences, optimal annealing temperature (TA), number of scorable bands and fragment size range in bp of each primer.	131
Table 3.4	Summary of +1 and +2 primers amplification.	141
Table 3.5.	Type I, Type II and Type III MSAP profiles.	143
Table 4.1.	DNA reference standards recommended for the estimation of nuclear DNA amounts in absolute units.	152
Table 4.2.	DNA content (pg) estimation in vanilla using different standards.	158
Table 4.3.	Replication ratio in different endopolyploidy levels of vanilla.	162
Table 4.4.	Percentage of nuclei and mean DNA content in 3, 6, 12 and 15 months of callus culture.	163
Table 5.1.	Primer design parameters for qPCR.	185
Table 5.2.	The components in an mRNA qPCR reaction.	188
Table 5.3.	The components in a miRNA qPCR reaction.	189
Table 5.4.	List of reference gene candidate for mRNA and miRNA RT-qPCR analysis.	192

Table 5.5.	Nanodrop and Qubit readings.	194
Table 5.6.	Statistics of reads generated by mRNA sequencing of *V. planifolia* callus before and after pre-processing.	196
Table 5.7.	Statistics for *de novo* transcriptome assembly.	196
Table 5.8.	BLAST search, GO mapping and annotation project statistics of vanilla transcriptome (a total of 22,025 sequences with DGEs) against the NCBI protein databases.	198
Table 5.9.	DGE summary of vanilla callus over 3, 6 and 12M prolonged culture durations.	201
Table 5.10.	Time course expression analysis results.	201
Table 5.11.	Histone protein families expressed in vanilla callus cultures.	204
Table 5.12.	List of transcription factor gene families and members expressed in vanilla callus cultures.	215
Table 5.13.	Summary of pre-analysis for miRNA sequencing of *V. planifolia* calli.	221
Table 5.14.	Summary of reads from raw data to annotated miRNA sequences of *V. planifolia.*	222
Table 5.15.	Expression frequencies of miRNA unique sequences.	224
Table 5.16.	DME summary of vanilla callus over 3, 6 and 12M prolonged culture durations.	224
Table 5.17.	Conserved domains from BLAST results of vpl_contig_9944 to putative methione synthase gene with gene ontology of transferase activity and binding.	228
Table 5.18.	Conserved domains from BLAST results of vpl_contig_24683 to putative retroviral	230

polyproteins with gene ontology of DNA integration.

Table 5.19. Conserved domains from BLAST results of vpl_contig_18270 to putative retroviral polyproteins with gene ontology of DNA binding. 232

Table 5.20. Summary of genes to be validated using qRT-PCR. 234

Table 5.21. Size ranges of expected amplicon size and PCR-amplified amplicon. 235

Table 5.22. Estimation of expression stability of reference gene / miRNA candidates across different prolonged culture duration of vanilla callus computed by five algorithms (RefFinder, Delta CT, BestKeeper, NormFinder and GeNorm). 241

Table 5.23. PCR efficiency results of mRNA reference and target genes. 243

Table 5.24. PCR efficiency results of miRNA reference and target genes. 244

Table 5.25. Correlation analysis of RNA-seq and RT-qPCR results. 251

Table 5.26. A summary of role of miRNA candidates (miR159, miR168 and miR319) in plants, compiled from literature review. 261

LIST OF APPENDICES

Appendix A.	The list of standard solutions, chemicals, enzymes and kits used in the study and their respective manufacturers.	272
Appendix B.	Equipment, apparatus and software used in this study.	275
Appendix C.	Chemical buffers used in this study and their buffer components.	277
Appendix D.	Formulation of media used in this study.	279
Appendix E.	List of adapter and primer sequences used in ISSR, MSAP and qPCR analysis.	280
Appendix F.	List of plant (*Viridiplantae*) with available miRNA sequences deposited in mirBase version 21.0, which was used for miRNA annotation in this study.	281
Appendix G.	Identification of conserved miRNA families in *Vanilla planifolia* with corresponding isoforms.	284
Appendix H.	Contig sequences of mRNA used in RT-qPCR analysis, in FASTA format.	298
Appendix I.	List of primer sequences used in qRT-PCR analysis.	305
Appendix J.	Reference stability pre-screening results for 86 potential miRNA candidates from RNA-seq data, computed by five algorithms (RefFinder, Delta CT, BestKeeper, NormFinder and GeNorm).	

LIST OF ABBREVIATIONS

%	Percentage
°C	Degree Celsius
18S	18S ribosomal RNA
2,4-D	2,4-Dichlorophenoxyacetic acid
ACT	β-actin
ALF	Alfafin-like factor
ANOVA	Analysis of variance
BAP	6-Benzylaminopurine
bHLH	Basic helix-loop-helix
BLAST	Basic local alignment search tool
BP	Biological Process
bp(s)	Base pair(s)
CBF	Core binding factor
CC	Cellular Components
cDNA	Complementary DNA
CIM	Callus induction media
Cq	Cycle of quantification
DEG(s)	Differentially expressed gene(s)
DEM(s)	Differentially expressed microRNA(s)
DGE	Differential gene expression
DME	Differential microRNA expression
DMNT	DNA (cytosine-5)-methyltransferase
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNA	Deoxyribonucleic acid
dNTP(s)	deoxynucleoside triphosphate (s)
EDTA	Ethylenediaminetetraacetic acid
EF1	Elongation factor-1 alpha
ERF	Ethylene-responsive transcription factors
et al.	et alia
FDR	False discovery rate
g	Gram (for weight)
g	Gravity (for speed)
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GO	Gene ontology
H_2O	Water
HATs	Histone acetyltransferases
HCl	Hydrochloric acid
HDACs	Histone deacetylases
IUPAC	International Union of Pure and Applied Chemistry

kb	Kilo base pair
L	Litre
LINE	Long interspersed nuclear elements
LRR	Leucine rich repeats
LSD	Lysine-specific demethylase
LTR	Long terminal repeats
M	Molar
MDS	Multidimensional scaling
MET	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase
MF	Molecular Function
mg	Milligram
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulphate
min	Minute (s)
mips-REdat	MIPS Repeat Element Database
miRNA(s)	microRNA(s)
ml	Millilitre
mM	Millimolar
mRNA	Messenger RNA
MS	Murashige and Skoog
MSAP	Methylation sensitive amplified polymorphisms
MTA-SAHN	5-methylthioadenosine S-adenosylhomocysteine nucleosidase
MyB	Myelobastosis
MyC	Myelocytomatosis
NAA	1-Naphthaleneacetic acid
NaAc	Sodium acetate
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NCBI	National Center for Biotechnology Information
ng	Nanogram
NGS	Next generation sequencing
NRT	No reverse transcriptase (control)
NTC	No template control
ORF	Open reading frame
PCR	Polymerase chain reaction
pg	Picogram
PGR(s)	Plant growth regulator(s)
Ppm	Parts per million
PTC	Plant Tissue Culture
qPCR	Quantitative polymerase chain reaction
RE	Restriction endonuclease
RIM	Root induction media

RIN	RNA integrity number
RNA	Ribonucleic acid
rpm	Revolutions per minute
rRNA	Ribosomal RNA
RT	Reverse transcription
RT(s)	Retrotransposon(s)
RT-qPCR	Reverse transcription quantitative real time polymerase chain reaction
S	Ribosomal Subunit
s	Second (s)
SAH	S-adenosyl-L-homocysteine hydrolase
SAM	S-Adenosylmethionine
S-AT	S-acyltransferase
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SIM	Shoot induction media
SINE	Short interspersed nuclear elements
SV	Somaclonal variation
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
TDZ	Thidiazuron
TE	Tris -EDTA
TE(s)	Transposable element(s)
TF(s)	Transcription factor(s)
Tris-HCl	Tris-hydrochloride
TUB	α-tubulin
UBQ	Ubiquitin
V	Volt
V. planifolia	*Vanilla planifolia*
v/v	By volume
w/v	By weight
µg	Microgram
µl	Microliter

Chapter I

INTRODUCTION

Chapter I

INTRODUCTION

1.1. Background and Objectives of Study

Vanilla is the second most expensive spice in the world, after saffron (Mani *et al.*, 2017). It is widely used in the food, cosmetic and medicinal industries (Busconi *et al.*, 2017). High demand for natural vanillin has led to the need for more planting materials for vanilla orchid. Traditional stem-cutting is not sustainable as continuous injury to the mother plant causes retardation to the vanilla plant growth and thus a reduction of yield, while the seed germination method is also not viable because the germination rate of vanilla seeds is low (Tan *et al.*, 2011a). Besides, continuous clonal propagation of *V. planifolia* through stem-cutting leads to monoculture, exposing the crop to severe damage when the vanilla is affected by pests and diseases (Gopinath, 1994). Meanwhile, the introduction of new genetic material is greatly constrained by several factors, such as its asexual clonal propagation, mostly due to self-pollination of the vanilla flowers, and the threat posed to wild populations of vanilla by land use pressures (Lubinsky, 2003).

Plant tissue culture plays an important role in the cultivation of vanilla by overcoming the problems mentioned above. Clonal propagation of disease-free, elite vanilla genotypes through *in vitro* micropropagation helps to ensure the performance and productivity of the vanilla crop. On the other hand, plant tissue culture could be used as a tool to make *in vitro* crop improvement of vanilla (Minoo *et al.*, 1997), by introducing genetic variability through induction of somaclonal variations. In addition, *in vitro* applications, such as *in vitro* vanillin production in callus cells through elicitations (Funk and Brodelius, 1990) or biotransformations (Velankar and Heble, 2004), has also highlighted the importance of tissue culture as a fundamental technique for biotechnological applications in vanilla.

A major challenge in vanilla tissue culture is the effective screening of somaclonal variations (SV). For a commercial scale of production through mass micropropagation, pre-screening for SV in an established tissue culture system is crucial to ensure minimal off-types was produced. While for *in vitro* crop improvement, a quick and effective selection process for somaclonal variations is also important (Minoo *et al.*, 1997). Not all of the SV could be detected early by morphological abnormalities, and these SV could result in a great impact on agricultural yield and crop potential, such as those shown in the cases of oil palm and banana industries reviewed in Krishna *et al.* (2016). In recent studies, it was suggested that complex epigenetics (Miguel and Marum, 2011) play an important role in contributing to the genetic, cytological and morphological phenotypic changes in SV.

In this study, the main aim was to investigate the occurrence of somaclonal variations in the tissue culture plants of the world most cultivated commercial vanilla species, *Vanilla planifolia* Andrews, using a cultivar grown in Malaysia. In this study, two types of regeneration system (the *direct* and the *indirect* nodal regeneration), using nodal explants, were established and somaclonal variations were evaluated at morphological, cytological, genetic, epigenetic DNA and transcriptomic levels. This research aimed to provide a scientific insight into the possible molecular mechanisms involved in the occurrence of SV in vanilla at different levels, which could be further applied to the development of an effective molecular pre-screening of clonally micropropagated vanilla samples and early selection of *in vitro* culture samples for genetic improvement of vanilla.

In order to achieve the overall aim of the study, the following specific objectives were undertaken:

1) To establish aseptic *in vitro* culture and efficient regeneration of *V. planifolia*.
2) To investigate the occurrence of somaclonal variations (SV) during direct and indirect nodal plant regeneration of *V. planifolia* at morphological, genetic and epigenetic levels, by using the inter-simple

sequence repeats (ISSR) and methylation-sensitive amplified polymorphisms (MSAP) marker techniques;

3) To determine the pre-existing ploidy levels in different vegetative explants of vanilla and the ploidy population changes in nodal-derived calli and calli maintained over prolonged culture periods of 3, 6, 12 and 15 months, using flow cytometry (FCM);
4) To study the global mRNA and miRNA expression profiles of vanilla calli maintained over prolonged culture duration using the next generation sequencing (NGS) and reverse transcription-quantitative real time PCR (RT-qPCR) technology.

1.2. Thesis outline and project overview

The overall workflow of the project is shown in **Figure 1.1**. In this chapter (**Chapter 1**), an overall introduction of project background, objectives of study and literature review of important topics (a review on somaclonal variations, a background overview of vanilla orchid biology and agriculture, a summary and review of vanilla micropropagation and an introduction to the principles of molecular techniques used in this study) were covered. In **Chapter 2**, establishment of aseptic cultures and nodal regenerations systems using a Malaysian *V. planifolia* cultivar was reported. **Chapter 3** reports the screening of nodal regenerants established in previous chapter for phenotypic, genetic and epigenetic variations. Inter-simple sequence repeats (ISSR) were used as genetic markers, while methylation-senstive amplified polymorphisms (MSAP) were used as epigenetic markers. Comparisons of genetic and epigenetic stability between the directly and indirectly regenerated vanilla regenerants were described. **Chapter 4** describes the cytological analysis using flow cytometry was carried out to investigate the genome content and ploidy levels of different vegetative explant tissues and *in vitro* callus cultures of vanilla over prolonged culture durations of 3, 6 and 12 months. In **Chapter 5**, *de novo* mRNA and miRNA sequencings were performed to further study the gene expression in callus regeneration pathway over prolonged culture periods, which is the

proposed cause of somaclonal variations, with the aim to investigate potential causal mechanisms involved in somaclonal variations in vanilla at the transcriptional level. Quantitative real time PCR was performed to validate candidate genes from the sequencing data. Finally in **Chapter 6**, a summary and discussion on different vanilla regeneration pathways and their applications, recommendations of several molecular and cytological SV detection methods which were proven useful, and suggestions for future research work and improvements were presented and discussed.

PROJECT OUTLINE OVERVIEW

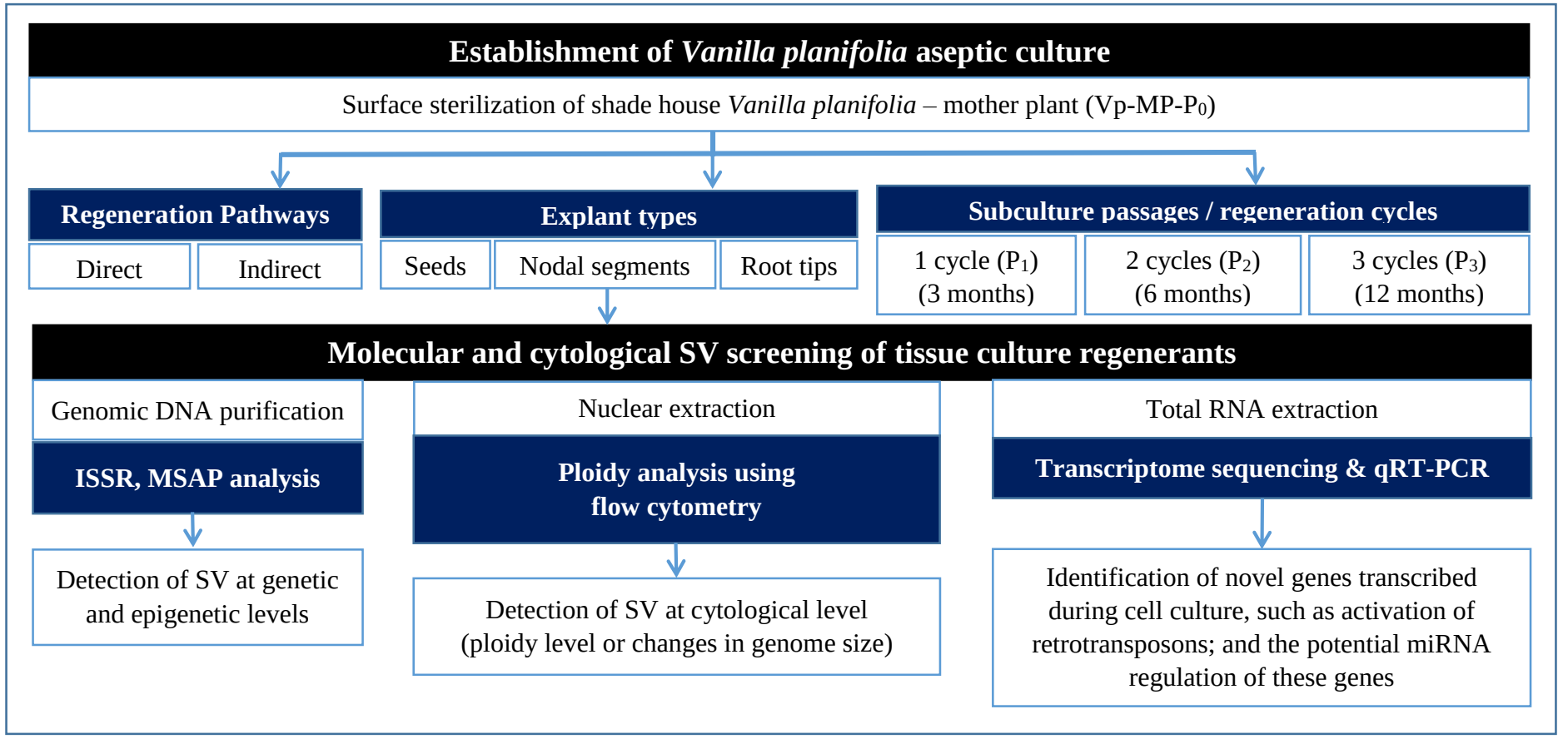


Figure 1.1. Overview of overall research project outline.

1.3. LITERATURE REVIEW

1.3.1. Introduction to vanilla orchids

1.3.1.1. Origin, uses and nutraceutical properties of vanilla

Vanilla is a flavoring traditionally recovered from the cured beans of the orchid genus *Vanilla* (Busconi *et al.*, 2017). It is the second most expensive spice in the world after saffron (Mani *et al.*, 2017). The genus Vanilla is indigenous to Central America, in particular Mexico, and comprises over 110 different species (Soto Arenas, 2003; Soto Arenas and Dressler, 2010); of which only two are currently cultivated for commercial purposes, namely *Vanilla planifolia* (*V. planifolia*) Jacks. ex Andrews and *Vanilla*×*tahitensis* (previously *V. tahitensis* Moore). Among these, *V. planifolia* is the most high-yielding vanilla species, which is most widely cultivated and used for the production of food additives, providing 95% of world vanilla production (Purseglove *et al.*, 1988; Busconi *et al.*, 2017). Vanilla and its extracts are very important and popular as natural flavouring in food, beverages and confectionery preparations (Anuradha *et al.*, 2013) and fragrance ingredient (Purseglove, 1972; Funk and Brodelius, 1994). Other than that, vanilla is well-known for its medicinal properties. In olden medicinal literature, vanilla was documented for its traditional uses as aphrodisiac (Rain and Lubinsky, 2011), anticonvulsive (Ojemann *et al.*, 2006; Zhou *et al.*, 1965) and ayurvedic medicines, as well as used in traditional treatments for fever, dysmenorrhea, ayurveda, and hysteria (Duke 2002). In eighteenth and nineteenth centuries, vanilla was included in the European and American pharmacopeia for its medicinal uses (King *et al.*, 1898; Bythrow, 2005). In the modern biomedical studies, antimicrobial (Fitzferald *et al.*, 2003), anticarcinogenic (involved in DNA strand repair and tumour inhibition) (Duran and Karran, 2003) and anticlastogen (involved in reduction of chromosomal damage caused by X-ray and ultraviolet (UV) (Keshava *et al.*, 1998) properties of vanilla have been reported. The MX-1520 is a modified pro-drug of vanillin and worked as an alternative drug to hydroxurea (which had side effects) for treating sickle cell disease (Zhang *et al.*, 2004). In addition, vanilla was applied as natural preservatives and veterinary medicine (for its sedative effects) (Bythrow, 2005; Duke, 2002). Heliotrope aroma (a vanilla-like scent)-exposed

MRI patients showed 63% less overall anxiety (Bythrow, 2005) and vanilla flavouring encouraged patients on dialysis to eat more (Bulpitt, 2005). Besides, the antioxidant and anti-inflammatory properties of vanilla and vanilla constituents (essential oil or extract) were reported (Kumar *et al.*, 2002; Teuscher *et al.*, 2005; Maurya *et al.*, 2007). In recent studies by Peretti and co-workers (2017), vanillin treatments were also shown to have positive effects in nerve injury treatments by promoting increase in intramuscular vascularization for the muscles, with a pro-inflammatory potential for tibialis anterior. Vanilla pods, vanilla extracts (the isolated extract from vanilla pods) and vanillin (the main flavour of the vanilla extract) constitute a multimillion-dollar industry (Schwab *et al.*, 2017).

1.3.1.2. Vanilla history and cultivation

The history of human utilization of vanilla flavour began with the Aztecs (in current day Mexico) during the 1300s (Schwab *et al.*, 2017). The Aztecs utilized the seedpods from vanilla orchids for flavour and fragrance (Schwab *et al.*, 2017). The earliest documented consumption of vanilla pods dates back to 1520, when the Spanish arrived in Mexico and transported the pods back to Spain (Schwab *et al.*, 2017). Subsequently, the demand for vanilla pods in Spain, France and the rest of Europe increased, and this demand led to domestication of the vanilla orchid, *V. planifolia* (Garrage and Moller, 2015). *V. planifolia* differs very little from its ancestors in the wild as the vanilla plants that are grown in La Reunion, Madagascar, Mauritius and Seychelles are derived from a single cutting of *V. planifolia* (Schwab *et al.*, 2017). Stem cutting propagation results in a lack of genetic variation in vanilla plants (Odoux and Grisoni, 2011).

The orchid, *V. planifolia* is a climbing perennial vine with a large, green, fleshy and succulent stem that is photosynthetic (Schwab *et al.*, 2017). It has oblong, smooth, bright green leaves and adventitious aerial roots that grow opposite of each leaf, aiding lateral support (Schwab *et al.*, 2017). The roots are symbiotically-associated with endotropic mycorrhizal fungi which received the sugars produced by *V. planifolia* during photosynthesis, and in return, helped the

vanilla plants to obtain water and minerals through its fungal mycelium (Anuradha *et al.*, 2013; Schwab *et al.*, 2017; Porras-Alfaro and Bayman, 2007).

1.3.1.3. Vanilla ethymology and taxonomy

Vanilla is indigenous to Central America. In Mexico, the vanilla fruit was called "*tlilxochitl*", which comes from the words "*tlili*" and "*xochitl*", meaning "black" and "pod" (Bouriquet, 1954; Purseglove *et al.*, 1981; Charles, 2012). When discovered by the Spaniards in South America, it was then called "vanilla" from the Spanish word "*vaina*" meaning "pod", which was in reference to the production of vanilla from the seed capsule (Charles, 2012; Christenhusz *et al.*, 2017). The latin word "*planifolia*" described the flatness ("*plani*") of the leaves "*folia*") (Charles, 2012).

The Vanilla plants belong to the Orchidacaeae family, one of the largest families of flowering plants (**Table 1.1**). Among the 800 genera and the 25,000 species within the Orchidaceae family, the genus *Vanilla* contains by itself more than 110 species (Bory *et al.*, 2008b). The detailed classification of vanilla plants into tribes, subfamilies and genera has been challenging (Schwab *et al.*, 2017). The increased use of DNA-based systematic sequence studies has governed the latest classification of the vanilla species (Schwab *et al.*, 2017). According to these studies, the vanilla orchid belongs to the genus Vanilla, tribe Vanilleae, subfamily Vanilloideae and family Orchidaceae (Cameron and Molina, 2006; Cameron *et al.*, 1999; Cameron, 2004, 2009). In the genus of *Vanilla* of about 110 vanilla species (**Table 1.2**), only three species were known to be important for commercial cultivation: *V. planifolia*, *V. tahitensis* and *V. pompon* (Schwab *et al.*, 2017).

Table 1.1. Taxonomic classification of *Vanilla planifolia.* Retrieved [Nov, 11, 2017], from the Integrated Taxonomic Information System (ITIS) on-line database, http://www.itis.gov.

Kingdom	Plantae
Subkingdom	Viridiplantae
Infrakingdom	Streptophyta
Superdivision	Embryophyta
Division	Tracheophyta
Subdivision	Spermatophytina
Class	Magnoliopsida
Superorder	Lilianae
Order	Asparagales
Family	Orchidaceae
Genus	Vanilla Mill.
Species	*Vanilla planifolia* Jacks ex. Andrews

Table 1.2. A list of vanilla species and their geographical locations (Soto Arenas and Cribb, 2010; Odoux and Grisoni, 2010; Azofeifa-Bolaňos *et al.*, 2017; Fraga *et al.*, 2017; Pansarin and Miranda, 2016)

V. abundiflora (Borneo)	*V. grandiflora* (Venezuela, Trinidad, Tobago,	*V. perrieri* (Madagascar)
V. acuminata (Gabon)	Guyana, Peru, Brazil)	*V. phaeantha* (Florida, Caribbean, Mexico, Venezuela)
V. africana (Tropical Africa)	*V. grandifolia* (Príncipe to Zaire)	*V. phalaenopsis* (Seychelles)
V. albida (Indonesia)	*V. griffithii* (West Malaysia)	*V. planifolia* (Mexico and Central America, widely cultivated)
V. andamanica (Andaman Islands)	*V. guianensis* (Guyana, Suriname, Brazil)	*V. platyphylla* (Indonesia: Sulawesi)
V. angustipetala (Brazil, Paraguay, Argentina)	*V. hallei* (Gabon)	*V. poitaei* (Caribbean)

Table 1.2. A list of vanilla species and their geographical locations (continued).

V. annamica (South China to Vietnam)
V. aphylla (Southeast Asia to Java)
V. appendiculata (Guyana, Surinam, Brazil, Peru)
V. arcuata (Brazil)
V. bahiana (Brazil: Bahia)
V. barbellata (South Florida through Caribbean)
V. bertoniensis (Paraguay)
V. bicolor (Caribbean to North South America)
V. borneensis (India, Thailand, Malaysia, Indonesia)
V. bradei (Brazil: São Paulo)
V. calopogon (Philippines: Luzon)
V. calyculata (Honduras, El Salvador, Colombia)
V. chalottii (Gabon)
V. chamissonis (Brazil)
V. claviculata (Caribbean)
V. columbiana (Colombia)
V. coursii (Madagascar)
V. crenulata (Tropical Africa)
V. cristagalli (N Brazil)
V. cucullata (Cameroon)
V. decaryana (SW Madagascar)
V. diabolica (Indonesia: Sulawesi)
V. dilloniana (S Florida to Caribbean)
V. dubia (Brazil: Minas Gerais)
V. dungsii (Brazil)
V. edwallii (Brazil to Argentina)
V. fimbriata (Brazil, Guyana)
V. francoisii (NE Madagascar)
V. gardneri (Brazil)
V. giulianettii (New Guinea)
V. hamata (Peru)
V. hartii (Mexico to French Guiana and Brazil)
V. havilandii (Borneo)
V. helleri (Nicaragua, Costa Rica)
V. heterolopha (Gabon, Congo)
V. hostmannii (Colombia, Guyana, French Guiana, Suriname, Venezuela, Brazil)
V. humblotii (Comoros)
V. imperialis (Tropical Africa to Tanzania and Angola)
V. inodora (Mexico and Central America)
V. insignis (Mexico and Central America)
V. kaniensis (New Guinea)
V. kempteriana (New Guinea)
V. kinabaluensis (Peninsular Malaysia to Borneo)
V. madagascariensis (North and Northwest Madagascar)
V. methonica (Colombia, Peru)
V. mexicana (Florida, Mexico to Venezuela)
V. montana (Malaysia)
V. moonii (Sri Lanka)
V. nigerica (South Nigeria to Cameroon)
V. ochyrae (Cameroon)
V. odorata (South Mexico to tropical South America)
V. organensis (Brazil: Rio de Janeiro)
V. oroana (Ecuador)
V. ovalis (Philippines)
V. ovata (Venezuela & Guyana)
V. palembanica (Sumatra)
V. palmarum (Amazon Basin)
V. parvifolia (South Brazil to Paraguay)
V. penicillata (Venezuela, Brazil)
V. polylepis (Kenya to S tropical Africa)
V. pompona (Mexico, Nicaragua, Costa Rica, West Colombia, Ecuador)
V. ramificans (New Guinea)
V. ramosa (Ghana to Tanzania)
V. ribeiroi (Brazil, Guyana, Colombia)
V. roscheri (Ethiopia to KwaZulu-Natal)
V. ruiziana (Peru, Bolivia)
V. savannarum (Cuba)
V. schwackeana (Brazil: Minas Gerais)
V. seranica (Indonesia: Seram)
V. seretii (West Coast tropical Africa)
V. siamensis (Yunnan to Thailand)
V. somai (China to Taiwan)
V. sotoarenasii (Costa Rica)
V. sprucei (Colombia)
V. sumatrana (Sumatra)
V. tahitensis (South Pacific)
V. trigonocarpa (Costa Rica to North Brazil)
V. utteridgei (New Guinea)
V. vellozii (Brazil, Paraguay, Argentina)
V. verrucosa (Argentina)
V. walkeriae (South India, Sri Lanka)
V. wariensis (New Guinea)
V. wightii (South India, Sri Lanka)
V. yersiniana (Thailand, Vietnam)
V. zanzibarica (East Africa)

1.3.1.4. Vanilla botany and cultivation

Vanilla plants are terrestrial, climbing and branch poorly in the shade unless the vines are exposed to the sun (Fouché and Jouve, 1999). They are grown in hot-humid tropical climates and have the agro-ecological requirements of: temperature at 20 to 32ºC; precipitation at an average rate from 2,000 to 3,000 mm per year; altitude from sea level to 600 m; shading at 50 to 70%; well-drained soil rich in humus; and with support trees to climb on (Busconi *et al.*, 2017). Optimal flowering of the vanilla plants and their subsequent pod production from successfully fertilized flowers, requires specific climatic conditions of a dry and cool season of at least two months (Hernandez and Lubinsky, 2011).

As a hemi-epiphytic orchid, the vanilla plant requires physical support, shade and organic material during cultivation (Busconi *et al.*, 2017). In its natural habitats, vanilla grows in forest composed of mostly secondary vegetation called as "*acahual"* (Havkin-Frenkel and Belanger, 2011). Plant species such as the "laurel" (*Litcea glaucescens*), "pata de vaca" (*Bahuinia divaricate*), "cojón de gato" (*Tabernaemontana* sp.), "cacahuapaxtle" or "balletilla" (*Hamelia erecta*) and "capulím" (*Eugenia capulia*), are commonly encountered in the *acahual*, and are used as support trees for vanilla (Curti-Diaz, 1995). Traditionally in Mexico, vanilla is also intercropped with other crops such as coffee or orange, where the vanilla benefits from the abundant organic matter and shade of the intercropped plant, while the crop diversification allowed growers to benefit from two products in one site (Havkin-Frenkel and Belanger, 2011). These vanilla plants are thus supported by trees that are used to provide shade to the coffee such as *Inga sp.* or other introduced species such as *Erythrina* sp. (Havkin-Frenkel and Belanger, 2011). In the case of an intensive monoculture crop system, vanilla is cultivated in either field or shade house, with "*pichoco"* (*Erthrina* sp.) or "*cocuite*" (*Gliricidia sepium*) as support trees.

1.3.1.5. Vanilla flowering and pollination

In general, the first flowering, or "rehearsal" ("*ensayo"*) of vanilla, happens three years after planting (Havkin-Frenkel and Belanger, 2011). The physiological cue to flower is induced by climatic, mechanical (Havkin-Frenkel and Belanger, 2011) or drought stress (Odoux and Grisoni, 2011). Cool temperatures "burns" the apical tip of the vanilla plant, killing it, and break the apical dominance of the plant while stimulating laterial floral buds to develop (Havkin-Frenkel and Belanger, 2011). The infloresence of vanilla is a raceme of which individual flowers are borne on stalks arranged around a central axis (Odoux and Grisoni, 2011) (**Figure 1.2**, **Figure 1.3a-b**). An individual plant usually forms 8 to 15 racemes or more (Odoux and Grisoni, 2011; Ranadive, 2005), depending on the quantity of flowering and rooting shoots, and the prevailing environmental conditions. The vanilla flowers are yellow, bisexual and usually sprout towards the top of the plant (**Figure 1.3**) (Schwab *et al.*, 2017).

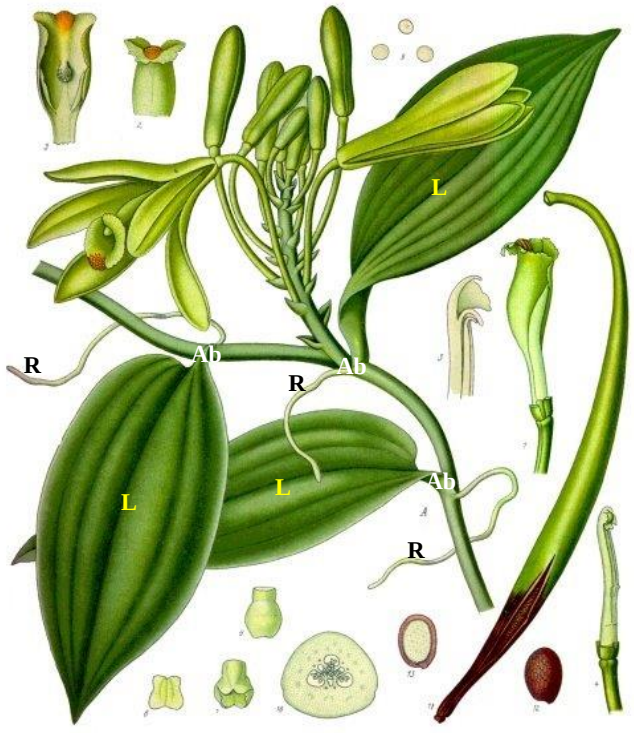


Figure 1.2. Schematic diagram showing plant forms of a vanilla orchid. Axillary buds, leaves and aerial roots are abundant at each nodes of the plant. R: aerial roots; Ab: axillary buds (wrapped in between leave bases and nodes); L: leaves. Adapted from IPNI (2017).



Figure 1.3. *Vanilla planifolia* vines, flowers and pods in KK vanilla plantation in Jerantut, Pahang, Malaysia. (a-b) Vanilla inflorescence: Before pollination, ovaries face upward with their flower, and then curve downwards (Havkin-Frenkel and Belanger, 2011), (c) Young flower buds and (d) mature flower of *V. planifolia* Andrews. (e) Hand pollination process of vanilla flower. (f) post-pollination of vanilla flower, flowers wilted after one day and successfully pollinated flower will form vanilla pods; (g) immature vanilla pods; and (h) vanilla vines in greenhouse of Bestino Natural Oils Sdn Bhd, Perak. [April 19, 2014; November 12, 2014]

Normally, only one flower per raceme opens per day (Havkin-Frenkel and Belanger, 2011). Flowering of vanilla occurs in sequential order, starting from the base of the raceme (Havkin-Frenkel and Belanger, 2011). The flowers are fully open at dawn and last 6 to 8 h before closing in the afternoon heat (Havkin-Frenkel and Belanger, 2011). In nature, vanilla flowers are pollinated by specific natural pollinators (Havkin-Frenkel and Belanger, 2011). The identities of the natural pollinators are still unclear. In Mexico, the bees (*Melipona beechii*), hummingbirds (*Cynniris* sp.) and bats pollinate vanilla (Havkin-Frenkel and Belanger, 2011). Other documented bee species include the shiny orchid bee *Euglossa viridissima* (Soto-Arenas 1999a, 2003; Hagsater *et al.*, 2005; Lubinsky *et al.*, 2006) and *Eulaema* sp. (Havkin-Frenkel and Belanger, 2011). However, the mechanisms by which these bees actually pollinate the vanilla flowers is yet to be documented, while the rate of natural pollination is very low (only about 1% of the fruits) (Havkin-Frenkel and Belanger, 2011).

Although vanilla flowers are bisexual, they are not able to self-pollinate (Schwab *et al.*, 2017). Therefore, vanilla plants are usually hand-pollinated (**Figure 1.3e**), which is a labour intensive process. Pollination requires outside action either by transferring pollen from an anther to the stigma (**Figure 1.4a**) using a soft tiny paintbrush or by lifting a thin-membrane that prevents self-fertilization, and subsequently pressing the flower's anther towards the stigma (Fouche and Jouve, 1999) (**Figure 1.4b**). The lack of specific flower pollinators was a key issue encountered when introducing *V. planifolia* to the rest of the world from Mexico (Schwab *et al.*, 2017). In 1841, Edmond Albius had developed an artificial pollination technique of vanilla flowers by using bamboo sticks (Schwab *et al.*, 2017). Even today, hand pollination techniques are used for the flower pollination in vanilla production, and these techniques have not developed much further since 1841 (Fouche and Jouve, 1999; Rodolphe *et al.*, 2011).

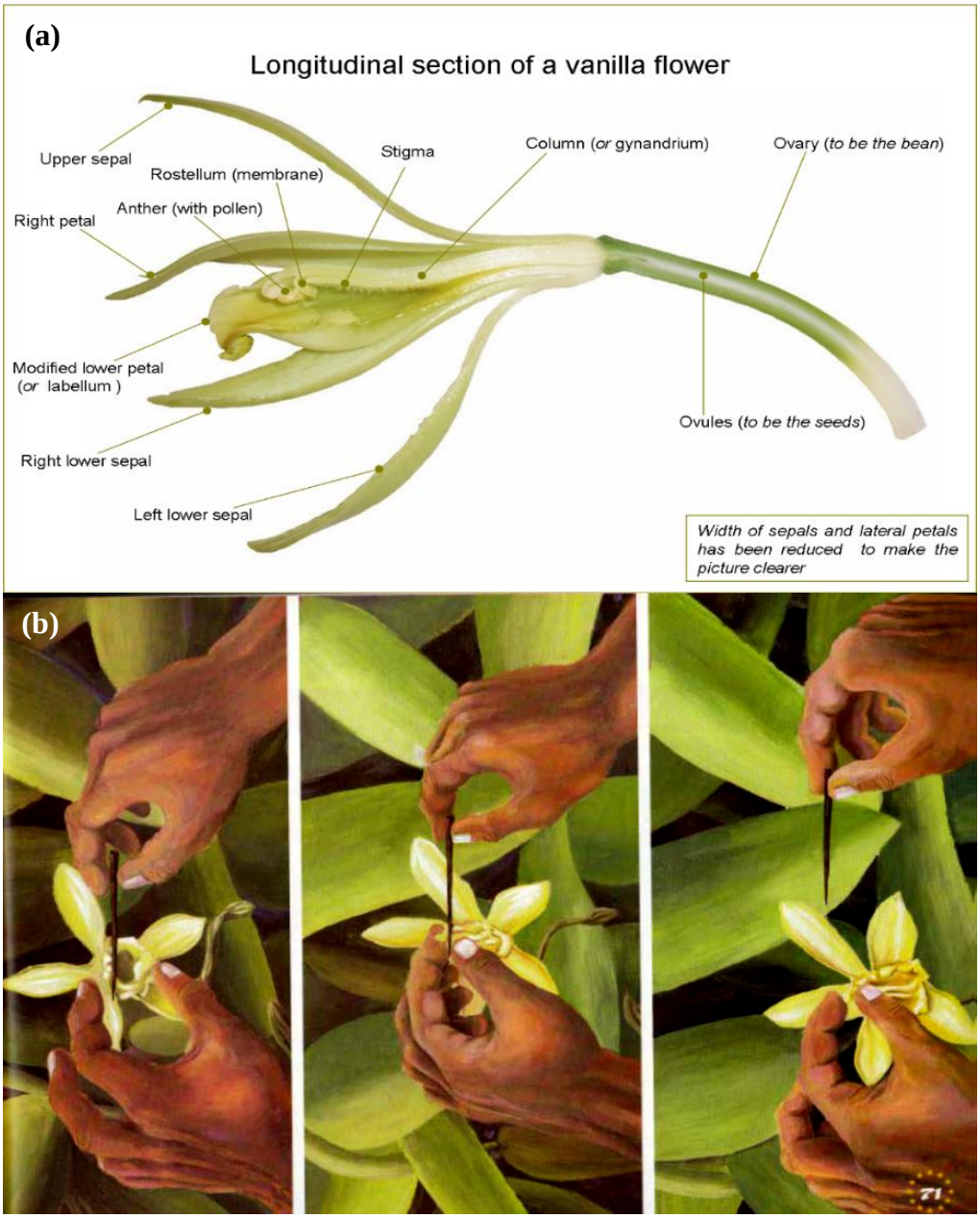


Figure 1.4. Illustration of (a) longitudinal section and (b) hand pollination process of vanilla flower. Adapted from Massey (2017).

1.3.1.6. Botany of green vanilla pod

The vanilla pod is a seed capsule generally referred to as a vanilla bean (Schwab *et al.*, 2017). The pod reaches its full size about 10 to 15 weeks after pollination (Schwab *et al.*, 2017). Fully matured pods are about 15 cm long and are pale green to yellow in color (Schwab *et al.*, 2017) (**Figure 1.5a**). In the transverse section view, a mature green vanilla pod is triangular and contains a central

cavity with a lot of minute black seeds (Schwab *et al.*, 2017) (**Figure 1.5b**). From the outer part to the inner cavity, the pod consists of the following tissues: epicarp, mesocarp and endocarp (Schwab *et al.*, 2017). The mesocarp which consists of parenchyma cells fills the majority of the fruit volume, while the cavity of the pod contains black seeds that are attached to a long narrow funicle (Odoux and Brillouet, 2009) (**Figure 1.5c-d**). Vanilla pods are harvested at the age of 8 to 9 months, which is before the pods begin to split from the end and turn yellow in color (Schwab *et al.*, 2017). The immature green vanilla pods are almost odorless because the key flavour components are stored as glucosides. Freshly harvested pods are processed by curing to stop the natural vegetative process (Schwab *et al.*, 2017) (**Section 1.3.1.7**). This process induces enzymes responsible for the formation of aromatic flavour constituents while preventing microbial growth, thus enabling the long-term preservation of vanilla pods (Odoux and Grisoni, 2011). Curing methods can be different according to the country of cultivation (Odoux and Grisoni, 2011). As a result, the curing process has a major influence on the variety, quality and aromatic profile of the pods that are traded (Rao and Ravinshankar, 2000).

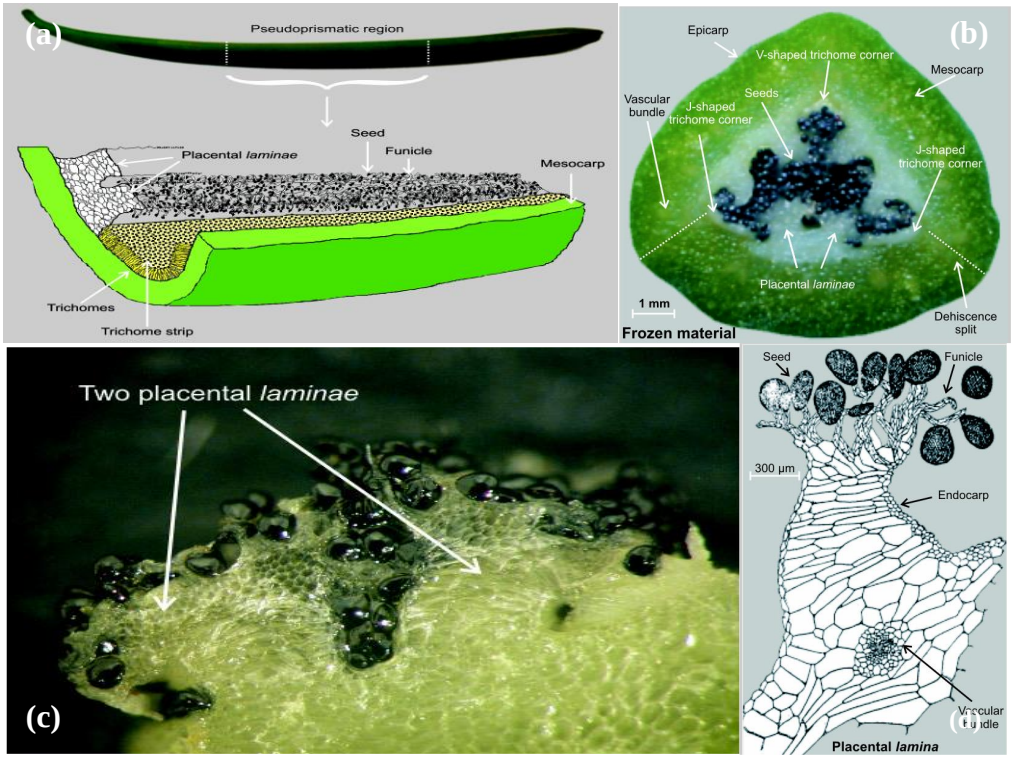


Figure 1.5. Vanilla seed pod. (a) Schematic drawing of longitudinally sectioned vanilla pod; (b) Split anatomy of vanilla pod; (c) two placental lamilae; and (d) its schematic diagram. Adapted from Odoux and Grisoni (2011).

1.3.1.7. When the greens become black – vanilla pods curing

The characteristic flavour of vanilla is due to the presence of a large number of aromatic compounds that are produced during various stages of curing or processing. The principal component of vanilla is “vanillin”, which contributes to one third of the total vanilla flavour, whereas other volatile and non-volatile compounds contribute to the remaining two thirds of the flavour (Ranadive, 1994). Green vanilla beans are almost odourless, the aroma in the pods or beans is developed through a labour-intensive curing process (**Figure 1.6**). In general, the curing process of vanilla beans involves wilting, sweating, drying and conditioning of the pods (Dignum *et al.*, 2001). Although every vanilla growing country has developed its own curing process, all the processes generally involve four common steps, which are scalding, sunning or sweating, drying and conditioning (Anuradha *et al.*, 2013) (**Figure 1.6**). The entire process generally takes six to eight months to be completed (Rao and Ravinshankar, 2000). The vanilla extract contains not only vanillin, but also a range of related phenyl-propanoid (C6-C3) compounds, which combine together to impart the unique subtlety to natural vanilla flavour (Clark, 1990). In green vanilla beans, only a little free vanillin is present, the aroma precursors such as β-D-glucosides come into contact with β-D-glucosidases to form vanillin only during the curing process (Walton *et al.*, 2003).

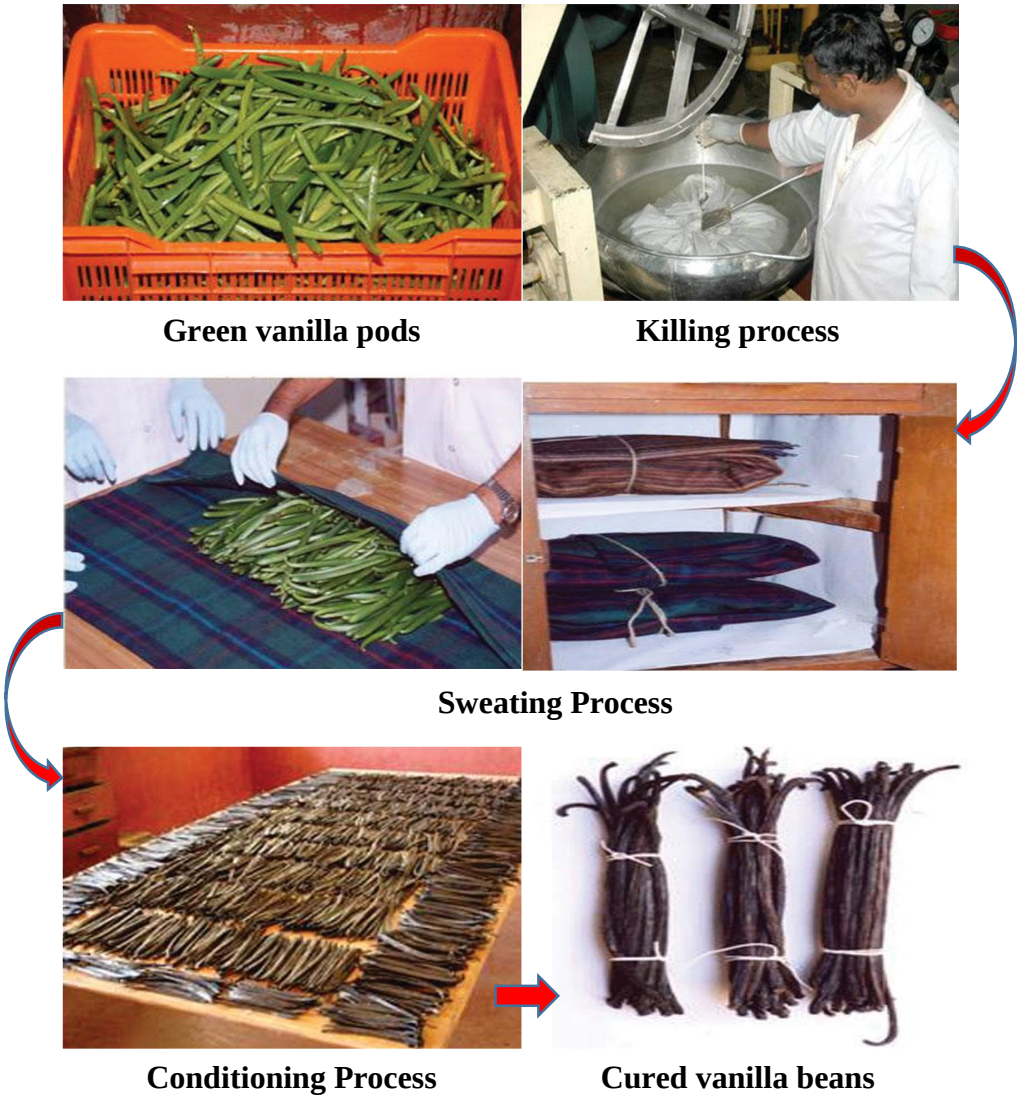


Figure 1.6. Main stages of the vanilla curing process, also defined as controlled ripening process. In hot water treatment of freshly-collected green vanilla beans, the process is called *killing* as it disrupts the respiratory function by destructing cell membranes and cell walls of the bean tissue and thus brining about their physiological death (Anuradha *et al.*, 2013). The *sweating* process which follows the killing process involves reduction of moisture levels in vanilla pods to the level sufficient for further enzyme activities while reducing the risk of microbial spoilage (Ranaddive, 1994). In general, in this step vanilla beans develop their characteristic color, aroma and flavouring properties (Anuradha *et al.*, 2013). The process lasts for 7-10 days and is carried out in sweat boxes, enclosed rooms and rarely, in ovens (Anuradha *et al.*, 2013). At the end of sweating period, the cured beans, which are now brown in color and aromatic, with 60-70% moisture, are further dried to 23-32% moisture to protect them from microbial spoilage and reduces undesirable enzyme acitivies and biochemical changes in the vanilla pods. Finally, at the conditioning process, the vanilla beans are stored in closed boxes for one to several months (Anuradha *et al.*, 2013). In this step, various chemical and biochemical reactions, such as esterification, etherification, oxidative degradation take place to produce various aroma constituents and further enchance the overall flavour quality of cured beans

(Anuradha *et al.*, 2013). The methods of curing followed are named after the places of its origin, such as Mexicam process, Bourbon process, Peruvia n process and Guyana process (Anuradha *et al.*, 2013). Adapted from Anuradha *et al.* (2013) and Lionnet (1959).

Madagascar is the largest vanilla pod producer in the world. Madagascar, together with La Réunion islands, account for nearly 75% of the vanilla pod market (Schwab *et al.*, 2017). Vanilla pods from Madagascar and other islands east of Africa (La Réunion, Mauritius, the Comoros and the Seychelles) have the status 'Bourbon', and are considered the best vanilla pods in the current market (Schwab *et al.*, 2017). The vanilla growers in this region also produce the majority of the world's natural vanilla extract, which is an alcohol extract from matured and cured vanilla pods (Schwab *et al.*, 2017). The vanilla extract isolated from the vanilla pods includes more than 200 aromatic compounds (Gallage and Moller, 2015). This intensive curing process is also one of the reasons why vanilla is the world third most expensive spice, making it a commercially valuable crop.

1.3.1.8. Diseases of vanilla orchids

The conditions of temperature and humidity under which vanilla grows tend to favour the development of pathogens, mainly fungi (Havkin-Frenkel and Belanger, 2011). Besides, the monoclonal propagation has also led to the susceptibility of vanilla plants to diseases. Diseases are one of the main factors that damage and reduce vanilla production (Havkin-Frenkel and Belanger, 2011). Vanilla is susceptible to various pathogen diseases, including root and stem root caused by *Fusarium oxysporum* F. sp. vanilla (**Figure 1.7a-b**), black rot (*Phytophtora* sp.), athracnose (*Colletotrichum* sp.) (**Figure 1.7c-d**), rust (*Uromyces* sp.) (**Figure 1.7e**), and various viral diseases such as Cybidium Mosaic Virus (CYMV), Vanilla Mosaic Virus (VMV) (**Figure 1.7f**), Vanilla necoris potyvirus (VNPV) and Odontoglossum Ringspot Virus (ORSV). Plant tissue culture is one of the approaches to mass propagate disease-free vanilla clones, as well as to increase the genetic diversity of vanilla through *in vitro* crop improvements.

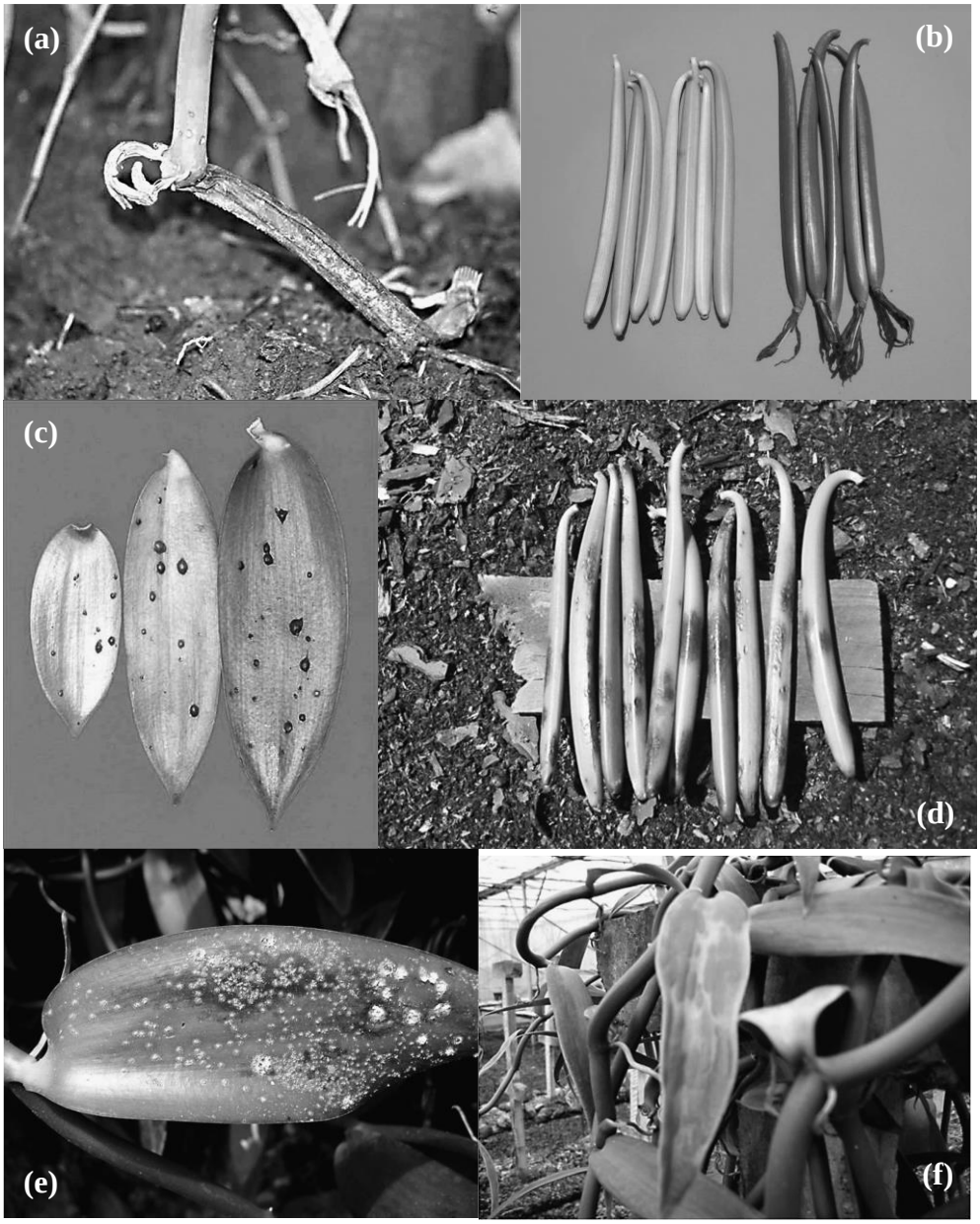


Figure 1.7. Vanilla diseases. (a) Stem rot disease caused by *Fusarium oxysporum* f. sp. vanilla, (b) yellowing and falling of fruits: diseased fruits (left) and healthy fruits (right), (c) Anthracnose disease on vanilla leaves: initial, intermediate and end stage, (d) damaged fruits, "pintos", caused by *Colletotrichum* sp., (e) symptoms of *Uromyces* sp., and (f) distortion and mosaic lesions infected by vanilla mosaic virus on a vanilla leaf. Images adapted from Hernández-Hernández (2011).

1.3.1.9. Vanilla macro- and micropropagation

Traditionally, vanilla is propagated by stem-cutting. Cuttings are made from highly productive and vigorous individuals that have never produced fruits (Ranadive, 2005). The cuttings used for propagation should not be a flowering shoot and should have at least 3 nodes, usually 6 to 8 nodes, 80-200 cm long and 1 cm in diameter, with viable axillary buds for producing new shoots from which the plant will grow (Ranadive, 2005). Shorter cuttings, 20 cm in length, will take three to four years to flower and fruit; while cuttings of 90-100 cm in length are usually preferable as they tend to flower earlier and will usually flower and fruit in one to two years (Odoux and Grisoni, 2011). Due to the succulent nature of the vanilla cuttings, they could be stored or transported for a period of up to two weeks, if required (Odoux and Grisoni, 2011).

1.3.1.9.1. *In vitro* conservation through shoot tips

Traditionally, vanilla germplasm is conserved in clonal repositories belonging to botanical gardens and scientific institutions. Traditional conservation systems are highly costly and limits the number of accessions that can be preserved. For the conservation of vanilla biodiversity, attempts have been made *in vitro* (Jarret and Fernandez, 1984; Minoo *et al.*, 2006b) through shoot tip culture.

1.3.1.9.2. *In vitro* germination and hybridization using seeds

Vanilla produces numerous minute seeds that do not germinate under natural conditions. Vanilla seeds, like other orchids, require the presence of specific mycorrhizal fungi in order to germinate in nature (Rasmussen *et al.*, 2015). For breeding purposes, vanilla can be grown from seed explant (Odoux and Grisoni, 2011). Hybridization and the production of plants from seeds have been carried out in Puerto Rico and Madagascar (Odoux and Grisoni, 2011). Tissue culture techniques can be used to successfully germinate the seeds. The vanilla seeds need to be disinfected and cultured in nutrient medium (Knudson, 1950). Germination of vanilla seeds is generally optimal when cultured in the dark at 32ºC, except that seeds of interspecific crosses between *V. planifolia* and *V. pompona* required a higher temperature of 34ºC for germination (Odoux and Grisoni, 2011). Protocols for seed and embryo culture of vanilla have been

standardized (Gu *et al.*, 1987; Knudson, 1950; Minoo *et al.*, 1997; Withner, 1955). Cytokinin is required during the germination of vanilla seeds, probably related to the utilization of lipids that constitute the primary storage material in most orchid seeds and it has been observed that unless storage lipid is utilized, germination does not continue (De Pauw *et al.*, 1995; Odoux and Grisoni, 2011).

1.3.1.9.3. Micropropagation

To date, *in vitro* regeneration of vanilla has also been developed through shoot tip and axillary bud (Geetha and Shetty 2000, George and Ravishanka 1997, Giridhar and Ravishankar 2004, Kalimuthu *et al.* 2006, Gopi *et al.* 2006, Lee-Espinosa *et al.* 2008, Tan *et al.* 2011b), root apex (Philip and Nainar 1986), leaf (Gu *et al.* 1987, Janarthanam and Seshadri 2008, Tan *et al.* 2011a) and seed-derived protocorm-like bodies (Malabadi and Nataraja 2007, Retheesh and Bhat 2011).

Micropropagation of vanilla using apical meristem was standardized for large-scale multiplication of disease-free and genetically stable plants (Kononowicz and Janick, 1984; Minoo, 2002; Minoo *et al.*, 2006; Minoo and Nirmal Babu, 2009). Successful plant regeneration from shoot and seed-derived callus was also reported in vanilla (Nirmal Babu *et al.*, 1997; Minoo, 2002), variability among callus-regenerated plants in vanilla has not been well studied (Odoux and Grisoni, 2011).

1.3.1.10. Global market and demands of vanilla

Vanilla is the only orchid of significant economic importance as an edible crop (Fouché and Jouve, 1999). The increasing popular demand for natural vanilla flavor has increased the global market for vanilla. Vanilla supply and prices are regularly affected by abiotic and biotic factors such as cyclone damage (in Madacasgar), pests and diseases (**Section 1.3.1.8**) and political regulatory factors (Brown, 2009). Additionally, the labor-intensiveness of growing and curing vanilla beans as described earlier has made them the second most expensive spice in the world, after saffron (Havkin-Frenkel and Belanger, 2011). According to official report from FAOSTAT, the highest vanilla price has gone

up to 1,272.52, 51,311.62 and 16,182.94 USD/tonnes for Indonesia, Madagascar and Mexico varieties respectively in the year 2003 (**Figure 1.8**, FAOSTAT, 2016).

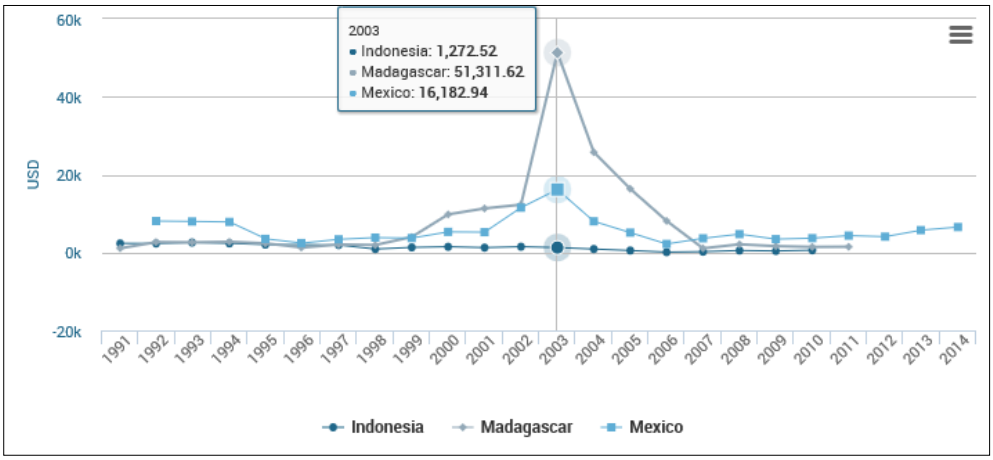


Figure 1.8. Producer prices of world top three vanilla producers (in USD per tonnes) from 1991-2014 (FAOSTAT, 2016).

The top five world players in vanilla production are Indonesia, Madacasgar, China, Papua New Guinea and Mexico, with Indonesia, producing up to 2.5k tonnes production in the past 20 years (**Figure 1.9a-b**, FAOSTAT, 2016). And in recent years, Tonga, Kenya and Uganda are also leading their way in production yield (**Figure 1.9c**). Madagascar produces around 75% of the world's supply of vanilla. A poor harvest and overstocking by importers has driven the price up by 129% over the last year to $224 per kg, making it 14 times more expensive than the third priciest spice (cardamom, $36 per kg) in the world. In the past ten years, although the harvest area increased, the growth rate of vanilla yield decreased (**Figure 1.10**), which suggested that agricultural improvements are needed to maximize the yield and land use. This has led to the importance to improve and mass produce high quality plant materials for the vanilla agro-industry. The use of clonal elite planting materials that ensures the maximum yield potential rather than relying on traditional stem-cutting methods, as well as the utility of *in vitro* improvement methods to create more elite vanilla varieties, could help in improving the vanilla agriculture.



Figure 1.9. World vanilla beans production from 1994 to 2014: (a) quantities, (b) distribution, and **(c) world countries delivering the five highest vanilla yield** (FAOSTAT, 2016).

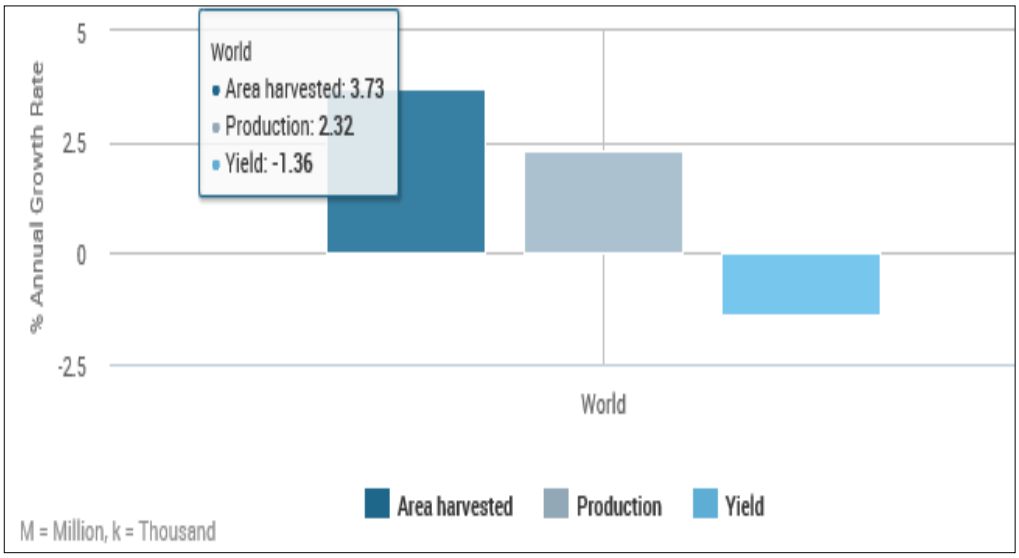


Figure 1.10. World area harvested, yield and production of vanilla from 2004 to 2014 (FAOSTAT, 2016).

1.3.2. Introduction to plant tissue culture (PTC) and somaclonal variation (SV)

1.3.2.1. Introduction to plant tissue culture (PTC)

Plant tissue culture (PTC), or micropropagation, is the technique to produce plants of desirable agricultural traits, from small pieces of plant (the explant) in relatively short periods of time (Hussain *et al.*, 2012). It enables the rapid multiplication of disease-free plants under a controlled aseptic environment (Hussain *et al.*, 2012). Depending on the plant species, an explant may be taken from many parts of the plant, namely the shoot tip, leaf, axillary bud, stem or root tissues of the source mother plant (Hussain *et al.*, 2012).

Generally, the development of plantlets *in vitro* are divided into three to five major stages (Murashige, 1974). At the stage I (the initiation and establishment of aseptic culture phase), a suitable plant part or explant, is sterilized and cultured aseptically onto a culture medium (Torres, 1989). At the stage II (or termed the multiplication phase), the number of disinfected explants from stage I could be rapidly increased through different types of multiplication pathway, such as the direct or indirect pathways, which were also divided into

two types, namely the somatic embryogenesis or organogenesis pathways, based on the nature of cell development (Adamczuk *et al.*, 2012) (**Figure 1.11**). Plant materials in the stage II may be recycled by subculturing the explants back onto a proliferation medium for further multiplications, or it could enter the stage III, which is the conditioning or pre-transplant phase (Torres, 1989). At this stage, the explants from stage II are conditioned or prepared for transplant: somatic embryos (from the somatic embryogenesis pathway) are germinated, and shoot (single or multiple; from the organogenesis pathway) are prepared for rooting (Torres, 1989). These three stages were first described by Murashige in 1974, and have hence been adopted for the propagation of a wide variety of plants at both research and commercial level (Torres, 1989). Two other stages can be added to the process, which includes the stage 0 which involves the preparation of the stock or mother plant from which the primary explants derived; and the stage IV, which is the acclimatization of the plantlets established *in vitro*, a process which plantlets established *in vitro* are transferred to an external environment (Torres, 1989).

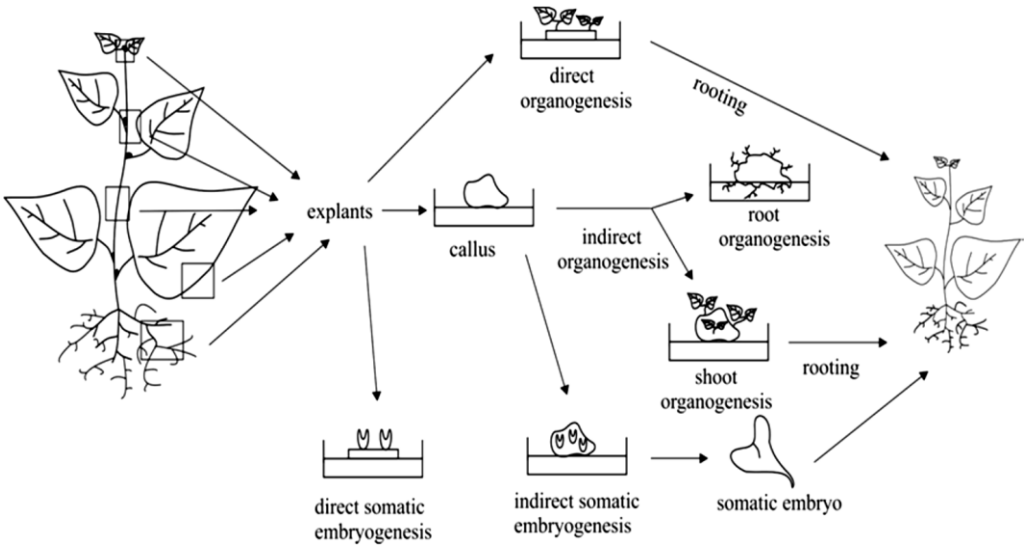


Figure 1.11. Routes of plant regeneration in tissue cultures *in vitro* from different explants. Adapted from Adamczuk *et al.* (2012).

One of the most important factors which governs the growth and morphogenesis of plant tissues in culture is the composition of the culture medium (Hussain *et al.*, 2012). The basic components in plant tissue culture media is made to be very similar to the basic nutrient requirements of whole

plants, which are generally made up of some or all of the following components: macronutrients, micronutrients, vitamins, amino acids or other nitrogen supplements, sugar(s), other undefined organic supplements, solidifying agents or support systems, and growth regulators (Torres, 1989; Hussain *et al.*, 2012). Several media formulations have been developed and are widely used for the majority of tissue culture work, such as those described by White (1963), Murashige and Skoog (1962), Gamborg *et al.* (1968), Nitsch and Nitsch (1969), Gautheret (1942), and Schenk and Hilderbrandt (1972). Murashige and Skoog's MS medium, Gamborg's B-5 medium and Schenk and Hilderbrandt's SH medium are high in macronutrients, while the other media formulations contain considerably less of macronutrients (Torres, 1989). Besides, there were also media specifically developed for orchids, such as the Knudson's C orchid medium (Knudson, 1950).

Among all the different media, Murashige and Skoog medium (MS medium; see **Appendix D** for its media composition) is most extensively used media in *in vitro* culture. The pH of the media plays an important role in affecting the plant growth and performances of the PGRs, and is usually adjusted to the value between 5.4 to 5.8 (Hussain *et al.*, 2012). Besides, the composition of the medium, in particular the nitrogen source and PGRs, has profound effects on the growth response of the initial explants by determining the developmental pathway of the plant cells and tissues in culture medium (Hussain *et al.*, 2012). The type and concentrations of plant growth regulators (PGRs) used are dependent mainly on the plant species, explant type and the objective of experiment (Sahni *et al.*, 2017). There are many types of PGRs that could be applied in plant tissue culture, which include the classical plant hormones (auxins, cytokinins, gibberellins, abscicic acid, ethylene) and natural growth substances with phytohormonal-like regulatory roles (polyamines, oligosaccharins, brassinosteroids, jasmonates, salicylates, dehydrodiconiferyl alcohol glucosides, systemin, turgorins and other natural stimulators and inhibitors) (Gaspar *et al.*, 1996). Among these, auxins and cytokinins are the most widely used PGRs in plant tissue culture, and their levels in plant cells determine the type of culture being established or regenerated (Sahni *et al.*, 2017). In general, higher auxin favors the formation of roots, whereas higher

cytokinins promotes shooting (Sahni *et al.*, 2017). A balance of both auxin and cytokinin leads to the development of callus, a mass of undifferentiated cells (Ikeuchi *et al.*, 2013) (**Figure 1.12**). In this study (**Chapter 2**), the role of auxins and cytokinins are explored in different explants of vanilla.

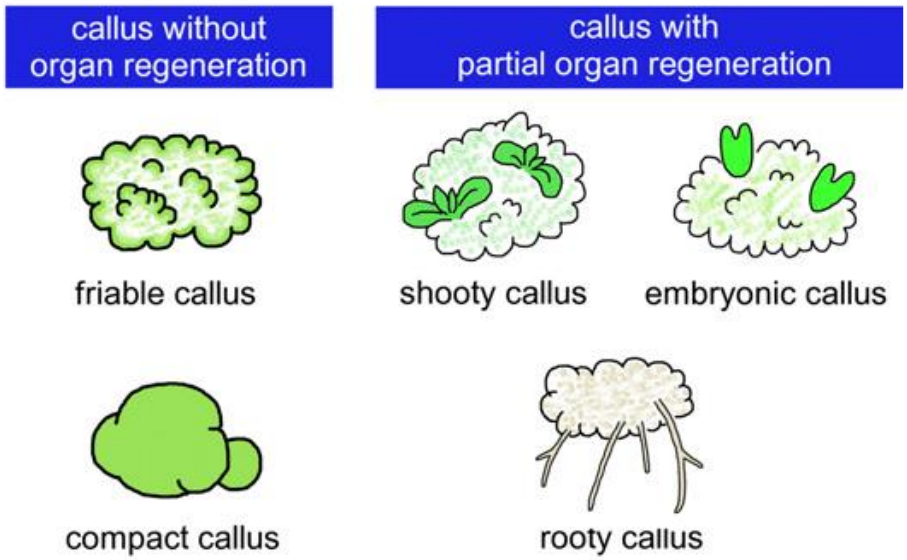


Figure 1.12. A schematic illustration of different types of callus that can be induced *in vitro*: friable, compact, shooty, embryonic and rooty calli Adapted from Ikeuchi *et al.* (2013).

1.3.2.2. Definitions and implications of SV in agriculture

Theoretically, plant cells *in vitro* and their regeneration into whole plants is an asexual process which involves only mitotic cell division; therefore, it should be 'clonal' and not cause any variation, producing genetically uniform plants (Larkin and Scowcroft, 1981). However, in practical situations, the occurrence of uncontrolled and random spontaneous variation during the culture process can occur (Bairu and Kane, 2011). This phenomenon is called somaclonal variation (SV).

Some 50 years ago, Braun (1959) first reported that *in vitro*-cultured cell clones display gene alterations. The term 'somaclonal variation' (SV) was first defined by Larkin and Scowcroft (1981) as genetic variation occurring in tissue culture-derived plants which originated by adventitious regeneration from somatic cells. Later this definition was again redefined as "phenotypic variation,

either genetic or epigenetic in origin, displayed among somaclones, namely plants derived from any form of cell culture involving the use of somatic plant cells" (Miguel and Marum, 2011). The general term "somaclones" refers to the plants derived from cell and tissue culture, while "somaclonal variants" refers to somaclones displaying variations (Bhatia, 2015). The variants obtained using callus cultures are referred as calliclones, while variants obtained from protoplast cultures are known as protoclones (Sahijram, 2015). Besides, SV is variation that is inherent to *in vitro* culture but not related to the genetic transformation techniques (Sági *et al.*, 1998).

Somaclonal variation is a double-edged sword whereby its presence in micropropagation programmes is inimical, while it can be gainfully exploited to create stable variations such as disease resistance, where other methods failed or are cumbersome to achieve (Sahijram, 2015). Plants regenerated from tissue and cell cultures show heritable variations for both qualitative and quantitative traits that can be advantageous in agriculture (Sahijram, 2015). Therefore, somaclonal variation could be exploited as a tool for the creation of desirable agronomic traits (Bairu *et al.*, 2011). Commercially useful crop varieties have been successfully generated from *in vitro* selection of somaclonal variants in numerous plants species, such as apple (*Malus sp.*), banana (*Musa sp.*), chilli (*Capsicum sp.*) and many others (**Table 1.3**) for superior horticultural traits such as disease resistance and increased crop yield.

Table 1.3. A list of commercially exploited horticultural crop varieties which originated from SV. Adapted from Krishna *et al.* (2016).

Horticultural crop	Characteristic of somaclone	References
Aglaonema	Cultivar 'Moonlight Bay' and 'Diamond Bay' from 'Silver Bay,' and 'Emerald Bay,' from 'Golden Bay'	Henny *et al.* (1992, 2003)
Apple (*Malus* x *domestic*a Borkh.)	Resistance to *Erwinia amylovora*	Chevreau *et al.* (1998)
Apple rootstocks M 26 and MM 106 (*Malus pumila* Mill.)	Resistance to *Phytophthora cactorum*	Rosati *et al.* (1990)
Apple rootstock Malling 7	Resistance to white root rot (*Dematophora necatrix*)	Modgil *et al.* (2012)
Anthurium sp.	'Orange Hot' derived from 'Red Hot' clone	Henny and Chen (2011)
Banana (*Musa acuminata* L.)	Semi-dwarf and resistant to *Fusarium* wilt TC1-229	Tang *et al.* (2000)
	Larger bunch size var. TC2-425; Resistant to *Fusarium oxysporum* f. sp. cubense (Foc) race 4; bunch 40 % heavier than cv. Formosana	Hwang (2002)
	Fusarium wilt-resistant somaclonal variants of banana cv. Rasthali	Ghag *et al.* (2014)
	Var. CIEN-BTA-03, resistant to yellow Sigatoka	Giménez *et al.* (2001)
	10 somaclones; GCTCV215-1 released for commercial planting	Hwang and Ko (1992, 2004)
	Var. CUDBT-B1, reduced height and early flowering	Martin *et al.* (2006)
	Var. Tai-Chiao No. 5, superior horticultural traits and resistance to *Fusarium* wilt	Lee *et al.* (2011)
Begonia (*Begonia* x *elatior*)	Plant morphology, number of flowers per plant, and flower size	Jain (1997)
Brinjal (*Solanum melongena* L.)	Stress-tolerant somaclone selection	Ferdausi *et al.* (2009)
Blackberry	Thornless var. 'Lincoln Logan'	Hall *et al.* (1986)
Capsicum (*Capsicum annuum* L.)	Yellow fruited var. Bell sweet	Morrison *et al.* (1989)
Calthea roseopicta	Developed common cultivars like Angela, Cora, Dottie, Eclipse and Saturn	Chao *et al.* (2005)
Carrot (*Daucus carota* L.)	Resistance to leaf spot (*Alternaria dauci*)	Dugdale *et al.* (2000)
	Resistant to drought	Rabiei *et al.* (2011)
Carnation (*Dianthus caryoplyllus* L.)	Resistant to *Fusarium oxysporum* f. sp. dianthi	Esmaiel *et al.* (2012)
Celery (*Apium graveolens* L.)	Fusarium resistant var. UC-TC	Heath-Pagliuso and Rappaport (1990)
	Multiple-resistant (insect resistance against *Spodoptera exigua* and disease resistance against Fusarium yellow) somaclones K-26, K-108 and K-128	Diawara *et al.* (1996)
Celosia argentea L.	Resistance to nematode	Opabode and Adebooye (2005)

Cereus peruvianus	Shoots with different areoles characteristics	Resende *et al.* (2010)
Chili pepper (*Capsicum annuum* L.)	Early flowering and increase of yield components	Hossain *et al.* (2003)
Chrysanthemum (*Dendranthema grandiflora*)	Variation in leaf, flower shape and petal size	Ahloowalia (1992)
	Daisy type chrysanthemum	Jevremović *et al.* (2012)
	Attractive variants with changed inflorescence colors	Miler and Zalewska (2014)
Citrus spp.	Resistant to *Phoma tracheiphila*	Deng *et al.* (1995)
	Salinity tolerance	Ben-Hayyim and Goffer (1989)
Cuphea viscosissima Jacq.	Significantly superior over the parents for mean plant height, leaf area, seed yield, per cent caprylic acid and lauric acid contents	Ben-Salah and Roath (1994)
Cymbopogon winterianus Jowitt	Aromatic grass var. CIMAP/Bio-13 with 50–60 % increased oil yield	Mathur *et al.* (1988)
	Increased total oil yield and quality with high geraniol content	Nayak *et al.* (2003)
	Cymbopogon martinii Increased oil content	Patnaik *et al.* (1999)
Dieffenbachia sp.	Novel and distinct foliar variegation with taller, larger canopy and longer leaves than 'Camouflage' parental plants	Shen *et al.* (2007)
Garlic (*Allium sativum* L.)	Consistently higher bulb yield than the parental clone	Vidal *et al.* (1993)
	Resistance against the pathogenic fungi '*Sclerotium cepivorum*'	Zhang *et al.* (2012)
Geranium spp.	Vigorous and attractive flower	Skirvin and Janick (1976)
	Isomenthone-rich somaclonal mutant	Gupta *et al.* (2001)
	Cv. 'CIM Pawan, a somaclone of the Bourbon type variety Bipuli, with more herbage and essential oil yield than Bipuli	Saxena *et al.* (2008)
Gerbera (*Gerbera jamesonii* Bolus)	Novel cultivars	Minerva and Kumar (2013)
Ginger (*Zingiber officinale* Rosc.)	Tolerant to wilt pathogen (Fusarium oxysporum f.sp. zingiberi Trujillo)	Bhardwaj *et al.* (2012)
Grapevine (*Vitis vinifera* L.)	Resistant to *Botrytis cinerea* and *Plasmopara viticola*	Kuksova *et al.* (1997)
Haemerocallis spp.	Dwarf, short flowers, male sterile var. Yellow Tinkerbell	Griesbach (1989)
Hedychium (ornamental ginger)	Ramata, dwarf and variegated cultivar	Sakhanokho *et al.* (2012)
Java citronella (*Cymbopogon winterianus*)	Somaclonal variant variety CIMAP/Bio-13, which yields 37 % more oil and 39 % more citronellon than the control variant	Mathur (2010)
Kiwi fruit (Actinidia deliciosa)	5 somaclones, derived from cv. Tamuri, tolerant to NaCl	Caboni *et al.* (2003)
Mango (*Mangifera indica* L.)	Resistant to *Colletotrichum gleosporiensis*	Litz *et al.* (1991)
Mint (*Mentha arvensis*)	Increased herb and oil yield	Kukreja *et al.* (1991; 2000)

Myrobolan (*Prunus cerasifera* Erhr)	Water logging-tolerant clone variant (S.4) of myrobolan rootstcock Mr.S 2/5 for peach cv. Sun Crest	Iacona *et al.* (2013)
Olive (*Olive europea*)	Bush olive somaclone (BOS), columnar olive somaclone (COS)	Leva *et al.* (2012)
Patchouli (*Pogostemon patchouli*)	Higher herb yield and essential oil content	Ravindra *et al.* (2012)
Pea (*Pisum sativum* L.)	Resistance to *Fusarium solani*	Horáček *et al.* (2013)
Peach (*Prunus persica* L.))	Somaclones S156 and S122 resistant to leaf spot, moderately resistant to canker in cvs. Sunhigh and Red haven	Hammerschlag and Ognjanov (1990)
	Resistant to root-knot nematode (*Meloidogyne incognita* Kofoid and White)	Hashmi *et al.* (1995)
	Somaclone S 122-1 was found resistant to bacterial canker (*Pseudomonas syringae* pv. *syringae*)	Hammerschlag (2000)
Pear (*Pyrus* sp.)	Resistant to *Erwinia amylovora*	Viseur (1990)
	Pear rootstock (*Pyrus communis* L.) 'Old Home 9 Farmingdale (OHF 333)' Tolerance to the fire blight	Nacheva *et al.* (2014)
Philodendron	Cultivars 'Baby Hope' from 'Hope'	Devanand *et al.* (2004)
Picrorhiza kurroa	Higher glycoside contents including kutkoside and picroside I in somaclone 14-P derived through *Agrobacterium rhizogenes* mediated transformed hairy root cultures of *P. kurroa*	Mondal *et al.* (2013)
Pineapple (*Ananas comosus* L., Merr.)	Spineless variant	Jaya *et al.* (2002)
	Cvs. P3R5 and Dwarf, variation in fruit color, growth habit, fruit size and length of plant generation cycle	Pérez *et al.* (2009, 2012)
Potato (*Solanum tuberosum* L	Non-browning var. White	Arihara *et al.* (1995)
	Somaclones for heat tolerance	Das *et al.* (2000)
	Somaclones IBP-10, IBP-27 and IBP-30, derived from cultivar Desiree, showed higher resistance to *Alternaria solani* and *Streptomyces scabiei*	Veitia-Rodriguez *et al.* (2002)
	Improved size, shape, appearance, starch content and starch yield	Thieme and Griess (2005)
	Superior processing attributes than cv. 'Russet Burbank'	Nassar *et al.* (2011)
	High-yielding genotype SVP-53	Hoque and Morshad (2014)
	Increased phytonutrient and antioxidant components over cv. 'Russet Burbank'	Nassar *et al.* (2014)
Quince A (*Cydonia oblonga*)	High soil pH	Dolcet-Sanjuan *et al.* (1992), Marino *et al.* (2000)
Stevia rebaudiana	High glycoside contents (steviol, stevioside, and rebaudioside)	Khan *et al.* (2014)
Strawberry (*Fragaria* sp.)	Resistant to *Fusarium oxysporum* f. sp. *fragariae*	Toyoda *et al.* (1991)
	Resistant to *Alternaria alternate*	Takahashi *et al.* (1993)
	Resistant to *Phytophthora cactorum*	Battistini and Rosati (1991)

	Improved horticultural traits	Biswas *et al.* (2009)
	Resistant to *Verticillium dahliae*	Zebrowska (2010)
	'Serenity', a paler skin-colored, late season, resistant to powdery mildew and *Verticillium* wilt somaclonal variant of the short-day cv. 'Florence'	Whitehouse *et al.* (2014)
Sweet potato (*Ipomea batatas* L. Lam.)	Tolerant to salinity	Anwar *et al.* (2010)
Sweet orange (*Citrus sinensis* (L.) Osb.)	Somaclone of OLL (Orie Lee Late) sweet orange; late maturing; suitable for fresh market or processing, exceptional juice quality and flavor	Grosser *et al.* (2015)
St. Augustine grass [*Stenotaphrum secundatum* (Walt.) Kuntze]	Freeze-tolerant somaclonal variant SVC3	Li *et al.* (2010)
Syngonium podophyllum Schott	22 cultivars, derived from original 'White Butterfly' clone, with distinct and stable foliage characteristics	Henny and Chen (2011)
Tomato (*Lycopersicon esculentum* L.)	High solid contents var. DNAP9	Evans (1989)
Tulip (*Tulipa* sp.)	''Bs6'', selected from among the micropropagated plants of the cultivar 'Blue Parrot' with red-violet colored longer flower and stem	Podwyszynska *et al.* (2010)
Torenia (*Torenia fournieri*)	Flower color somaclonal variants	Nhut *et al.* (2013)
Turmeric (*Curcuma longa* L.)	High essential oil yielding somaclones	Kar *et al.* (2014)
	Turmeric somaclone resistant to *Fusarium oxysporum* f.sp. Zingiberi	Kuanar *et al.* (2014)
Indian ginseng (*Withania somnifer*a (L.) Dunal)	Withanolide (12-deoxywithastramonolide)-rich somaclonal variant	Rana *et al.* (2012)

On the other hand, in the cases of mass propagation of elite plants and plant culture for genetic transformation where clonality is required, somaclonal variation is undesirable. For instance, somaclonal variations have been major problems in the banana and oil palm industry. In banana industry, SV is genotype dependent. Distinct off-types for banana included dwarfism and disease resistance. The 'Grand Naine' group had dwarf, giant and massada (mosaic leaf variegation) somaclonal variants (Vidhya and Ashalatha, 2002; Walther *et al.*, 1997; Martin *et al.*, 2006; Uma *et al.*, 2002), while the 'Cavendish' group had variants for plant stature, abnormal leaves, pseudostem pigmentation, persistence of flowers and split fingers (Israeli *et al.*, 1991) (**Figure 1.13**) were derived from somaclonal variation of tissue culture. In oil palm, the mantled flower abnormality (**Figure 1.14**) had caused great economic losses in the oil palm industry, as the somaclonal variation took up to years to be detected (Jaligot *et al.*, 2011). This has raised the importance of studying SV in perennial fruit crops, rather than focusing primarily on herbaceous plant species.



Figure 1.13. Somaclonal variants of banana showing (a) abnormal, spindle-shaped leaf lamina, (b-c) leaf texture/colour variants and (d) abnormal split-fingers at banana brunch. Adapted from Sahijram (2015).



Figure 1.14. Somaclonal variants of oil palms with mantled fruits forms: normal fruit form (left) and mantled fruit form (right). Adapted from Ong-Abdullah *et al.* (2015) and Smulders and Klerk (2011).

Vanilla shared some common characteristics with oil palm as it is also a perennial crop, where the plant would only flower and produce vanilla pods, after reaching maturity at least 3 years after field plantation (Castro-Bobadilla and García-Franco, 2007; Diez *et al.*, 2016). Thus, effective screening of SV during *in vitro* genetic improvement process, or early detection for clonal fidelity during mass propagation is important in vanilla to prevent economic losses. To achieve such economic gain, it is important to understand the mechanisms underlying and factors which affects the occurrence of SV.

1.3.2.3. Possible mechanisms associated to SV

To date, the mechanism associated to the occurrence of SV is still poorly understood. Although it has been studied extensively, the causes remain largely theoretical or unknown (Skirvin *et al.*, 1993, 1994). Cultured plant cells are believed to experience genomic stresses as a result from different physical and chemical factors during *in vitro* culturing environment such as the influence of PGRs, temperature changes, wounding response, disturbed circadian rhythm, nutrient and other chemical supplements and photoperiods (Neelakandan and Wang, 2012). Coupled with the developmental events of dedifferentiation and regeneration, these changes are manifested at the gene expression level in transcription factors (TF), structural genes and protein kinases, contributing to the adaptation of explant to stresses and the reorientation of its developmental program (Neelakandan and Wang, 2012). In *in vitro* culture, both genetic changes, such as chromosomal changes, DNA sequence alterations, transpositions and amplifications; and the later discovered epigenetic changes, such as DNA methylation, small RNA-mediated gene regulation and chromatin modification, could occur (Neelakandan and Wang, 2012). The spectrum of these changes can therefore potentially give rise to phenotypic changes among the regenerants, causing the production of somaclonal variants (Neelakandan and Wang, 2012) (**Figure 1.15**).

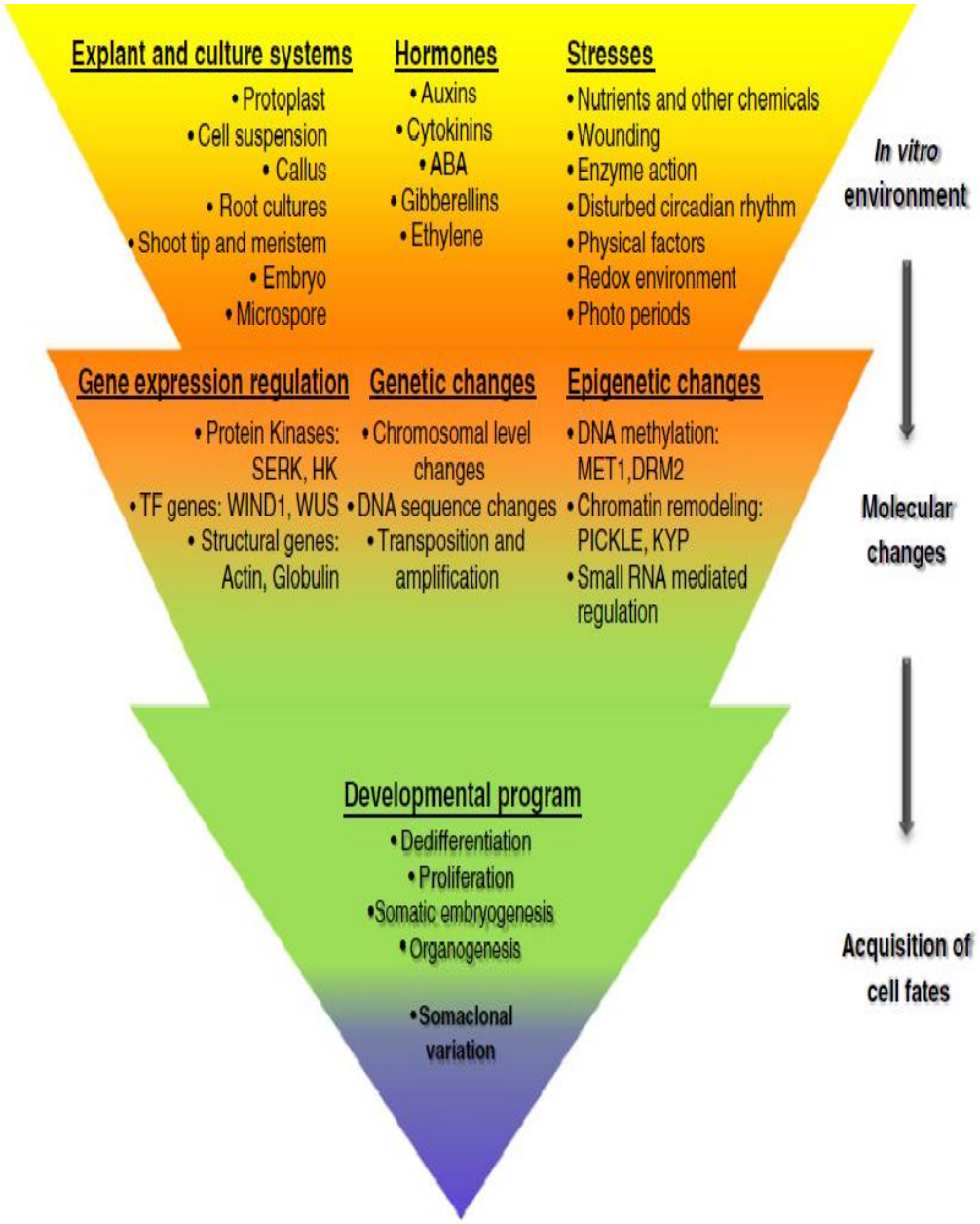


Figure 1.15. ***In-vitro*** **culture and molecular changes caused in the SV process.** Adapted from Neelakandan and Wang (2012).

1.3.2.3.1. Genetic mechanisms associated to SV

Genetic variations observed in *in vitro* regenerated plants are largely stochastic, unpredictable and non-reproducible, and have been classified as originating from the source or explant ('pre-existing') or induced during the culture process ('*de novo*') (Larkin and Scowcroft, 1981). Various types of genetic changes have been observed, including polyploidy, aneuploidy, (point) mutations and new insertions of (retro) transposons (**Table 1.4**) (Smulders and Klerk, 2011). For example, in the flowering plant *Saintpaulia*, tissue culture regenerants show altered variegation and flower color phenotype because of the activation and mobility of VGs1 transposon (Sato *et al.*, 2011a). Percentages of 'pre-existing' mutated cells in the source tissue from different plant parts and the '*de novo*' variations generated during culturing and regeneration process were estimated using quantitative real-time PCR technique and compared. It was demonstrated that heritable variations in regenerated plants are indeed predominantly associated with '*de novo*' tissue culture-induced changes (Sato *et al.*, 2011b). Label portions are known to exist in the genome rendering it susceptible to variation. These portions get modulated when cells undergo the 'stress' during TC, resulting in higher rearrangement and mutation rates than other portions of the genome (Thomas *et al.*, 2002). Occurrence of hot-spots of mutation and recurring-menus of alternative alleles is consistent with this response being limited to a sub-fraction of the genome. Other factors such as ploidy level, transposable elements, genomic status of donor plant and DNA methylation status may be responsible for the induction of SV (Sahijam *et al.*, 2003). Recent literature suggests that hypo- and hyper-methylation of DNA triggering genome-wide changes may be the underlying cause (James *et al.*, 2007).

Table 1.4. Tissue culture-induced transposable elements in plants. Adapted from Neelakandan and Wang (2012).

No.	Name	Species	Type of transposon	Class	Activation	References
1	*Ac1*	Maize	DNA transposon	II	Low frequency transposition in *in vitro* culture-regenerated plants	Peschke *et al.* (1987) and Brettell and Dennis (1991)
2	*BARE-1*	Maize	DNA transposon	II	Proposed to be transpositionally active in *in vitro* culture	Planckaert and Walbot (1989)
3	*dTstu1*	Maize	DNA transposon	II	Low frequency transposition in *in vitro* culture-regenerated plants	Peschke and Phillips (1991)
4	*Karma*	Tobacco	Retrotransposon	I	Tissue culture	Pouteau *et al.* (1991)
5	*LIb*	Tobacco	Retrotransposon	I	Tissue culture	Hirochika (1993)
6	*LORE1*	Tobacco	Retrotransposon	I	Tissue culture, wounding, methyl jasmonate, and some fungal elicitors	Hirochika (1993)
7	*Lullaby*	Rice	Retrotransposon	I	Tissue culture, Introgression, one copy in chromosome 7 active in Nipponbare, reverts in the regenerated plant	Hirochika *et al.* (1996)
8	*MERE1*	Rice	Retrotransposon	I	Tissue culture	Hirochika *et al.* (1996)
9	*mPING*	Rice	Retrotransposon	I	Tissue culture	Hirochika *et al.* (1996)
10	*Mu*	Carrot	En/Spm like element	II	Long-term ([4 years) sub-culturing in suspension cultures	Ozeki *et al.* (1997)
11	*nDaiZ*	Rice	Tourist like MITE	II	Induced by anther culture, interspecific hybridisation	Kikuchi *et al.* (2003)
12	*Pong*	Rice	DNA transposon	II	Active in indica cell culture line, also in interspecific hybridization	Jiang *et al.* (2003)

13	*Popin*	Rice	Retrotransposon	I	Activated in R1 plants, not in cultured cells or R0 plants	Komatsu *et al.* (2003)
14	*RYS1*	Rye	Fold back transposon		Activated by *in vitro* stress	Alves *et al.* (2005)
15	*Spm*	Maize	Rice Pong homologues	II	Transcriptionally active in *in vitro* culture, somaclonal variants in anther/ callus cultures and lines	Barret *et al.* (2006)
16	*Tdc1*	Sweet potato	Retrotransposon	I	High transposition in callus and meristem culture	Yamashita and Tahara (2006)
17	*Tnt1*	Rice	MITES	II	One copy in Chromosome 5 activated by tissue culture, only in a few genotypes	Huang *et al.* (2009)
18	*Tos10*	Rice	Retrotransposon	I	Transposed in callus	Picault *et al.* (2009)
19	*Tos17*	Maize	Novel maize specific TE, small size (43 bp)		Tissue culture	Rhee *et al.* (2009)
20	*Tos19*	Medicago	Retrotransposon	I	Low copy number, activated in tissue culture and shows preferential gene insertion	Rakocevic *et al.* (2009)
21	*Tto1*	Potato	Stowaway MITE	II	Tissue culture	Momose *et al.* (2010)
22	*Tto2*	Lotus	Retrotransposon	II	Activated in plants regenerated from dedifferentiated cells	Fukai *et al.* (2010)
23	*VGs1*	*Saintpaulia* sp.	DNA transposon	II	Tissue culture	Sato *et al.* (2011)
24	*ZmTPAPong like-1 and 3*	Barley	Retrotransposon	I	Activated by tissue culture involving dedifferentiation	Evrensel *et al.* (2011)

1.3.2.3.2. Epigenetic mechanisms associated to SV

Epigenetic changes are caused by changes in the expression of the information in the DNA brought about by alterations in DNA methylation, histone modifcation and small RNA-directed DNA methylation (RdDM) (Smulders and Klerk, 2011) (**Figure 1.16**). These modifications may influence gene transcription (Smulders and Klerk, 2011). Epigenetic changes could be temporary, where the plants may revert back to normal phenotypes; but it could be also long lasting by which the changes are transferred during sexual propagation (Brettell and Dennis, 1991).

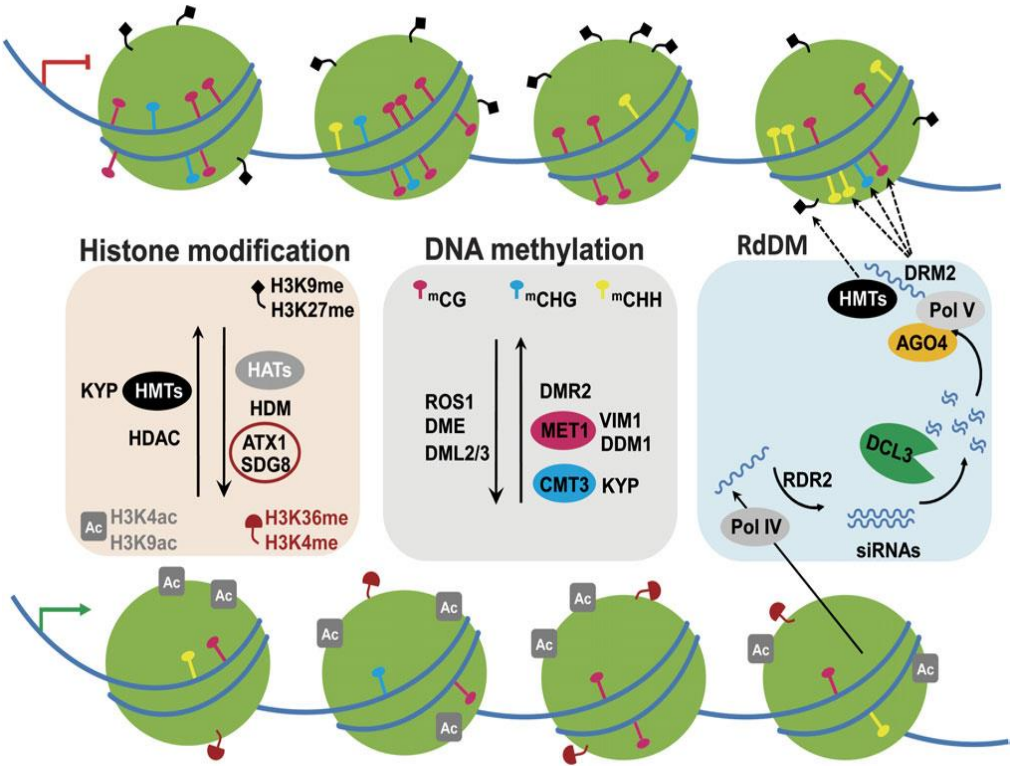


Figure 1.16. Schematic illustration of changes in chromatin structure via DNA methylation, histone modification, and small RNA-directed DNA methylation (RdDM). Repression of gene transcription is associated with methylation of DNA by maintenance and *de novo* DNA methyltransferases. CG methylation (mCG) is maintained by MET1 but it is also controlled by VIM1 and the chromatin remodeller DDM1. CHG methylation (mCHG) is maintained by CMT3 and the HMT KYP, and at some loci is redundantly controlled by CMT3 and DRM2. CHH methylation (mCHH) is maintained by the RdDM pathway, and de novo methylation of DNA in all of these sequence contexts is generally established by DRM2. In addition, the chromatin repressed state is associated with histone methylation at H3K9 and H3K27 by HMTs. Upon

acetylation at H3K4 and H3K9 by HATs and methylation at H3K4 and H3K36 by HMTs (ATX1 and SDG8), as well as DNA demethylation by the glycosylase/lyase proteins ROS1, DME, DML2, and DML3, chromatin is modified and gene transcription may be activated. Transposable elements are kept silent by the RdDM pathway involving the generation of transcripts by PolIV that are converted into dsRNAs by RDR2 and cleaved by DCL3 into 24 nucleotide long siRNAs. Antisense siRNAs are loaded onto AGO4, which guide PolV to homologous DNA, which is then methylated in all cytosine sequence contexts by de novo DNA methyltransferases (DRM2). AGO4, ARGONAUTE 4; ATX1, ARABIDOPSIS TRITHORAX-LIKE PROTEIN 1; CMT3, CHROMOMETHYLASE 3; DCL3, DICER-LIKE 3; DDM1, DECREASE IN DNA METHYLATION 1; DME, DEMETER; DML2/3, DEMETER-LIKE 2/3; DRM2, DOMAINS REARRANGEDMETHYLTRANSFERASE 2; HAT, histone acyltransferase; HDAC, histone deacetylase; HDM, histone demethylase; HMT, histone methyltransferase; KYP, KRYPTONITE; MET1, DNA METHYLTRANSFERASE 1; RDR2, RNA-DEPENDENT RNA POLYMERASE 2/3; ROS1, REPRESSOR OF SILENCING 1; SDG8, SET DOMAIN GROUP 8; VIM1, VARIANT IN METHYLATION 1. Adapted from Smulders and Klerk (2011).

The evaluation of epigenetic modifications in plants cultured *in vitro* has up to now been mostly focused on the analysis of DNA methylation because this is one of the best described epigenetic mechanisms. In the last couple of years, with the development of high-throughput genome technologies, a number of options have become available to profile the epigenome of several organisms including plants such as *Arabidopsis*. These new methods are able to provide a detailed characterization of genomic DNA methylation and histone modifications at an unprecedented resolution, which can be integrated with transcriptomics data including the smRNA transcriptome (Zhang *et al.*, 2007; Lister *et al.*, 2008, 2009; Wang *et al.*, 2009b). For example, a few studies have reported the detection of modifications in histones and smRNA levels in plant cells cultured *in vitro* (**Table 1.5**) (Miguel and Marum, 2011).

Table 1.5. Modifications in histones and small RNA levels detected in plant cells/tissues cultured *in vitro*. Adapted from Miguel and Marum (2011).

Species	***In vitro* cultured cells/tissues**	**Histone and small RNA level modifications**	**Reference**
Arabidopsis thaliana	Cell suspension cultures	Loss of H3 methylation on Lys9 and, in some cases, gain of H3 trimethylation at Lys4 Increased levels of 21 nucleotide siRNAs	Tanurdzic *et al.* (2008)
Arabidopsis thaliana	Cell suspension cultures	Loss of acetylated H3 and H4 and trimethylated Lys4 H3	Berdasco *et al.* (2008)
Nicotiana tabacum	Protoplasts	Increased levels of acetylated H3; modification of Lys9-methylated H3	Williams *et al.* (2003)
Solanum tuberosum	Cell suspension cultures	Multiacetylation of H3.1, H3.2 and H4	Law and Suttle (2005)
Zea mays	Callus cultures	Increase of ubiquitinated H2A in callus derived from root differentiation zone compared with callus derived from other root zones	Alatzas and Foundouli (2006)

1.3.2.4. Possible factors influencing rate of SV

The extent of genomic instability depends on a range of factors, including explant source, genotype, nature or concentration of the exogenous growth hormone used in nutrient media, *in vitro* system, genome size, age of the culture and presence of an intermediate callus phase (Bairu *et al.*, 2011).

1.3.2.4.1. Explant source

In general cases, highly differentiated tissues such as roots, leaves and stems, when used as an explants or explant-source, produce more variation than axillary buds and shoot tips, which have pre-existing meristems (Duncan, 1997). However, there are also exceptions, for example, shoot tip culture of banana give rise to higher rate of somaclonal variation compared to cell suspension cultures initiated from immature male flowers (Shchukin *et al.*, 1997, 1998; Cote *et al.*, 2000), which may be due to the presence of pre-existing variation such as endopolyploidization, somatic mutations, etc. in the explants used to initiate tissue culture (Veilleux and Johnson, 1998).

1.3.2.4.2. Genotype

Different genomes could respond differently to the plant tissue culture (PTC) stress-cased variations, indicating that SV might have genotypic components (Bairu *et al.*, 2011). The differences in stability are most likely contributed by the differences in genetic make-up whereby some components of the plant genome are less stable during the culture process (‘labile genome portions’) (Bairu *et al.*, 2011). One of the suggested explanations for this phenomenon is the presence of repetitive DNA sequences, which can differ in quality and quantity between plant species (Lee and Phillips, 1988). For example, in banana, one of the most studied model plants for SV, different varieties (such as 'New Guinea Cavendish' and "Williams') had different levels of susceptibility to SV (Stover, 1987; Israeli *et al.*, 1991; Sahijram *et al.*, 2003). Similar cases were also reported in different cultivars of wheat (Mehta and Angra, 2000), coffee (Etienne and Bertrand, 2003) and strawberry (Popescu *et al.*, 1997).

1.3.2.4.3. Plant growth regulators (PGRs)

An efficient micropropagation requires optimal concentration and precise ratios of auxins and cytokinins (Skoog and Miller, 1957), while the genetic composition of a cell population can be influenced by the relative levels of both auxins and cytokinins (D'amato, 1975). Past research studies suggested that plant growth regulators (PGRs) are involved in the occurrence of SV in three main contexts. Firstly, the primary events that triggers morphogenesis, as controlled by exogenous PGRs, might have induced variability through cell-cycle disturbance (Peschke and Phillips, 1992). Secondly, PGRs preferentially increase the rate of divisions of pre-existed genetically abnormal cells (Bayliss, 1980). Thirdly, sub- and supra-optimal levels of PGRs and synthetic hormones have been shown to link with SV to a greater extent (Vidal and De García, 2000; Martin *et al.*, 2006). This includes cytokinins, such as BAP, at high levels such as 15 mg/L in banana cultivar 'Williams' (Giménez *et al.*, 2001) or 30 mg/L in rice callus (Oono, 1985); or auxins, such as synthetic auxin 2,4-D (Nehra *et al.*, 1992; Gesteira *et al.* 2002; Jin *et al.*, 2008). It was reported that 2,4-D is often associated with genetic abnormalities such as polyploidy and endoreduplication (Bouman and De Klerk, 2001; Adhmed *et al.*, 2004 Mohanty *et al.*, 2008). However, there were also exceptions in this aspect. For example, in some species or cultivar, exposure to high concentration of PGRs, such as up to 53.28 µM (12 mg/L) of BA and 55.80 µM (12 mg/L) of kinetin did not cause any SV in banana cultivar 'Nanjanagudu Rasabale' (Venkatachalam *et al.*, 2007). In view of the contradictory reports, the roles of type and concentration of PGRs particularly cytokinins on incidence of SV in different species or cultivar remains a subject for debate and warrants further stringent experiments for confirmation (Bairu *et al.*, 2011).

1.3.2.4.4. Culture time and long-term cultures

The age of cultures has been shown to enhance variability among regenerated plants, and it is positively related to the number of generations or subculture cycles (Reuveni and Israeli, 1990; Rodriges *et al.*, 1998; Bairu *et al.*, 2006, 2011). For example, in banana cv 'Nainco' shoot tip cultures, although no somaclonal variants were observed at the end of three subcultures, SV appeared after the

fifth subculture and their frequency increased after the subsequent five, seven, nine and eleven subcultures (Rodrigues *et al.*, 1998). It is said that *in vitro* culture conditions and rapid multiplication of PTC may affect genetic stability of the explants by increasing the mutation rate per generation of cells, and thus accumulating it over time (Duncan 1997).

1.3.2.4.5. Type of TC-induced response

The type of TC-induced response also plays an important role in the induction of SV. The presence of a disorganized growth phase is considered as one important origin and cause of SV (Rani and Raina, 2000). Generally, the more the organizational structure of the plant is broken down, the greater the chance of mutations occurring (Araújo *et al.*, 2001; Cooper *et al.*, 2006). Callus cultures are characterized by extensive dedifferentiation induced by high concentration of PGRs, in contrast to regeneration of shoots from pre-existing meristems, which may influence the induction of SV (Bairu *et al.*, 2011), as shown in the callus culture of various plant species such as *Oryza sativa* (Araújo *et al.*, 2001), *Fragaria* sp. (Popescuu *et al.*, 1997) and *Citrus* sp. (Hao and Deng, 2002). However, exceptions were also observed, such as in *Musa* species, of which the shoot-tip derived cultures exhibited a higher rate of SV than callus (Shchukin *et al.*, 1997, 1998).

1.3.2.5. Tools used for SV detection

As reviewed (**Section 1.3.2.3**), the basic cause of SV may be attributed to changes at genetic level (DNA amplification and deletion, mobility of transposable elements), epigenetic level (DNA methylation, histone modifications and small RNA regulations) and cytological level (karyotype: chromosome number and structure, chromosome rearrangements, somatic crossing over, sister chromatid exchange). With the advances in molecular technologies, SV can be characterized based on using a wide range of techniques and tools (**Table 1.6**). Different levels of characterization have advantages and disadvantages (summarized in **Table 1.7**) and thus, different techniques are often combined and complemented in one study. In this study, we combined the

use genetic (ISSR) and epigenetic markers (MSAP), ploidy analysis (flow cytometry) and transcriptome profiling (NGS sequencing and RT-qPCR) to study the SV occurrence in tissue culture of vanilla.

Table 1.6. Techniques used for detecting cell culture-induced genomic changes (Neelakandan and Wang, 2011).

Type of variations	Technique
Chromosome level changes	Flow cytometry Fluorescent in situ hybridization (FISH)
DNA sequence changes	Restriction fragment length polymorphism (RFLP) Inter-simple sequence repeats markers (ISSR) Random amplified polymorphic DNA (RAPD) Microsatellites or simple sequence repeat (SSR) Polymerase chain reaction–restriction fragment length polymorphism (PCR–RFLP) Amplified fragment length polymorphism (AFLP) Next generation sequencing of genomic DNA libraries
DNA sequence-specific changes	Inter-retrotransposon amplified polymorphism (IRAP) Sequence-specific amplification polymorphism (SSAP) Transposon display
RNA expression changes	Differential display Suppression subtractive hybridization EST array or Gene Chip hybridization cDNA macroarray hybridization Genome-scale microarray hybridization Differential display Suppression subtractive hybridization EST array or Gene Chip hybridization cDNA macroarray hybridization
Protein expression changes	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
Isozyme pattern changes	Starch gel electrophoresis and staining
Metabolite changes	Nuclear magnetic resonance spectroscopy (NMR)
DNA methylation changes	Methylation-sensitive Restriction fragment length polymorphism (metRFLP) Methylation-sensitive restriction fragment

	length polymorphism (metRFLP) Methylation-sensitive amplification polymorphism (MSAP) Based on treatment with the restriction enzyme McrBC, which only cleaves DNA at 5-methylcytosine DNA degradation followed by reversed-phase high-performance liquid chromatography (RP-HPLC) DNA degradation followed by high performance capillary electrophoresis (HPCE) Based on bisulfite treatment, which changes un-methylated cytosines into uracil, and sequencing Transposon methylation display (TMD)
Chromatin condensation	Fluorescent in situ hybridization (FISH)
Chromatin modification	Chromatin IMMUNOPRECIPITATION (ChIP) analysis
Small RNA expression changes	Cloning and northern analysis Cloning and qRT-PCR MicroRNA array hybridization Cloning and deep sequencing of small RNA libraries

Table 1.7. Strengths and weaknesses of different marker systems for the assessment of clonal fidelity (Krishna *et al.*, 2016).

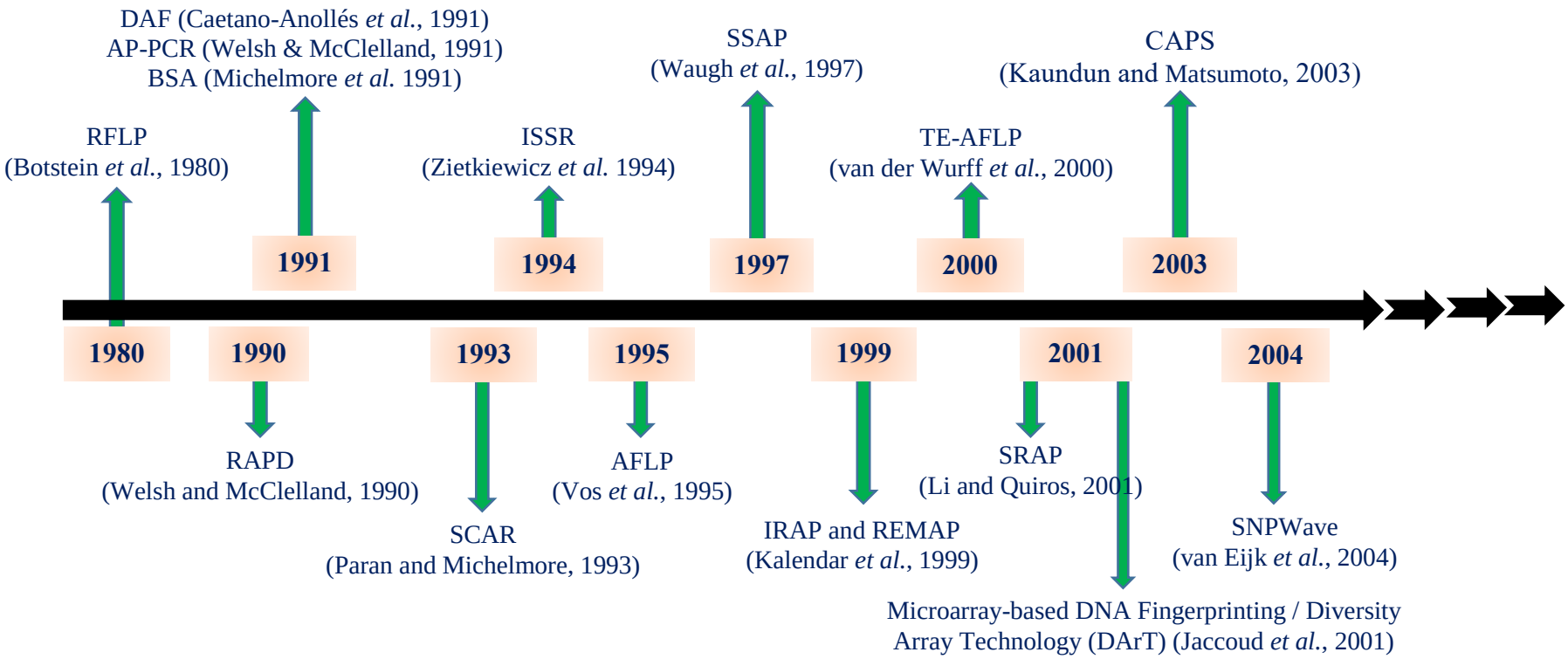
Advantages	Disadvantages
Morphological traits	
Visual differentiation Does not require any laboratory facility Suitable for preliminary detection	Sensitive to ontogenic changes and other environmental factors Limited in numbers Time-consuming
Cytological markers (flow-cytometry)	
Sample preparation and analysis is convenient and rapid in case of in flow-cytometry Rapid and efficient method for routine large-scale studies of ploidy level Unfailing detection of even the smallest modifications in chromosome number	Cytosolic compounds may interfere with quantitative DNA staining in flow-cytometry Absence of a set of internationally agreed DNA reference standards in case of in flow-cytometry Time-consuming chromosome counting
Isozyme markers	
Codominant expression Ease of performance	Sensitive to ontogenic changes and other environmental factors Limited in numbers Not all of these reagent systems work efficiently with all plant species Tissue-specific expression
DNA markers	
Codominant expression Any source DNA can be used for the analysis Phenotypically neutral Not sensitive to ontogenic changes and other environmental factors Capability to detect culture-induced variation both at the DNA sequence and methylation pattern levels	RAPD markers are dominant and do not permit the scoring of heterozygous individuals. Besides, they exclusively identify sequence changes Possible non-homology of similar sized fragments as ISSR is a multilocus technique Disadvantages of AFLPs include the need for purified, high molecular weight DNA, the dominance of alleles and the possible non-homology of comigrating fragments belonging to different loci Involvement of high development costs in SSR markers if adequate primer sequences for the crop species of interest are unavailable. Further, mutations in the primer annealing sites may result in the occurrence of null alleles (no amplification of the intended PCR product), which may lead to errors in scoring.

1.3.3. Molecular Tools applied in this study

1.3.3.1. Inter-simple sequence repeats (ISSR) markers for genetic studies

1.3.3.1.1. DNA fingerprinting and ISSR marker

DNA markers have proved valuable in crop breeding, especially in studies on genetic diversity and gene mapping. Broadly, DNA-based molecular markers can be categorized as hybridization-based or PCR-based (Sarwat *et al.*, 2011). The former employ restriction enzyme-digestion of DNA, followed by hybridization with a known labelled probe, which helps in identification of profile; while PCR-based markers involve primers to amplify DNA sequences (Sarwat *et al.*, 2011). A comprehensive review of most of the DNA-based markers is provided by a number of authors (Sarwat *et al.*, 2011; Gupta *et al.*, 2008; Jones *et al.*, 2009; Arif *et al.*, 2010). **Figure 1.17** shows a lineology or time-course study of some popular markers from their emergence till date. The commonly used polymerase chain reaction (PCR)-based DNA marker systems are random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and simple sequence repeats (SSRs) or microsatellites (Reddy *et al.*, 2002). The major limitations of these methods are low reproducibility of RAPD, higher cost of AFLP and the need to know the flanking sequences to develop species specific primers for SSR polymorphisms. The inter-simple sequence repeats (ISSR)-PCR is a technique that overcomes most of these limitations (Zietkiewicz *et al.*, 1994; Gupta *et al.*, 1994; Wu *et al.*, 1994; Meyer *et al.*, 1993). It has been used by the research community in various fields of plant improvements (Godwin *et al.*, 1997). The technique is useful in areas of genetic diversity, phylogenetic studies, gene tagging, genome mapping and evolutionary biology in a wide range of crop species (Sarwat *et al.*, 2011). In this method, SSR sequences are used as primers to amplify mainly the inter-SSR regions. SSRs or microsatellites are short tandem repeats (STRs) or variable number of tandem repeats (VNTRs) of 1 to 4 bases of DNA ubiquitously present in eukaryote genomes (Tautz and Renz, 1984). They are dispersed throughout the genome and vary in the number of repeat units.



.

Figure 1.17. Lineology of DNA-based markers (Time Clock) (Sarwat, 2012)**.** AFLP: Amplified fragment length polymorphism. AP-PCR: Arbitrarily primed PCR, BSA: Buld segregant analysis. CAPS: Cleaved amplified polymorphic sequences. DAF: DNA amplification fingerprinting. IRAP: Inter-retrotransposon amplified polymorphism. ISSR: Intersimple sequence repeats. RAPD: Random amplification of polymorphic DNA. REMAP: Retro-transposon-microsatellite amplified polymorphism. RFLP: Restriction fragment length polymorphism. SAMPL: Selectively amplified microsatellite polymorphic loci. SCAR: Sequence characterized amplified region. SRAP: Sequence-related amplified polymorphism. SSAP: Sequence-specific amplification polymorphism.

1.3.3.1.2. Principles of ISSR marker

Inter simple sequence repeat (ISSR) technique is a PCR based method, which involves amplification of DNA segment present at an amplifiable distance in between two identical microsatellite repeat regions oriented in opposite direction (**Figure 1.18**). The technique uses microsatellites, usually 16–25 bp long, as primers in a single primer PCR reaction targeting multiple genomic loci to amplify mainly the inter-SSR sequences of different sizes. The microsatellite repeats used as primers can be di-nucleotide, tri-nucleotide, tetra-nucleotide or penta-nucleotide. The primers used can be either unanchored (Gupta *et al.,* 1994; Meyer *et al.,* 1993; Wu *et al.*, 1994) or more usually anchored at 3' or 5' end with 1 to 4 degenerate bases extended into the flanking sequences (Zietkiewicz *et al.,* 1994). This technique combines most of the benefits of AFLP and microsatellite analysis with the universality of RAPD. ISSRs have high reproducibility possibly due to the use of longer primers (16 to 25 mers) as compared to RAPD primers (10- mers) which permits the subsequent use of high annealing temperature (45 to 60ºC) leading to higher stringency. The studies on reproducibility show that it is only the faintest bands that are not reproducible. About 92–95% of the scored fragments could be repeated across DNA samples of the same cultivar and across separate PCR runs when detected using polyacrylamide (Fang & Roose, 1997; Moreno *et al.*, 1998). As low as 10 ng of template DNA is needed for ISSR analysis. The annealing temperature depends on the GC content of the primer used and usually ranges from 45 to 65ºC.

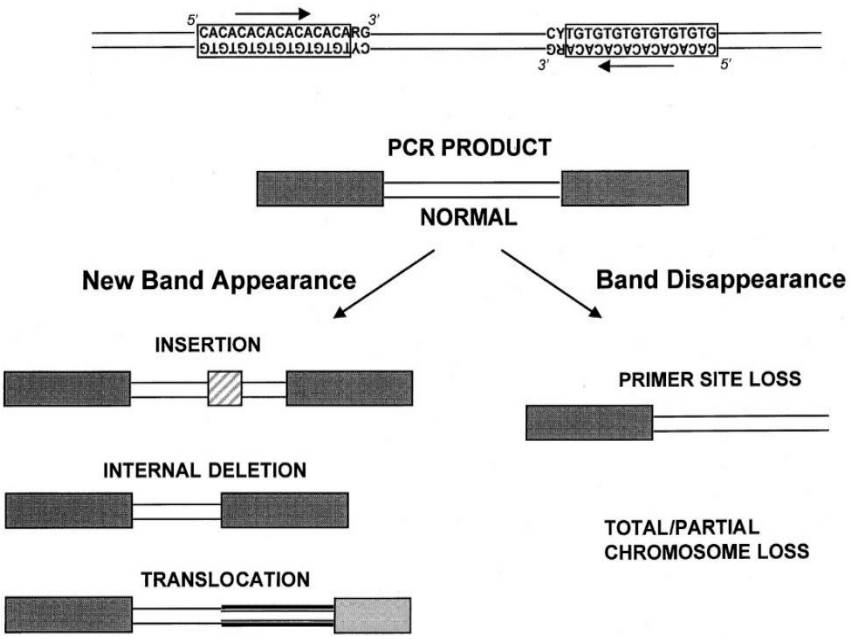


Figure 1.18. Schematic representation of the inter-simple sequence repeat (Inter-SSR-PCR or **ISSR-PCR)**. Adapted from Boland and Ricciardiello (1999). The PCR product is obtained by amplification with a single primer homologous to dinucleotide repeats and anchored at 3' by two nonrepetitive sequences (Boland and Ricciardiello, 1999). Genomic instability is deduced by the appearance or disappearance of bands when comparing PCR banding of normal and mutant samples (Boland and Ricciardiello, 1999). The new bands could be larger version of the original "normal" amplicon due to an insertion in the inter-SSR site, a smaller band due to a deletion or a rearrangement, creating a novel inter-SSR site (Boland and Ricciardiello, 1999).

ISSRs segregate mostly as dominant markers following simple Mendelian inheritance (Gupta *et al.*, 1994; Tsumura *et al.*, 1996; Ratnaparkhe *et al.*, 1998; Wang *et al.*, 1998). However, they have also been shown to segregate as co-dominant markers in some cases thus enabling distinction between homozygotes and heterozygotes (Wu *et al.*, 1994; Akagi *et al.*, 1996; Wang *et al.*, 1998; Sankar & Moore, 2001).

1.3.3.1.3. Applications of ISSR marker

ISSR technique is also very simple, fast, cost-effective, highly discriminative and reliable. At present, ISSR marker has been successfully applied to detect the genetic similarities or dissimilarities in micropropagated material in various plants. Molecular profiling using ISSR has recently been done to assess genetic stability of various plant species such as *Gerbera sp.* (Bhatia *et al.*, 2009), *Camellia sp.* (Devarumath *et al.*, 2002), *Pisum sativum* (Kuznetsova *et al.* 2005), *Codonopsis lanceolata* (Guo *et al.*, 2006), *Oryza sativa* (Ngezahayo *et al.*, 2007) and *Musa spp.* (Ray *et al.*, 2006).

1.3.3.2. Methylation-sensitive Amplified Polymorphisms (MSAPs) markers for epigenetic studies

1.3.3.2.1. Cytosine methylations in plants

The DNA of most eukaryotic organisms contains 5-methylcytosine (mC) residues, which not only represent important epigenetic information involved in the regulation of gene expression during various cellular developmental processes, such as cell differentiation, imprinting or chromosome inactivation (Rai *et al.*, 2010; Renfree *et al.*, 2013; Sharp *et al.*, 2011); but also contributes to genome stability by repressing harmful genetic elements (Yoder *et al.*, 1997). In human, epigenetic dysregulation may be changed by diet or stressful external conditions (Lim and Song, 2012; Solis *et al.*, 2012), causing the occurrence of human developmental diseases and cancer; while in plants, these epialleles-induced mutations (epimutations) may be induced by abiotic (Bednarek *et al.*, 2017) or biotic stresses (Dowen *et al.*, 2012), producing heritable phenotypic diversity to support adaptation and evolution (Durand *et al.*, 2012). In mammals, DNA is predominantly methylated in CG dinucleotides (Su *et al.*, 2011a, b; Ramsahoye *et al.*, 2000; Ichiyanagi *et al.*, 2013); however, in plant DNA, a cytosine in any sequence context can be methylated (Fulnecek *et al.*, 1998), with decreased frequency for CG, CHG, and CHHmotifs (H= A or T or C) (Cokus *et al.*, 2008).

To identify changes in DNA methylation connected with biological processes or different types of treatments, a convenient method for mC detection needs to be used. Recently, several genome-wide methods were developed to analyze mC (Lister and Ecker, 2009), based on bisulfite modification (Cokus *et al.*, 2008) and immunoprecipitation (Down *et al.*, 2008). However, as these methods are dependent on detailed pre-knowledge of the genome sequence, Methylation Sensitive Amplified Polymorphism (MSAP) are preferred in studies involving in non-model organisms, such as vanilla. MSAP approach is capable of quantifying events such as *de novo* methylation, demethylation, methylation and preservation of methylated status within *Hpa*II-*Msp*I restriction sites, by quantifying events affecting CG and CHG sequences and thus providing important insights into the role of epigenetics of abiotics or biotic stresses in plants (Bernarek *et al.*, 2017).

1.3.3.2.2. Principles of MSAP markers

The MSAP procedure was first introduced by Reyna-Lopez and co-workers (1997). In their modification of the original amplified fragment length polymorphism (AFLP) protocol (Vos *et al.*, 1995), frequent cutter MseI was replaced by methylation-sensitive MspI and HpaII restriction enzymes (**Figure 1.19**).

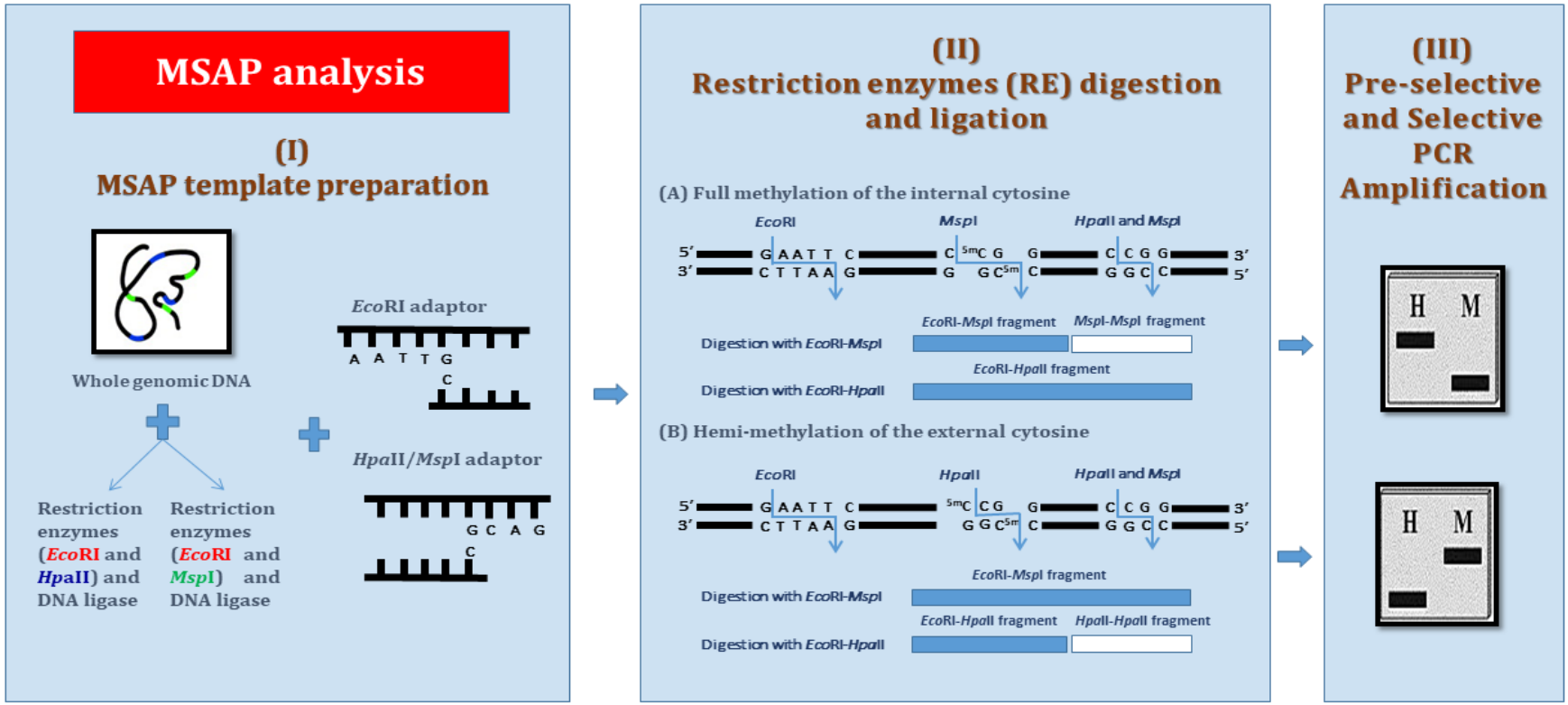


Figure 1.19. Schematic diagram of MSAP process (Fulneček and Kovařík, 2014)**.** Each purified genomic DNA sample is firstly digested with a rare-cutter (*EcoR*I) endonuclease (RE), and two combinations of frequent cutters in separate reactions: First combination with a methylation-sensitive (*Hpa* II), and second combination with a methylation-insensitive (*Msp* I) endonuclease, then the resulting DNA fragments are ligated to specific adaptors. Subsequently, the ligated DNA fragments are used as templates in a preselective PCR reaction using adaptor-specific primers. Followed, the resulting PCR products are then used as DNA template in a selective PCR reaction using one (+1), two (+2) or three (+3) selective

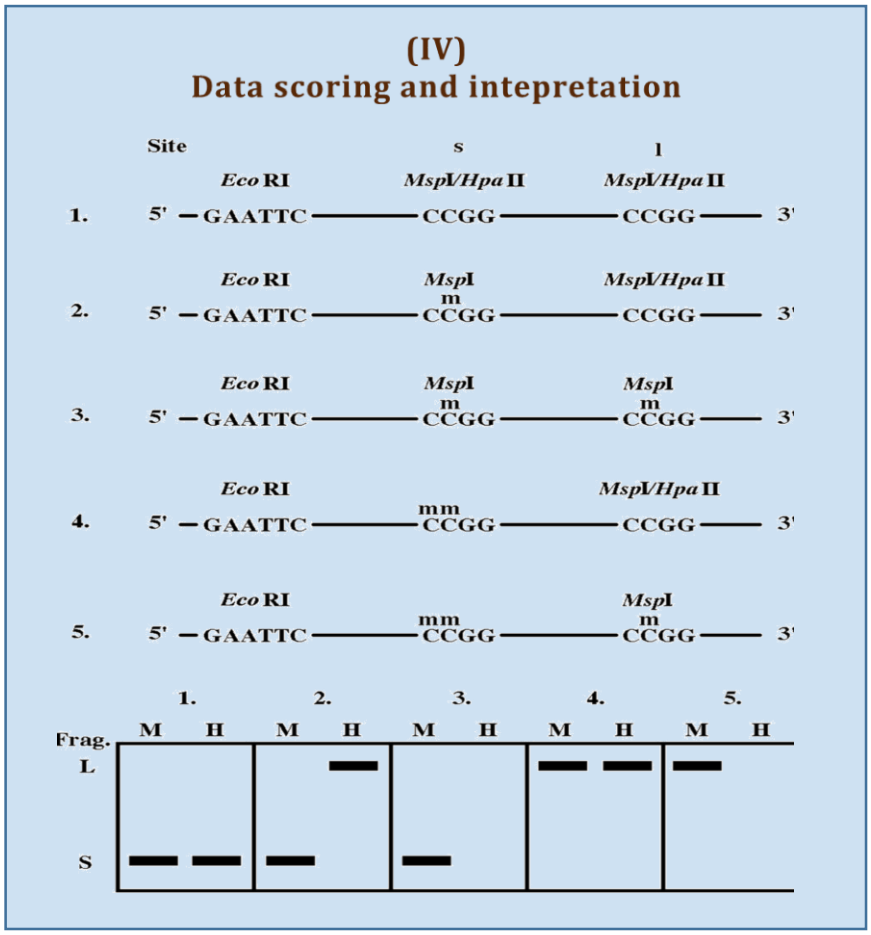


Figure 1.19. (continued) nucleotides. Finally, the selective PCR products are loaded into a gel system for scoring and data analysis. **(IV) Data scoring and interpretations:** A DNA methylation event was detected when bands present in the profile from the reaction *Eco*RI (E) + MspI (M) were absent from the reaction *Eco*RI + *Hpa*II (H) (Peraza-Echeverria *et al.*, 2001). This indicated that the internal cytosine was methylated (5'-Cm-CGG-3'). The contrary situation, where a band was present in H but absent in M, indicated that the external cytosine of one DNA strand was methylated (5'-^{m}CCGG-3'). This is regarded as the hemimethylated state. A DNA methylation event was considered to be a polymorphism when at least one plant differed from the others in the methylation pattern, for example, a band was present in both M and H in one plant but was only found in M in other plants. When, in one or some samples, no bands were detected in both digestions this was not considered as a polymorphism due to the fact that it is not possible to discard the presence of an AFLP event. Events in which there was an obvious difference in band intensity were also considered as DNA methylation polymorphisms The proportion of polymorphisms (*P*) was estimated by division of the number of DNA methylation polymorphisms by the total number of bands (Peraza-Echeverria *et al.*, 2001). At a particular locus, there are four possible types of MSAP profile (Type I to IV): Type I whereby AFLP bands were present in both *Eco*RI/*Hpa*II and *Eco*RI/*Msp*I lanes; type II whereby bands were present in *Eco*RI/*Msp*I lanes, but not in *Eco*RI/*Hpa*II lanes; and type III whereby bands were present in *Eco*RI/*Hpa*II, but not in *Eco*RI/ *Msp*I lanes.

1.3.3.2.3 Applications of MSAP for methylation analysis in PTC

The MSAP method is independent on the availability of genome sequence information and has been used frequently to analyze DNA methylation changes in plants and animals (Paun *et al.*, 2010; Xu *et al.*, 2012; Yang *et al.*¸ 2011). In *in vitro* culture systems, MSAP has been successfully applied in the detection of SV caused by epigenetics in various plant species, namely the bonnet bellflower (*Codonopsis lanceolate* Benth. et Hook. f) (Guo *et al.,* 2007), freesia (*Freesia hybrida* Klattl) (Gao *et al.*, 2010), wild barley (*Hordeum brevisubulatum* (Trin.) Link) (Li *et al.*, 2007), rice (*Oryza sativa* L.) (Wang *et al.*, 2013), dessert banana (*Musa acuminate* L.) (Peraza-Echeverria *et al.*, 2001; James *et al.*, 2004), grapevine (*Vitis vinifera*) (Baránek *et al.*, 2010), garlic (*Allium sativum* L.) (Gimenez *et al.*, 2016), cassava (*Manihot esculenta*) (Kitimu *et al.*, 2015) and moth orchids (*Doritaenopsis* sp.) (Park *et al.*, 2008).

1.3.3.3. Flow cytometry (FCM) for cytological studies

1.3.3.3.1. Principles of Flow cytometry (FCM)

First developed by Dillar and co-workers from the National Flow Cytometry Resource lab (USA) in 1969, flow cytometry (FCM) is a laser-based technology that could be utilized to measure characteristics of biological particles at whole cell or specific cellular constituent level, such as nuclei or organelles (Jaroszeski and Radcliff, 1999). It scans single particles or cells as they flow in a liquid medium past an excitation light source (Jaroszeski and Radcliff, 1999).

Flow cytometers can be described as four interrelated systems (**Figure 1.20**), the fluidic system which transports particles in samples through the instrument during FCM analysis; an illumination system for particle interrogation; the optical and electronics systems that direct, collect and translate scattered and fluorescent light signals of illuminated particles; and finally a data storage and computer control system which interprets and collates the translated light and electrical signals into meaningful data for storage and subsequent

analysis (Jaroszeski and Radcliff, 1999). The underlying principle of flow cytometry is that light is scattered and fluorescence are emitted as light from the excitation source strikes the moving particles (Jaroszeski and Radcliff, 1999). Light scattering and fluorescence is measured for each individual particle that passes the excitation source. Scattering and emission data can be used to examine a variety of biochemical, biophysical, and molecular aspects of particles (Jaroszeski and Radcliff, 1999).

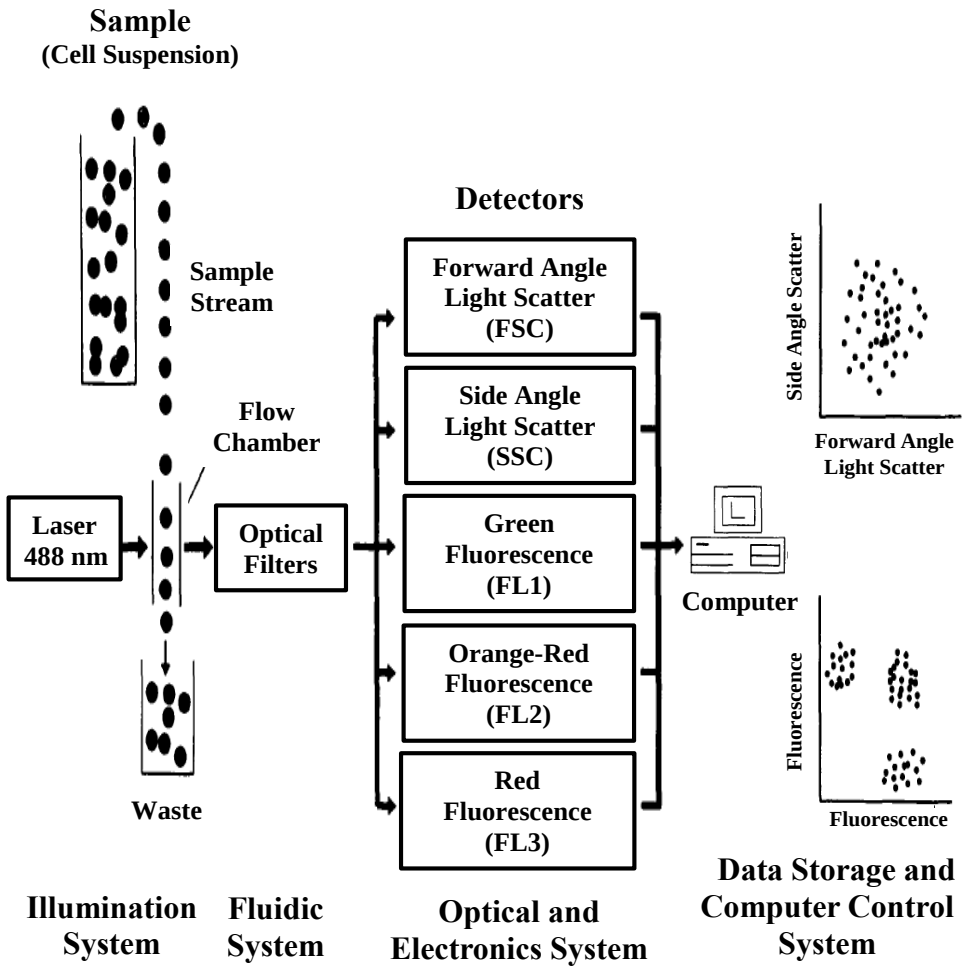


Figure 1.20. A schematic of the primary components that comprise a flow cytometer (Jaroszeski and Radcliff, 1999). Dark arrows indicated the flow of particles and information. Firstly, particles or cells were transported through a fluidic system transports and past a focused laser beam which is generated by an illumination system. Particle interrogation the took place, one cell at a time, in a flow chamber. The resulting scattered light and fluorescence are gathered by an optical and electronics system that then translates the light signals into information and were saved by the data storage and computer control system. After data from a sample have been stored, retrospective graphical data analysis could be performed with the use of flow cytometry analysis software.

Forward and side scattered lights, and the fluorescent signals were the main information collected in FCM. Single or multiple fluorescent could be detected by the use of different fluorescent channels (FL1 to FL4) (Jaroszeski and Radcliff, 1999). Forward scattered light is a result of diffraction, which provides basic morphological information, such as relative cell size, which is referred to as forward angle light scatter (FSC) (Jaroszeski and Radcliff, 1999). Light that is scattered at 90º to the incident beam is the result of refracted and reflected light (Jaroszeski and Radcliff, 1999). This type of light scatter is referred to as side angle light scatter (SSC), which is an indicator of granularity within the cytoplasm of cells, as well as surface/membrane irregularities or topographies (Jaroszeski and Radcliff, 1999). The correlation of FSC and SSC light signal measurements allows for the discrimination of various cellular subpopulations in a heterogeneous sample and identification of viable, less viable (dead or apoptotic cells), and necrotic cells (Jaroszeski and Radcliff, 1999).

1.3.3.3.2. Applications of FCM in plants

FCM is a powerful tool for many scientific disciplines, because it allows characterization of cells or particles within a sample (Jaroszeski and Radcliff, 1999). Because of its high-throughput, accuracy and resolution, negligible destructiveness to plants, as well as low operating cost per sample, FCM has replaced other methods of ploidy and/or genome size estimation in a large number of laboratories (Vrána *et al.*, 2014).

In plants, DNA flow cytometry has been applied in a vast array of applications ranging from basic research to breeding, including the use for nuclear genome size estimation (Hülgenhof *et al.*, 1988), ploidy screening (De Laat *et al.*, 1987), mixoploidy and aneuploidy detection (Roux *et al.*, 2001, 2003), assessment of the degree of polysomaty (Barow and Meister, 2003), cell cycle kinetics (Sandoval *et al.*, 2003) and reproductive pathways (Matzk *et al.*, 2000). In this research study, FCM is applied for the estimation of genome size and ploidy in SV screening of vanilla tissue culture materials, which can be especially unstable in its DNA content. Using FCM, the endoreduplication intensity in different plant organs and tissues can also be analysed.

Nuclear genome size could be estimated using plant FCM by the inclusion of nuclei from an internal standard with known genome size, which were isolated, stained and analysed simultaneously with the samples of interest (Dolezel, 1991). The absolute DNA amount of a sample could then be calculated based on the values of the G1 peak means and measured with the relative fluorescence intensity of the stained nuclei on a linear scale (Galbraith *et al.*, 1998).

Traditional chromosome counting method is not only laborious and prone to errors when chromosomes are small, but is also often limited to a relatively low number of cells analyzed (Bohanec, 2003). FCM, on the other hand, allows detailed information about the existence of mixoploid tissues and their proportions, in a much higher throughput (Bohanec, 2003). When optimized protocols are used, a high number of samples, 200 or more per day, can be accurately analyzed (Bohanec, 2003).

1.3.3.3.3. Endopolyploidy in plants

In the classical cell cycle, the nuclear DNA contents vary only within the range of 2C and 4C, where C is the haploid DNA content per nucleus (Yang and Loh, 2004). When mitotic DNA replication in somatic cells is not followed by cell division (a process called endoreduplication), variation of cellular ploidy levels (designated as somatic polyploidy or endopolyploidy) can result (Larkins *et al.*¸ 2001). Endoreduplication is defined as a process by which cells successively replicate their genomes without mitosis, resulting in an increase in nuclear DNA ploidy (Wildermuth *et al.*, 2017).

Various forms of endopolyploidy have been described across different eukaryotic and bacterial lineages (Leitch and Dodsworth, 2017). In plants, the occurrence of endomitotic cycle (**Figure 1.21**) is rare, and the most common forms of endopolyploidy are endocycles and partial endocycles (Leitch and Dodsworth, 2017). In a normal mitotic cell cycle, the common cell cycle progression is the G1-S-G2-M cycle. The mother cell prepares itself during the G1 and G2 phases, replicates its DNA during the DNA synthesis (S) phase, and

splits the DNA equally between two daughter cells during the mitosis (M) phrase. In an endoreplication cycle, the cell cycle progression consists of only the G and S phrases, and therefore the cells just cycle between the S and gap (G) phases, leading to a doubling of the DNA amount after each complete endocycle (e.g. 2C → 4C → 8C → 16C) (**Figure 1.21**) and generating polyploidy cells with multiple copies of the original DNA content (Shu *et al* .2018). In a partial endocycle (also called progressively partial endoreduplication or PPE), only part of the genome (referred to as P and corresponding to the amount of DNA in the replicated part of the 2C genome) is replicated during the S phase (Leitch and Dodsworth, 2017). This leads to cells containing less than double the amount of DNA at the end of each endocycle and with successive DNA amounts of each PPE increasing from 2C → 2C+P → 2C+3P →2C+7P, and so on (Bory *et al.*, 2008; Leitch and Dodsworth, 2017). In plants, PPE has so far only been reported in orchids (*Orchidaceae*), where it has been identified in representatives of all subfamilies except *Apostasioideae* (Trávníček *et al.*, 2015; Leitch and Dodsworth, 2017).

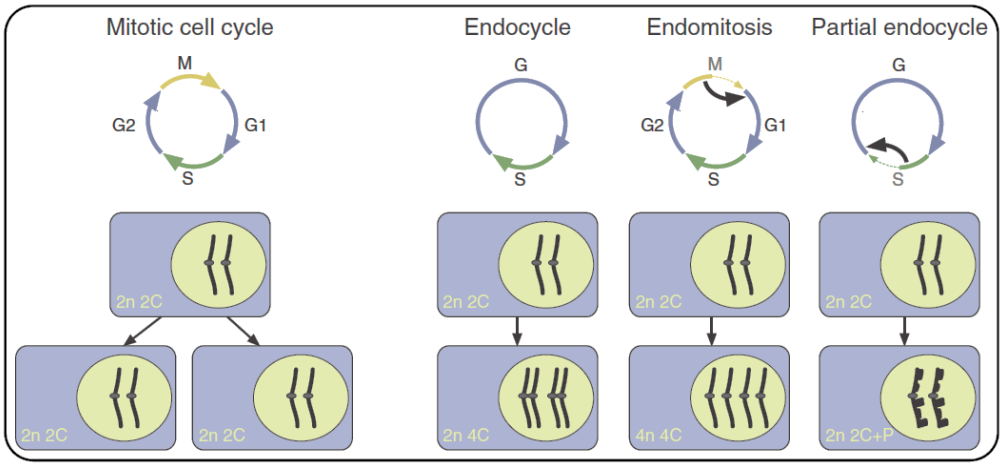


Figure 1.21. Endopolyploidy in plants. The canonical mitotic cell cycle is shown on the left, which includes different phases: synthesis (S phase) where DNA is replicated and mitosis (M phase) where chromosomes segregate and cells divide to produce two daughter cells, with each preceded by a gap phase (G1 or G2) preceding each. In the endocycle, mitosis (M phase) is absent, resulting in one G phase and S phase and a doubling of DNA (4C) per endocycle but not chromosome number (2n). In endomitosis, cells enter mitosis and chromosomes segregate but exit before cell division (i.e. partial M phase), resulting in a doubling of chromosome number as well as DNA (4C, 4n). Finally, in partial endocycles (also known as partial progressive endoreduplication –

PPE), cells skip M phase, but during S phase, they only replicate specific parts of the chromosomal DNA, resulting in cells with partial increases in DNA content (2C + P where P is the part of the 2C genome that has been replicated). Adapted from Leitch and Dodsworth (2017).

Endopolyploidy were found to be correlated to the plant morphogenesis in cell culture (Yang and Loh, 2004). In tobacco, it was reported that the morphogenetic response of the tissues culture was related to the nuclear DNA content variation within stem explants of different ages (Altamura *et al.*¸1987). In *Oncidium* Gower Ramsey, a hybrid orchid, only the root tips, cut surfaces of stem segments and young leaves were able to form callus in tissue culture, while other explants, such as old leaves and the roots without meristem tips, could not form any callus (Chen and Chang, 2000). Molecular data also showed that the nuclear DNA modulation was closely related to the acquisition of embryogenic competence in cultured carrot hypocotyls (Geri *et al.*, 1999). In cabbage plants, the number of endocycles was found to be tissue-specific and was characteristic of the developmental stage (Kudo and Kimura, 2001a, 2001b, 2002). These studies suggested that pattern of endopolyploidy may represent the characteristic of the developmental and physiological properties of the tissue. Endopolyploid cells were found also in protocorms, young leaves and root tips from *S. plicata* seedlings in tissue culture (Yang and Loh, 2004).

1.3.3.4. Transcriptome profiling

Transcriptome profiling represents the profile of a complete set of RNA molecules from any genome at any time or condition, which plays essential roles in several biological processes, including untranslated RNA species such as microRNAs (Pereira *et al.*, 2017).

1.3.3.4.1. Next-generation transcriptome sequencing (NGS)

Next generation transcriptome sequencing (or RNA-seq) is the state-of-art technique for transcriptomic studies, which can be applied for studies of both species with (reference-based) and without (*de novo*) a reference genome. In addition, it allows the global detection of novel genes and isoforms, genomic

alterations, splice and chimeric variants, gene fusions and gene expression quantification (Conesa *et al.*, 2016).

Although sequencing technologies has become more affordable (**Figure 1.22**), the studies require careful considerations of various experimental aspects in order to choose the best approach given a specific biological question, limited budget or challenging sample material (da Fonseca *et al.*, 2016). Researchers should also be aware of the assumptions and limitations of each method used (da Fonseca *et al*, 2016).

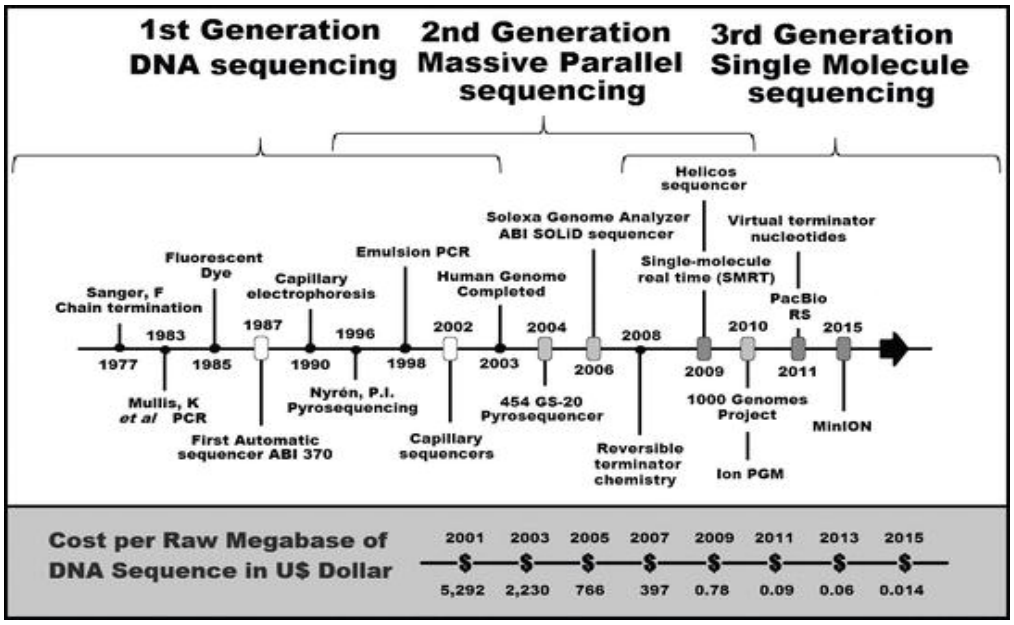


Figure 1.22. Timeline of DNA sequencing technology evolution, from Sanger to NGS, and the cost per raw megabase of DNA sequenced. Adapted from Pereira *et al.* (2017). Dots: milestones. Rectangles: Sequencing equipment - white: first-generation sequencing; light grey: second-generation sequencing; dark grey: third-generation sequencing.

Extensive reviews on the experimental considerations for carrying out an RNA-seq experiment, especially for non-model organisms, have been discussed by various authors (da Fonseca *et al.*, 2016; Ungaro *et al.*, 2017; Conesa *et al.*, 2016). A typical RNA-seq experiment consists of a good experimental design, cost estimation, sample preparation, library construction, sequencing and data analysis (Kukurba and Montgomery, 2015). Parameters to consider before beginning an RNA-seq experiment include the number and type of replicates (technical or biological replicates), sequencing platform (such as Illumina or Ion

Torrent), library preparation method (such as rRNA depletion or mRNA enrichment; strand-specific or not; single or paired end), throughput, read length, sequencing depth and coverage (Conesa *et al.*, 2016).

Principles of Illumina sequencing platform

Illumina is one of the second-generation sequencing platforms developed by Bentley and co-workers (2008), after the Roche 454 (Weller *et al.*, 2008). It consists of binding the DNA in a flow-cell through adapters, followed by parallel massive amplification that occurs in clusters for each DNA strand that was originally bound in the flow-cell, called the bridge-amplification (**Figure 1.23**) (Pereira *et al.*, 2017). The Illumina sequencing process generates paired-ends sequences, which are an advantage over other sequencing methodologies, as they improve the accuracy of mapping, mainly in repetitive regions or where DNA rearrangements or gene fusions occur. It also uses the 'reversible terminator chemistry', which is a modified fluorescent dNTPs that reversibly block DNA synthesis, thus the addition of each nucleotide can be synchronized and monitored by a charge-coupled device (CDD) sensor (Bentley *et al.*, 2008). It is one of the most accurate and with lowest error rate of sequencing methodologies used at present. However, it generally requires higher DNA concentration (Pereira *et al.*, 2017).

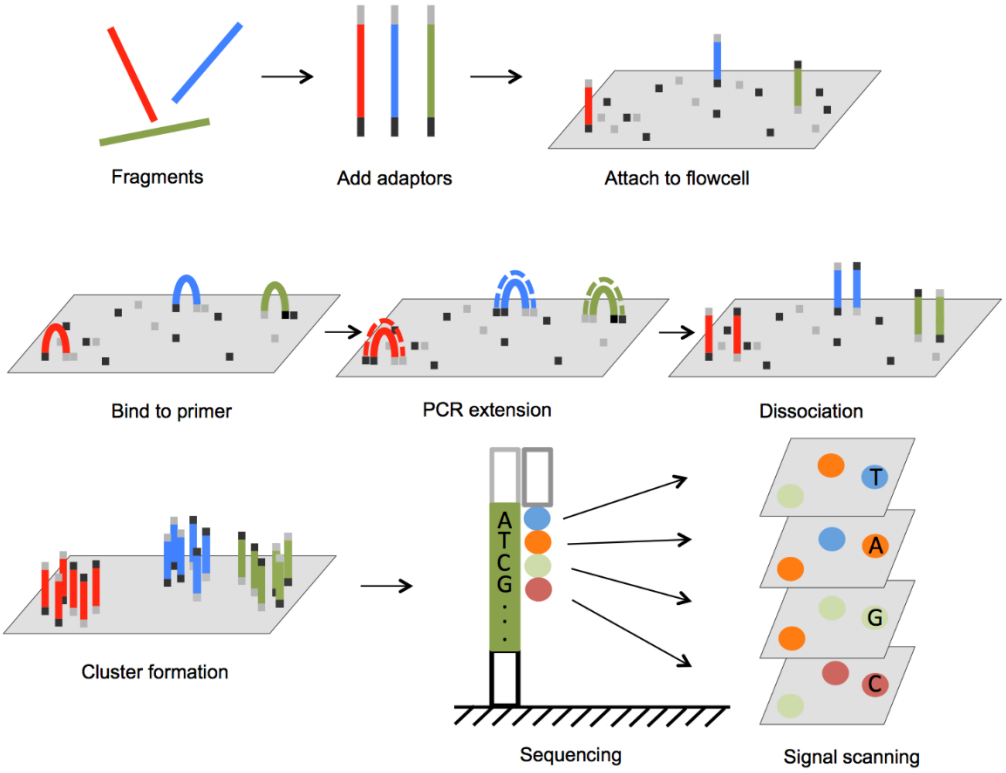


Figure 1.23. Outline of Illumina sequencing by synthesis (SBS) process. Adapter from Lu *et al.* (2016). Firstly, adaptors are annealed to the ends of sequence fragments. Secondly, fragments bind to primer-loaded flow cell and bridge PCR reactions amplify each bound fragment to produce clusters of fragments. Then, during each sequencing cycle, one fluorophore attached nucleotide is added to the growing strands. Laser excites the fluorophores in all the fragments that are then being sequenced and an optic scanner collects the signals from each fragment cluster. Followed which, the sequencing terminator is removed and the next sequencing cycle starts.

The Illumina sequencers have been modified and improved over the years. There are several different models of Illumina sequencers in the market, which includes the benchtop sequencers (such as the iSeq 100 system, MiniSeq system, MiSeq series and NextSeq Series), and the production-scale sequencers (namely the NextSeq Series, HiSeq Series, HiSeq X Series and NovaSeq 6000 system), each differing in its run time, maximum output, maximum reads per run and maximum read length, for the use of different applications, such as large (human, plant, animal) or small (microbes and viruses) whole-genome sequencing, exome sequencing, target gene sequencing (amplicon and gene panel), whole transcriptome sequencing, miRNA and small RNA analysis,

DNA-protein analysis, methylation sequencing and metagenomics sequencing (Illumina, 2017; Fonseca *et al.*, 2017). In this study, the NextSeq500 high throughput sequencer was employed for both mRNA and miRNA sequencing of vanilla.

1.3.3.4.2. Reverse transcription quantitative real-time PCR (RT-qPCR)

1.3.3.4.1.1. Nomenclature and standardization of terminologies

Terminologies used in this study followed the guidelines as proposed in the Minimum Information for Publication of Quantitative Real Time PCR Experiments (MIQE) (Bustin *et al.*, 2009), such as the abbreviation of qPCR used for quantitative real-time PCR, RT-qPCR used for reverse transcription-qPCR and the RT-PCR referred as its conventional meaning, which was the reverse transcription-PCR. Genes used for normalization were referred to as *reference genes*, instead of *housekeeping genes* (Bustin *et al.*, 2009). The *quantification cycle* (Cq), which were according to the RDML (Real-Time PCR Data Markup Language) data standard (http://www.rdml.org), instead of threshold cycle (Ct), crossing point (Cp) or take-off point (TOP) used by competing manufacteres of real-time instruments, were used to describe the fractional PCR cycle for quantification (Bustin *et al.*, 2009; Lefever *et al.*, 2009).

1.3.3.4.1.2. Principles of RT-qPCR: mRNA and miRNA

RT-PCR approaches: Conventional (mRNA) vs ligation-based (miRNA)

In the conventional two-step cDNA synthesis, RT-PCR uses RNA as the starting material. RNA is first transcribed into complementary DNA (cDNA) by reverse transcriptase from total RNA or messenger RNA (mRNA) (Bustin, 2005). Subsequently, the cDNA is used as a template for the quantitative polymerase chain reaction (qPCR) reaction (Bustin, 2005) (**Figure 1.24)**. The RT of plant miRNAs, differs from the conventional RT approach, consists of an additional pre-ligation step of the miRNAs with pre-adenylated adaptor. This is because

plant microRNAs (miRNAs), differed from animal microRNAs, possess a 2'-O-methyl group on the ribose of the 3' terminal nucleotide, which functions to protect the mature microRNAs from uridylation and truncation that leads to degradation (Yu *et al.*, 2005; Yu and Chen, 2010; Zhao *et al.*, 2012). Thus, in this study, the ligation-mediated RT approach is applied for the RT of miRNAs, instead of the conventional cDNA synthesis approach used for RT of mRNAs.

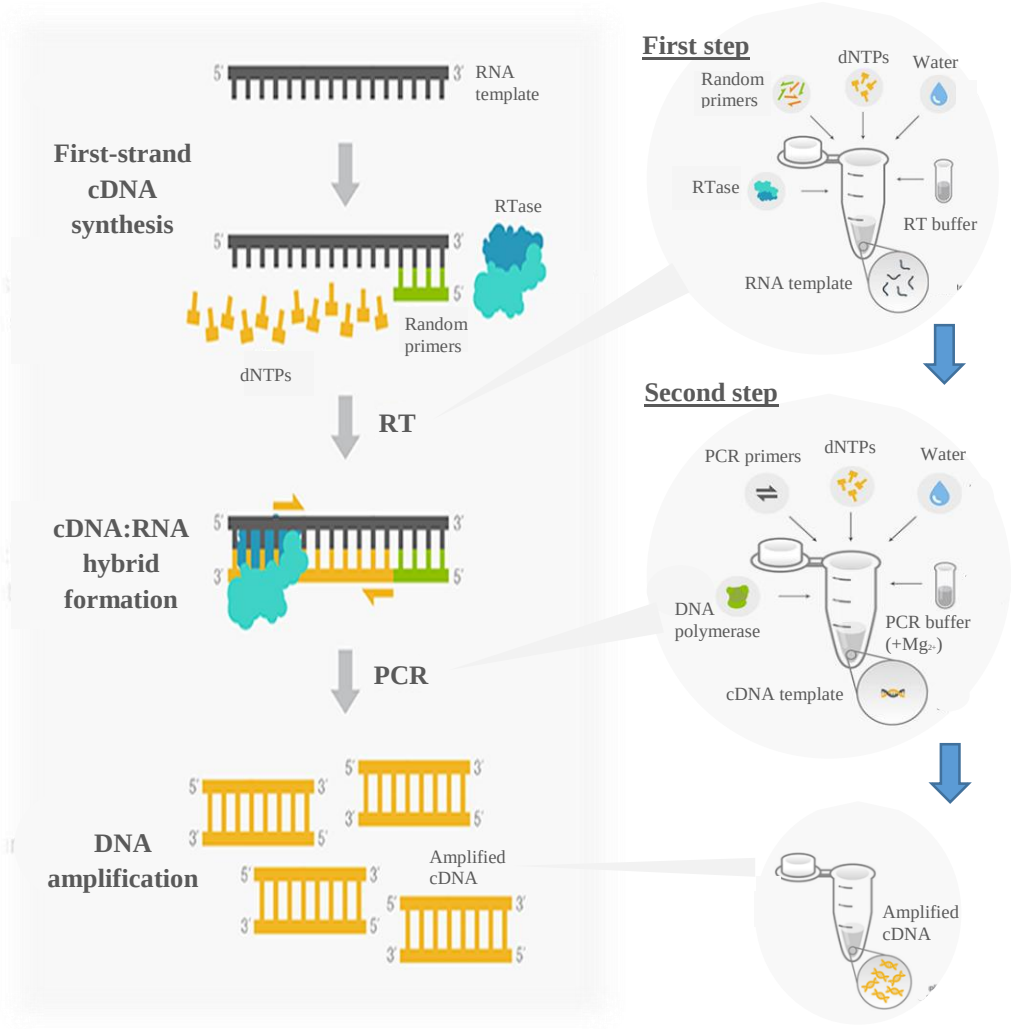


Figure 1.24. Two-step reverse transcription polymerase chain reaction (RT-PCR). Adapted and modified from Thermo Scientific (2017). RT: Reverse transcription; RTase: Reverse transcriptase.

Quantitative detection of qPCR

After reverse transcription, the cDNA library is quantitatively detected in a qPCR system. The quantitative endpoint for real-time PCR is the threshold cycle

(C_T) (Schmitgen and Livak, 2008). The C_T is defined as the PCR cycle at which the fluorescent signal of the reporter dye crosses an arbitrarily placed threshold (**Figure 1.25**) (Schmitgen and Livak, 2008). The quantitative detection using qPCR can be done probe-based or dye-based. Probe-based detection includes hydrolysis probes, hybridization probes, molecular beacons or Scorpions™; while the dye-based detection includes SYBR®, YOYO™, BOXTO, Amplifluor™, quencher-labelled primers or LUM™ primers (Wong *et al.*, 2015; Sigma-Aldrich, n.). In this study, the SYBR® dye-based DNA binding detection method was used. The SYBR® dyes bind non-specifically to double-stranded DNAs (Qiagen, 2017).

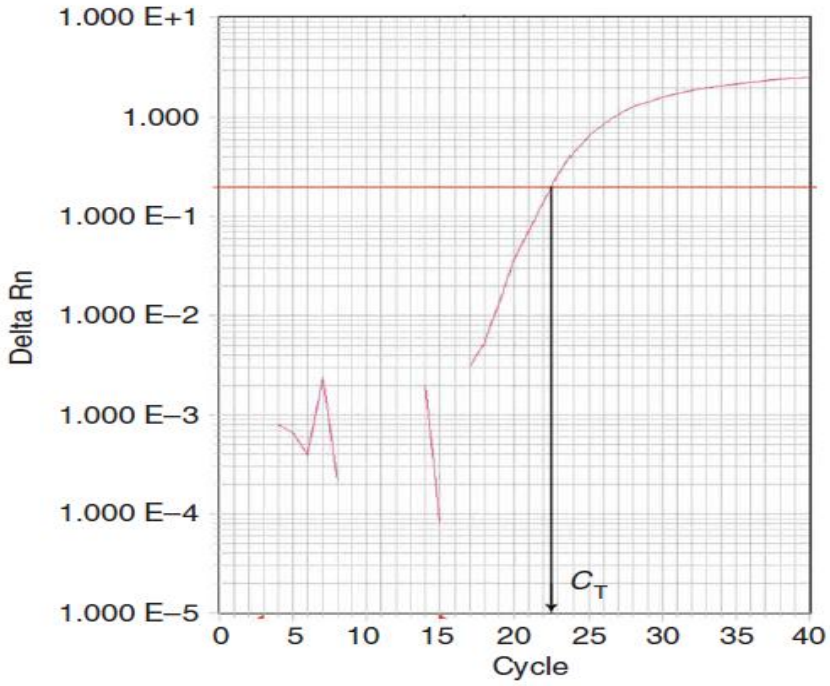


Figure 1.25. Real-time PCR output and calculation of C_t. Adapted from Schmitgen and Livak (2008). By presenting data as the C_T, one ensures that the PCR is in the exponential phase of amplification (Schmitgen and Livak, 2008). The numerical value of the C_T is inversely related to the amount of amplicon in the reaction (i.e., the lower the C_T, the greater the amount of amplicon) (Schmitgen and Livak, 2008).

There are a few important considerations to take note of during quantitative detection of qPCR, namely the determination of primer specificity and the selection of suitable reference genes (for relative quantification method).

Determination of primer specificity

The specificity of primers can be indicated through the analysis of a melt (or dissociation) curve (**Figure 1.26A)**: the final PCR product is denatured at a temperature gradient from 50 to 95ºC to detect the crease in fluorescence as the dye dissociates (Qiagen, 2017). As a result, PCR amplicons of different lengths and sequences would melt at different temperatures, resulting in the observation of distinct melt peaks (**Figure 1.26B**) (Pryor and Witter, 2006).

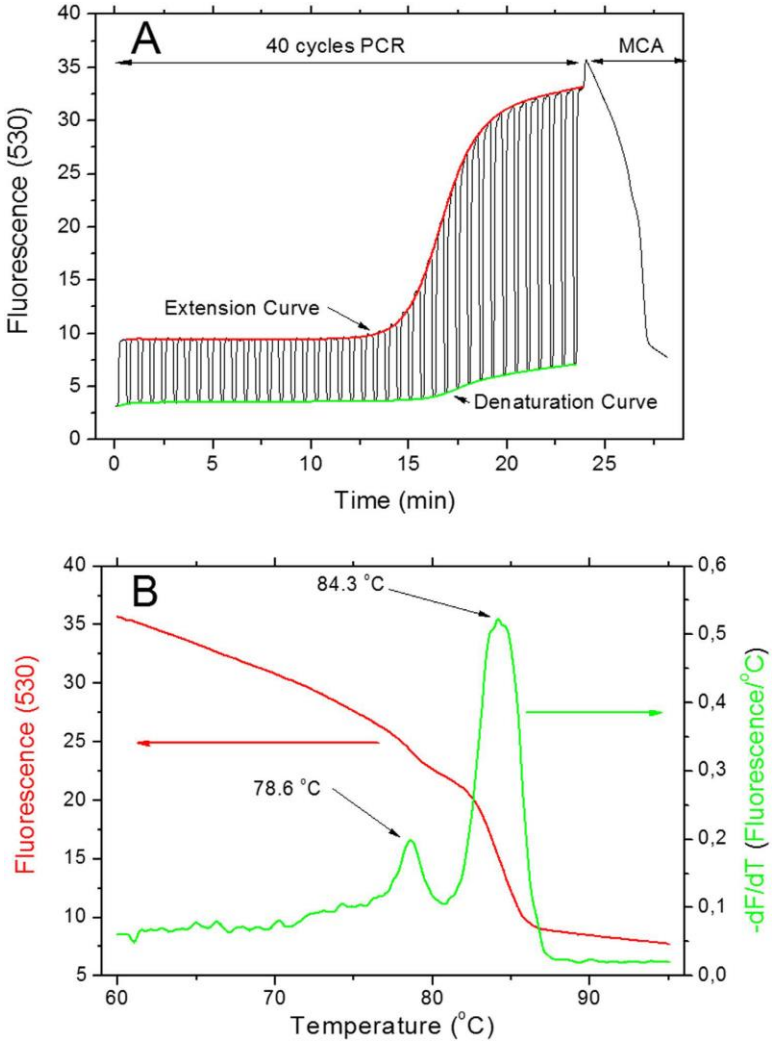


Figure 1.26. Melt Curve Analysis (MCA) of qPCR. Adapted from Ahberg and Neužil (2015): (A) Amplification curve from a typical 40-cycle PCR data set. Fluorescence amplitude was recorded at the end of the extension (**red**) and denaturation (**green**) step, as shown. The normalized amplification curves are shown as inset. At the end of the PCR, MCA was performed. (B) Results of the MCA analysis showing DNA templates with two different melting temperatures, 78.6ºC and 84.3ºC.

Normalization with Reference genes

In a relative quantification method, the RT-qPCR data needs to be normalized with suitable internal reference gene(s) for an accurate and reliable analysis of target gene expression (Czechowski *et al.* 2005). Normalization is an essential process to correct non-specific variations arising due to the differences in template quality and amount used that can thus affect the RT-qPCR amplification efficacy. Besides, normalization also allows the direct comparison of normalized transcript expression levels between samples (Maroufi *et al.* 2010).

There are various softwares available for the validation of reference gene stability (Vandersompele *et al.*, 2009), such as geNorm (Vandesompele 2002); BestKeeper (Pffafl *et al.*, 2004) and Normfinder (Andersen *et al.*, 2004). In geNorm, reference genes are ranked based on stepwise elimination of least stable gene and the expressions stability is defined as average pairwise variation (standard deviation of log transformed ratios) of a given gene with all other candidate reference genes (Vandersompele *et al.*, 2009). In Bestkeeper, the reference genes are ranked based on repeated pairwise correlation and regression analysis of a given gene with all other tested candidate reference genes (Vandersompele *et al.*, 2009). In Normfinder, statistical linear mixed-effects modelling is used (Vandersompele *et al.*, 2009).

1.3.3.4.1.3. Applications of RT-qPCR

RT-qPCR has been widely applied in both research and industrial uses of different fields, such as clinical diagnostic or risk assessments of infectious diseases (Espy *et al.*, 2006; Mackay, 2004; Dreier *et al.*, 2007), cancer (Aldape *et al.*, 2002; Mocellin *et al.*, 2003; Bernard and Wittwer, 2002) and genetic disorders (Mahdieh and Rabbani, 2013), environmental remediation and food safety test (Fillion, 2012), screening for phytopathogen diseases (Abd-Elsalam, 2003; Tomlinson *et al.*, 2007), detection of exogenous genes in genetically modified organisms (GMOs) (Holst-Jensen *et al.*, 2003), single-cell gene expression analysis or dianogsis (Ståhlberg and Bengtsson, 2010; Ok *et al.*, 2016) and forensic genotyping (Alonso and Fernándex, 2007; Ricci *et al.*, 2006).

1.3.3.4.3. Plant microRNAs: Principles and bioinformatics

1.3.3.4.3.1. Definitions, databases and nomenclature

MicroRNAs (miRNAs) are a class of single-stranded RNAs of 19 to 25 nucleotides in length that are generated by the precise excision of the initial duplex from the stem of a stem-loop precursor RNA (Axtell and Meyers, 2018). The initial duplex consists of the strand that will become the functional miRNA and a partner that is less frequently bound to an AGO protein, the miRNA* ("microRNA-star") (Axtell and Meyers, 2018). Plant miRNAs generally target protein-coding mRNAs or long, non-coding RNAs. In plants, RNA targets of miRNAs are repressed, a result of either or both mRNA destabilization and translational repression (Axtell and Meyers, 2018). Plant miRNAs were first identified in early 2002 (Llave *et al.*, 2002a; Park *et al.*, 2002; Reinhart *et al.*, 2002). Like their animal counterparts, they are short sequences mostly 21 nt long, processed from longer precursor sequences. While animal precursor sequences usually have a length of 70–80 nt, plant miRNA precursor sequences are much more variable and range from 50 nt to more than 350 nt.

1.3.3.4.3.2. Biogenesis and functions of miRNAs in plants

The biogenesis of miRNA sequences has been largely elucidated. The mature miRNA (often designated as miR) is transcribed from a characteristic stem–loop sequence (called as pre-miR) by RNA polymerase II (RNA polII), which in turn may be excised from a longer primary transcript (or pri-miR) (**Figure 1.27**) (Xu *et al.*, 2019).

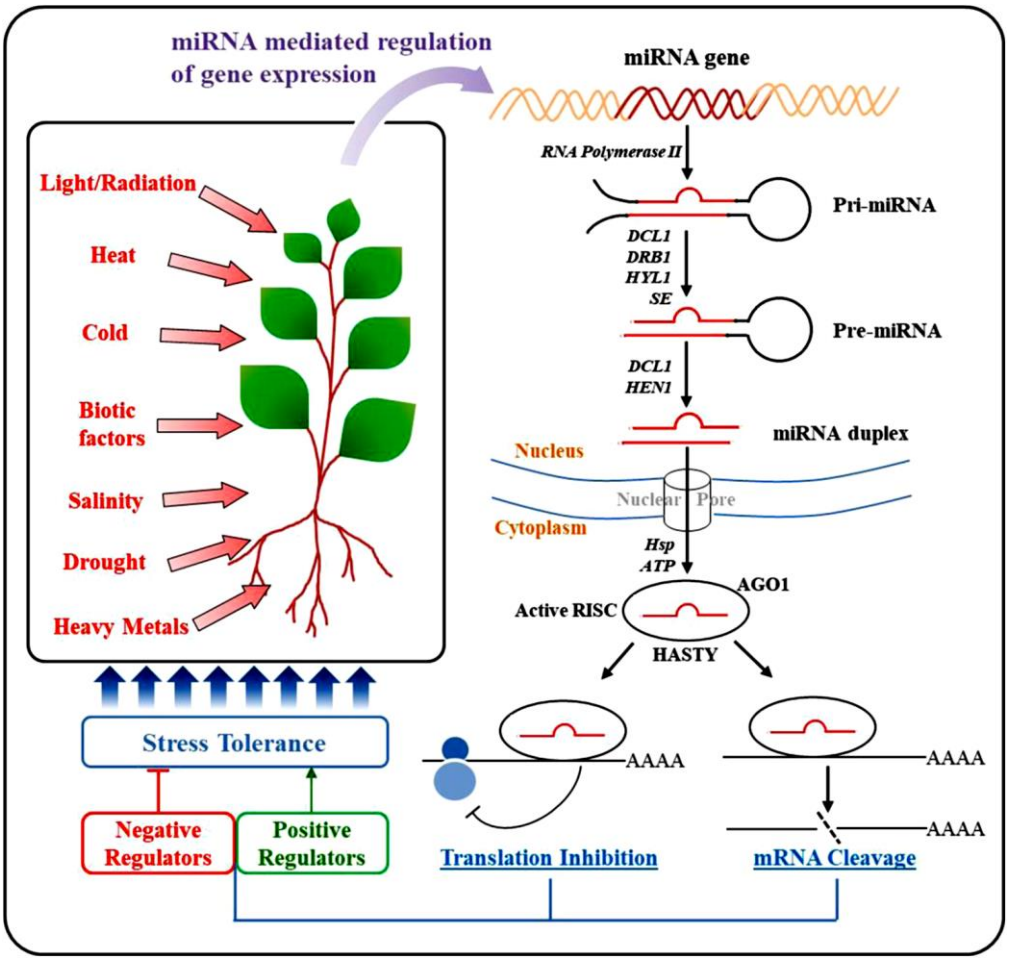


Figure 1.27. The miRNA biogenesis and general mechanism of miRNA-mediated regulation of stress-responsive gene expression in plants to attain stress tolerance. DCL1: Dicer-Like 1 ribonuclease, DBR1: Double-Stranded RNA Binding-1, HYL1: Hyponastic Leaves-1, SE: Serrate, HEN1: Hua-Enhancer1, Hsp: Heat shock protein, AGO1: Argonaute, RISC: RNA-induced silencing complex. The plant miRNA transcripts have characteristic features of RNA polII activity including 5'-methylated cap and 3'-poly(A) tail. The local hairpin structure within pri-miR is recognized by DCL1/HYL1/Serrate proteins to generate precursor miRNA (pre-miR). DCL1 acts on pre-miR and release miRNA duplexes. These duplexes are modified by HEN1 that methylates 3'-OH residues. HASTY, a homolog of Exportin-5, exports out modified miRNA duplexes to cytoplasm. These duplexes further recruit several proteins, including AGO1, to form a functional miRNA-protein complex termed miRNA-induced silencing complex (mi-RISC). The mi-RISC assembly binds to target transcripts and leads to either cleavage or translational repression. The single stranded miRNAs that lack any methylation or uridylation are degraded by SDN1 protein. Adapted from Xu *et al.* (2019) and Naqvi *et al.* (2012).

Plant miRNAs play important roles not only in the growth and developmental processes, but also in the plant response to environmental

conditions (Xu *et al.*, 2019). It has been experimentally proven that small RNAs (including miRNAs) are involved in the signalling of regulatory-signalling events for plant growth regulators which profoundly affects overall growth and development of plants (Xu *et al.*, 2019). Apart from growth and development, increasing evidences have shown that the regulation of miRNA levels can reprogram the responses of plants to their dynamic environments with biotic and abiotic stresses (**Figure 1.27**) (Brant and Budak, 2018; Nogoy *et al.*, 2018). Numerous miRNAs have been characterized as responsive to biotic (Islam *et al.*, 2018) and abiotic stresses including drought (Zhang *et al.*, 2018), salinity (Feng *et al.*, 2017), nutrient deficiency (Yu *et al.*, 2018), cold (Megha *et al.*, 2018), heat (Li *et al.*, 2017a, b), exposure to UV radiation (Casadevall *et al.*, 2013; Dotto *et al.*, 2018), metal stress (Gupta *et al.*, 2014), sulfur dioxide (SO_2) stress (Li *et al.*, 2016a, b), continuous light (Matsunaga *et al.*, 2015; Ragupathy *et al.*, 2016), elevated CO_2 levels (May *et al.*, 2013), and oxidative stress (Iyer *et al.*, 2012) from different plant species.

1.4.3.4.3.2. Principles of bioinformatics for plant microRNAs

Target prediction of plant miRNAs

Nucleotide pairing between plant miRNAs and target mRNAs leads to target cleavage (**Figure 1.28**) (Mallory and Bouché, 2008). Target prediction of plant miRNA can be carried out in two approaches. The first effective approach was the computational prediction based on the perfect or near-perfect complementarity properties of miRNA to their targets (Rhoades *et al.*, 2002). In this approach, a sequence pattern complementary search for transcript targets is carried out for a given miRNA sequence, allowing few mismatches, which are usually two to three mismatches (Bonnet *et al.*, 2006). Jones-Rhoades and Bartel (2004) developed a sensitive procedure for identifying authentic targets, through a penalty scoring systems that takes into account the formation of bulged (G:U base pairs), gaps and mismatched nucleotides when the targets are base-paired with their miRNAs. In this scoring system, mismatched patterns are detected in the miRNA:mRNA duplexes within 20-base sequence window, with 0 points being assigned to each complementary pair, 0.5 points being assigned to each

complementary pair, 0.5 points to each G:U wooble, 1 point to each non-G:U wooble mismatch, and 2 points to each bulged nucleotide in either RNA strand (Jones-Rhoades and Bartel, 2004). In general, a score of 3.0 or less was demonstrated for the identification of authentic targets with very high confidence (Jones-Rhoades and Bartel, 2004) This scoring scheme was implemented in this study with *psRobot* module (Wu *et al.*, 2012) for identifying miRNA targets in transcript sequences obtained from transcriptome sequencing. Many miRNA computational prediction tools integrate target detection to support the result of miRNA prediction, of which the predicted targets have been verified experimentally (**Table 1.8**).

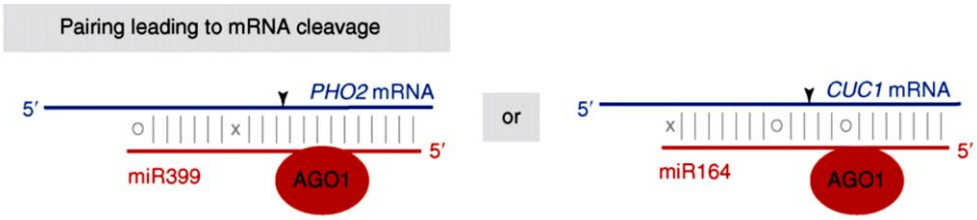


Figure 1.28. Nucleotide pairing between plant miRNAs and target mRNAs. Figure shows the examples of representative pairings (*miR399-PHO2* and *miR164-CUC1*) that occur between miRNAs (**red**) and mRNAs (**blue**) which lead to the AGO1-mediated cleavage (circle indicate the G-U woobles). Xs indicate nucleotide mismatches, and bulge nucleotides are indicated on the mRNA strand. Black arrowheads denote the position of cleavage. The miRNA complementary sites are located in the ORFs of their targets, but a few target sites have been reported in the 5' and 3' UTR sequences (Jones-Rhoades and Bartel, 2004). Adapted from Mallory and Bouché (2008).

Table 1.8. Functions of conserved plant microRNAs (miRNAs) for which computational prediction and experimental evidence have been published. Adapted from Bonnet *et al.* (2006) and Griffiths-Jones *et al.* (2006).

.

Function	miRNA families	Target gene(s)	miRNA conservation	References
Floral timing and leaf development	156	SPL transcription factors (TF)	A G O P Sb So	Kasschau *et al.* (2003)
Floral development	171	SCL-like TFs	A O P S	Llave *et al.* (2002b); Schwab *et al.* (2005)
Floral development and vegetative phase change	172	AP2-like TFs (AP2, TOE1, TOE2, GL15)	A G O P Sb Z	Aukerman and Sakai (2003); Chen (2004); Lauter *et al.* (2005); Kasschau *et al.* (2003);
Expression of auxin response genes, developmental defects	160	160: Auxin response factors (ARF10, ARF16, ARF17)	160: A G M O P Sb Z	Jones-Rhoades and Bartel (2004); Guo *et al.* (2005), Mallory *et al.* (2005); Wang *et al.*
	164	164: NAC1	164: A O P Sb Z	(2005); Kasschau *et al.* (2003); Allen *et al.*
	167	167: ARF6, ARF8	167: A G O P So Sb Z	(2005)
	393	393: TIR1	393: A M O P Sb Z	
	390	390: ARF3, ARF4 via TAS3 (ta-siRNA)	390: A O P	
Organ separation and number	164	CUC TFs	A O P Sb Z	Laufs *et al.* (2004); Mallory *et al.* (2004a); Baker *et al.* (2005)
Organ polarity, vascular and meristem development	165/166	HD-ZIP III TFs (PHB, PHV, leaf 1)	A G M O P Sb Z	Emery *et al.* (2003); Floyd and Bowman (2004); Juarez *et al.* (2004); Kim *et al.* (2004); Mallory *et al.* (2004b); McHale and Koning (2004); Williams *et al.* (2005b); Zhong and Ye (2004)
Floral and leaf patterning	159/319	159: MYB TFs 319: TCP TFs	A G M O P S S Z	Palatnik *et al.* (2003); Achard (2004); Millar and Gubler (2005)
Regulation of the miRNA pathway	162	162: DCL1	162: A M O P Z	Xie *et al.* (2003); Vaucheret *et al.* (2004);
	168	168: AGO1	168: A G O P So Sb Z	Allen *et al.* (2005)
	403	403: AGO2	403: A P	
Sulfate assimilation	395	ATP sulfurylases (APS1, APS3, APS4)	A O P S Z	Jones-Rhoades and Bartel (2004)
Lignin formation?	397	Laccases	397: A O P	Jones-Rhoades and Bartel (2004); Schwab *et*
	408			*al.* (2005); Lu *et al.* (2005b)

Oxidative stress	398	Copper superoxide dismutases CSD1, CSD2	A G O P	Jones-Rhoades and Bartel (2004)
Phosphate homeostasis	399	Phosphate transporter, E2 ubiquitin-conjugating enzyme	A M O P Sb Z	Jones-Rhoades and Bartel (2004); Chiou *et al.* (2006); Fujii *et al.* (2005)
Unknown	161	161: PPR	161: A	Jones-Rhoades and Bartel (2004); Allen *et al.* (2005); Vazquez *et al.* (2004); Rhoades *et al.* (2002); Sunkar *et al.* (2004); Sunkar & Zhu (2005); Peragine *et al.* (2004); Yoshikawa *et al.* (2005); Park *et al.* (2002)
	163	163: SAM-dependent methyl transferase	163: A	
	173	173: PPR via TAS1 & TAS2 (ta-siRNA)	173: A	
	390	390: Receptor-like kinase	390: A O P	
	394	394: F-box	394: A O P Sb Z	
	396	396: Growth response factors (GRF1, GRF2, GRF3, GRF7, GRF8, GRF9), rhodanese	396: A G O P So Sb Z	

The second approach analyses the expression patterns of the miRNAs and their potential targets. In this study, both target prediction through search for sequence pattern complementary between miRNA and its potential target with score systems by Jones-Rhoades and Bartel (2004), followed by expression analysis from mRNA sequencing data with double confirmation by RT-qPCR was carried out to look product of degradation consecutive to the cleavage of the target miRNA (Llave *et al.*, 2002b; Jones-Rhoades *et al.*, 2004; Lu *et al.*, 2005b).

miRNA Databases

With the advent of sophisticated bioinformatics tools, publicly available miRNA databases coupled with experimental methods including direct cloning, deep/degradome sequencing have facilitated the identification of plant miRNAs along with their target identification and functional characterization (Sheng *et al.*, 2018). The number of identified miRNAs in plants is on gradual rise as reflected by increasing entries in latest versions of miRBase (**Table 1.9**).

Table 1.9. Plant microRNA (miRNA) distribution.

Plant Species	miRNA families	miRNA genes
Acacia auriculiformis	7	7
Acacia mangium	3	3
Aegilops tauschii [ASM34733v2]	88	173
Amborella trichopoda [AmTr_v1.0]	124	129
Aquilegia caerulea	45	45
Arabidopsis lyrata [v1.0]	205	384
Arabidopsis thaliana [TAIR10]	326	428
Arachis hypogaea	23	32
Asparagus officinalis [Aspof.V1]	101	101
Avicennia marina	2	3
Brachypodium distachyon [Bd21]	317	525
Brassica napus [BrassicaDB-20080411]	90	92
Brassica oleracea	10	11
Brassica rapa [Brapa_1.1]	96	157
Bruguiera cylindrica	4	4
Bruguiera gymnorhiza	4	4
Camelina sativa [Cs]	91	112
Carica papaya	79	81
Chlamydomonas reinhardtii [CHLRE3.1]	50	86
Citrus clementina	5	5
Citrus reticulata	4	4
Citrus sinensis [Csi_valencia_1.0]	151	246
Citrus trifoliata	6	6
Cucumis melo	120	120
Cucumis sativus [ASM407v2]	5	5
Cunninghamia lanceolata	4	4
Cynara cardunculus	48	57
Digitalis purpurea	13	13
Elaeis guineensis	6	6
Eugenia uniflora	39	74
Festuca arundinacea	15	15
Fragaria vesca [v1.1]	118	131
Glycine max [Glycine_max_v2.0]	684	756
Glycine soja	13	13
Gossypium arboreum	1	1
Gossypium herbaceum	1	1
Gossypium hirsutum	78	80
Gossypium raimondii [Graimondii2_0]	296	296
Helianthus annuus	6	7
Helianthus argophyllus	3	3
Helianthus ciliaris	3	3
Helianthus exilis	2	2
Helianthus paradoxus	3	3

Helianthus petiolaris	3	3
Helianthus tuberosus	16	16
Hevea brasiliensis [Hevbra_1.0]	30	30
Hordeum vulgare [Hvul_cvHarunaNijo3HBAC01]	69	71
Linum usitatissimum	124	124
Lotus japonicus [lj2.5]	299	365
Malus domestica	308	322
Manihot esculenta [Cassava4]	174	175
Medicago truncatula [JCVI_Mt3.5.2]	672	756
Nicotiana tabacum	162	164
Oryza sativa [MSU7]	604	738
Paeonia lactiflora	15	15
Panax ginseng	29	32
Phaseolus vulgaris [PhaVulg1_0]	8	10
Physcomitrella patens [v3.0]	247	298
Picea abies	594	600
Pinus densata	29	29
Pinus taeda	35	36
Populus euphratica	4	4
Populus trichocarpa [Poptr2_0]	352	401
Prunus persica [Prunus_persica_NCBIv2]	180	214
Rehmannia glutinosa	32	37
Saccharum officinarum	16	16
Saccharum sp.	19	20
Salicornia europaea	31	31
Salvia miltiorrhiza	1	1
Salvia sclarea	18	18
Selaginella moellendorffii	58	64
Solanum lycopersicum [SL2.50]	112	147
Solanum tuberosum [SolTub3.0]	224	343
Sorghum bicolor [Sorghum_bicolor_NCBIv3]	205	241
Theobroma cacao	82	82
Triticum aestivum	122	125
Triticum turgidum	1	1
Vigna unguiculata	18	18
Vitis vinifera [12X]	163	186
Vriesea carinata	33	65
Zea mays [B73_RefGen_v4]	174	325

Data retrieved from the miRNA registry release 22 (http://mirbase.org/cgi-bin/browse.pl; retrieved on [June 01, 2019]). Genomes that have been fully sequenced are indicated in bold type with its genome sequence database version indicated in brackets.

CHAPTER II

IN VITRO REGENERATION OF *VANILLA PLANIFOLIA* ANDREWS

Chapter II

In Vitro regeneration of *Vanilla planifolia*

2.1. Introduction

The high demand for natural vanillin has led to the need for expansion in land area for planting high quality vanilla plants, which in turn has led to increased demands for this plant as a planting stock. Like other orchids, the seeds of *Vanilla planifolia* (*V. planifolia*) Andrews rarely germinate because of the presence of germination inhibitors (Dequaire 1976) and its dependency on the presence of mycorrhizal fungi (Arditti and Ghani 2000). Thus, traditionally, *V. planifolia* is propagated by stem-cuttings. However, this method is labour intensive, time consuming (Kalimuthu *et al.* 2006, Janarthanam and Seshadri 2008) and not economical, as the collection of the stem cuttings from mother plants may eventually cause retardation to their growth and reduce their yield potential (Davidonis and Knorr 1991, Geetha and Shetty 2000, Kalimuthu *et al.* 2006). Therefore, plant tissue culture plays an important role in producing pathogen-free and high quality clones of plants, overcoming the seed germination difficulties of orchids in nature. Clonal propagation of elite genotypes through *in vitro* micropropagation helps to ensure the performance and productivity of the crop. Besides, the growing interests to increase vanilla yield through biotechnological applications, such as *in vitro* vanillin production in callus cells through elicitations (Funk and Brodelius, 1990) or biotransformations (Velankar and Heble, 2004), or *in vitro* crop improvements (Minoo *et al.*, 1997) also highlighted the importance of tissue culture as a fundamental technique for vanilla biotechnology.

To date, *in vitro* regeneration of vanilla has also been developed through shoot tip and axillary bud (Geetha and Shetty 2000, George and Ravishanka 1997, Giridhar and Ravishankar 2004, Kalimuthu *et al.* 2006, Gopi *et al.* 2006,

Lee-Espinosa *et al.* 2008, Tan *et al.* 2011b), root apex (Philip and Nainar 1986), leaf (Gu *et al.* 1987, Janarthanam and Seshadri 2008, Tan *et al.* 2011a) and seed-derived protocorm-like bodies (Malabadi and Nataraja 2007, Retheesh and Bhat 2011). *V. planifolia* could be propagated through both direct and indirect regeneration. Micropropagation through direct organogenesis using nodes as explants has been reported to be an appropriate technique for *V. planifolia* (Lee-Espinosa *et al.* 2008). This technique has successfully produced approximately 100,000 to 150,000 plantlets per explant from each mother plant within 12 months (Lee-Espinosa *et al.* 2008). In contrast to direct regeneration, the induction of callus for orchids has been very difficult due to its slow growing and necrotic tendency of orchid callus (Huan *et al.* 2004, Long *et al.* 2010). In recent years, only a few studies have been reported for orchid callus cultures, such as *Cymbidum* (Chang and Chang 1998) and *Dendrobium* (Roy and Banerjee 2003). Hence, only a few studies have been reported on micropropagation via indirect organogenesis in *V. planifolia*, where poor callusing was reported (Gu *et al.* 1987, Kononowicz and Janick 1984). Earlier study revealed poor callusing for *V. planifolia* through roots (Philip and Nainar 1986). In recent years, Tan *et al.* (2011a) also reported that only 35% and 10% of callus was recorded in nodal and leaf explants respectively when cultured on MS medium containing 2.0 mg/L 1-naphthaleneacetic acid (NAA) and 1.0 mg/L 6-Benzylaminopurine (BAP).

The overall aim of study reported in this chapter is to aseptic *in vitro* culture and efficient regeneration of *V. planifolia* Andrews for further studies of somaclonal variations in experiments reported in the later chapters. In order to achieve the overall aim of the study, the following specific objectives were carried out: (i) establishment of aseptic cultures of vanilla from the shade house, and (ii) optimization of vanilla regeneration protocols using different source of explants (node, root, leaf and seed) and plant growth regulator (PGR) treatments.

2.2. Materials and Methods

2.2.1. Instrument and apparatus preparation

A list of instrument and apparatus that were used in this study is provided in **Appendix B.** All surgical instruments such as scalpels and forceps, glassware, and other apparatus were pre-disinfected by soaking with liquid disinfects, 10% (v/v) Clorox® or Cleaning Agent Decon® 90 (Decon Laboratories Limited, UK) overnight and rinsed under running tap water, followed by autoclaving for 15 min at 121°C, 15 psi; and dried in an oven at 70ºC for overnight. All procedures for *in vitro* culturing were carried out in a laminar air flow cabinet under strict aseptic conditions. An Esco Airstream Horizontal Laminar Flow Cabinet (ESCO®, USA) with Ultra Low Penetration Air (ULPA) filter was used, to ensure all sterilized materials remained contamination-free. The transfer area in laminar air cabinet was cleaned with 70% (v/v) ethanol (freshly prepared every month), followed by germicidal ultraviolet (UV) light exposure for at least an hour prior to usage (Sathyanarayana and Varghese, 2007). Sterile tools, such as forceps, scalpels and disposable blades were sterilized regularly using glass bead sterilizers (Innotech Biosystems International, USA) for 30 seconds at 200 to 250ºC for 30 seconds.

2.2.2. Chemicals, media and stock plant growth regulators preparation

A list of plant tissue culture sterilization agents, culture media, PGRs and vitamin is listed in **Appendix A**. Full-strength and half-strength Murashige and Skoog (MS) media were prepared using 4.43 g/L and 2.215 g/L of MS-basal salt (*Phyto*technology Laboratory®, USA), respectively, with the addition of 3.0% (w/v) sucrose (Fisher Scientific, USA) and 3.5% (w/v) Phytagel™ (Sigma-Aldrich, USA) unless otherwise stated. For liquid medium, the addition of Phytagel™ was omitted. The MS media were adjusted to pH 5.8 ± 0.2 with 1 N NaOH or 1 N HCl before autoclaved for 15 min at 121°C, 15 psi. All plant growth regulator stock solutions were prepared at a concentration of 1.0 mg/mL,

filtered through 0.2 µm filter (Minisart®, Sartorius) in the laminar air flow cabinet, and stored in the 4ºC fridge for up to six months. Light-sensitive hormones, such as TDZ, were stored in the dark at 4ºC.

2.2.3. Culture condition

All cultures were incubated at 25 ± 5°C and relative humidity (RH) of 60 ± 18 % under 16:8 light-dark cycle with a light intensity of 40 µmol m^{-2} s^{-1} provided by cool white fluorescent lights, unless otherwise stated. The temperature and RH of culture rooms were monitored every 25 min with a HOBO Wireless Data Logger System (OnSet, USA).

2.2.4. Source of plant materials

Explants used for tissue culture (**Figure 2.1b-d**) were obtained from the *V. planifolia* plants grown in the shade house of University of Nottingham Malaysia Campus (**Figure 2.1a**). The original source of *V. planifolia* cuttings were obtained from the nursery of Kuasa Sekata Sdn Bhd, Selangor, Malaysia, and vanilla seed pods (**Figure 2.1f-g)** were obtained from the KK vanilla plantation from Jerantut, Pahang. **Figure 2.1e** showed the flowers morphology of *V. planifolia*, Andrews of which the vanilla pods were formed from.



Figure 2.1. Source of vanilla tissue culture mother plants for nodal cuttings and seed pods. (a) mother plants of *V. planifolia* grown in the shade house of University of Nottingham Malaysia campus; *in vitro V. planifolia* plantlets at (b) 7-day and (c) 3-week post-sterilization: scale bar = 1.0 cm; and (d) vanilla plantlets from which explant sources were obtained for regeneration optimization experiments; (e) vanilla flower (photo credits to Mr. Yong DY) and (f-g) seed pods from KK vanilla plantation in Jerantut, Pahang: scale bar = 2.0 cm.

2.2.5. Establishment of *V. planifolia* aseptic cultures

Juvenile nodal segments (1.5 to 2.0 cm) of *V. planifolia* were surface disinfected according to the protocol by Tan *et al.* (2010a) with slight modification. The nodal segments were first cleaned thoroughly by gently brushing under running tap water for 10 min and washing in liquid detergent containing one to two drops of Tween 20 in 1.0 L distilled water, followed by three rinses with sterile distilled water. Further surface disinfection was carried out by submerging and shaking the explants in 0.2% (v/v) carbendazim (Mei Zim, Malaysia) for 20 min, rinsing thrice with sterile distilled water, immersing and shaking in 70% (v/v) ethanol for 2 min followed by washing thrice with sterile distilled water. The cleaned explants were then surface sterilised in 20% (v/v) Clorox® (commercial bleach containing 5.25% w/v sodium hypochlorite) with the addition of a few drops of Tween 20 for 20 min with constant shaking inside a laminar air flow cabinet and finally rinsed thrice with sterile distilled water before cultured on full-strength Murashige and Skoog (MS) solid medium containing 1.0 mg/L 6-Benzylaminopurine (BAP) for one month. After one month, nodes were cultured in full-strength MS liquid medium without plant growth regulators for a week. Shoots regenerated from clean nodal explants free of microorganism contamination were used for subsequent callus and shoot regeneration experiments.

This surface sterilization protocol was further optimized for various carbendazim concentrations used (0.2, 0.4, 0.6, 0.8 and 1.0%) and incubation durations (20 min, 60 min, 120 min and 24 h). The aseptic culture established in this protocol were used as explants for further hormonal optimization. The axenic status of cultured tissues was confirmed by using liquid broth culture assay. Samples of tissues were cultured in liquid MS for seven days to detect the presence of microbial contaminants. Data of the percentages of contamination and phytotoxicity were recorded after one month from culture initiation. Phytotoxicity was scored by stress symptoms, including browning, bleaching and death of explant tissue.

2.2.6. Vanilla nodal, leaf and root tips culture

Juvenile vanilla nodal, leaf, root tip explants from three-month-old *in vitro* vanilla plantlets were cultured on half-strength MS media supplemented with 1.0 g/L myo-inositol, 1.0 mg/L thiamine-HCl and various combinational concentrations of NAA, 2,4-D and BAP for callus induction or shoot initiation (see **Table 2.1**). For nodal segments culture (NT), nodal segments containing one intact axillary bud, with its attached leaf removed, were cultured. For root tips regeneration (RT), 0.5 ± 0.1 cm of aerial root tips were cultured. For leaf apices and bases (LT), 0.5-1.0 cm in diameter from leaf tip and leaf bases respectively were used and the explants were placed with the adaxial surface in contact with the solid culture medium (Arditti, 2006). The percentage of explants forming callus/shoot, the average size of callus formed and number of shoots formed were recorded after three months of culture. For each experimental treatment, three replicates each consisting of five units were carried out. The percentage of callus and shoot induced, callus sizes and weights, and number of shoots formed per explant were measured after 90 days of culture.

Table 2.1. Hormonal treatments for the culture of different vanilla explants.

Explants	**Hormones (concentration range) and combinations**
Nodal segments (NT)	0, 0.5, 1.0, 2.5, 5.0 mg/L BAP in combination with 0, 0.5, 1.0, 2.5 mg/L 2,4-D or NAA
Root tips (RT)	RT-1: 0, 0.5, 1.0, 2.5, 5.0 mg/L BAP or NAA or 2,4-D; or 0.1, 0.25, 0.5, 1.0 mg/L TDZ
	RT-2: 0.5, 1.0 mg/L BAP or 0.25, 0.5 mg/L TDZ in combination with 0, 0.5, 1.0, 2.0 mg/L NAA or 2,4-D
Leaf (LT)	0, 0.5, 1.0, 2.5, 5.0 mg/L BAP in combination with 0, 0.5, 1.0, 2.5 mg/L 2,4-D, NAA or TDZ

2.2.7. Shoot proliferation and maintenance, and rooting

Successfully generated calli (30 ± 5 mg) were transferred to full-strength MS medium supplemented with 1.0 mg/L BAP for shoot initiation and elongation under light condition, in accordance to Tan *et al.* (2011a)'s protocol. Elongated shoots were maintained on the same medium every four weeks for three months. The percentage of callus cultures forming shoots, number of shoots formed per callus and average shoot length were recorded after six months of culture. Shoots of 2–4 cm in height were transferred for root induction on full-strength MS basal medium supplemented with NAA according to Tan *et al.* (2011a)'s protocol.

2.2.8. Vanilla seed culture

Five seven-month-old vanilla pods (**Figure 2.1f-g**), with an average pod lengths of 15.40 ± 0.82 cm, were surface sterilized by dipping the whole pod into pure ethanol, followed by 3-5 seconds of Bunsen burner flame. After that, the seed pods were sterilized in 100% Clorox for 15 min with constant agitation, and air dried in laminar air flow cabinet. Dried pods were cut longitudinally to reveal the seeds. The vanilla seeds were first suspended in autoclaved water, and diluted (Hoang *et al.*, 2017) to an average of 100 ± 10 seeds per 100 µl. Estimation of seed number was carried out by suspending 10 µl of seed suspension onto a haemacytometer and counted under a light microscope. A 300-µl seed suspension aliquot was dispensed onto each culture jar containing growth medium, spread evenly and allowed to dry in laminar air flow. Three different orchid mediums, namely Murashige and Skoog (MS) medium, Knudson C orchid medium and Gamborg B5 medium, each added with 0.5 mg/L TDZ, were tested for germination. Each medium consists of 3 plates per seed pod. Seed germination and the number of PLB formation were recorded. Seed germination percentage was calculated by the following formula:

$$\text{Percentage seed germination} = \frac{\text{Number of seeds germinated}}{\text{Total number of seeds}} x\ 100$$

(Roy *et al.*, 2011; Balilashaki *et al.* 2015)

2.2.9. Morphological observation and histological analysis

Morphogenesis observation of callus (derived from nodes and roots), protocorm-callus (from seeds), single or multiple shoot regeneration, and any phenotypic abnormalities from the *in vitro* regenerants, were observed macroscopically, and microscopically using dissecting microscope, Nikon SMZ1500. For histological analysis, plant cells were fixed for 24-48 hours at room temperature in a Glutaldehyde-Paraformaldehyde-Caffeine (GPC) (Sigma-Aldrich, USA) fixative solution, then dehydrated in series of ascending ethanol series at 30-100% (v/v) and infiltrated with basic resin (Leica Historesin Embedding Kit, Germany) for 1-2 weeks. Subsequently, the plant cells were embedded and polymerized in a mold overnight. Serial sectioning was then performed at 4 µm using Leica RM2165 rotary microtome (Leica Microsystems, Germany). The samples were stained with toluidine blue staining reagent (Sigma-Aldrich, USA) for checking. Finally, the slides were stained with periodic acid, Schiff's reagents and Naphthol Blue Black reagent (Sigma-Aldrich, USA) and mounted onto Cytoseal^{TM} 60 mounting medium. The sample slides were viewed under light microscope, Nikon Eclipse 80i (Nikon, USA) and all microscopic photos were captured with camera through the NIS-Element 3.0 software (Nikon, USA).

2.2.10. Statistical analysis

The experiments were conducted in a completely randomized block (CRB) design with five replicates of three culture units for each treatment. Data in percentage were transformed (arcsine) using logistic regression analysis, at the assumptions of generalized linear models. All data were also subjected to one-way analysis of variance (ANOVA) using GenStat 18th Edition software. Duncan's multiple range test (DMRT) with significant difference level at $p \leq 0.05$ was then used to analyse the differences between the treatments.

2.3. Results

2.3.1. Establishment of vanilla aseptic cultures

Using the surface sterilization method developed by Tan *et al.* (2011a), the vanilla nodal cultures had an average of 80-90% bacterial contamination and 50-60% fungal contamination. Bacterial contamination was mostly observed within two weeks, but fungal contamination was detected even after one month (**Figure 2.2**). Therefore, observations at post-surface sterilization were carried out for at least a month to ensure the cultures were truly contamination-free. Increased carbendazim concentration to 0.4% for 20 min has successfully controlled the fungal contamination, while further increase in carbendazim concentration at 0.6% and above caused phytotoxicity (bleaching, browning and eventually death effect) to the explants (**Table 2.2**). Bacterial contamination (up to 90%) were only controlled when the duration of carbendazim treatment was prolonged to 60 min. Overall, 0.4% carbendazim for 60 min was the optimal protocol for reducing fungal and bacterial contamination in aseptic culture of vanilla.

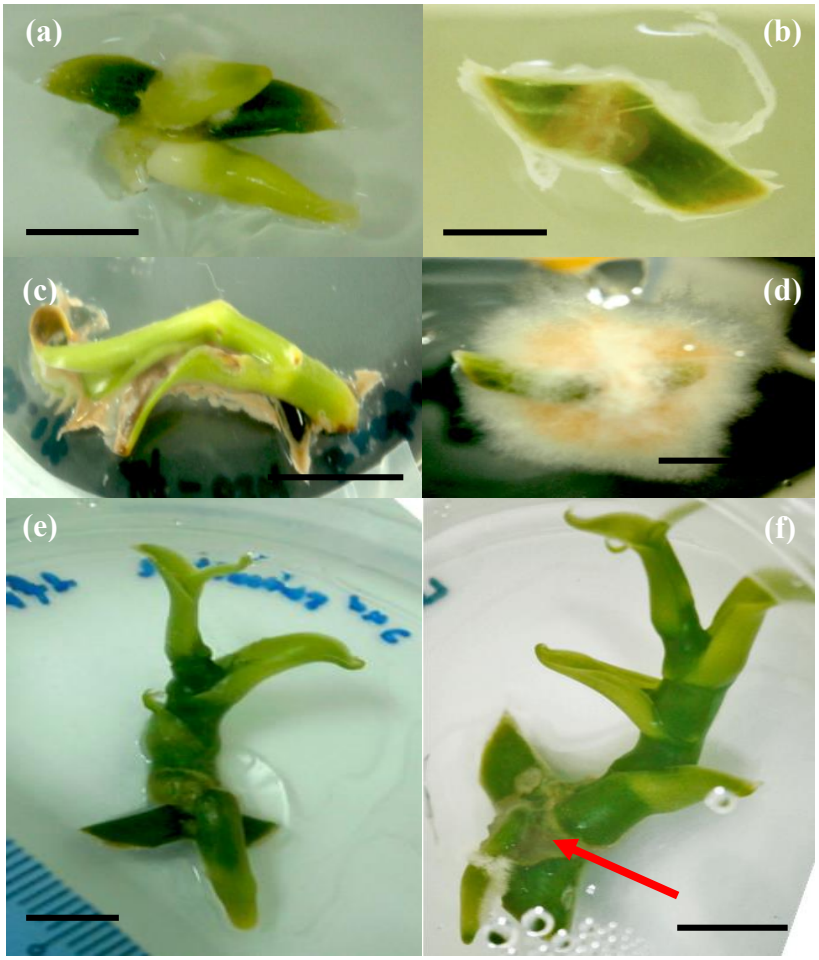


Figure 2.2. Vanilla post-surface sterilization. Fungal and bacteria contamination early detections: Clean and healthy growing node (a), bacterial contaminated node (b) and shoot tip (c), and fungal contaminated node detected within 1-2 weeks post surface sterilization. Some contaminations could only be detected after one month: Clean (e) and fungal contaminated (f) nodal-derived shoots after one month post surface sterilization. **Red** arrow indicates fungus. Scale bar = 1.0 cm

Table 2.2. Results of surface sterilization using carbendazim, after one month.

Carbendazim Percentage	Bacterial Contamination (%)	Fungal Contamination (%)	Survival (%)	Phytotoxicity (%)
0.2%	90.00 ± 11.55	50.00 ± 8.16	0.00 ± 0.00	0.00 ± 0.00
0.4%	90.00 ± 8.16	0.00 ± 0.00	10.00 ± 8.16	0.00 ± 0.00
0.6%	90.00 ± 11.55	0.00 ± 0.00	0.00 ± 0.00	10.00 ± 8.16
0.8%	80.00 ± 16.33	0.00 ± 0.00	10.00 ± 8.16	10.00 ± 8.16
1.0%	90.00 ± 11.55	0.00 ± 0.00	0.00 ± 0.00	20.00 ± 14.14
Carbendazim Duration	**Bacterial Contamination (%)**	**Fungal Contamination (%)**	**Survival (%)**	**Phytotoxicity (%)**
20 min	90 ± 8.16	0.00 ± 0.00	10.00 ± 8.16	0.00 ± 0.00
60 min	10 ± 8.16	0.00 ± 0.00	90.00 ± 14.14	0.00 ± 0.00
120 min	20 ± 16.33	0.00 ± 0.00	80.00 ± 8.16	0.00 ± 0.00
Overnight (16-18h)	20 ± 14.14	0.00 ± 0.00	80.00 ± 16.33	0.00 ± 0.00

Values expressed as means with standard deviations.
Each mean was based on 4 replicates; each replicate consists of 4 cultures.

Figure 2.3a shows the liquid cultures of clean and contaminated vanilla. Gram staining test (**Figure 2.3b**) showed purple color stained bacteria, indicating that the bacteria isolated were likely to be Gram-positive.

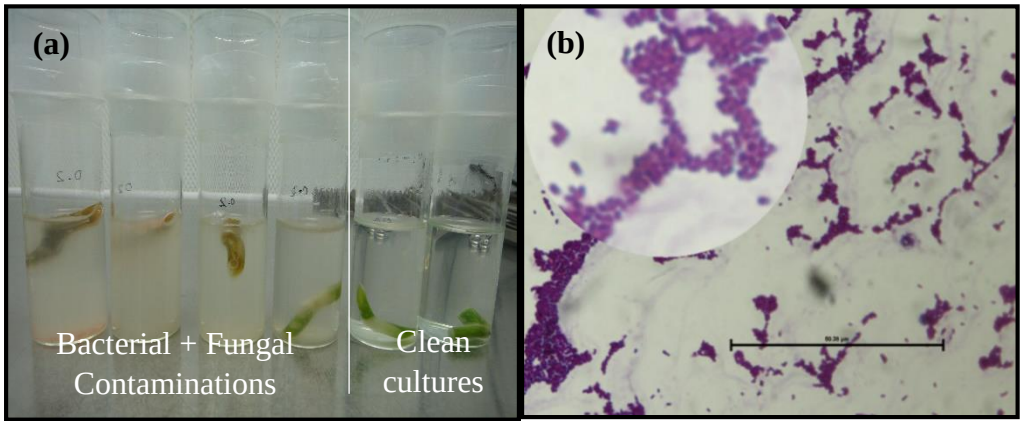


Figure 2.3. (a) Liquid culture test of vanilla explants after one-month post-surface sterilization, and (b) Gram stain test of bacterial isolates isolated from vanilla cultures. Scale bar = 50.38 µm.

2.3.2. Nodal culture (NT)

Direct and indirect shoot regenerations were observed (**Table 2.3**).

Table 2.3. The influence of BAP, TDZ, NAA or 2,4-D hormonal treatments on nodal explants when cultured on semi-solid MS medium, after 90 days of culture.

PGR(s)	Concentration (mg/L)	Total response	% callus response	Average shoot number
No PGR (Control)		100.00	0.00^{a}	1.0^{a}
BAP only	0.5	100.00	0.00^{a}	6.0^{d}
	1.0	100.00	0.00^{a}	8.0^{f}
	2.5	100.00	0.00^{a}	2.0^{b}
	5.0	100.00	0.00^{a}	0.0^{a}
BAP + NAA	0.5 + 0.5	83.33	0.00^{a}	4.0^{cd}
	0.5 + 1.0	83.33	0.00^{a}	3.5^{c}
	0.5 + 2.5	83.33	50.0^{c}	9.0^{fg}
	1.0 + 0.5	50.00	0.00^{a}	1.0^{a}
	1.0 + 1.0	66.67	0.00^{a}	1.5^{ab}
	1.0 + 2.5	66.67	66.67^{c}	10.5^{hi}
	2.5 + 0.5	66.67	0.00^{a}	9.8^{gh}
	2.5 + 1.0	16.67	50.00^{c}	7.0^{ef}
	2.5 + 2.5	50.00	50.00^{c}	3.5^{c}
BAP + 2,4-D	0.5 + 0.5	66.67	0.00^{a}	1.0^{a}
	0.5 + 1.0	0.00	0.00^{a}	0.0^{a}
	0.5 + 2.5	0.00	0.00^{a}	0.0^{a}
	1.0 + 0.5	16.67	50.00^{c}	7.0^{ef}
	1.0 + 1.0	16.67	50.00^{c}	8.5^{fg}
	1.0 + 2.5	0.00	16.67^{ab}	1.5^{ab}
	2.5 + 0.5	16.67	0.00^{a}	1.0^{a}
	2.5 + 1.0	0.00	0.00^{a}	0.0
	2.5 + 2.5	16.67	0.00^{a}	1.0^{a}
BAP + TDZ	0.5 + 0.25	50.00	41.67^{bc}	11.5^{i}
	0.5 + 0.5	83.33	50.00^{c}	9.7^{gh}
	0.5 + 1.0	100.00	50.00^{c}	5.6^{de}
	0.5 + 2.5	83.33	0.00^{a}	0.0^{a}
	1.0 + 0.25	100.00	0.00^{a}	9.5^{g}
	1.0 + 0.5	66.67	25.00^{b}	8.5^{fg}
	1.0 + 1.0	50.00	25.00^{b}	6.0^{d}
	1.0 + 2.5	100.00	0.00^{a}	2.5^{bc}
	2.5 + 0.25	83.33	0.00^{a}	1.0^{a}
	2.5 + 0.5	83.33	0.00^{a}	1.0^{a}
	2.5 + 1.0	70.00	25.00^{b}	3.0^{bc}
	2.5 + 2.5	60.00	0.00^{a}	1.5^{ab}

Values expressed as means of 3 replicates; each replicate consisted of 4 cultures.
Different letters within column (lowercase) indicate a significant difference ($p \leq 0.05$) according to Duncan's multiple range.

2.3.2.1. Direct shoot regeneration

Nodal explants went through direct regeneration when treated with BAP alone. Without phytohormones (control treatment), vanilla nodes produced an average of only one shoot per node (**Figure 2.4a**, **Table 2.3**). Increasing BAP concentrations promoted multiple shoot responses of vanilla. Across the BAP concentrations, optimal shoot response was achieved with the treatment of 1.0 mg/L BAP (**Figure 2.4b**), at an average of 8.0 shoots per node after 90 days of culture (**Table 2.3**). Higher BAP concentrations caused abnormality in vanilla nodes. At normal conditions, shoot multiplication occurred only at the axillary buds of basal nodes (bn). However, at 5.0 mg/L of BAP, shoots exhibited abnormal phenotypes, all axillary buds on vanilla aerial nodes germinated and formed shoot primordia (**Figure 2.4c**). These eventually failed to develop into proper shoots that could be regenerated.

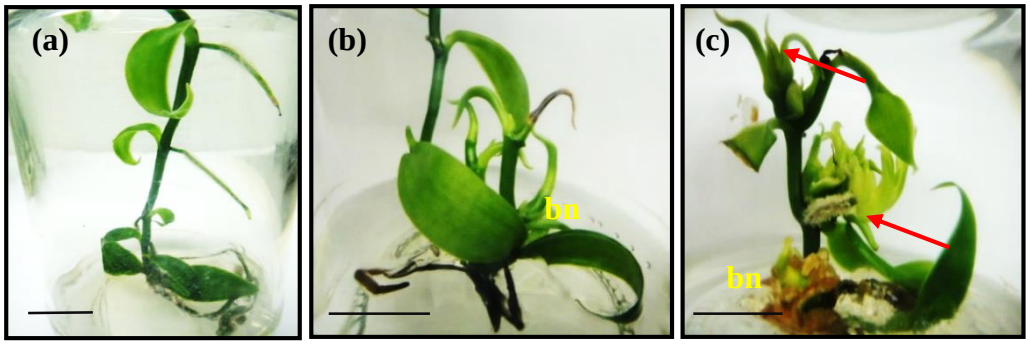


Figure 2.4. Direct single, multiple and abnormal shoot regenerations of vanilla nodal cultures. Nodal segments when treated with (a) 0.0 (b) 1.0 and (c) 5.0 mg/L BAP for 4 months; bn: basal node; **red** arrow indicates the germination of axillary buds at multiple aerial nodes other than the basal node. Bar = 1.0 cm.

2.3.2.2. Indirect regeneration

The addition of NAA to BAP induced callusing in vanilla. The optimal treatment among NAA treatments was 2.5 mg/L with 1.0 mg/L BAP, with the highest callusing rate (66.67%) and the highest number of shoots formed after 6 months (an average of 10.5 shoots per callus) (**Table 2.3)**.

The callusing rate of vanilla increased with NAA concentrations, when treated in combination with 2.5 mg/L BAP. At lower levels of NAA (0.5 mg/L,

Figure 2.5a), direct multiple shoot multiplications were promoted mostly without going through the callus phase; at the medium levels of NAA (1.0 mg/L, **Figure 2.5b**), green calli were formed; while at the higher levels of NAA (2.5 mg/L, **Figure 2.5c**), calli formed were yellowish and compact.

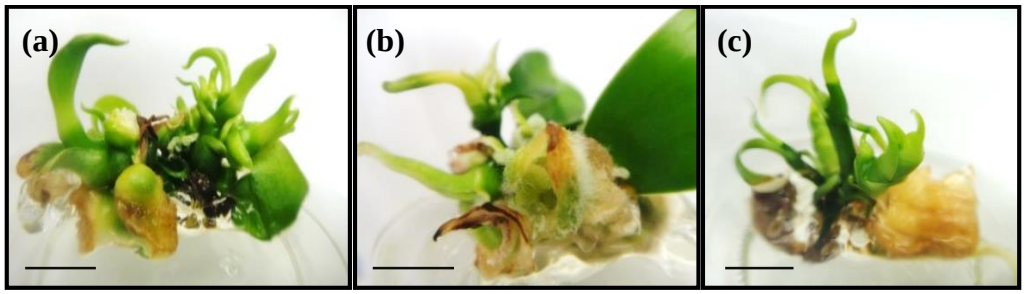


Figure 2.5. Different shooting or callusing responses in BAP treatment with various NAA levels. Nodes treated with 2.5 mg/L BAP in combination with: (a) 0.5 mg/L formed multiple shoots, (b) 1.0 mg/L formed green calli, and (c) 2.5 mg/L NAA formed yellowish and compact calli. Bar = 0.5 cm.

Varied BAP levels also affected the callus morphology of vanilla, when treated in combination with 2.5 mg/L NAA. At lower levels of BAP concentration (0.5 mg/L, **Figure 2.6a**), calli were less de-differentiated and darker green in colour; at medium level of BAP concentration (1.0 mg/L, **Figure 2.6b**), calli were more embryogenic-looking and light green in colour. While at higher level of BAP concentration (2.5 mg/L, **Figure 2.6c**), calli appear to be more de-differentiated, compact and yellow in colour.

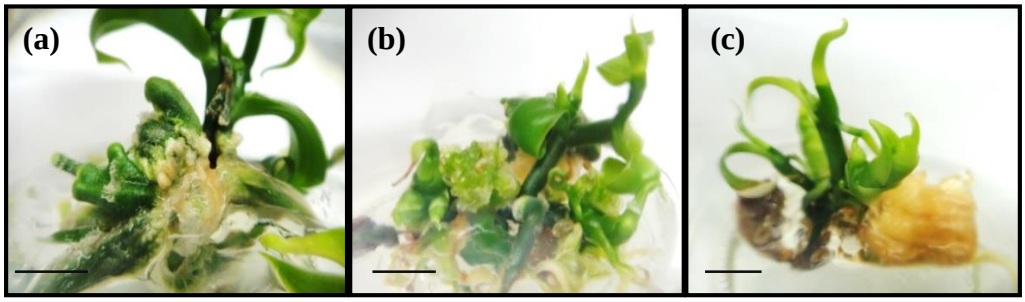


Figure 2.6. Different callus morphologies in NAA treatments with various BAP levels. Nodal segments when treated with 2.5 mg/L NAA and BAP at (a) 0.5 mg/L, calli were dark green and less de-differentiated; (b) 1.0 mg/L, calli were light green and embryogenic-looking; and (a) 2.5 mg/L, calli were yellow, compact and differentiated. Bar = 0.5 cm.

The addition of 2,4-D also induced callusing in vanilla. Callus formation is the highest when nodes were treated with 0.5 mg/L or 1.0 mg/L of 2,4-D in

combination with 0.5 mg/L BAP (**Table 2.3**). However, 2,4-D concentrations at 1.0 mg/L and above, caused explants browning over time, showing phytotoxicity effect to the vanilla plants. Even though calli were formed, the original explants showed stress symptom (**Figure 2.7**), with its internode and leaves wilted and turned brown eventually.

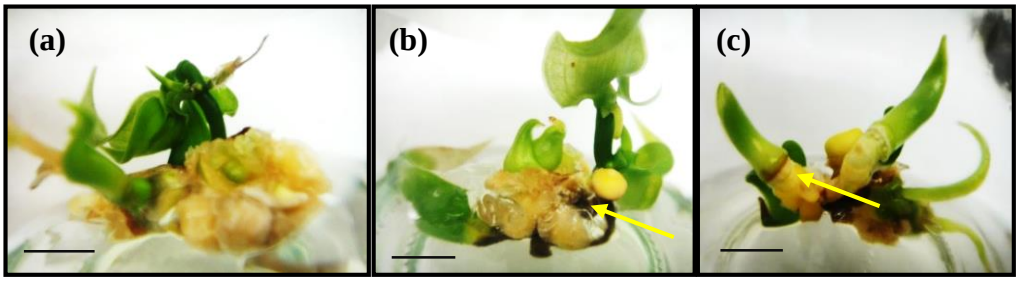


Figure 2.7. Callus formation from 2,4-D treatments. Nodes treated with (a) 0.5mg/L, (b) 1.0 mg/L and (c) 2.5 mg/L 2,4-D, in combination with 1.0 mg/L BAP, for 4 months. **Arrow** shows the browning effects due to phytotoxicity of 2,4-D. Bar = 0.5 cm.

2.3.3. Root tips culture

In vitro culture of vanilla root tips formed callus under the influence of cytokinins, instead of auxins. In the control treatment (without the addition of any phytohormones), callusing rate was zero but root elongated. It was shown that high callusing rates (86.67 to 100.00%) were achieved with cytokinin treatments alone (TDZ or BAP) of root tips (**Table 2.4, Figure 2.8**), while under auxin treatments alone (NAA or 2,4-D), root tips neither elongated nor callused. When cytokinin was interacted with auxin treatments (**Table 2.5**), the combination of auxins and cytokinins produced callus ranging from 20.00 to 100.00%; however, the average size of callus formed in auxin-cytokinin combinational treatments were much smaller than the cytokinin-only treatments after 90 days. This further suggested that in vanilla root tips, exogenous application of auxins (NAA or 2,4-D) exhibited inhibitory effects on callus that requires elevated cytokinin levels to reach a balanced auxin-cytokinin ratio for forming callus.

Table 2.4. The influence of sole applications of BAP, TDZ, NAA or 2,4-D hormonal treatments on callus response of root tips explants when cultured on semi-solid MS medium, after 90 days of culture.

Plant growth regulators (PGRs)		Callus Response	
PGR Type(s)	**Concentration (mg/L)**	**Total response (%)**	**Average Size (mm^2)**
No PGR (Control)		0.0^{a}	-
BAP only	0.5	100.00^{c}	21.83^{a}
	1.0	93.33^{bc}	20.50^{a}
	2.0	100.00^{c}	12.03^{a}
TDZ only	0.1	100.00^{c}	18.17^{a}
	0.25	100.00^{c}	18.47^{a}
	0.5	90.00^{b}	12.67^{a}
	1.0	86.67^{b}	14.73^{a}
NAA only	0.5	0.00^{a}	-
	1.0	0.00^{a}	-
	2.0	0.00^{a}	-
2,4-D only	0.1	0.00^{a}	-
	0.5	0.00^{a}	-
	1.0	0.00^{a}	-
	2.0	0.00^{a}	-

Values expressed as means of 3 replicated experiments; each replicate consisted of 10 cultures

Different letters within column (lowercase) indicate a significant difference ($p \leq 0.05$) according to Duncan's multiple range test

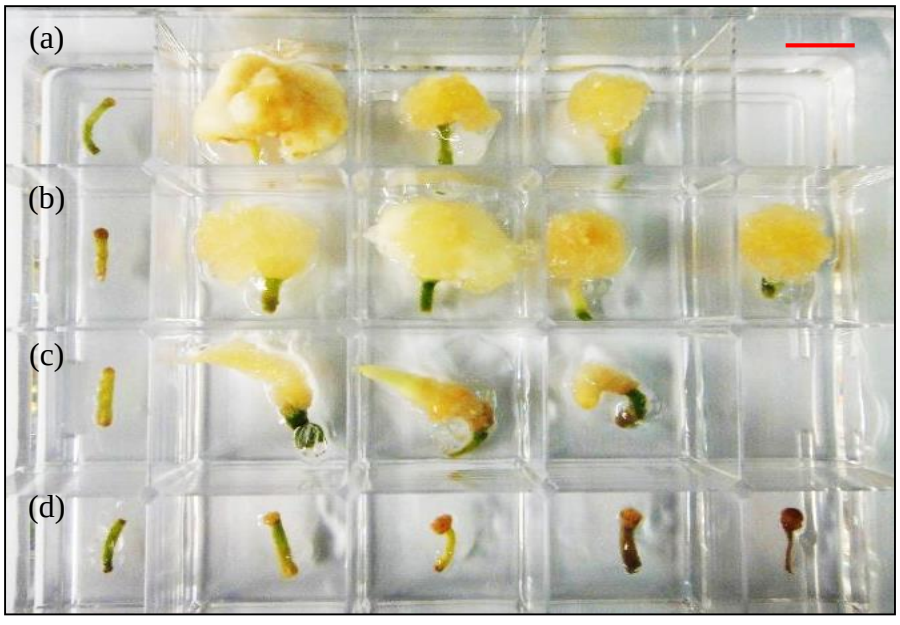


Figure 2.8. Response of *in vitro* root tips culture to different hormones. Concentrations from low to high, from left to right: (a) BAP (0.5, 1.0, 2.0 mg/L), (b) TDZ (0.1, 0.25, 0.5, 1.0 mg/L), (c) NAA (0.5, 1.0, 2.0 mg/L), (d) 2,4-D (0.1, 0.5, 1.0, 2.0 mg/L) after 90 days. Bar = 0.5 cm.

Table 2.5. The influence of BAP or TDZ in combination with 2,4-D and NAA hormonal treatments on callus response of root tips explant after 90 days of culture.

PGRs concentration (mg/L)			Callus Response	
BAP	**NAA**	**2,4-D**	**Total response (%)**	**Average Size (mm^2)**
0.5	-		100.00^{h}	21.83^{fg}
	0.5	-	90.00^{fgh}	15.33^{def}
	1.0		76.67^{efgh}	5.47^{ab}
	2.0		26.67^{abc}	3.57^{ab}
	-	0.1	0.00^{a}	-
		1.0	0.00^{a}	-
		2.0	0.00^{a}	-
1.0	-		93.33^{gh}	20.50^{fg}
	0.5	-	76.67^{efgh}	5.40^{ab}
	1.0		83.33^{efgh}	19.27^{fg}
	2.0		50.00^{bcde}	8.30^{abc}
	-	0.1	56.67^{cdef}	19.57^{fg}
		1.0	66.67^{defgh}	2.77^{a}
		2.0	36.67^{bcd}	4.83^{ab}
TDZ	**NAA**	**2,4-D**	**Total response (%)**	**Average Size (mm^2)**
0.25	-		100.00^{h}	18.47^{efg}
	0.5	-	100.00^{h}	16.50^{ef}
	1.0		76.67^{efgh}	23.87^{g}
	2.0		26.67^{abc}	15.60^{def}
	-	0.1	76.67^{efgh}	17.47^{efg}
		1.0	36.67^{bcd}	2.87^{a}
		2.0	0.00^{a}	-
0.5	-		90.00^{fgh}	12.67^{cde}
	0.5	-	83.33^{efgh}	12.30^{cde}
	1.0		50.00^{bcde}	5.60^{abc}
	2.0		83.33^{efgh}	10.10^{bcd}
	-	0.1	60.00^{cdefg}	4.73^{ab}
		1.0	20.00^{ab}	8.60^{abc}
		2.0	76.67^{efgh}	2.70^{a}

Values expressed as means of 3 replicated experiments; each replicate consisted of 10 cultures

Different letters within column (lowercase) indicate a significant difference ($p \leq 0.05$) according to Duncan's multiple range test.

In terms of shoot induction and proliferation, however, low levels of cytokinin did not promote shoot proliferation in callus after three months of culture, even under light condition. In order for the callus cells to regenerate into shoots, cytokinin levels need to be increased to 0.5 or 1.0 mg/ml BAP. This highlighted the role of higher concentrations of cytokinin in addition to the presence of light for shoot regeneration of *V. planifolia* callus from root tips. The calli formed from root tips were mostly compact, irregular and light yellow in colour (**Figure 2.9a-c**), turning from light green to dark green in the progress of differentiating into PLB masses (**Figure 2.9d-f**). Under histological analysis, it was observed that after two weeks of incubation in callus inducing medium, root caps in root tips started to disintegrate (**Figure 2.10a**). At the initial callus stage, cells were globular and highly vacuolar (**Figure 2.10b**). Heart-shaped embryo-like structures with meristematic globular cells (**Figure 2.10c**, **d**) started to form in 3-months-old callus after incubated in shoot regeneration media. Polysaccharide accumulation (dark purple stained granules in **Figure 2.10c-h**) was observed as the callus cells began to regenerate and shoot primordia (**Figure 2.10f**) was formed with vascular bundle formation (**Figure 2.10g-h**) in the more differentiated protocorm-like structures.

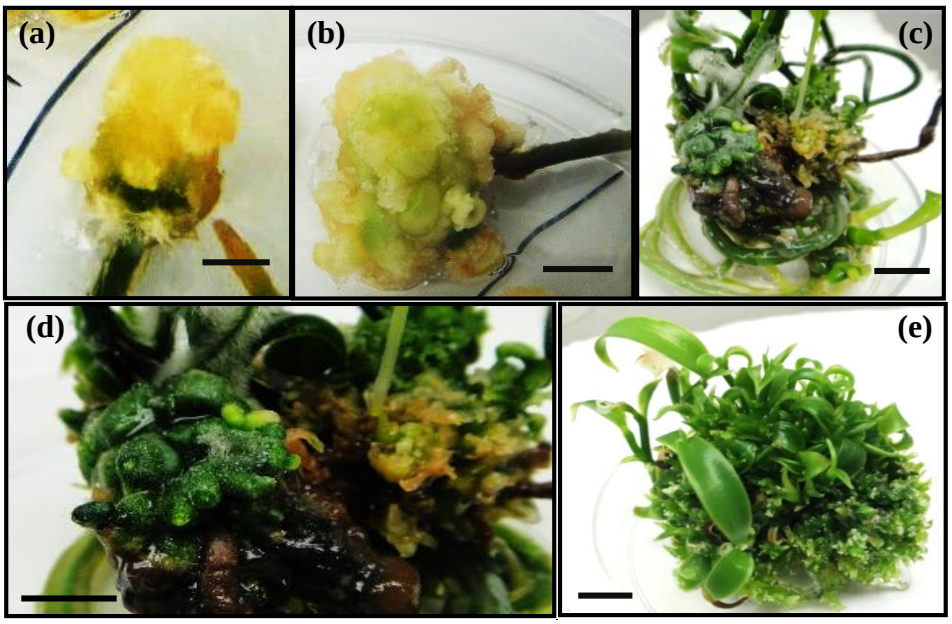


Figure 2.9. The progressive development of root-derived callus to PLB and multiple shoots mass of root tips of *V. planifolia*: (a) 90-days-old and (b) 150-days-old root-derived callus; (d-f) Differentiating shoot primodia (d,e) and shoots (f) derived from root callus culture. Bar = 0.5 cm.

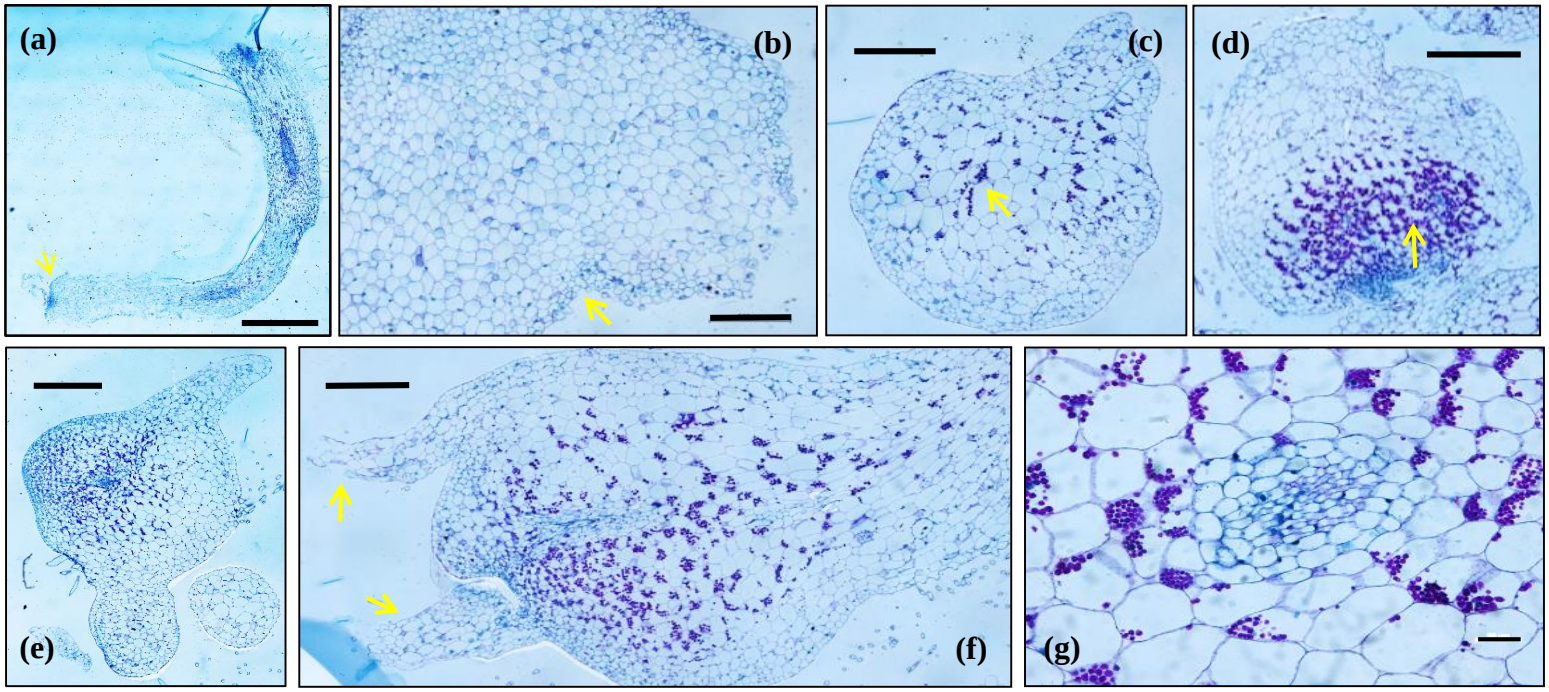


Figure 2.10. Light micrograph showing the sequence of development of callus from cross section of vanilla root tip: (a) Root tip (two weeks after incubation in callus inducing medium), arrow showing root tip, bar = 2500 µm; (b) Heart shape embryo-like structure, arrow showing tunica corpus organization, bar = 500 µm; (c - d) Meristematic cells with changes in starch accummulations (dark purple stained granules, as indicated by arrows), bars = 500 µm; (e) Meristematic structures forming shoot apex, bar = 500 µm; (f) shoot apex with leaf primodia and development of vascular tissues in protocorm-like structure, arrow shows distinct leaf primordia formation, bar = 500 µm; (g) vascular bundles formation in protocorm-like body structure, bar = 50 µm.

2.3.4. Leaf culture

Unfortunately, the leaf explants remained green with no growth responses observed in leaf explants despite being cultured in a range of hormonal treatments after 3 months. The leaf explants were totally not responsive. Both leaf tips and leaf bases were used but no responses were observed in either of the explants.

2.3.5. Seeds culture

Seed germination of vanilla was only observed on Knudson C medium, out of the three media tested. Out of a total of approximately 1,500 sowed seeds on Knudson C medium supplemented with 0.5 mg/L TDZ, only thirty-three embryos (2.2%) germinated after 16 months of seed culture (**Table 2.6**). Out of the 33 germinated embryos, only five of them (0.003%) managed to form actively proliferating protocorms (**Figure 2.11**). The protocorms were successfully maintained on hormone-free medium and continued to proliferate and formed an average of 11.4 ± 6.80 shoots, after 6 months of culture on hormone-free medium (**Table 2.6**).

Table 2.6. Vanilla seed germination, protocorm formation and number of shoot formation.

Culture duration	Seed germination	Protocorm formation	Number of shoots formed after 6 months
16th month	33 (2.200%)	5 (0.003%)	11.4 ± 6.80
18th month	14 (0.009%)	0 (0.000%)	0.00 ± 0.00

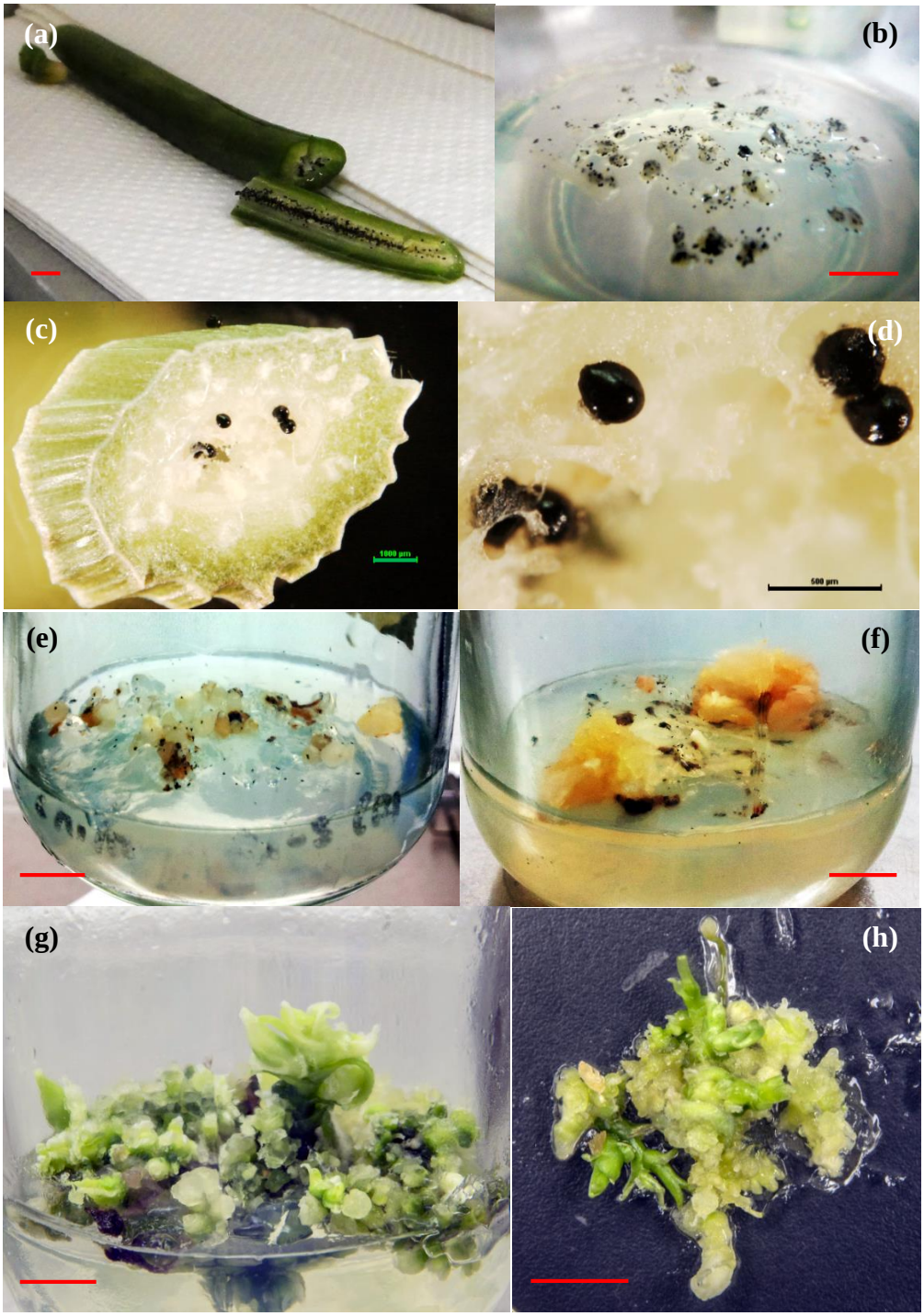


Figure 2.11. Vanilla seeds germination and PLB formation. (a, b) Surface sterilized vanilla pod (inner view, cultured on media); (**c-d**) Microscopic view of vanilla seeds (**e-f**) vanilla seed germination on Knudson C media with 0.5 mg/L TDZ after 16 months; Callus embryos germinated and **(g-h)** Protocorms generated from seed culture and maintained on hormone-free MS media. Red line denotes 1.0 cm, green line denotes 1000 µm (0.1 cm), and black line denotes 500 µm (0.05 cm).

2.4. Discussion

2.4.1. Establishment of aseptic culture of vanilla

Diseases are one of the main factors that damage and reduce vanilla production and productivity. The conditions of temperature and humidity under which vanilla grows tend to favour the development of pathogens, mainly fungi (Havkin-Frenkel and Belanger, 2011). Vanilla plantations were often devastated by root and stem rot (*Fusarium oxysporum* f. sp. *vanillae)*, black rot (*Phytophtora* sp.) and anthracnose (*Colletotrichum* sp.) (Havkin-Frenkel and Belanger, 2011). These airborne fungi could be trapped easily in the spaces between the axillary bud and leaves of vanilla nodes, which were not easily removable and could be easily carried into *in vitro* culture systems. This could be problematic as some fungus could even appear after one month of culture (**Figure 2.2f**). Fungal contamination was the first major problem faced in this study during the establishment of sterile culture from the shade house grown vanilla, and was successfully eliminated with increased carbendazim concentration. Carbendazim is a broad-spectrum benzimidazole fungicide, the breakdown product of thiophanate methyl and benomyl, having strong inhibition on *Fusarium sp.*, which was also the main fungal pathogen in vanilla culture, and also other fungal pathogens such as *Botrytis cinerea* (Bi *et al.*, 2009; Debergh *et al.*, 1993).

The second major issue faced in this experiment was the high bacterial contamination, up to 90%. The bacterial contamination did not all appear immediately after surface sterilization, but some could appear after two weeks to a month, and thus could potentially be endophytic bacteria. Vanilla, like other orchids, has systemic endophytes which could help in their growth and development especially during seedling development (Faria *et al.*, 2013). These bacteria were harmless inside the plants, even at the initial stages of *in vitro* culture system. However, as presented by Herman (1989), the usually non-pathogenic or even growth-promoting bacteria can become detrimental under the special growth conditions of *in vitro* culture, which was termed as

'vitropaths'. Our study supported this idea as it was observed thoroughout our studies that although some endophytic bacteria that occurred in vanilla did not affect plant growth, occasionally some bacteria tend to outgrow the plant, and eventually killed the plant and colonized the entire medium, thus affecting other downstream processes of the tissue culture process. Prolonged exposure to carbendazim from 20 min to an hour successfully reduced the endophytic bacterial contamination down to 10%. This study showed that although carbendazim is a fungicide, it also had bactericidal effects. This observation was supported by the reports on direct antibiotic effects of carbendazim by Garcia *et al.* **(**2002) and Wang *et al.* (2009c, 2009d, 2011).

Besides, not only having both fungicidal and bactericidal effects, carbendazim is also known to have enhancing effect on plant growth when used in low concentrations. Tripathi and Ram (1982) were the first to report enhancing effect of carbendazim on callus growth, root induction and shoot differentiations of *Daucus carota* culture. Furthermore, Shields and co-workers (1984) reported that carbendazim was also not phytotoxic for callus and root cultures of *Nicotiana tabacum* at low dosages. Debergh and co-workers (1993) stated a cytokinin-like activity of carbendazim, with enhanced shoot height and increased number of shoots in *Cordyline terminalis* and had no phytotoxicity effect of up to 160 µg/ml. Our study supported the statement that a low dosage (0.4% and below) and overnight incubation with carbendazim was harmless to vanilla.

Further characterization through isolation of pure bacterial colonies and Gram-staining test showed purple colonies (**Figure 2.3b**), signifying that the endophytic bacteria in vanilla cultures were Gram-positive. In vanilla, *Bacillus amyloliquefaciens* (*B. amyloquefaciens*), which was also Gram-positive (Fan *et al.*, 2012), was a systemic bacterial endophyte and a defensive mutualist of vanilla, which tend to densely colonize the shoot meristems (**Figure 2.12**) and stomatal areas of stems and leaves, as isolated and microscopically observed by White and co-workers (2014). However, further confirmation approaches, such as 16S DNA barcoding test (Barghouthi, 2011), need to be carried out to truly identify the bacteria species in order to further understand the endophytic

bacteria in this vanilla cultivar. There were controversial discussions on the topic of endophytic bacteria in plant tissue culture systems, where some researchers suggested that growth-promoting or stabilizing endophytic bacteria should be kept co-existing with the plants, not to be eliminated, in order to improve the plant survival and growth during critical phases of plant tissue culture (Quambusch *et al.*, 2014). For instance, Pirttilä and co-workers (2004) observed that freeing buds of *Pinus sylvestris* L. of bacterial endophytes resulted in an altered morphology, which could be restored by adding endophytic products to the plant medium. Probably further research investigations into beneficial endophytic bacteria could be done for boosting the vigour of micropropagated vanilla.

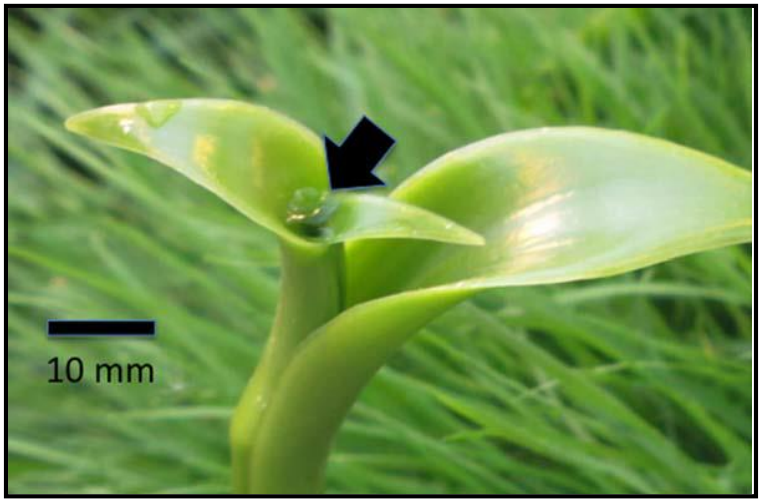


Figure 2.12. Shoot tip of a vanilla plant bearing *Bacillus amyloliquefaciens*. Arrow shows liquid exudate at the shoot opening of infected plant (White *et al.*, 2014).

2.4.2. Nodal culture – The direct and indirect regeneration

The initiation of shoots or callus from nodal explants was influenced by the growth regulator treatments (**Table 2.3**).

2.4.2.1. Direct shoot regeneration

In general, nodal culture of vanilla was very responsive to direct regeneration. The treatment, 1.0 mg/L BAP has been the most widely reported optimal treatment for direct regeneration of vanilla (Sreedhar *et al.*, 2007, Tan *et al.*, 2011a). Our findings agreed with the cited studies, as among the BAP treatments (0 to 5.0 mg/L), 1.0 mg/L BAP gave the optimal number of shoot formation (an average of 8.0 shoots per node).

2.4.2.2. Indirect shoot regeneration

In the combination with 2,4-D, 1.0 mg/L BAP and 1.0 mg/L 2,4-D gave an average of 50.0% callus response and an average of 8.5 shoots. At concentrations above 1.0 mg/L, 2,4-D caused browning of the original explants under tissue culture stress (**Figure 2.7**). This could be caused by the herbicidal properties of 2,4-D (Phua *et al.* 2016). The auxin 2,4-D is an effective plant cell regulator in inducing cell elongation and enlargement commonly used for callus induction (Phua *et al.* 2016). In the combination with NAA, 2.5 mg/L NAA and 1.0 mg/L BAP was the optimal hormone combination for callus induction (66.67% callus response) and its subsequent shoot regeneration (an average shoots number of 10.5). The results agreed with those of the studies by Tan *et al.* (2011a).

From the study, it was shown that among the different nodal protocols used, the performance of nodal culture was rather consistent with previous research work by other researchers (Tan *et al.*, 2011a; Janathanam and Seshadri, 2008). Indirect regeneration protocol was able to produce more shoots than the direct regeneration protocol. However, the genetic fidelity of the regenerated shoots needed to be analysed. The cultures established in this section were subjected to further analysis of somaclonal variations in the subsequent chapters. The treatments 1.0 mg/L BAP and 2.5 mg/L NAA in combination with 1.0 mg/L BAP were selected for direct and indirect regeneration, respectively, for nodal culture.

2.4.3. Root tips culture as the potential alternative regeneration

During the *in vitro* nodal cultures of *V. planifolia* plantlets on MS medium supplemented with 1.0 mg/L BAP, callus and PLB formation was sporadically observed in aerial root tips submerged in medium (**Fig. 2.13**). Therefore, further regeneration experiments were carried out using BAP, in combination with other auxins or cytokinins, to explore the potential of this explants in vanilla micropropagation.

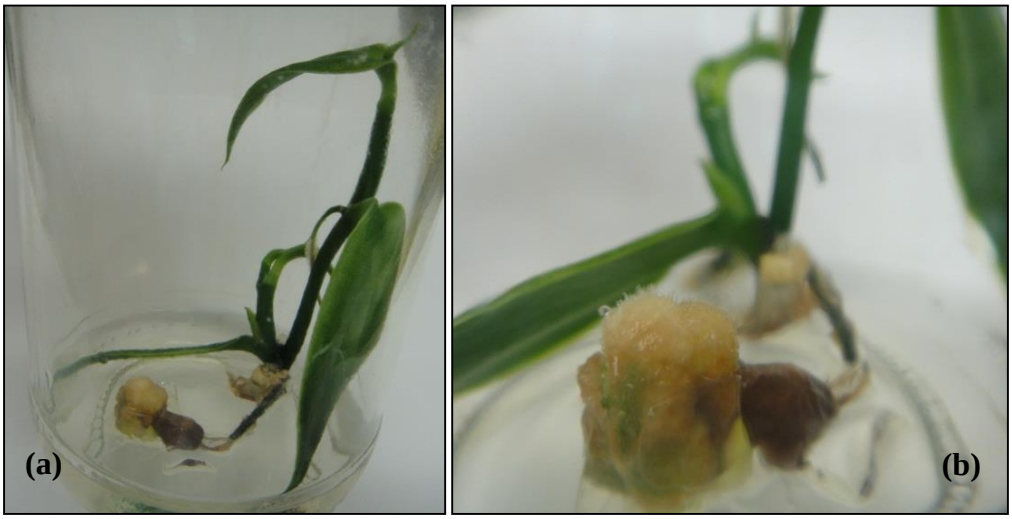


Figure 2.13. Sporadical callus formation (a-b) in aerial root tips of *V. planifolia* root tips partially submerged in MS medium supplemented with 1.0 mg/L BAP.

Root tips culture could be a useful alternative propagation method to nodal culture, with the major advantage to propagate the plants without sacrificing the monopodial mother plant, as opposed to the current nodal regeneration method (Philip and Nainar, 1988). Aerial root tips are also explants of plentiful supply, with at least one root tip present on each node of a healthy vanilla plant (**Figure 1.2**) (Arditti, 2009). The aerial roots of vanilla, like other epiphytic orchids, are rather unique, with the entire length of the root, except the tip, covered by dead and spongy-like velamentous tissues which function for water and nutrient uptake, as well as for protection against UV-B damage,

underneath which lies the cortex region containing chloroplasts (Chomicki *et al.* 2014, Zotz and Winkler 2013). The orchid roots may form as much as half the total plant mass, with roles in absorption of mineral, carbon fixation, water conservation and anchorage (Zhang *et al.* 1995).

Plant hormones are important endogenous factors for the integration of various environmental inputs and regulation of plant development (Benková and Hejátko 2009). In this research, hormonal study has highlighted the important role of cytokinin in callus induction of *V. planifolia* root tips. From the first experiment (**RT-A**, **Table 2.4**), it was shown that high callusing rate (86.67 to 100.00%) were achieved with cytokinin treatments alone (TDZ or BAP) of root tips, while under auxin treatments alone (NAA or 2,4-D), root tips neither elongated nor callused. In the second experiment (**RT-B**, **Table 2.5**), the role of cytokinins in vanilla callus development is further supported. Although the combination of auxins and cytokinins produced callus at the rate of 20.00 to 100.00%, the average sizes of callus formed in auxin-cytokinin combinational treatments were much smaller than the cytokinin-only treatments after 90 days. This further supports that in vanilla root tips, exogenous application of auxins (NAA or 2,4-D) exhibited inhibitory effects on callus that required elevated cytokinin levels to reach a balanced auxin-cytokinin ratio for forming callus.

In general, an intermediate ratio of auxin to cytokinin promotes callus induction, while the presence of high ratio of was known to promote callusing in *in vitro* plants (Skoog and Miller 1957). A comparison with callus induction using nodal explants of *V. planifolia* (**NT**, **Table 2.3**) has shown that callus formation in nodal explants was formed at higher rate in high auxin to cytokinin ratios, which differed from the response of root tips culture which was formed at higher cytokinin to auxin ratios. This phenomenon unique to root tips could be explained by the possible presence of high endogenous auxin levels in excised root tips, which has instead promoted root elongation (at an average growth of up to 2.0 cm in four weeks in control treatments.

Vanilla aerial roots, like the aerial root tips of other Crassulacean acid metabolism (CAM) orchids, might contain high endogenous auxin, cytokinin or

abscisic acid contents (Zhang *et al.* 1995), which function to allow for its elongation, formation of lateral roots, gravitropism and control of root apical meristems (Aloni *et al.* 2006). This hypothesis could be supported by a study conducted by Neelannavar (2006), where the scanning of IAA using ultraviolet (UV), thin layer chromatography (TLC), gas liquid chromatography (GLC) and gas chromatography-mass spectrometry (GC-MS) in vanilla root tips extracts showed higher levels of auxin in root tips from old aerial roots and also young culture root tips of vanilla than other tissues.

The interaction between auxin and cytokinin is important in callus development and differences in pre-existing hormone levels could be the reason of inconsistency in callus responses. From this experiment, it was suggested that certain auxin to cytokinin ratio plays a role in callus response of *V. planifolia in vitro* cultures and the application of cytokinin is required for callus formation in *V. planifolia*. The different levels of pre-existing endogenous auxin in root tips of *V. planifolia* could counter-react with the application of exogenous auxins. When TDZ or BAP is applied with NAA or 2,4-D, different auxin to cytokinin ratios gave different levels of callus response. More consistent results are achieved when TDZ or BAP is applied alone, as compared to their combinations with auxins 2,4-D and NAA.

In the past research, vanilla root culture was rarely reported. Philip and Nainar (1986) reported the development of direct shoot regeneration from root tips of vanilla. However, the research paper had focused on development, but not on the practical procedure of micropropagation (Arditti, 2006). Callus induction in root tips had not been reported. We proposed that this is because previous studies had focused on the use of high auxins, while our findings have shown that cytokinin, applied solely and/or in combination with lower auxin ratio, efficiently generate calli in vanilla, as compared to the PGR-free medium or medium with high auxin-to-cytokinin ratio. In this study, formation of root tips-derived calli initiated shortly after 90 days, which was about the same time period as the callusing rate of vanilla nodes, but at a much higher rate of callusing. This showed that root tip could be a potential explant to be utilized when callus regeneration pathway is required.

In this study, we reported a novel approach to efficiently induce indirect regeneration in vanilla root tips, where a rate of up to 100% callusing rate was reported using cytokinin-medium, and the callus response rate were also higher, with callus initiation observed within two weeks to one month, as compared to 3 months as observed in nodal explants by previous studies. This finding will serve as a potential tool for the mass propagation of *V. planifolia* to be utilized both commercially and for further applications such as vanilla crop improvement or elicitations of vanillin, as well as for studies of callus and PLB development. However, the use of root tips for commercial clonal propagation needs further validation for somaclonal variations as it goes through the callus-phase.

2.4.4. Low germination rate in seeds culture

Seed germination and survival rates of orchid seeds are naturally low. In this study, the vanilla seeds germination occurred only after 16 months of culture. Although a total of 33 seeds germinated (**Figure 2.11e-f**), only five of them (0.003%) successfully developed into green, actively-dividing protocorms (**Figure 2.11g-h**). The three media tested were based on literature review, where Knudson (1950) used Knudson C medium for vanilla seeds germination, while Parra-Quezada and co-workers (1990) claimed that MS medium performed better than Knudson for vanilla seed germination and seedling growth. The presence of cytokinin for germination of vanilla seeds is considered to be related to the utilization of lipids that constitute the primary storage material in most orchid seeds and it has been observed that unless the storage lipid is utilized, germination does not continue (De Pauw *et al.*, 1995). Out of the three media tested, germination and growth of vanilla PLB was best achieved with Knudson C medium. However, due to small replication size, the effect of media types and hormone on vanilla seeds germination and growth could not be concluded. Further studies on seeds culture could be done with larger sample size and inclusion of 2.0 g/L tryptone (Minoo, 2002), which were proposed to have promoting effect on the size and development of protocorm.

2.5. Conclusions and future prospects

Establishment of sterile culture was achieved with slight modifications of the existing protocol by Tan *et al.* (2011a). The main problems faced, fungal pathogen and endophytic bacteria contamination, were resolved. Fungal contamination was eliminated and bacterial contamination rate was reduced from 90% to 10% when carbendazim concentration and duration was increased from 0.2% for 20 min to 0.4% for an hour. The low phytotoxicity rate also showed that this treatment did not cause negative effect on the plant growth, which was supported by numerous literature (Shields *et al.* 1984; Tripathi and Ram 1982; Debergh *et al.* 1993). Further research investigations into beneficial endophytic bacteria isolated from vanilla culture are suggested for future improvements on the vigour of micropropagated vanilla.

In this study, we also reported a novel approach to efficiently induce indirect regeneration in vanilla root tips, where a rate of up to 100% callusing rate was reported using cytokinin-containing medium, and the callus response rates were also higher, with callus initiation observed within two weeks to one month, as compared to 3 months as observed in nodal explants. We propose that root tips could be an alternative explant to be utilized for vanilla mass propagation or for further applications, such as vanilla crop improvement or elicitations of vanillin, as well as for studies of callus and PLB development. However, the use for commercial clonal propagation needs further validation for somaclonal variations as this root tips regeneration protocol goes through the callus phase.

Both seeds and leaf culture have not been successful in this study. Seed germination occurred only after 16 months. Seeds germinated on Knudson C medium with 0.5 mg/L TDZ successfully produced five protocorms which formed an average of 11.4 ± 6.80 shoots after three months of germination. Although a total of 47 embryos germinated after 18 months, only 5 survived. Further studies on seeds culture with larger sample size and inclusion of 2.0 g/L tryptone to promote size and development of protocorm are proposed to improve this study. Leaf cultures were not responsive, which could be genotype-specific.

Nodal culture of vanilla was successfully established using protocols reported by Tan *et al.* (2011a). Optimal direct shoot regeneration through nodal culture was established with 1.0 mg/L BAP, with an average of 8.0 shoots formed after 90 days of culture; while optimal callus regeneration protocol was established with 2.5 mg/L NAA and 1.0 mg/L BAP at the callusing rate of 66.67% and an average shoot regeneration of 10.5 per node. The treatment with 1.0 mg/L BAP and 2.5 mg/L NAA were select as the optimal treatment representative for indirect shoot regeneration, while treatment with 1.0 mg/L BAP was selected as the representative treatment for direct shoot regeneration of vanilla, which were used for further analysis of somaclonal variations in the subsequent chapters.

CHAPTER III

MORPHOLOGIC, GENETIC AND EPIGENETIC STABILITIES OF VANILLA PLANTS DERIVED FROM DIRECT AND INDIRECT NODAL AND ROOT REGENERATION AS REVEALED BY ISSR AND MSAP

Chapter III

Morphologic, genetic and epigenetic stabilities of vanilla plants derived from direct and indirect nodal regenerations as revealed by ISSR and MSAP

3.1. Introduction

Somaclonal variations are a common phenomenon in micropropagation, such as observed in commercially important banana and oil palm industries. Somaclonal variation has not been well-studied in the micropropagation of *V. planifolia.* Sreedhar and co-workers (2007) conducted a study on multiple shoots regenerated from axillary bud explants established for 10 years in the lab and screened the plants using 45 RAPD and 20 ISSR primers, and concluded that there were no genetic polymorphisms. However, studies later have shown that there are more factors that could contribute to somaclonal variations, such as epigenetic alternations and cytological aberrations. From **Chapter 2**, two types of regeneration pathways were optimised for vanilla using the nodes as explants, which include a direct and an indirect regeneration method. The indirect regeneration system produced a significantly higher number of regenerated shoots (at an average of 10.5 shoots per node in 6 months), as compared to the direct regeneration system (which produced an average of 8.0 shoots per node in 6 months). In this study, the ISSR and MSAP marker systems were used to evaluate the genetic and epigenetic fidelity of vanilla plants produced from these two regeneration systems, in order to determine the suitability of the micropropagation protocols.

The inter-simple sequence repeats (ISSR) marker system is a well-established and effective genetic marker system widely employed in genetic diversity studies. In general, the ISSR technique combines most of the benefits of amplified fragment length polymorphisms (AFLP) and microsatellite analysis with the universality of RAPD, while this method does not require pre-knowledge of genome sequences and is relatively simple and cost-effective, not requiring sophisticated detection systems (Reddy *et al.*, 2002). It has been applied in SV screening of many commercially important crop species, including *Oryza sativa* (Ngezahayo *et al.*, 2007), *Musa spp.* (Ray *et al.*, 2006; Lakshmanan 2007) and *Zea mays* (Ramakrishnan *et al.*, 2014).

The methylation-sensitive amplified polymorphisms (MSAP) marker system is a modified protocol from the AFLP system, which incorporated the use of methylation-sensitive restriction enzymes for detection of cytosine methylation in plants (Fulneček and Kovařík, 2014). It is also a relatively simple and cost-effective approach which could be incorporated into routine commercial screening of SV of many commercially important plant species, such as *Coffea arabica* (Landey *et al.*, 2015), *Malus pumila* (Hao *et al.*, 2001), *Saccharum officinarum* (Francischini *et al.*, 2017), *Oryza sativa* (Wang *et al.*, 2009a, 2013). Epigenetic polymorphisms have been an overlooked aspect in vanilla SV screening as so far there isn't any report on using epigenetic markers to screen SV in vanilla.

Both genetic and epigenetic mutations are important, contributing to economically important traits. Taking oil palm (*Elaeis guineensis*) as an example, the mantled flower caused by SV was caused by epigenetic misregulations (Jaligot *et al.*, 2000, 2004), and had caused severe crop losses which were only detected when the plant reached maturity. Vanilla shared the common with oil palm that it is also a perennial crop, where the plant would only flower and produce vanilla pods, after reaching maturity at least 3 years after field plantation (Castro-Bobadilla and García-Franco, 2007; Diez *et al.*, 2016). Therefore, both genetic and epigenetics are important aspects to consider when screening for SV in vanilla.

The main aim of this project is to investigate the occurrence of somaclonal variations (SV) during direct and indirect nodal plant regeneration of *V. planifolia* at morphological, genetic and epigenetic levels, by using the inter-simple sequence repeats (ISSR) and methylation-sensitive amplified polymorphisms (MSAP) marker techniques.

The specific objectives in order to achieve the main aims are: (i) morphological pre-screening for phenotypic abnormalities in *in vitro* cultured vanilla; (ii) isolation of high molecular weight genomic DNAs from vanilla mother plants (MPs) and *in vitro* vanilla cultures; (iii) optimization of the ISSR and MSAP protocols for genetic and epigenetic analysis; and (iv) molecular screening for genetic and epigenetic instabilities in tissue culture vanilla plantlets regenerated from direct and indirect regeneration pathways using optimized ISSR and MSAP protocols.

3.2. Materials and methods

3.2.1. Chemical reagents and apparatus

The list of standard solutions, enzymes and kits that were used in this study is stated in **Appendix A.** The list of equipment and apparatus that were used in this study is stated in **Appendix B**. Molecular grade chemicals, which are DNAse and RNAse-free, such as nuclease-free water, were used.

3.2.2. Oligonucleotides Synthesis and Primers Design

The adaptors and primer oligonucleotides used in this study were synthesized and de-salted by a commercial service provider company (Integrated DNA Technologies (IDT), Singapore (Sg)). The sequences and source references are listed in **Appendix E**. Primer sequences of inter-simple sequence repeat (ISSR) markers were obtained from published literature (Sreedhar *et al.*, 2007; Jhang *et al.* 2010). For methylation-sensitive amplified polymorphism (MSAP) markers,

all possible *Eco*RI and *Hpa*II/*Msp*I +1 and +2 primers were pre-screened (**Appendix E)**.

3.2.3. Plant materials

Methods of the establishment of tissue culture regenerations are described in **Chapter 2**. Vanilla calli were induced from nodal explants on callus inducing medium (CIM) (MS medium containing 2.5 mg/L NAA and 1.0 mg/L BAP) for indirect regeneration of vanilla (hence denoted as "iVp"); while the direct multiple shoot regeneration was established on shoot inducing medium (SIM) (MS medium containing 1.0 mg/L BAP) for direct regeneration of vanilla (diVp). In the genetic screening study, 20 iVp and 3 diVp regenerants from tissue culture of vanilla were screened for SV using ISSR markers. Among the 20 iVp regenerants, samples are randomly selected, except for samples 18-20, which included morphologically distint vanilla plants (as described in **Section 3.3.1**) with long narrow leaf morphology, small broad leaf morphology and abnormal shoot branching respectively. In the epigenetic screening study, both iVp and diVp regenerants from three biological replicates (denoted as "iVp1 to 3" and "diVp1 to 3" respectively) were randomly selected and screened for SV using the MSAP markers. For the iVp regenerants, three shoots were randomly selected from each callus (denoted as "iVp1-1 to 3", "iVp2-1 to 3" and "iVp3-1 to 3" for calli samples 1 to 3 respectively). In both genetic and epigenetic studies, the mother explant source was denoted as "MpVp"; and *V. graffiti*, a closely-related vanilla species were used as control for marker analysis (denoted as "Vg")."

3.2.4. Overall Methodology

An overall flow of research methodology in this part of the thesis study is presented in **Figure 3.1.** Genomic DNA extraction, quantification and quality check were performed prior to molecular marker analysis. High molecular weight genomic DNA obtained was then subjected to subsequent inter-simple sequence repeats (ISSR) and methylation-sensitive amplified polymorphisms (MSAP) analysis.

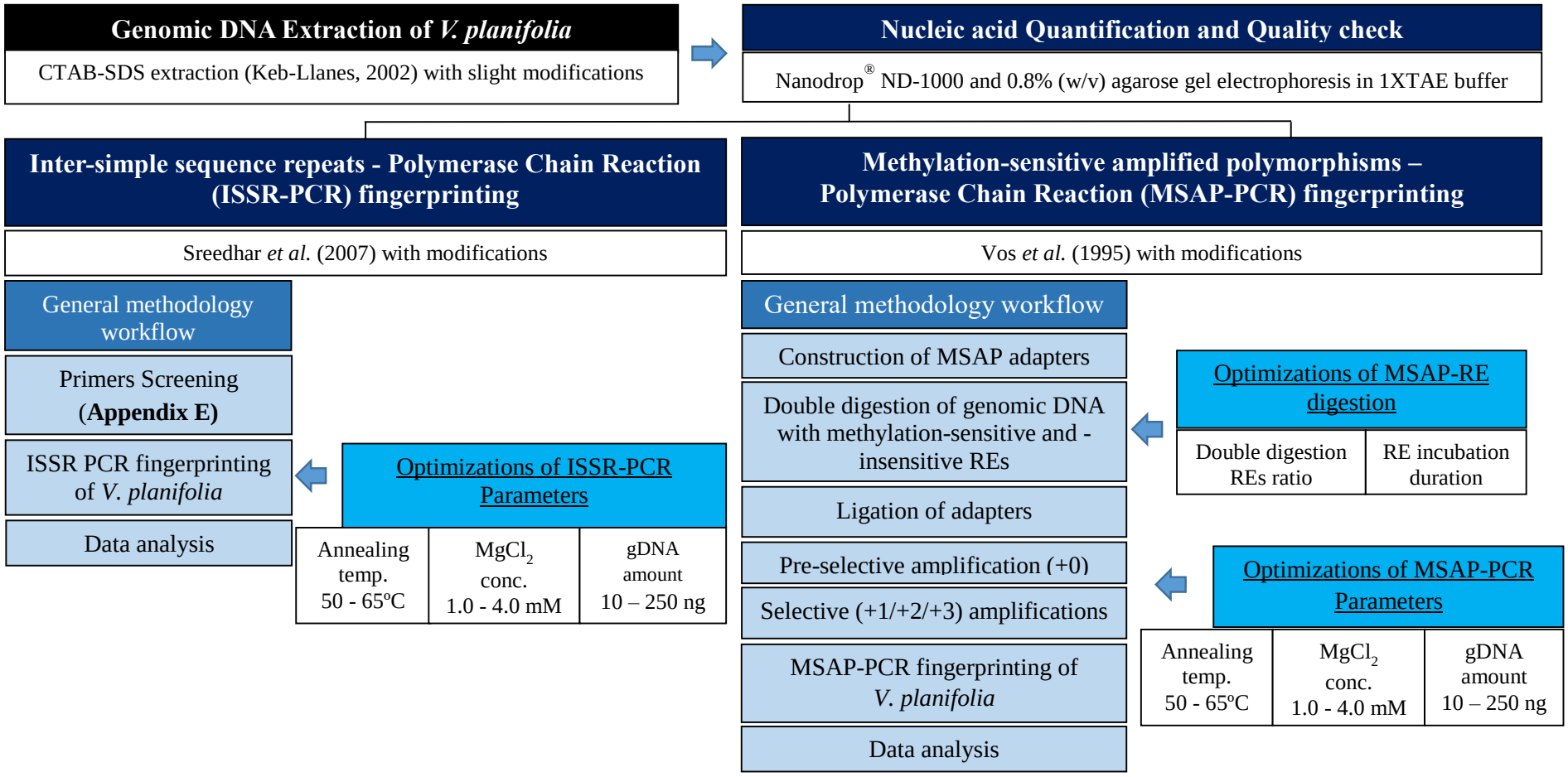


Figure 3.1. Overview of genetic and epigenetic molecular markers analysis workflow of vanilla.

3.2.5. Genomic DNA extraction of *V. planifolia*

Conventional DNA extraction of *V. planifolia* explant samples was carried out using the CTAB-SDS method developed by Keb-Llanes (2002) with slight modification. Leaves, nodes and newly differentiating calli, PLBs and leaf tissues were ground into fine power in liquid nitrogen using mortars and pestles, followed by incubation in extraction buffer (Buffer composition as listed in **Appendix C**) at 65°C for 30 min. Cold potassium acetate (Sigma Aldrich, USA) was added, mixed thoroughly and centrifuged at 13,000 g for 15 min at 4°C. Supernatant was then transferred to a clean Eppendorf tube, and added with equal volume of Phenol: Chloroform Isoamyl alcohol (PCI) (25:24:1), mixed thoroughly and spun down at 13,000 g for 20 min at 4ºC. 0.6 vol of 3.0 M sodium acetate (pH 5.2) (Sigma Aldrich, USA) and 5.4 vol cold isopropanol were added to precipitate the DNA in supernatant with incubation on ice for 30 min. The mixture was centrifuged at 13,000 g for 15 min to obtain pellet, which was then washed with 1 ml of 70% ethanol and vacuum dried. Purified DNA pellet was subsequently re-suspended in 50 µl $TE_{0.1}$ buffer. RNase treatment was performed after the extraction process to further purify the DNA from RNA contamination. 1.0 µl of RNAse A (DNase and protease-free) (Thermo Scientific, USA) was added to each purified DNA sample, and incubated at 37°C for 30 min in a thermoblock.

3.2.6. Nucleic acid quantification and quality check

The DNA was quantified spectrophotometrically using Nanodrop® ND-1000, while DNA integrity was checked by running samples on 0.8% agarose gel (Sambrook and Russell, 2001).

3.2.7. Inter-simple sequence repeats - Polymerase Chain Reaction (ISSR-PCR) fingerprinting

ISSR-PCR amplifications were performed in a total volume of 20.0 µl per reaction, which included 20 ng DNA, 1X PCR buffer, 0.2 mM each dNTPs, 3.0

mM $MgCl_2$, 0.5 µM of each primer and 1 U GoTaq^{TM} Flexi DNA polymerase (Promega, USA). PCR amplification was performed at initial step of 4 min at 94°C, 40 cycles of 1 min at 94°C, 1 min at 51.8°C, 2 min at 72°C and a final step of 10 min at 72°C. The amplified products were separated on a 2.0% (w/v) agarose gel in 1X TBE buffer, pre-stained with 3X SyBr Safe, and photographed under UV light by using a gel documental system.

3.2.7.1. Optimization of PCR conditions

Two main factors of ISSR-PCR, namely the concentration of $MgCl_2$ and the annealing temperature of the primers, were investigated. Optimization of $MgCl_2$ concentration was performed at a gradient concentration of 1.0, 1.5, 2.0, 2.5, 3.0, 3.5 and 4.0 mM; while the annealing temperature gradient was performed at the gradient temperature range of 50 to 60°C; to determine the best salt concentration and annealing temperature for the PCR. The optimal PCR conditions were used for analysis of vanilla mother plants and its tissue culture regenerants.

3.2.7.2. Statistical analysis

The ISSR-PCR products were scored manually either as present (1) or absent (0) in a binary matrix. Only consistent, reproducible, distinct and well-resolved fragments were scored. Smear and weak bands were excluded. The potential of marker for estimating genetic variability of *V. planifolia* was recorded by counting of bands. Primer banding characteristics including total number of bands (TB), number of polymorphic bands (PB) and the percentage of polymorphic bands (PPB) were obtained. The performance of ISSR markers was measured using three parameters: polymorphic information content (PIC), marker index (MI) and resolving power (RP). The PIC value for each locus was calculated using the following formula (Roldan-Ruiz *et al.*, 2000):

$$\mathrm{PIC}_i = 2f_i(1 - f_i)$$

Where PIC_i is the polymorphic information content of the locus i, f_i is the frequency of the amplified fragments and 1 - f_i is the frequency of non-amplified

fragments. The frequency was calculated as the ratio between the number of amplified fragments at each locus and the total number of accessions (excluding missing data). The PIC of each primer was calculated using the average PIC value from all loci of each primer. Effective multiplex ratio (EMR) was calculated using the formula:

$$\mathrm{EMR} = n\ x\ \beta$$

Where *n* is the average number of fragments amplified by accession to a specific system marker (multiplex ratio) and β is estimated from the number of polymorphic loci (PB) and the number of non-polymorphic loci (MB):

$$\beta = \frac{\mathrm{PB}}{\mathrm{PB} + \mathrm{MB}}$$

The marker index (MI) was calculated to characterize the capacity of each primer to detect polymorphic loci among the genotypes. Marker index for each primers was calculated as a product of PIC and EMR (Varshney *et al.*, 2007):

$$\mathrm{MI} = \mathrm{EMR}\ x\ \mathrm{PIC}$$

3.2.7.3. Analysis of genetic diversity

The data matrix of ISSR marker was then converted into genetic similarity matrix using Jaccard coefficient (Jaccard, 1908), and the genetic parameters: percentage of polymorphic band(s) (PPB), observed number of alleles (N_a), the effective number of alleles (N_e), Nei's genetic diversity (H), Shannon's information index (I), total genetic diversity (H_t), genetic diversity within population (H_s), Nei's genetic differentiation index among populations (G_{st}), and gene flow estimates between populations (N_m) were determined with POPGENE version 1.32 (Yeh, 1999). Genetic divergence between somaclone populations of *in vitro* vanilla (SV) and their mother plants (MP) was investigated using Nei and Li 's unbiased genetic distances and genetic identities (Yeh, 1999), which was utilized for the construction of a dendrogram using UPGMA (Unweighted Pair Group Arithmetic Mean Method). Bootstrap analysis (Felsenstein, 1985) using UPGMA search with 1,000 replicates was performed to obtain the confidence of the tree and was viewed with FigTree software version 1.4.3

(Rambaut, 2009). Analysis of molecular variance (AMOVA) (Excoffier *et al.*, 1992) was used to calculate variation among and within the vanilla populations using GenAlEx (Peakall and Smouse, 2006) version 6.503.

3.2.8. Methylation-sensitive amplified polymorphisms - Polymerase Chain Reaction (MSAP-PCR) fingerprinting of *V. planifolia*

The MSAP-PCR fingerprinting was performed with reference to the AFLP protocol developed by Vos *et al.* (1995), with slight modification. The restriction enzymes used in this study were *Eco*RI (as rare cutter), and *Hpa*II and *Msp*I (as frequent cutters). MSAP profiles of *V. planifolia* (*Vp*) tissue culture regenerants were compared with its mother plant (MP) from which the samples were regenerated from. A close species, *Vanilla graffiti* (*Vg*) was used as a control for the markers.

3.2.8.1. Construction of MSAP adapters

To construct double stranded MSAP adapters, the *Eco*RI forward and reverse adapters were mixed to a final concentration at 5 µM, and the *Hpa*II/*Msp*I forward and reverse adapters were mixed to a final concentration of 50 µM; and each of the mixture was heated to 95°C for 5 min to denature, and allowed to cool slowly at a decrement of 1°C per min to allow complete renaturation. This step was performed using a PCR thermocycler.

3.2.8.2. DNA double digestion with methylation-sensitive and –insensitive REs

To detect MSAP, two digestion reactions were set up at the same time for each genomic DNA sample. In the first reaction, approximately 500 ng of the genomic DNA was digested with 2.5 U of Fastdigest® *Eco*RI (Thermo Scientific,

USA) and 2.5U of Fastdigest® *Msp*I (Thermo Scientific, USA) in a final volume of 50µl reaction mixture containing 1X Fastdigest® reaction buffer and 0.1 mg/L Bovine Serum Albumin (BSA) (Promega, USA), and incubated at 37°C for 20 min. This digestion condition was an optimized condition from a preliminary test as stated in Section 2.2.3. The second digestion reaction was carried out in the same condition as above, however Fastdigest® *Hpa*II (Thermo Scientific, USA) was used instead of *Msp*I.

3.2.8.3. Adapter Ligation

Each digestion mixture was then ligated to double-stranded adapters to the ends of the restriction fragments, by adding 1.25 µg of each digested mixture to 5U of T4 DNA ligase (Thermo Scientific, USA) in a 50µl reaction containing 1X T4 DNA ligase buffer, 0.1 mg/ml BSA (Promega, USA), 5.0 µM of *EcoR*I adapter and 50.0 µM of *Hpa*II/*Msp*I adapter. The digestion-ligation (dig-lig) mixture was then incubated at 37°C for five hours to allow complete ligation. Finally, the dig-lig mix was diluted 10 fold with nuclease-free water prior to the next step of PCR amplification.

3.2.8.4. Pre-selective amplification (+1)

The pre-selective amplification was performed in a Nexus Gradient thermocycler (Eppendorf, Germany) with the following cycle regime: an initial hold of 2 min at 72°C, followed by 35 cycles of 30 sec at 94°C, 60 sec at 56°C and 6 sec at 72°C, 2 min at 72°C and a final hold of 10 min at 60°C. The PCR reaction was prepared in a total volume of 25 µl reaction mixture consisting of 1X GoTaq™ Flexi Green buffer (Promega, USA), 1.5 mM of $MgCl_2$, 0.2 mM of dNTPs (Promega, USA), 0.25 µM of each primer, 5.0 µl of diluted dig-lig mixture and 1.5 U of GoTaq ™ Flexi Taq Polymerase (Promega, USA). The primer sequences are listed in **Appendix E**. 15.0 µl of the +1 reaction was run on a 3.0% (w/v) agarose gel.

3.2.8.5. Selective amplification (+2)

The cycle conditions included an initial denaturation of 94°C for 2 min; a first cycle of 94°C for 30 sec, 65°C for 30 sec and 72°C for 2 min followed by 12 cycles of a stepwise decrease of the annealing temperature in each cycle by 0.7°C; 24 cycles of 94°C for 30 sec, 51.8°C for 30 sec, 72°C for 2 min; and a final elongation at 72°C for 10 min. The PCR reaction was prepared in a total volume of 25 µl reaction mixture consisting of 1X GoTaq™ Flexi Green buffer (Promega, USA), 2.0 mM of $MgCl_2$, 0.5 mM of dNTPs (Promega, USA), 1.25 µM of each primer, 5.0 µl of diluted pre-selective amplification reaction mixture and 1.5 U of GoTaq® Flexi Taq Polymerase (Promega, USA). The primer sequences are listed in **Appendix E**. 15.0 µl of the +2 reaction was run on a 3.0% (w/v) agarose gel.

3.2.8.6. Optimization of RE digestion and PCR conditions

The RE double digestion condition was optimized for its incubation time (20 min and 16 h) and ratio of EcoRI/MspI enzymes (1:1, 1:2, 1:3, 1:4, 0:2, 2:1, 3:1 and 4:1) in a single double digestion reaction. Three main factors of MSAP-PCR, namely the concentration of template DNA and $MgCl_2$, and the annealing temperature of the primers, were investigated using randomly selected pre-selective primer pair. Optimization of DNA concentration was performed at 6.25 ng, 12.50 ng, 25.00 ng, 37.50 ng and 50.0 ng; while the annealing temperature gradient was performed at the temperature range of 50-60°C to determine the best annealing temperature for the PCR.

3.2.8.7. Screening of MSAP pre-selective (+1) and selective (+2) primers

Sixteen pre-selective amplification primer pairs and 64 selective amplifications (+2) primer pairs were screened and compared for the choice of best primers to be used for real samples analysis. The optimized RE and PCR conditions, and

shortlisted primers from pre-primer screening were used for analysis of vanilla mother plants and its tissue culture regenerants.

3.2.8.8. Gel Electrophoresis

The amplified products were separated on a 3.0% (w/v) agarose gel in 1X TBE buffer, pre-stained with 3% (v/v) SyBr Safe (Invitrogen, USA). After gel electrophoresis, the gels were viewed under the UV transilluminator and images were captured using the FluorChem® HD 2 Gel Imaging systems (Alpha Innotech, USA).

3.2.8.9. MSAP data scoring and analysis

Only bright bands that could be clearly distinguished as present or absent on the gel for all cultures were scored. Polymorphisms were scored visually as present (1) or absent (0). The genetic DNA polymorphisms between *V. planifolia* (Vp) tissue culture regenerants were analyzed by comparing the *Msp*I (methylation-insensitive RE)-digested MSAP-PCR profile with its mother plant (MP) genetic profile. The analysis was performed using the Multi-variate statistical package (MVSP) software package. Data were compiled as a binomial matrix with individual and polymorphic loci. Cluster analysis was carried out using a distance matrix based on the genetic distance of Nei and Li (1979) and the unweighted pair group method with arithmetic average (UPGMA).

3.2.8.10. MSAP data interpretation

The epigenetic DNA polymorphisms of Vp tissue culture regenerants were analyzed by comparing the same regeneration line between *Msp*I-digested and *Hpa*II-digested MSAP-PCR profile (**Table 3.1**). A DNA methylation event was detected when bands present in the profile from the reaction *Eco*RI (E) + MspI (M) were absent from the reaction *Eco*RI + *Hpa*II (H) (Peraza-Echeverria *et al.*, 2001). This indicated that the internal cytosine was methylated (5'-C^{m}CGG-3'). The contrary situation, where a band was present in H but absent in M, indicated

that the external cytosine of one DNA strand was methylated (5'-^{m}CCGG-3'). This is regarded as the hemimethylated state. A DNA methylation event was considered to be a polymorphism when at least one plant differed from the others in the methylation pattern, for example, a band was present in both M and H in one plant but was only found in M in other plants. When, in one or some samples, no bands were detected in both digestions this was not considered as a polymorphism due to the fact that it is not possible to discard the presence of an AFLP event. Events in which there was an obvious difference in band intensity were also considered as DNA methylation polymorphisms The proportion of polymorphisms (*P*) was estimated by division of the number of DNA methylation polymorphisms by the total number of bands (Peraza-Echeverria *et al.*, 2001).

Table 3.1. MSAP-profiled methylation types and the scoring scheme (Sun *et al.*, 2014).

Type	Restriction pattern		Methylation status[c]	Scoring scheme		
	H[a]	M		unmethylated	^{HMe}CG/ ^{Me}CG	^{HMe}CCG
I	+[b]	+	unmethylated	1	0	0
II	-	+	^{HMe}CG/^{Me}CG	0	1	0
III	+	-	^{HMe}CCG	0	0	1
IV	-	-	^{Me}CCGG/ ^{HMe}CH^{Me}CGG/ ^{Me}C^{Me}CGG/ mutation	0	0	0

[a] H and M indicate the enzyme combinations of EcoRI/HpaII and EcoRI/MspI; [b] +: band presence; -: band absence; [c] HMe: hemi-methylated on only one strand; Me: fully methylated on both strands.

3.3. Results

3.3.1. Morphological abnormality of *in vitro* cultured *V. planifolia*

Approximately 50% of the callus-derived vanilla plants display morphological abnormalities. The abnormalities were characterized by several morphological traits including long narrow leaf morphology (~20%), small broad leaf morphology (~15%) and abnormal shoot branching (~5%). **Table 3.2** summaries the morphological selection criteria for normal vanilla and vanilla somaclones (**Figure 3.2**), which based on the leaf width to length ratio, leaf width to stem width ratio and shoot branching.

Table 3.2. Morphological description of tissue culture vanilla somaclones.

Type of somaclone	Morphological selection criteria
Normal vanilla plant (Broad leaf) (**Figure 3.2c**)	Leaves are broad and relatively larger than other leaf somaclones, with average leaf width to length ratio of 1:2 and leaf width larger than stem width; single terminal shoot apex and one node branches at a time.
Small broad leaf (**Figure 3.2b**)	Similar average leaf width to length ratio with normal vanilla plant (broad leaf), however, average leaf sizes are relatively smaller than normal broad leaf vanilla (**Figure 3.2e**), with leaf width is equal or smaller than stem width.
Long narrow leaf (**Figure 3.2d**)	Leaf morphologies are relatively longer and narrower than normal broad leaf vanilla (**Figure 3.2f**), with average leaf width to length ratio equal or less than 1:5.
Abnormal shoot branching	Multiple shoot apices at the shoot tip terminal (**Figure 3.2g**) or abnormal branching at every node (**Figure 3.2h**).

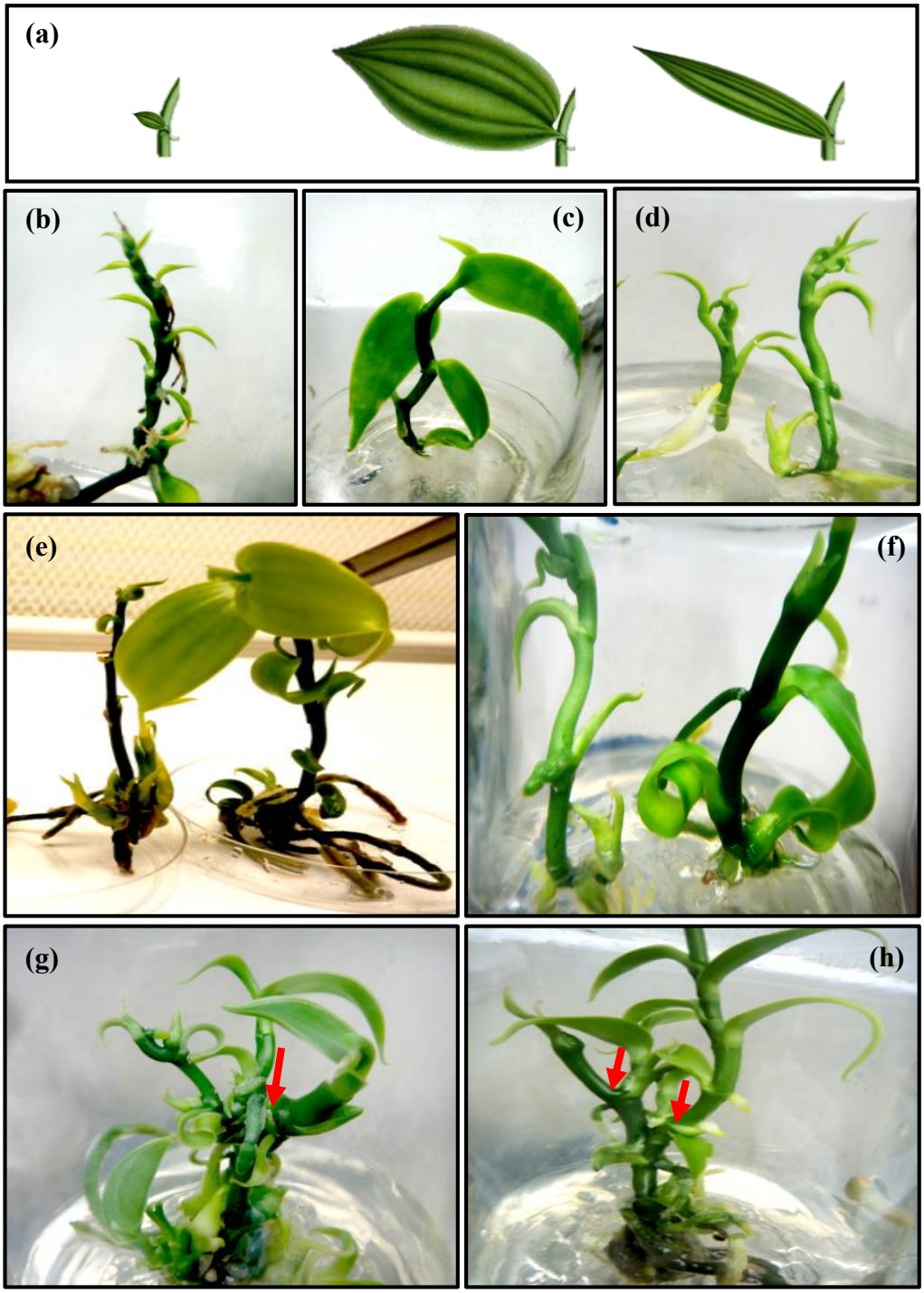


Figure 3.2. Somaclones with distinct morphological appearances: (a) Illustrative diagram showing morphologies of small broad, normal broad and long narrow leaves (from left to right); (b-d) Tissue culture plantlets with small broad, normal broad and long narrow leaves in culture jars; (e) side-by-side comparison between morphology of tissue culture plantlets with small broad leaves (left) and normal broad leaves (right); (f) side-by-side comparison between morphology of tissue culture plantlets with long narrow leaves (left) and normal broad leaves (right); (g) abnormal branching with multiple shoot apices (indicated by arrow); and (h) abnormal branching at every node (indicated by arrows).

3.3.2. DNA Extraction of *V. planifolia*

The extraction method modified from Keb-Llanes (2002) yielded an average of 500 ng to 1.0 µg of purified genomic DNA per 100 mg of ground samples, with purity of A_{260}/A_{280} and A_{260}/A_{230} in the ranges of 1.6-2.2 and 1.8-2.5 respectively. **Figure 3.3** shows an example of high molecular weight genomic DNA gel extracted from 10 selected somaclone samples regenerated from indirect nodal regeneration. Lambda-DNA (**λ-DNA**) at the range of 25 – 200 ng were used to estimate the concentrations of DNA samples, as a complement to concentration readings obtained by Nanodrop spectrophotometer. Only DNA purity A_{260}/A_{280} that falls between 1.8 to 2.2 and with intact genomic band on 0.8% agarose were used for subsequent PCR experiments.

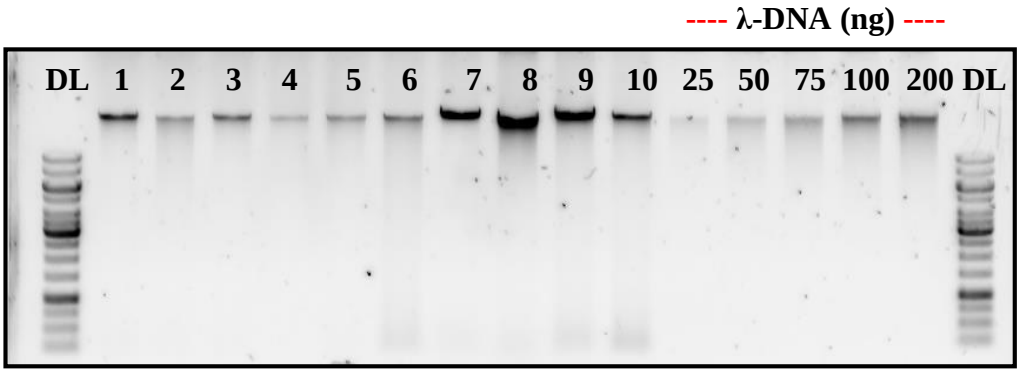


Figure 3.3. Gel image of genomic DNA extracted from somaclones samples, in comparison with λ-DNA in the dilution of 25, 50, 75, 100 and 200 ng, on 0.8% agarose gel in 1x TAE buffer. DL: 1kb DNA ladder (Thermo Scientific, USA); 1-10: Genomic samples (2 µl).

3.3.3. ISSR-PCR Fingerprinting

3.3.3.1. Optimization of ISSR-PCR Parameters

The optimal annealing temperature (T_A) for each ISSR primer was summarized in **Table 3.3**. The optimal DNA amount and $MgCl_2$ concentration for ISSR-PCR was 25.0 ng and 3.0 mM respectively. DNA amount did not affect the banding pattern, except when below 25.0 ng and more than 250 ng (**Figure 3.4a**). Highest number of scorable bandings were achieved at the $MgCl_2$ range of 2.5 to 3.5 mg/L, above and below which number of scorable bands were reduced (**Figure 3.4b**).

Table 3.3. Vanilla ISSR primers amplification properties: primer sequences, optimal annealing temperature (TA), number of scorable bands and fragment size range in bp of each primer.

Primer Name	Primer Sequence (5' - 3') [(a)]	T_A (°C)	Fragment size range (bp)	TB	PB	PPB	PIC	EMR	MI
UBC811	$(GA)_8C$	51.8	600–3000	5	2	40.00	0.09	2.47	0.22
UBC824	$(TC)_8G$	51.8	600–2000	6	5	83.33	0.16	11.30	1.75
UBC826	$(AC)_8C$	51.8	700–2500	7	4	57.14	0.11	1.96	0.21
ISSR2	$(CA)_8RY$	53.1	500-3000	9	8	88.89	0.26	4.79	1.27
ISSR4	$(GA)_8YC$	50.3	200-2500	18	15	83.33	0.19	2.75	0.53
ISSR6	$(AG)_8YT$	50.3	500-2500	8	8	100.00	0.18	6.91	1.24
ISSR7	$(AG)_7YC$	50.3	250-900	8	5	62.50	0.13	3.26	0.44
ISSR10	$(GT)_8YG$	53.1	400-600	5	3	60.00	0.17	2.92	0.49
ISSR23	$(CT)_8RC$	53.1	500-3000	12	10	83.33	0.24	3.30	0.78
ISSR30	$HVH(TG)_7$	50.3	700-3000	9	8	88.89	0.29	3.48	1.02
ISSR32	$C_3(GT)_7$	54.4	400-3000	5	2	40.00	0.05	2.30	0.10

[(a)] Nucleotide ambiguity codes: **Y** represented C/T, **R** represented A/G, **H** represented A or C or T, **V** represented A or C or G; **TB:** Number of total bands; **PB**: Number of polymorphic bands; **PPB**: Percentage of polymorphic bands; **PIC**: Polymorphic information content; **MI**: Marker index.

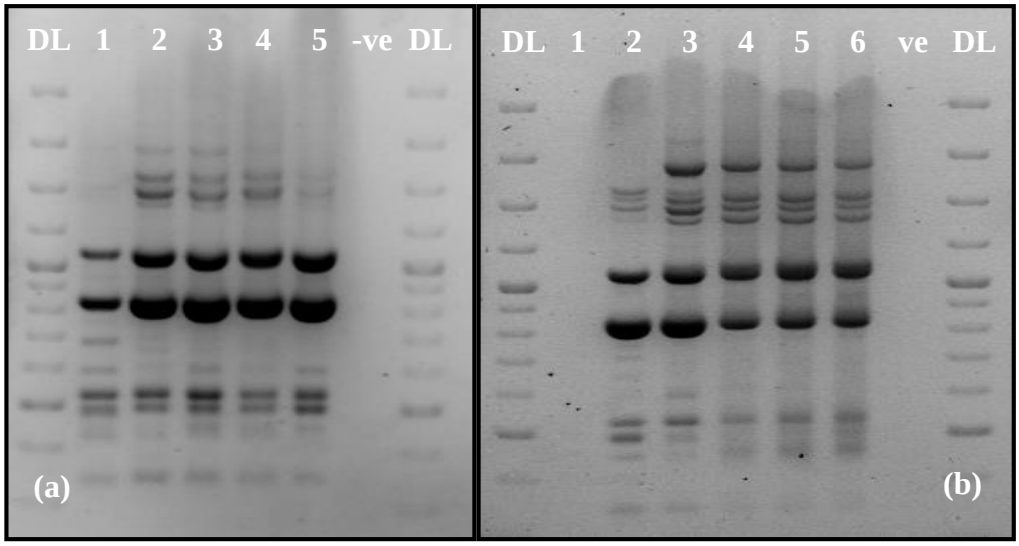


Figure 3.4. Optimization of ISSR-PCR parameters: **(a)** DNA amount: Lane 1-5: 25.0, 50.0, 75.0, 100.0 and 250.0 ng; and **(b)** $MgCl_2$ concentrations: Lane 1-6: 1.5, 2.0, 2.5, 3.0, 3.5 and 4.0 mM. Lane DL: 1kb DNA ladder; Lane -ve: Negative control (without DNA).

3.3.3.2. SV screening with ISSR-PCR

A total of 8 ISSR primers consisting of 100 loci were selected for genetic screenings of the 20 indirect nodal (callus) regenerants and 3 direct nodal regenerants from tissue culture of vanilla. Genetic polymorphisms were detected by presence or absence of ISSR amplified band. In overall, through random sampling, 95% of the callus regenerants were genetically polymorphic. **Figure 3.5 - 3.7** shows the examples of ISSR2, ISSR4 and ISSR6 respectively, when screened on the 20 selected calli regenerants. Arrows indicate distinct polymorphisms between the calli regenerants SV6 and SV18 with the other regenerants. From the dendrogram (**Figure 3.8**) clustering, it was also shown that samples SV6 and SV18 were the most polymorphic amongst all regenerants, while the regenerants from direct regenerations (diVps) clustered closely with the original mother plant (MP). Among the three samples showing morphological abnormalities (samples 18-20), calli regenerant SV18 (long narrow leaf morphology) showed the highest level of genetic polymorphisms (**Figure 3.8**). On the other hand, calli regenerant SV6 did not show visible abnormality at the morphological pre-screening stage, but were highly polymorphic, as shown using ISSR markers (**Figure 3.8**).

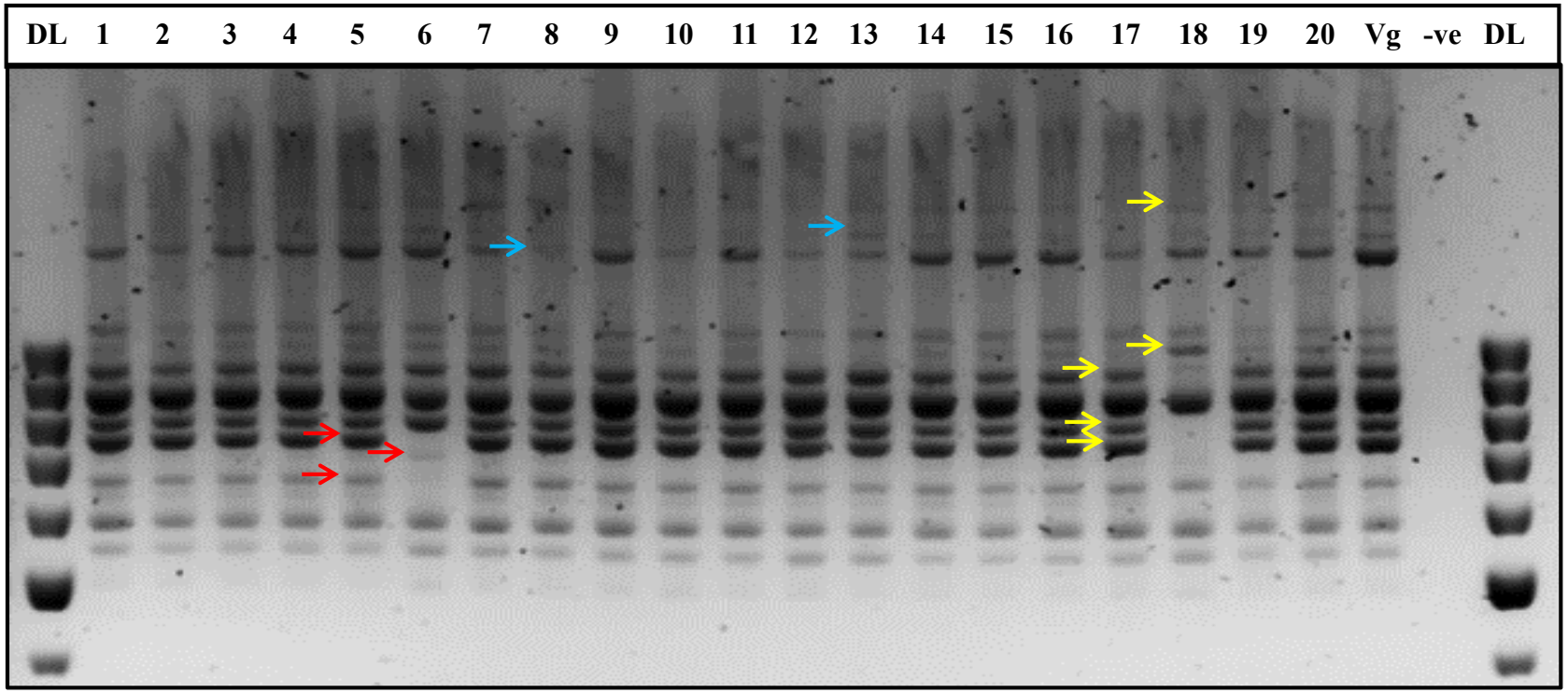


Figure 3.5. Screening of 20 selected somaclones regenerated from callus using the ISSR primer ISSR2. DL: 100 bp DNA Ladder; 1-20: Somaclones sample 1-20; Vg: *V. graffitti*; (-): negative control. Red arrow (→): Missing / additional bands for sample 6, Yellow arrow (→): Missing / additional bands for sample 18. Blue arrow (→) : Other polymorphic bands.

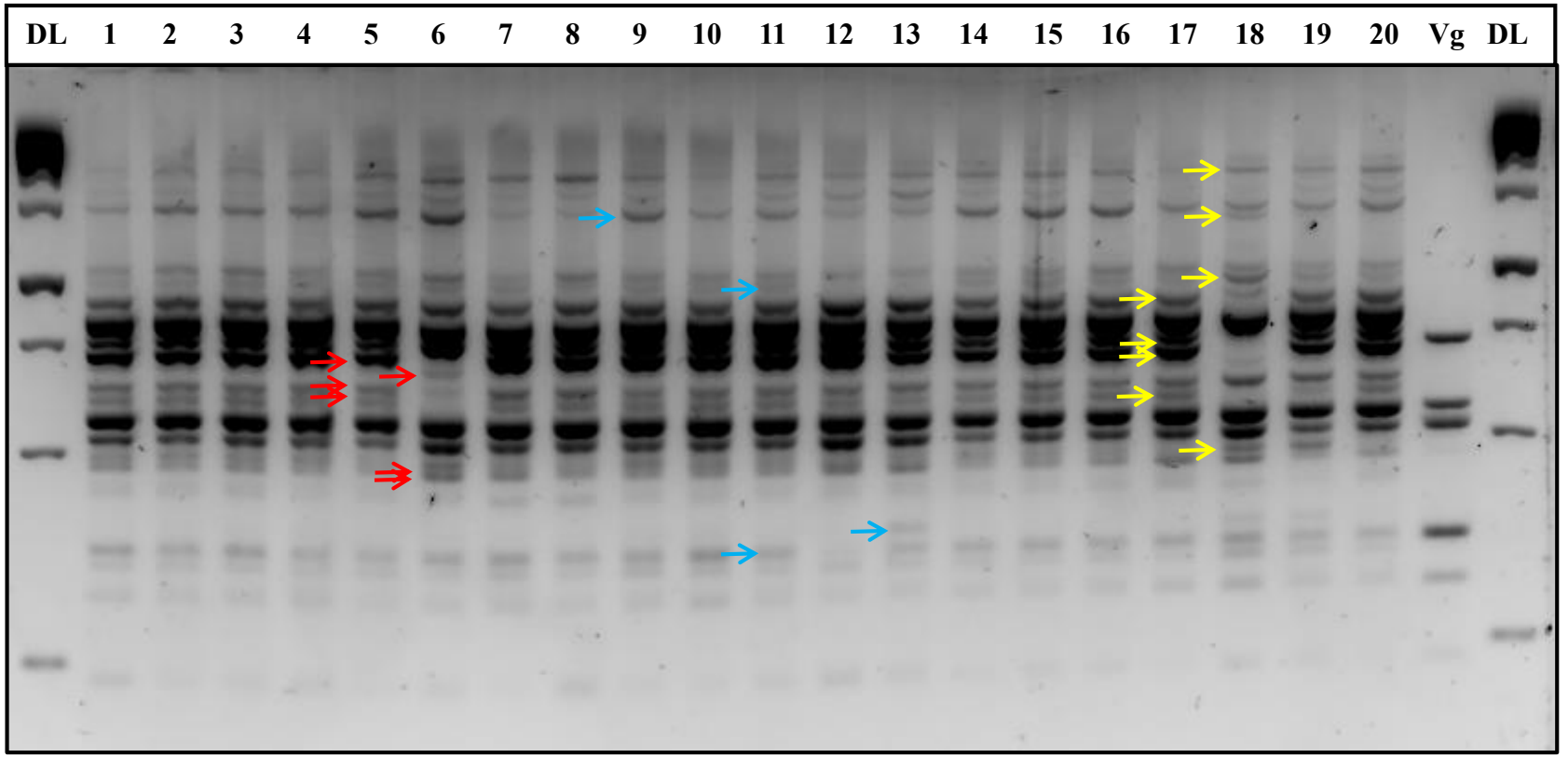


Figure 3.6. Screening of 20 selected somaclones regenerated from callus using the ISSR primer ISSR4. DL: 1 kbp DNA Ladder; 1-20: Somaclones sample 1-20. Red arrow (→): Missing / additional bands for sample 6, Yellow arrow (→): Missing / additional bands for sample 18. Blue arrow (→): Other polymorphic bands.

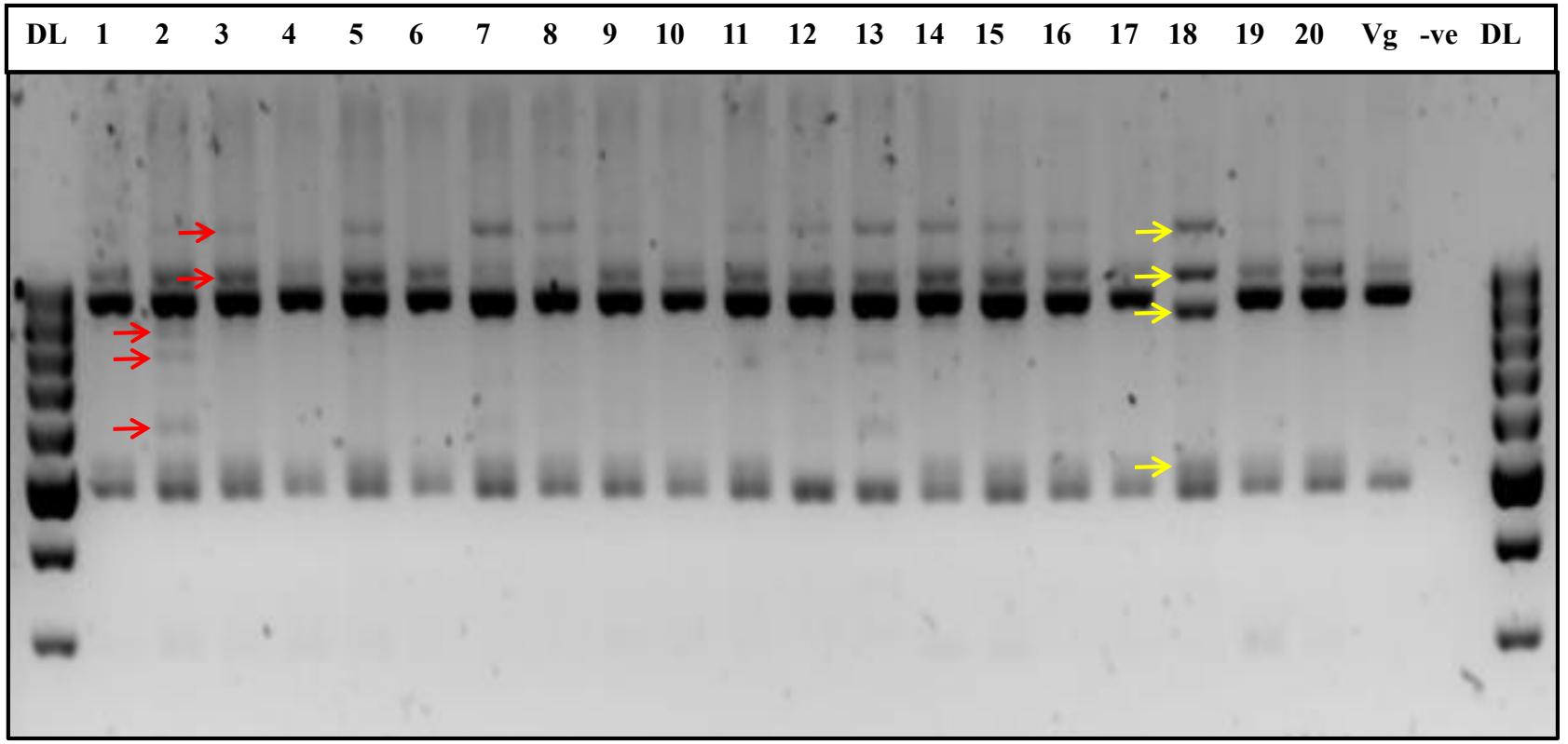


Figure 3.7. Screening of 20 selected somaclones regenerated from callus using the ISSR primer ISSR6. DL: 100 bp DNA Ladder; 1-20: Somaclones sample 1-20; Vg: *V. graffitti*; (-): negative control. Red arrow (→): Missing / additional bands for sample 4, Yellow arrow (→): Missing / additional bands for sample 18.

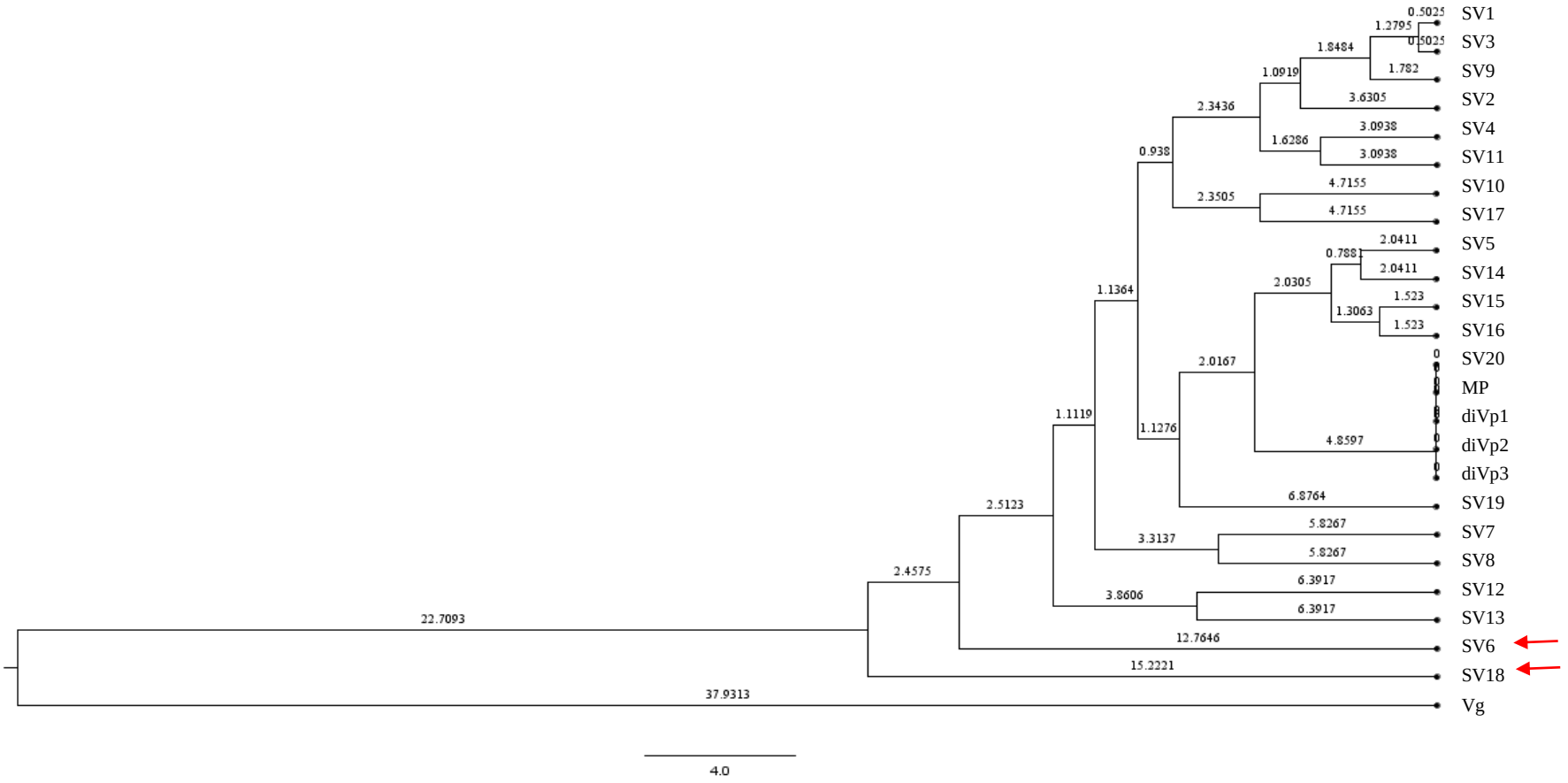


Figure 3.8. Dendrogram for ISSR analysis of somaclones. Arrows indicated the genetically most diverse somaclone lines SV6 and SV18.

3.3.4. MSAP-PCR Fingerprinting

3.3.4.1. Optimization of MSAP-RE digestion conditions

The restriction digestion profiles of *V. planifolia* genomic DNA using *Eco*RI and *Msp*I for different incubation time and different double digestion RE ratios were shown in **Figure 3.9a** and **b** respectively. Both RE incubation time (20 min and 16 h) did not show any differences in the digestion profile of the genomic DNA. Therefore, 20 min was chosen as the optimal RE digestion time for subsequent experiments. Different *Eco*RI and *Msp*I enzymes ratios also did not change the digestion profile of the genomic DNA, therefore, the subsequent experiment maintained at using the 1:1 ratio for *Eco*RI: *Msp*I digestion.

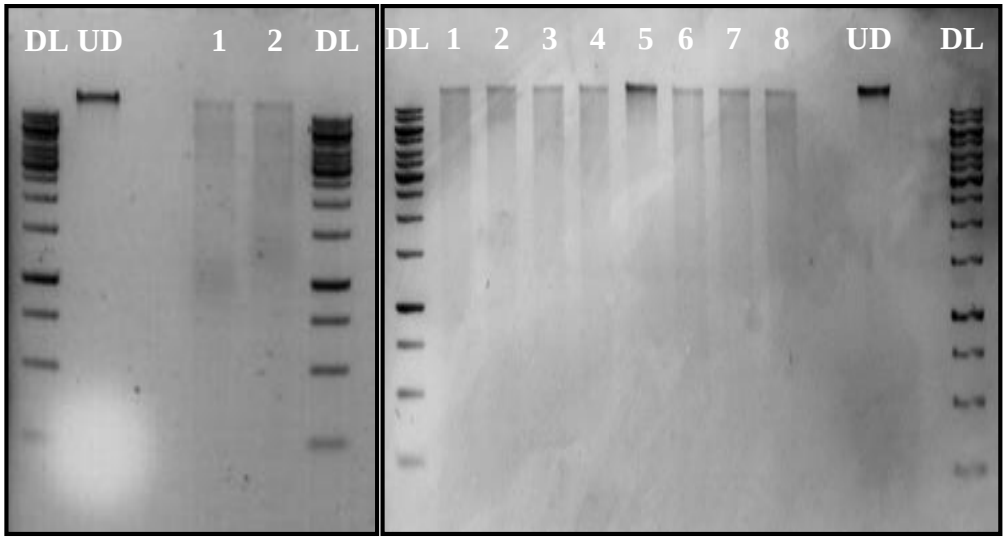


Figure 3.9. Optimization of RE digestion parameters for MSAP analysis. Effect of **(a)** different RE incubation time: genomic DNA digested with *Eco*RI/*Msp*I for 20 min (Lane 1) and 16 h (Lane 2); **(b)** Genomic DNA digestion using *Eco*RI:*Msp*I at ratio of 1:1, 1:2, 1:3, 1:4, 0:2, 2:1, 3:1 and 4:1 (Lane 1-8 respectively). DL: 1kb DNA ladder; UD: Undigested genomic DNA (Control).

3.3.4.2. Optimization of MSAP-PCR condition

PCR amplification profiles using a gradient of annealing temperature and DNA amounts were shown in **Figure 3.10a** and **b** respectively. The annealing temperature 51.8°C was chosen as the optimal temperature as it gave the highest number of scorable bands (**Figure 3.10a**). Similar amplification intensity and number of amplified bands were observed in DNA concentrations at the range from 12.5 to 250.0 ng, therefore the lowest amount of DNA (12.50 ng) was used for the subsequent experiments (**Figure 3.10b**).

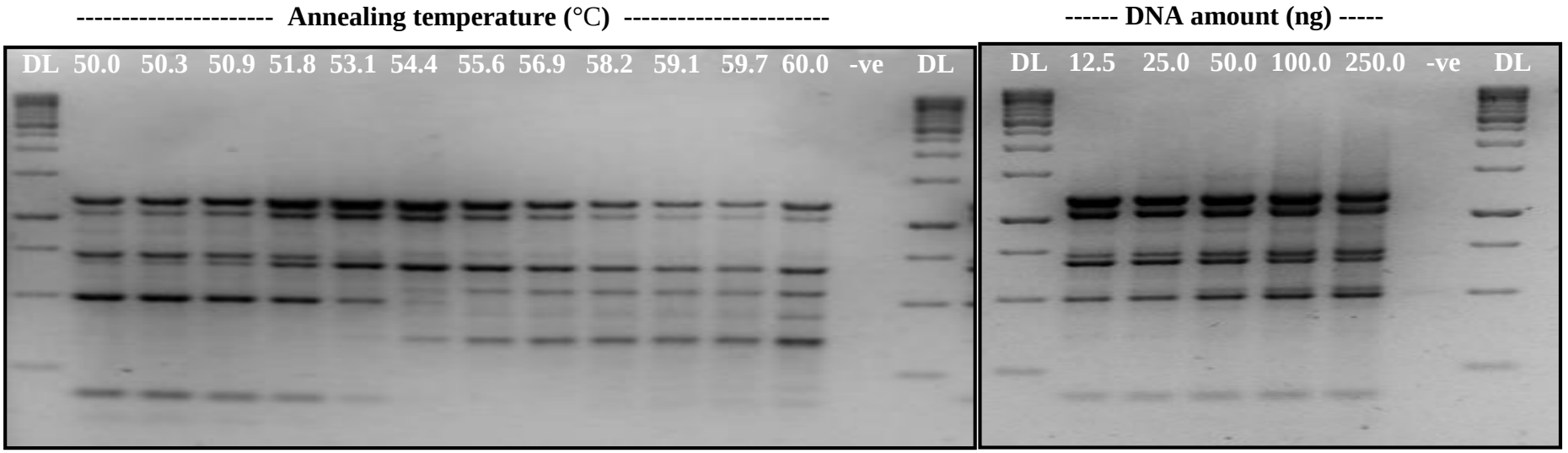


Figure 3.10. Optimization of MSAP-PCR parameters: **(a)** Annealing temperatures (50.0, 50.3, 50.9, 51.8, 53.1, 54.4, 55.6, 56.9, 58.2, 59.1, 59.7 and 60.0°C); and **(b)** DNA amount (12.5 ng, 25.0 ng, 50.0 ng, 100.0 ng, and 250.0 ng). Lane DL: 1kb DNA ladder; Lane -ve: Negative control (without DNA).

3.3.4.3. Selection between selective +1 or +2 primers

The amplification profiles of PCR primer screening using the pre-selective (+1) and selective (+2) primers are shown in **Figure 3.11** and **3.12** respectively. The comparison for number of scorable bands between +1 and +2 primers was summarized in **Table 3.4**. From the table, it was shown that the highest number of scorable bands was 17 for +1 primers, and 12 for +2 primers. In average, the +1 primers produce 8.875 scorable bands per primer, while the +2 primers an average of 4.24 scorable bands per primer. Therefore, further selective amplification using the +2 primers was not necessary for the vanilla samples as it limited the number of scorable bands.

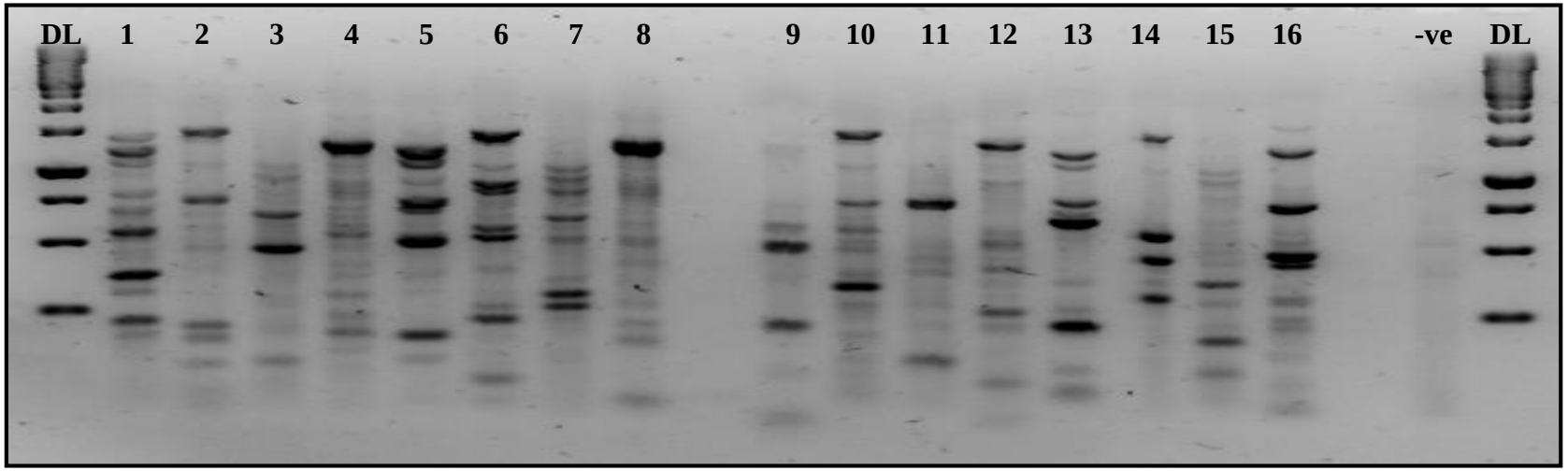


Figure 3.11. Primers screening for *EcoRI*/Hpa*II* (E/H) selective amplification (+1) primer pairs. DL: 1 kb DNA ladder; Lane 1-16: Selective (+1) primers $E_1MH_{1\text{-}4}$, E_2MH_1-$_4$, $E_3MH_{1\text{-}4}$ and $E_4MH_{1\text{-}4}$ respectively.

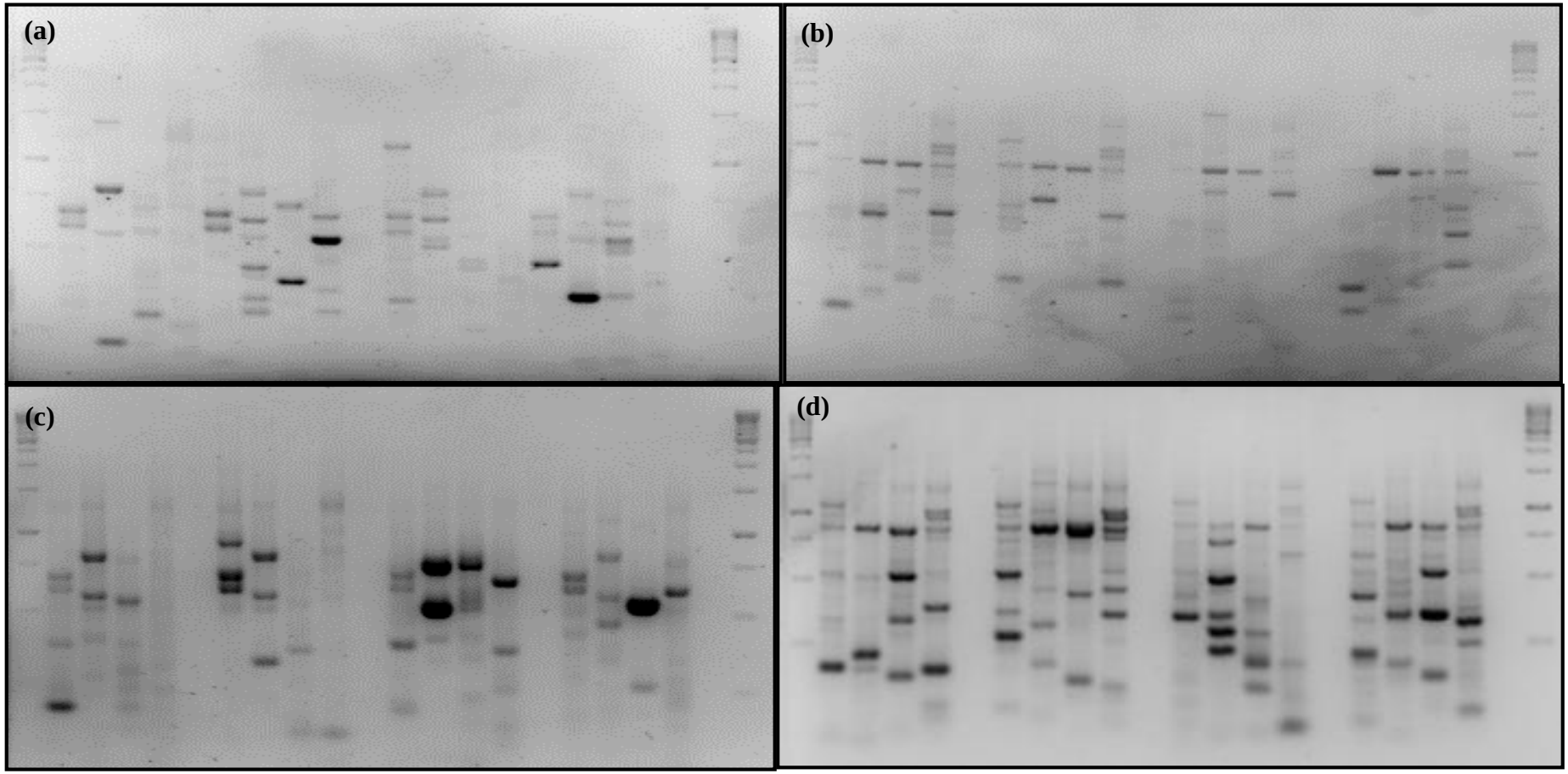


Figure 3.12. Primers screening for *EcoRI*/Hpa*II* (E/H) selective amplification (+2) primer pairs: (a) E_1MH_1, (b) E_1MH_4, (c) E_2MH_1 and (d) E_2MH_4 respectively.

Table 3.4. Summary of +1 and +2 primers amplification.

Primer	+1/+2	No. of scorable bands	Primer	+1/+2	No. of scorable bands
E_1MH_1	+1	10	$E_{1\text{-}3}MH_{4\text{-}1}$	+2	5
E_1MH_2	+1	7	$E_{1\text{-}3}MH_{4\text{-}2}$	+2	3
E_1MH_3	+1	4	$E_{1\text{-}3}MH_{4\text{-}3}$	+2	1
E_1MH_4	+1	17	$E_{1\text{-}3}MH_{4\text{-}4}$	+2	5
E_2MH_1	+1	8	$E_{1\text{-}4}MH_{4\text{-}1}$	+2	2
E_2MH_2	+1	11	$E_{1\text{-}4}MH_{4\text{-}2}$	+2	2
E_2MH_3	+1	9	$E_{1\text{-}4}MH_{4\text{-}3}$	+2	3
E_2MH_4	+1	8	$E_{1\text{-}4}MH_{4\text{-}4}$	+2	8
E_3MH_1	+1	4	$E_{2\text{-}1}MH_{1\text{-}1}$	+2	4
E_3MH_2	+1	12	$E_{2\text{-}1}MH_{1\text{-}2}$	+2	5
E_3MH_3	+1	11	$E_{2\text{-}1}MH_{1\text{-}3}$	+2	6
E_3MH_4	+1	7	$E_{2\text{-}1}MH_{1\text{-}4}$	+2	3
E_4MH_1	+1	10	$E_{2\text{-}2}MH_{1\text{-}1}$	+2	4
E_4MH_2	+1	5	$E_{2\text{-}2}MH_{1\text{-}2}$	+2	4
E_4MH_3	+1	8	$E_{2\text{-}2}MH_{1\text{-}3}$	+2	1
E_4MH_4	+1	11	$E_{2\text{-}2}MH_{1\text{-}4}$	+2	4
$E_{1\text{-}1}MH_{1\text{-}1}$	+2	2	$E_{2\text{-}3}MH_{1\text{-}1}$	+2	4
$E_{1\text{-}1}MH_{1\text{-}2}$	+2	4	$E_{2\text{-}3}MH_{1\text{-}2}$	+2	3
$E_{1\text{-}1}MH_{1\text{-}3}$	+2	3	$E_{2\text{-}3}MH_{1\text{-}3}$	+2	1
$E_{1\text{-}1}MH_{1\text{-}4}$	+2	0	$E_{2\text{-}3}MH_{1\text{-}4}$	+2	2
$E_{1\text{-}2}MH_{1\text{-}1}$	+2	3	$E_{2\text{-}4}MH_{1\text{-}1}$	+2	2
$E_{1\text{-}2}MH_{1\text{-}2}$	+2	7	$E_{2\text{-}4}MH_{1\text{-}2}$	+2	4
$E_{1\text{-}2}MH_{1\text{-}3}$	+2	2	$E_{2\text{-}4}MH_{1\text{-}3}$	+2	3
$E_{1\text{-}2}MH_{1\text{-}4}$	+2	5	$E_{2\text{-}4}MH_{1\text{-}4}$	+2	1
$E_{1\text{-}3}MH_{1\text{-}1}$	+2	5	$E_{2\text{-}1}MH_{4\text{-}1}$	+2	6
$E_{1\text{-}3}MH_{1\text{-}2}$	+2	5	$E_{2\text{-}1}MH_{4\text{-}2}$	+2	4
$E_{1\text{-}3}MH_{1\text{-}3}$	+2	4	$E_{2\text{-}1}MH_{4\text{-}3}$	+2	4
$E_{1\text{-}3}MH_{1\text{-}4}$	+2	0	$E_{2\text{-}1}MH_{4\text{-}4}$	+2	6
$E_{1\text{-}4}MH_{1\text{-}1}$	+2	4	$E_{2\text{-}2}MH_{4\text{-}1}$	+2	8
$E_{1\text{-}4}MH_{1\text{-}2}$	+2	3	$E_{2\text{-}2}MH_{4\text{-}2}$	+2	7
$E_{1\text{-}4}MH_{1\text{-}3}$	+2	5	$E_{2\text{-}2}MH_{4\text{-}3}$	+2	4
$E_{1\text{-}4}MH_{1\text{-}4}$	+2	0	$E_{2\text{-}2}MH_{4\text{-}4}$	+2	9
$E_{1\text{-}1}MH_{4\text{-}1}$	+2	4	$E_{2\text{-}3}MH_{4\text{-}1}$	+2	5
$E_{1\text{-}1}MH_{4\text{-}2}$	+2	5	$E_{2\text{-}3}MH_{4\text{-}2}$	+2	7
$E_{1\text{-}1}MH_{4\text{-}3}$	+2	3	$E_{2\text{-}3}MH_{4\text{-}3}$	+2	5
$E_{1\text{-}1}MH_{4\text{-}4}$	+2	8	$E_{2\text{-}3}MH_{4\text{-}4}$	+2	7
$E_{1\text{-}2}MH_{4\text{-}1}$	+2	6	$E_{2\text{-}4}MH_{4\text{-}1}$	+2	7
$E_{1\text{-}2}MH_{4\text{-}2}$	+2	2	$E_{2\text{-}4}MH_{4\text{-}2}$	+2	6
$E_{1\text{-}2}MH_{4\text{-}3}$	+2	1	$E_{2\text{-}4}MH_{4\text{-}3}$	+2	6
$E_{1\text{-}2}MH_{4\text{-}4}$	+2	7	$E_{2\text{-}4}MH_{4\text{-}4}$	+2	12

3.3.4.4. SV screening with MSAP-PCR

MSAP analysis of tissue culture regenerants was carried out using selected +1 primer sets (**Table 3.5**; **Figure 3.13**). From the cluster analysis, it was shown that the directly regenerated vanilla (diVp) regenerants were epigenetically closest to the mother plant (Vp-Mp), whereby the cluster between the MP and indirectly-regenerated Vanilla (iVp) regenerants were epigenetically more distant-related, as revealed with primer sets, E_3MH_1 E_3MH_4 and E_4MH_2 (**Figure 3.13**). The close species, *V. grafitti* (*Vg*) was the most distantly related sample with the other *V. planifolia*, which has demonstrated its purpose as a control in showing the MSAP marker efficiency (**Figure 3.13**). In total, the E_3MH_1 E_3MH_4 and E_4MH_2 primer sets produce 41 scorable loci, of which 21 were polymorphic and 20 were monomorphic. From the data, three types of MSAP bands were observed (**Table 3.5**). These types of bands were classified according to the following scheme. Type I whereby AFLP bands were present in both *Eco*RI/*Hpa*II and *Eco*RI/*Msp*I lanes; type II whereby bands were present in *Eco*RI/*Msp*I lanes, but not in *Eco*RI/*Hpa*II lanes; and type III whereby bands were present in *Eco*RI/*Hpa*II, but not in *Eco*RI/ *Msp*I lanes. For the direct regeneration, all samples exhibited Type I MSAP profile, while for indirect regenerated samples, a mixture of all types could be observed (**Figure 3.13**).

Table 3.5. Type I, Type II and Type III MSAP profiles.

Locus	Band Size range (bp)	iVp-1-1	iVp-1-2	iVp-1-3	iVp-2-1	iVp-2-2	iVp-2-3	iVp-3-1	iVp-3-2	iVp-3-3	diVp-1	diVp-2	diVp-3
E_3MH_{1-3}	900	I	I	II	I	I	II	I	I	I	I	I	I
E_3MH_{4-1}	1200-1500	III	III	I	III	I	I	II	I	II	I	I	I
E_3MH_{4-3}	1000-1200	III	III	II	I	I	I	I	I	I	I	I	I
E_4MH_{2-2}	1200-1500	III	I		I	I	I	I	I	I	I	I	I
E_4MH_{2-3}	1000-1200	III	I	III	I	I	I	I	I	I	I	I	I
E_4MH_{2-4}	1000	III	I		I	I	I	I	I	I	I	I	I
E_4MH_{2-5}	900	I	I	III	I	I	I	I	I	I	I	I	I
E_4MH_{2-9}	600-700	I	I	II	I	I	I	I	I	III	I	I	I
E_4MH_{2-11}	500-600	I	I		I	I	I	I	I	III	I	I	I
E_4MH_{2-12}	400-500	I	I		I	I	I	I	I	I	I	I	I

iVp: indirectly-regenerated vanillas, diVp: directly-regenerated vanillas.

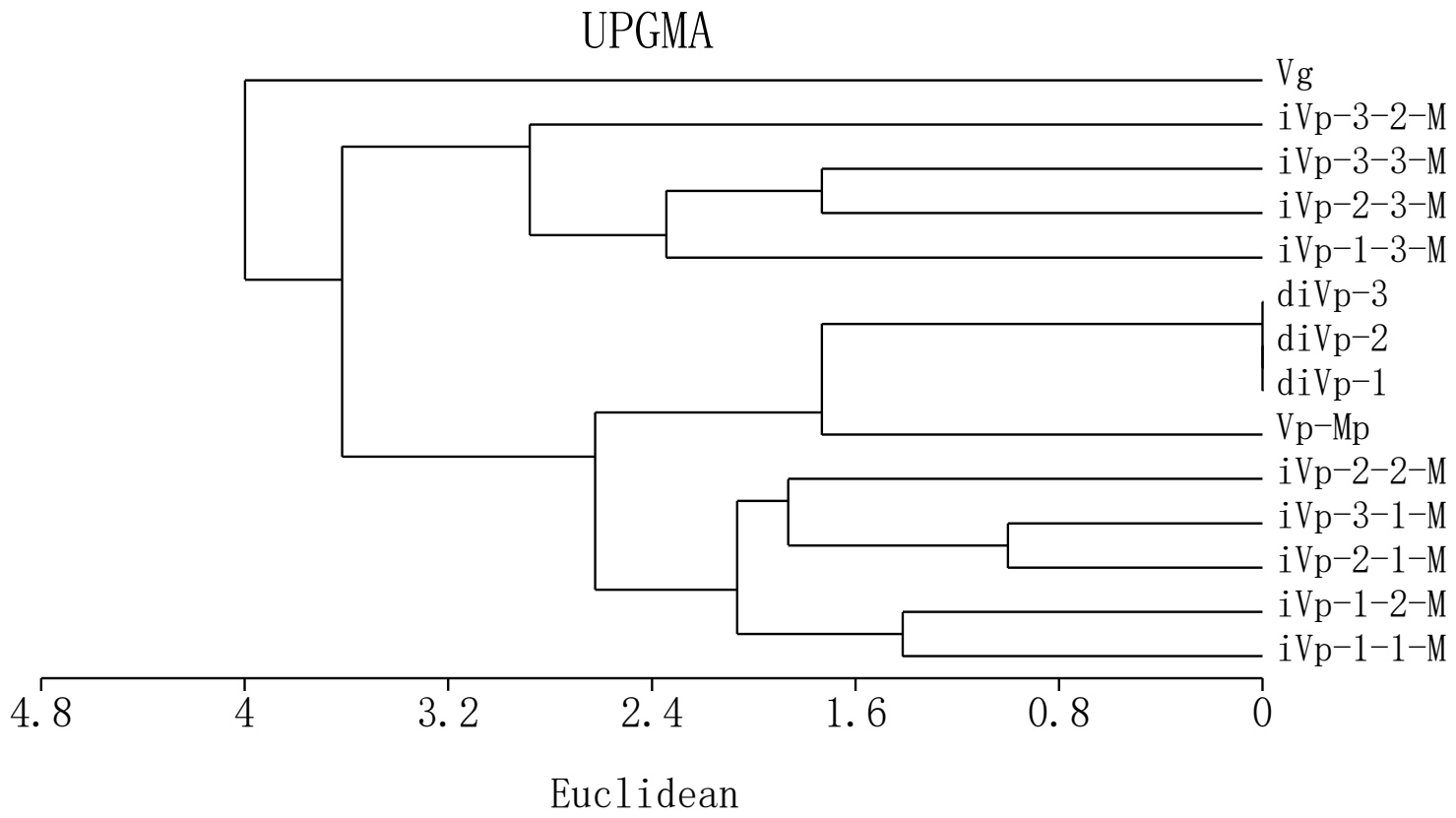


Figure 3.13. Dendrogram of genetic distances between vanilla tissue culture regenerants. iVp: indirectly-regenerated vanillas, diVp: directly-regenerated vanillas, Vp-MP: vanilla mother plant, Vg: *V. graffiti*, a closely-related vanilla species used as control for marker analysis.

3.4. Discussions

In this study, the ISSR marker system is an effective genetic screening tool. Our findings confirm the usefulness of ISSR markers in the analysis of somaclonal variation of micropropagated vanilla plantlets, similar to works on crops such as rye (*Secale cereale* L.) (Linacero *et al.* 2011), hydrangea (*Hydrangea macrophylla*) (Liu *et al.* 2011) and lilies (*Lilium orientalis*) (Liu and Yang 2012; Yin *et al.* 2013).

The ISSR results showed that the direct nodal regeneration retained genetic fidelity of vanilla plants, and thus is more suitable for clonal mass propagation of elite vanilla materials. This agreed with the previous studies by Sreedhar and co-workers (2007), who reported that vanilla direct nodal regenerants maintained genetic fidelity over ten years of *in vitro* culture.

In contrary to the direct regeneration, the indirect nodal regeneration resulted in genetic heterogeneity of vanilla regenerants, which agreed with a recent study by Ramírez-Mosqueda and Iglesias-Andreu (2015), who reported genetic heterogeneity in callus-derived vanilla regenerants, by using a different range of hormonal treatments (TDZ and 2,4-D). This finding agreed with studies in other plant species whereby callus-derived regenerants were genetically distinct (SV), such as *Oryza sativa* (Hirochika *et al.*, 1996), *Dendranthema grandifolia* (Miller and Zalewska, 2014), *Solanum melongena* (Nascer and Mahmood, 2014), *Pogostemon patchouli* (Ravindra *et al.*, 2012), *Solanum tuberosum* (Thieme and Griess, 2005; Munir *et al.*, 2011), *Curcuma longa* L. (Kar *et al.*, 2014), *Plumbago zeylanica* Linn. (Sivanesan, 2007), *Curcuma aromatic* (Mohanty *et al.*, 2008), *Amorphophallus rivieri* Durieu (Hu *et al.*, 2011), *Plantago major* (Ghorbanpour and Khadivi, 2015), *Beta vulgaris* L. (Gurel and Gurel 2014) and hybrid poplar (Son *et al.*, 1993).

The analysis on methylation patterns in *V. planifolia* regeneration systems through epigenetic marker MSAP, leads to an interesting observation: direct regeneration of *V. planifolia* is not only genetically, but also epigenetically more stable. While for the indirect regeneration of *V. planifolia*, both genomic and epigenomic instability were observed. The first observation that direct regeneration of *V. planifolia* through nodal culture produces genetically stable tissue culture plantlets agrees with what has been reported by Sreedhar *et al.* (2007) using the RAPD and ISSR marker systems; however, epigenetic analysis was not done in this or any other previous studies. Therefore, this current research has provided a novel insight whereby directly-regenerated Vanilla tissue culture regenerants are stable even at epigenomic level. In this study, there isn't a specific pattern to which the methylation occurs, which shows that this epigenetic phenomenon could occur by random. DNA methylation in tissue culture is also involved in the regulation of various genes transcription, such as the activation of retrotransposon and microRNAs (Saze *et al.,* 2012), such as studies in oil palm and wheat (Cantu *et al.*, 2010).

Mode of regeneration is one of the causes of SV in plants. Both culture initiation and subsequent subculture exposure explant to oxidative stresses (Krishna *et al.,* 2008). The findings in this study agreed with other studies that, callus formation, like protoplast culture, are considered "extreme" procedure which imposed stresses to the plant cells (Smulders and de Klerk, 2011), thus cultures that go through a callus phase theoretically promote a higher mutation rate (Zayova *et al.*, 2010).

The low genetic variability of vanilla makes it susceptible to pests and diseases which leads to a decrease in production, thus the use of indirect morphogenic routes could be applied for subsequent genetic improvement works for broadening genetic basis of vanilla. The protocol established may be of interest to broaden the limited genetic base of vanilla (Minoo *et al.*, 2006; Retheesh and Bhat, 2011).

Not only the number of subculture but their duration also contributes to enhancing the rate of somaclonal variations, especially cell suspension and callus cultures (Bairu *et al.,* 2006; Sun *et al.,* 2013). Studies have shown that

somaclonal variation is more apparent in plants regenerated from long-term cultures (Etienne and Bertrand 2003). Rival *et al.* (2013) noticed that *in vitro* proliferation induces DNA hypermethylation in a time-dependent fashion and changes in DNA methylation is involved in modulating the expression of embryogenic capacity of oil palm during tissue culture.

3.5. Conclusions and future prospects

Like oil palm, *V. planifolia* is a slow growing perennial crop, which takes a minimum of two to three years to reach maturity. Therefore, any mutations that occur could take up to years to be seen, and this could lead to severe yield loss. With respect to that, phenotypic studies on the somaclonal variation are also difficult as the tissue culture plantlets usually exhibits normal phenotypes and any potential mutations that occurs after the plant reaches maturity takes years to see. Thus, molecular markers play an important role to enable early pre-screening for somaclonal variants, and also to predict the variations that might occur by isolation and sequencing of mutated gene sequences.

This study has provided an insight onto the genetic and epigenetic stability of two different regeneration systems (direct and indirect) in *V. planifolia*. It has been shown, using both ISSR and MSAP markers, that *in vitro* plantlets grown via direct shoot regeneration remained genetically and epigenetically stable, while the indirectly-regenerated plantlets exhibited both genetic and epigenetic instability. The study suggested that the direct nodal regeneration is more suitable for clonal mass propagation of elite vanilla genotypes, while the indirect nodal regeneration could be a potential tool for broadening genetic variability of vanilla through *in vitro* improvements. In subsequent chapters, the vanilla callus cultures were further studied at cytological and transcriptomic levels to understand the possible mechanisms underlying the SV in vanilla calli.

CHAPTER IV

SYSTEMIC VARIATIONS OF CELLULAR ENDOPOLYPLOIDY IN DIFFERENT TISSUES OF *VANILLA PLANIFOLIA* AND ITS IMPLICATIONS IN MICROPROPAGATION

Chapter IV

Systemic variations of cellular endopolyploidy in different tissues of *V. planifolia* and its implications in micropropagation

4.1. Introduction

Formerly, plant tissue culture was considered as a method for cloning a particular genotype or a technique to obtain rapid rate of asexual propagation. However, phenotypic changes were often observed amongst regenerated plants. At first, these were seen as non-heritable variants as a consequence of "epigenetic" events (Larkin and Scowcroft, 1981). Subsequently, it was corroborated that many of the variations were associated with chromosomal aberrations (Edallo *et al.*, 1981, Rice, 1982, Karp, 1995). Cytological aberrations in somaclonal variations (SV) are mainly correlated with numerical or structural chromosomal changes (Molina and García, 1998). Since the 1980s, many authors (Pring *et al.*, 1981; Mc Coy *et al.*, 1982; Rhodes *et al.*, 1986; Lee and Phillips, 1987; Fluminhan *et al.*, 1996) have already documented polyploidy, aneuploidy, translocations, inversions, deletions, fragmentation, heteromorphic pairs, ring chromosomes and single gene recessive mutations in tissue culture regenerants.

The frequency of chromosomal changes is affected by different aspects of tissue culture, such as species (Bayliss, 1973; Novak, 1981; Kidwell *et al.*, 1993; Kirti *et al.*, 1991; Linacero *et al.*, 1992a, b), genotype (Amberger *et al.*, 1992, Cheng *et al.*, 1992, Li *et al.*, 1989, Molina *et al.*, 1992, Puolimatka *et al.*, 1993; Thorn, 1992), culture medium (Amberger *et al.*, 1992, Ghaemi *et al.*, 1993; Linacero *et al.*, 1992a, b; Wilkinson *et al.*, 1987), genotype-culture medium interaction (Wilkinson *et al.*, 1987), the origin of meristematic organised growth

from disorganised growth and the time spent in this state (Rhodes *et al.*, 1986, Lee *et al.*, 1987, Stelly *et al.*, 1989, Qureshi *et al.*, 1992, Karp, 1995) and the type and concentration of plant growth regulators (Bai and Knott, 1993).

With the advent of flow cytometry (FCM) for plants studies, the screening of large populations of regenerants for chromosomal abberrations has become feasible and may be used either for industrial or fundamental applications. Thus, over these last three decades (De Laat *et al.*, 1987; Ramulu and Dijkhuis 1986), the ploidy level characterization of plant materials, such as those kept in gene banks, has gained pace. Examples include banana (*Musa* spp.) (Pillay *et al.*, 2006), Agave (Palomino *et al.*, 2003), Medicago (Brummer *et al.*, 1999), Helleborus (Zonneveld, 2001), Dioscorea (Egesi *et al.*, 2001), quinoa (de la Cruz *et al.*, 2007) and bambara groundnut (Koně *et al.*, 2007). Flow cytometry is not only fast and cost-effective for analyzing large number of plants (Ochatt *et al.*, 2011) but it also allows the studying of different types of tissue and cell layers (Leus *et al.*, 2009).

In plant tissue culture system, there are two important cytological aspects to consider for clonality, namely the original genome size and pre-existing ploidy levels of the explants used for tissue culture, as well as the changes of the two parameters after tissue culture. Increased knowledge of the degree of endopolyploidy in the explant tissue source will be highly valuable for the maintenance of the original ploidy level in culture (Kudo and Kimura, 2002). Knowledge about the endopolyploidization processes in plants is especially important for *in vitro* culture since the differences in endopolyploidization levels in explant cells may be reflected in chromosome variation in callus culture, a possible reason for somaclonal variations in regenerated plants (Gilissen *et al.*, 1996; Fras and Maluszynska, 2004; Kolano *et al.* 2008). During long-term callus cultures, genetic and epigenetic changes occur (Bhatia *et al.*, 2015). The callus could then be composed of a population of cells with genotypic differences (Bhatia *et al.*, 2015).

The main aim of this study is to determine the pre-existing ploidy levels in different vegetative explants of vanilla and the ploidy population changes in nodal-derived calli and calli maintained over prolonged culture periods of 3, 6,

12 and 15 months, using flow cytometry (FCM). The specific objectives in order to achieve the main aim were: (i) establishment of a series of subculture passages for vanilla *in vitro* callus cultures for the subsequent FCM analysis; (ii) evaluation of suitable reference standard(s) to be used for the estimation of genome size and ploidy levels in vanilla nuclei samples; (iii) examination of the original genome size of vanilla mother plant (MP) and the pre-existing ploidy variations in different vegetative tissues of vanilla MP; and (iv) determination of the changes in mean genome size and ploidy population of the vanilla calli over prolonged culture periods up to 15 months.

4.2. Materials and methods

4.2.1. Chemicals reagents and apparatus

A list of standard solutions and equipment that were used in these experiments is stated in **Appendix A** and **Appendix B**, respectively. Cell culture-grade chemicals, such as Tween 20 (Sigma-Aldrich, cat no. P2287), were used to essentially prevent high sample background noise in flow cytometry analysis. Detailed chemical formulas for the preparation of chemical buffers for plant nuclei extraction are described in **Appendix C**. All buffers were filtered through 0.22 µm filter to remove small particles. Prepared buffers were stored at -20°C (except Otto buffers I and II, which were prepared fresh) in 5 ml aliquots and were never refrozen after thawing.

4.2.2. Source of plant materials

4.2.2.1. Internal reference standard plants

The internal reference standard plant candidates (**Table 4.1**) that were used for estimation of genome size in flow cytometric analysis were kindly provided by Prof. Ing. Jaroslav Doležel (The Centre of Plant Structural and Functional Genomics (CIRAD), France) and were grown in pots from seeds (**Figure 4.1**). All plants were maintained in shade house of University of Nottingham, Malaysia Campus.

Table 4.1. DNA reference standards recommended for the estimation of nuclear DNA amounts in absolute units (Dolezel *et al.*¸ 2005).

Plant species and cultivar	2C DNA content (pg DNA)
Raphanus sativus L. 'Sava'	1.11
Solanum lycopersicum L. 'Stupcke polni' rane	1.96
Glycine max Merr. 'Polanka'	2.50
Zea mays L. 'CE-777'	5.43
Pisum sativum L. 'Citrad'	9.09
Secale cereal L. 'Dankovske'	16.19
Vicia faba L. 'Inovec'	26.90

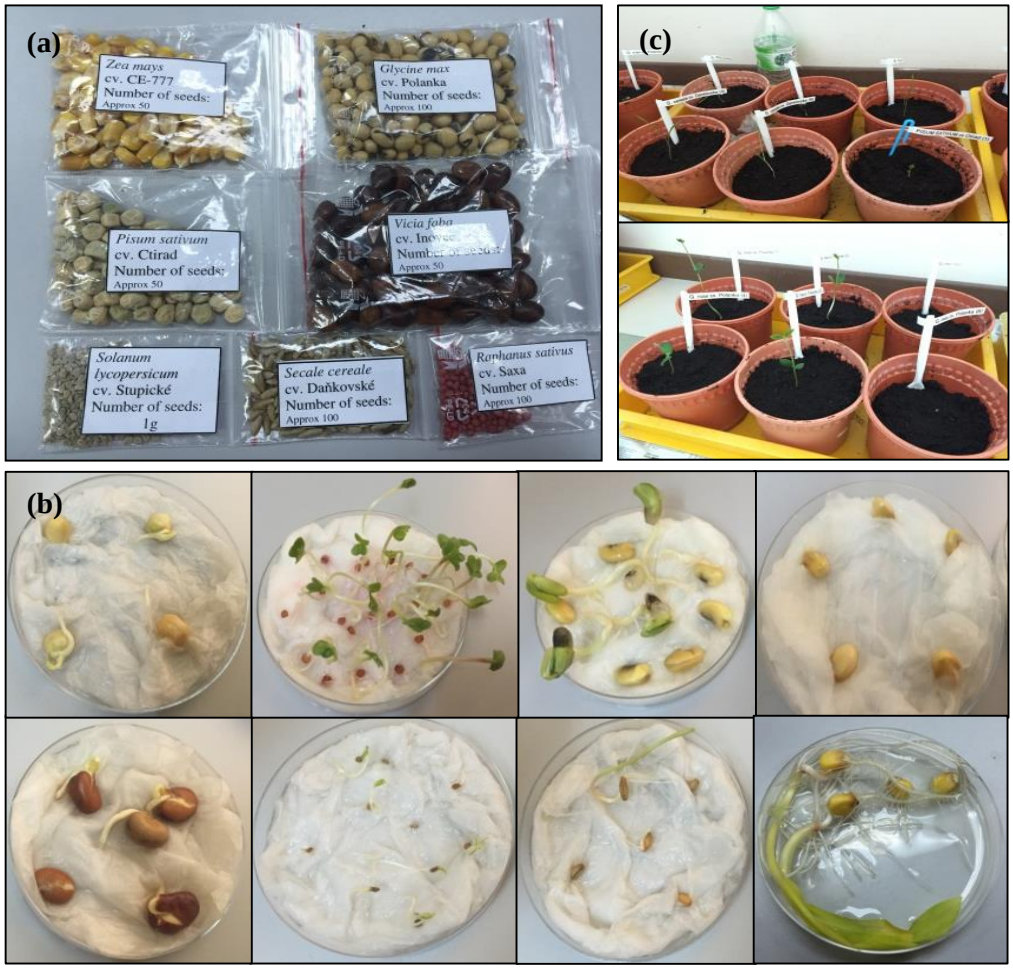


Figure 4.1. Seed germination of reference standard plants: (a) Seeds, **(b)** germinated seeds, and **(c)** 2-week-old seedlings of reference standards transplanted into pots with soil.

4.2.2.2. Vegetative tissues of vanilla

The vegetative tissues used in this study were shoot tip, axillary bud, leaf tip, leaf mid, leaf base, root tip and root base explants of *V. planifolia* mother plants harvested from the shade house at UNMC, Malaysia. The axillary bud, leaf and root samples were collected at node 2 to 6 of three individual vanilla plants. For leaf samples, leaf tip and leaf base were excised at diameter 5mm from the tip and base of each leaf respectively, while leaf mid was excised for an area of 5 x 5 mm^2 from the centre of each leaf. For root samples, root tip and root base were excised 5mm from the tip and base of each root respectively. For shoot tip sample, primordial and developing leaves embedding the shoot tips were carefully removed, and the shoot tip was harvested at the terminal (3 – 5 mm) portion of the vanilla shoots comprising the meristem. Each analysis comparison was carried out in triplicates of biological samples.

4.2.2.3. Vanilla calli under prolonged culture duration

Tissue culture explants of *V. planifolia* were harvested from the AAR-UNMC Biotechnology Research Centre (BRC) tissue culture laboratory in University of Nottingham, Malaysia Campus. Nodal cultures were established using the optimal treatments found in **Chapter 2**. Vanilla calli were induced from nodal explants on callus inducing medium (CIM) (MS medium containing 2.5 mg/L NAA and 1.0 mg/L BAP) for indirect regeneration of vanilla. The primary calli (**Figure 4.2**) were induced after 3 months, hence denoted as "3M". The calli induced were maintained by sub-culturing onto fresh CIM medium every month. At the 6th, 12th and 15th month, the calli were excised and transferred to fresh CIM medium for further proliferation, and the secondary calli generated from primary vanilla calli during these passages were denoted as "6M", "12M" and "15M" respectively (**Figure 4.2**). Nuclei samples were extracted from the 3M, 6M, 12M and 15M vanilla calli and subjected to FCM analysis. Each analysis comparison was carried out in triplicates of biological samples.

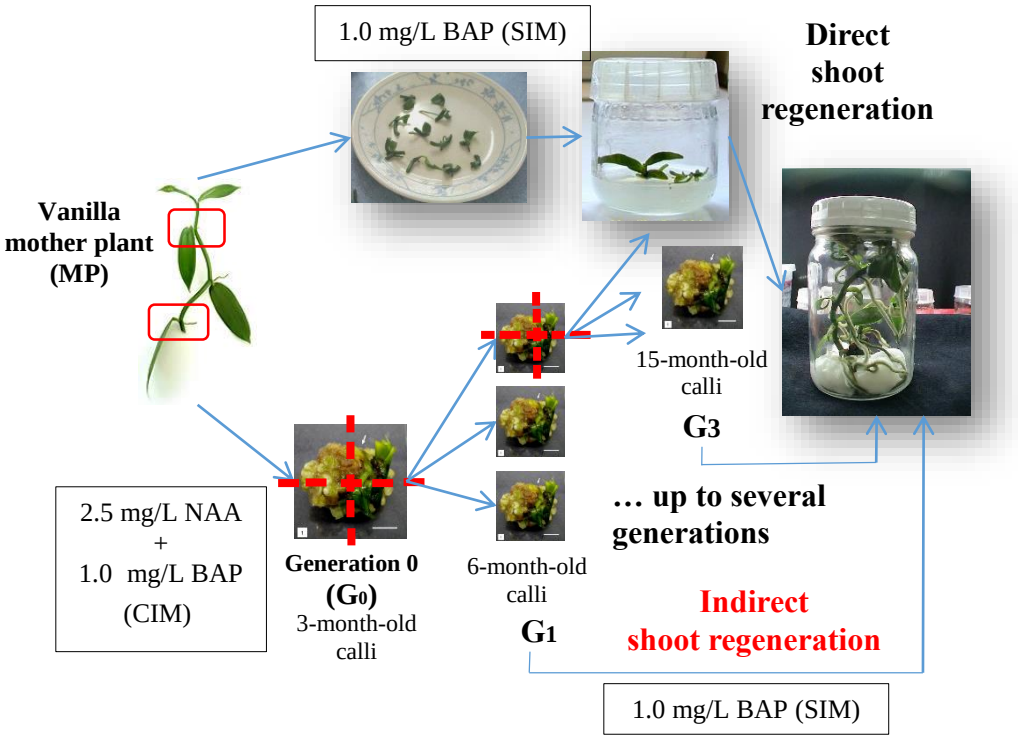


Figure 4.2. Diagrammatic representation of the direct and indirect plant regeneration pathways for vanilla. In a direct regeneration system of vanilla (top of diagram), shoots are regenerated directly from nodal explants without going through callus phrase. In an indirect regeneration of vanilla (bottom of diagram), callus culture was generated from the nodal explants and is maintained through subculturing onto fresh medium on monthly basis. The callus formed from the original nodal explants at generation 0 (G_0) is called the primary callus. Secondary callus cultures are initiated from pieces of tissue dissected from the primary callus. Such a culture could be kept for indefinite periods by successive subculturing. In this study, the vanilla G_0 primary callus developed after 3 months, and the G_1, G_2 and G_3 secondary calli were induced at the 6th, 12th and 15th months. These callus cultures were denoted as 3M, 6M, 12M and 15M.

4.2.3. Plant nuclei extraction and sample preparations

Approximately 50 ± 5 mg of young vanilla tissues (axillary buds, calli, shoot tips, root tips or leaves) and 10 ± 5 mg of leaf tissues of reference standard plants were used for sample preparations. Nuclei suspensions were prepared according to Galbraith *et al.* (1983). Plant tissue were chopped in Tris-$MgCl_2$ buffer (200 mM Tris, 4 mM $MgCl_2 \cdot 6H_2O$, 1.0% (v/v) Triton X-100, pH 7.5) (modified from Pfosser *et al.*, 1995), followed by the recovery of 1.0 ml of the

nuclei suspension which were then filtered through a 45-µm nylon filter mesh to remove cell fragments and large debris. Nuclei were stained with 50 µg mL^{-1} propidium iodide (PI) (Sigma Aldrich, USA), and 50 µg mL^{-1} RNase A (Thermo Scientific, USA) was simultaneously added to the nuclear suspension to prevent the staining of double-stranded RNA. Samples were then incubated on ice, orbital shaken at 200 rpm for 1 hour prior to staining and analyzed 15 min after staining with PI.

Selection of internal standards. The most suitable internal standards were selected from a list of established plant DNA reference standards (**Table 4.2**). Suitability of standard were selected based on the following criteria: (1) Genome size of standard does not differ from those of the sample more than two-fold. This reduces the risk of errors due to instrument nonlinearity; (2) Standards with too similar nuclear DNA content should not be used to avoid peak overlapping. This depends on the resolution of the DNA content histograms, the smallest reasonable difference being 15-20% (Dolezel *et al.*, 2005).

4.2.4. Flow cytometric analyses

Stained plant nuclei (0.5 - 1.0 ml) were loaded onto the sample injection port (SIP) and analysed in a BD Accuri™ C6 flow cytometer (BD Biosciences, USA). Prior to analysis, performance of the flow cytometer was validated using 6- and 8-peak fluorescent bead mixtures, according to the manufacturer's instructions. Fluorescence signals (pulse area measurements) were screened using the FL-2 (585/20 nm) band-pass filter. Analysis of the homogenates was based on light-scatter and fluorescence signals produced from 20 mW laser illumination at 488 nm. Data acquisition and analysis were done in the BD Accuri™ C6 Software Version 1.0.264.21 (BD Biosciences, USA) and ModFit LT™ Software Version 5.0 (Verity Software House (VSH), USA). Threshold levels were set empirically to eliminate the detection of irrelevant debris that are often found in plants homogenates. The threshold values were pre-set at 10,000 for FSC, with a secondary threshold of 1,000 for FL-2 (Galbraith, 2009), and later adjusted accordingly while acquiring data after observing the position

of the nuclei on the bivariate dot plots such that all the nuclei are on scale with the least amount of debris appearing in these plots. Flow rates were routinely operated at slow speed (14 µl per minute) to obtain high resolution histograms. Nuclei populations were gated in a logarithm plot of FL2-A versus FL2-H (**Figure 4.3)**, and 10,000 gated events were collected for each sample.

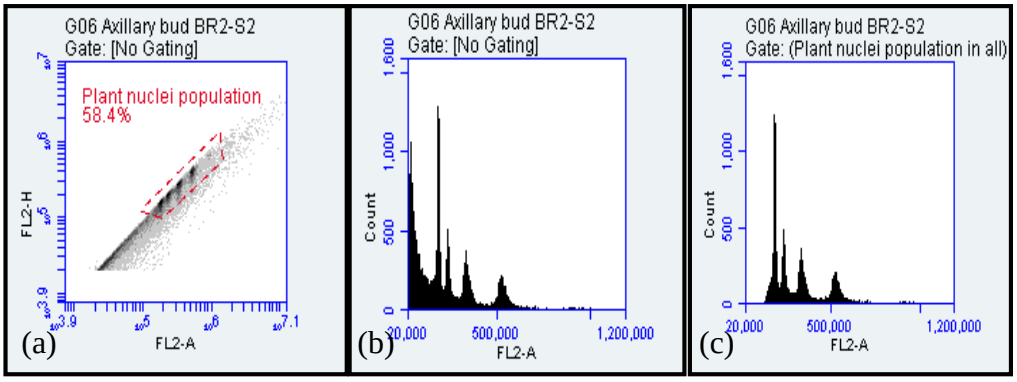


Figure 4.3. Bi-parametric dot plot of FL2-A versus FL2-H (a) **used for gating plant nuclei populations**: FL2-A histogram plot (b) before and after (c) population gating. Background debris were omitted to produce clear FL2 histogram. **Red** box showed the gated plant nuclei population which consisted of 58.4% of entire event populations.

Each experiment was performed and repeated in triplicate on different days, as recommended by Suda (2004) and Doležel and Bartos (2005) to avoid error caused by instrument drift (Loureiro *et al.*, 2006a). In each replicate, at least 5,000 nuclei were analyzed.

4.2.5. Formulas for genome size calculation

Main formulas were used for this study:

Coefficient of variations, CV (%) was calculated using the following formula:

$$\text{Half peak CV (\%)} = \frac{42.46\ x\ \text{width of peak at half the peak height}}{\text{peak position}} \quad (1)$$

Nuclear genome size of *V. planifolia* samples was estimated using *Zea mays* cv. (2C = 5.43, Doležel *et al.*, 1992) as internal reference standard, according to the following formula:

$$V.planifolia \text{ 2C nuclear DNA content (pg)} = \frac{\frac{G_0}{G_1}\text{peak mean of } V.planifolia}{\frac{G_0}{G_1} peak\ mean\ of\ standard} x\ 5.43 \quad (2)$$

Conversion of mass value into base-pair numbers was achieved according to the factor 1 pg = 978 Mbp (Doležel *et al.*, 2003). (3)

4.2.6. Statistical Analyses

Statistical analyses were performed using ANOVA in GenStat software version 18 to assess the differences among different nuclei extraction methods. Binomial data collected on the number of events was subjected to logistic regression in the generalized linear model procedure in GenStat software version 18. When treatments were significantly different (p-value ≤ 0.05), a Holm-Sidak multiple comparison test was used for pair-wise comparison.

4.3. Results

4.3.1. Genome size estimation of vanilla mother plant (MP) using internal standards

The relative genome size of vanilla was estimated to be 2C = 5.50 ± 0.27 pg, using *P. sativum* as internal control as the 2C DNA content was the closest to the 2C of vanilla (**Table 4.2**). Some tissue types showed the presence of multiple endopolyploid peaks with irregular replication ratios, which overlapped with the ploidy peaks of some internal standards (OL, **Table 4.2**). For instance, the 8C, 4C and 2C peaks of *S. lycopersicum*, *G. max* and *Z. mays* overlapped with the 2C peak in vanilla, while the 16C of *S. lycopersicum* and 2C peak of *P. sativum* overlapped with the 2C+P peak of vanilla. *R. sativus* L. Sava was used for ploidy study as its 2N and 4N peaks did not overlap with any endopolyploidy peaks of vanilla (**Figure 4.4**).

Table 4.2. DNA content (pg) estimation in vanilla using different standards.

Standard Plants	2C (pg)	Relative Vp DNA content estimation (pg ± SD)			
		2C	2N+P	2C+3P	2C+5P
R. sativus L. 'Sava'	1.11	4.97 ± 0.70	9.55 ± 0.69	10.89 ± 0.50	18.44 ± 0.50
S. lycopersicum L. 'Stupcke polni' rane	1.96	OL ($8C_{Le}$)	OL ($16C_{Le}$)	13.48 ± 0.43	22.72 ± 0.71
G. max Merr. 'Polanka'	2.50	OL ($4C_{Gm}$)	8.29 ± 0.18	13.00 ± 0.30	22.83 ± 0.19
Z. mays L. 'CE-777'	5.43	OL ($2C_{Zm}$)	8.01 ± 0.16	12.99 ± 0.32	22.42 ± 0.45
P. sativum L. 'Citrad'	8.85	5.50 ± 0.27	OL ($2C_{Ps}$)	12.57 ± 0.16	25.94 ± 0.21
S. cereal L. 'Dankovske'	16.19	5.65 ± 0.44	8.03 ± 0.42	12.68 ± 0.27	21.73 ± 0.25
V. faba L. 'Inovec'	26.90	4.35 ± 0.74	7.88 ± 0.52	12.77 ± 0.31	22.20 ± 0.26

OL: overlapped with internal standard; $2C_{Zm/Ps}$: 2C peak of *Z. mays* or *P. sativum*; $4C_{Gm}$: 4 peak of *G. max*; $8C_{Sl}$: 8C peak of *S. lycopersicum*.

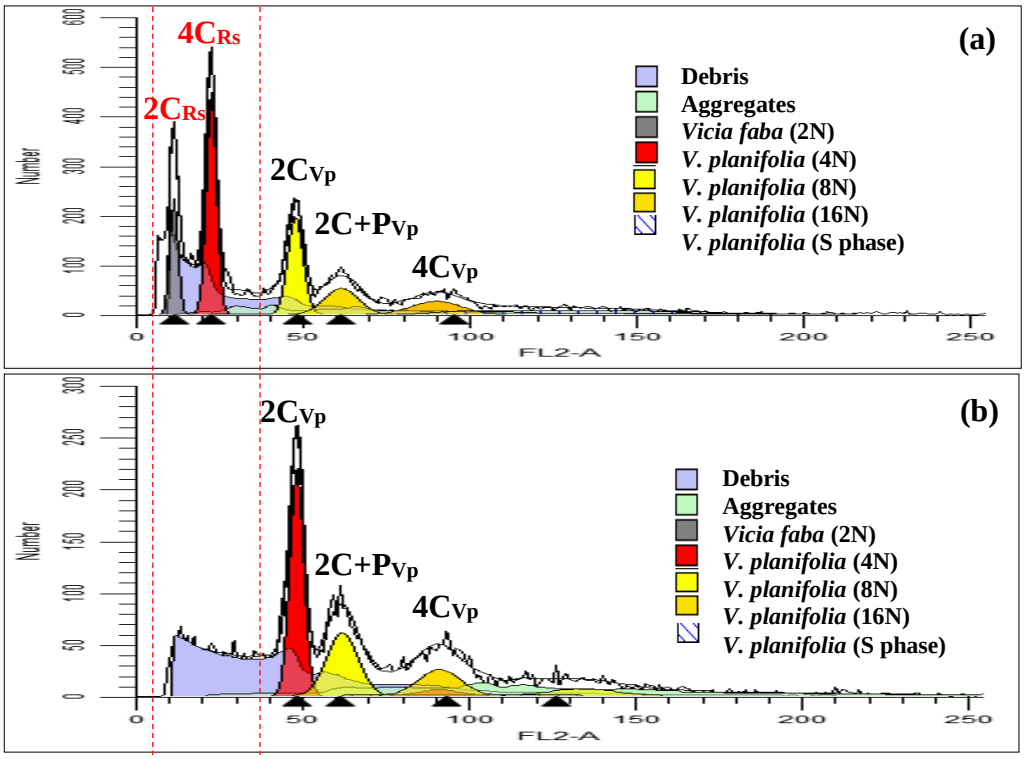


Figure 4.4. (a) *V. planifolia* leaf tip PI-stained nuclei (a) with and (b) without *R. sativus* as an internal control. The **red** lines show the positions of *R. sativus* 2N and 4N peaks. Rs: *R. sativus*, Vp: *V. planifolia*. 2C, 4C, 8C: diploid, tetraploid and octaploid peaks. The 2C peak corresponded to the G_0+G_1 peak of the cell cycle, while at twice the intensity, the 4C peak corresponded to the G_2+M phase of the cell cycle (Bory *et al.*, 2008).

4.3.2. Systemic endopolyploidy vs explant type in vanilla

Polysomaty of *V. planifolia* in different vegetative tissues (leaf, root, axillary bud) was observed (**Figure 4.5-4.6**). Fluorescence histogram revealed endoreplications resulting in 6 main peaks with irregular replication ratios that corresponded to the 1C, 2C, 4C, 8C, 16C and 32C nuclei (**Figure 4.5**). The marginal replication ratio, which is the ratio between peak positions, was irregular, being 1.33 ± 0.02, 1.42 ± 0.02, 1.60 ± 0.01, 1.76 ± 0.01 and 1.69 ± 0.03 (instead of 2) between respective peaks. These replication ratios were constant throughout all different tissues (**Table 4.3**).

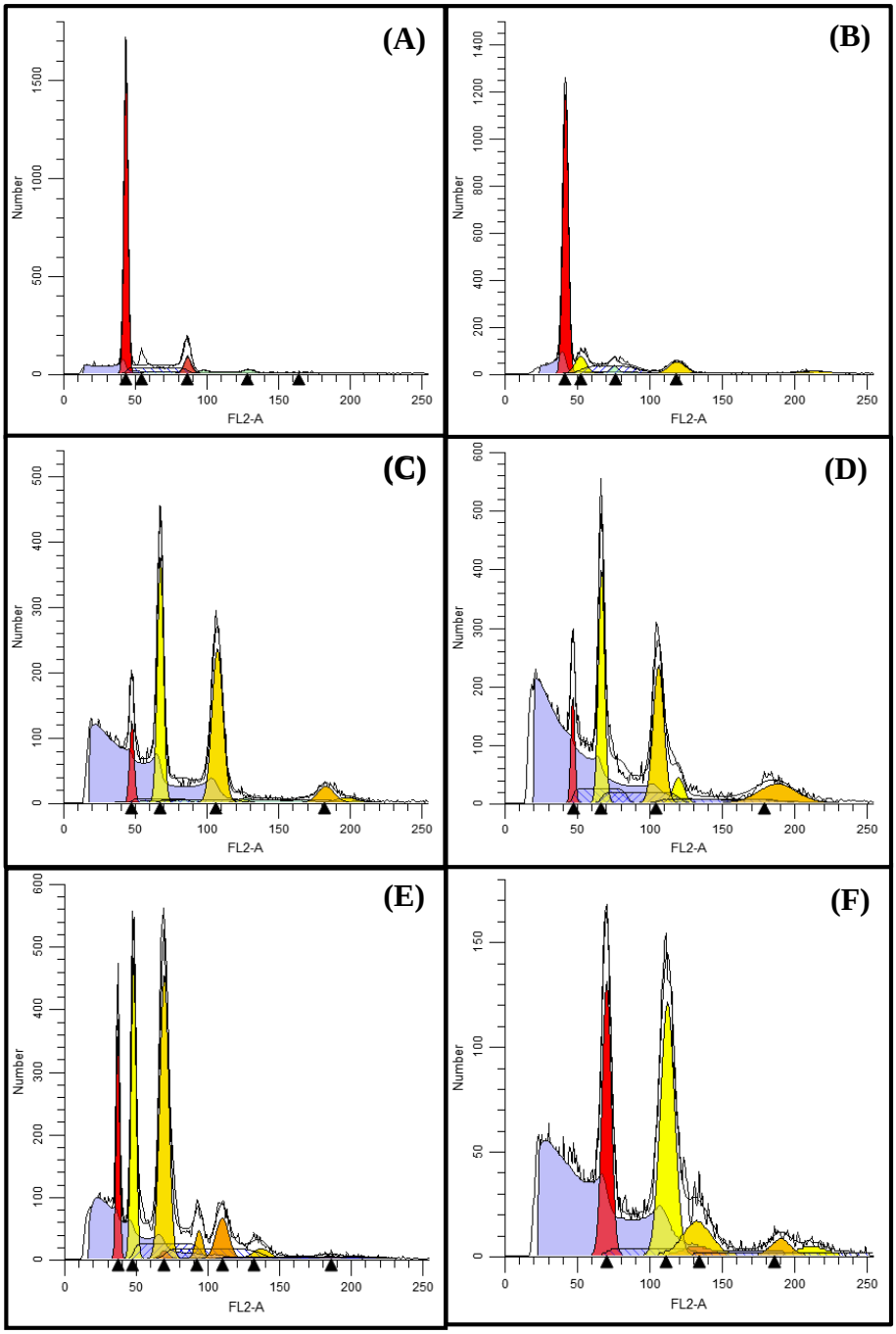


Figure 4.5. Nuclear DNA content and distributions of endopolyploid nuclei in vegetative tisssues of *V. planifolia*. A: Shoot tip, B: Axillary bud, C: Leaf tip, D: Leaf base, E: Root tip, F: Root base. The Y-axis represents the number of nuclei (events); the X-axis represents the relative DNA content (FL-2).

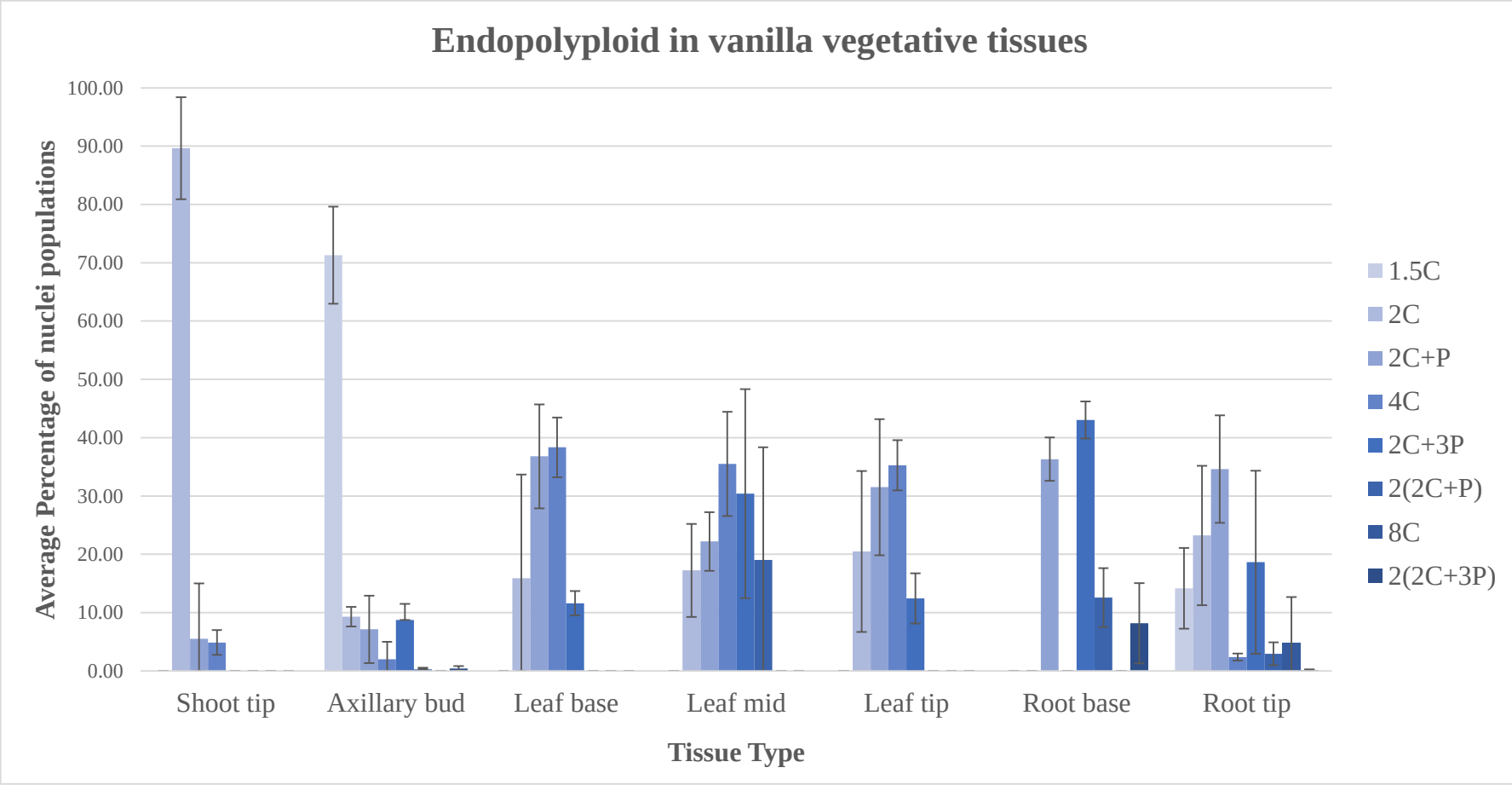


Figure 4.6. Variation in ploidy levels of nuclei populations from different tissues of vanilla.

Table 4.3. Replication ratio in different endopolyploidy levels of vanilla.

Replication	Ratio
1.5C to 2C	1.31 ± 0.03
2C to 2C+P	1.42 ± 0.02
2C to 4C	1.92 ± 0.05
2C+P to 2C+3P	1.60 ± 0.01
4C to 8C	1.95 ± 0.04
2C+3P to 2(2C+3P)	1.83 ± 0.05

Leaf tissues generally underwent an average of four to five endocycles, with 2C, 2C+P, 4C, 2C+3P and 2(2C+P) DNA peaks detectable by FCM (**Figure 4.6**). In most cases, the 2C peak of leaf nuclei was very low and thus not distinguishable in FCM (**Figure 4.4**). Root tip and root base of vanilla had two distinct ploidy patterns (**Figure 4.5E-F**); root tips had 7 to 8 endopolyploid peaks ranging from 1.5C to 2(2C+3P), while root bases had only 4 peaks corresponding to 2C+P, 2C+3P, 2(2C+P) and 2(2C+3P) being detected in FCM. Shoot tips showed a stable dominant (89.65 ± 8.75%) 2C peak (**Figure 4.5A**) which corresponded to the G_1/G_0 peak in cell cycle, and low amounts of 2C+P (5.48 ± 9.50%) and 4C (4.87 ± 2.11%). Differing from the shoot tips, the axillary buds contained a dominant and distinct 1.5C peak (which has the replication ratio of 1.31 ± 0.03 to the 2C peak) (**Figure 4.5B**), and low amounts of each endopolyploid peak: 2C (9.29 ± 1.68 %), 2C+P (7.12 ± 5.79%), 4C (2.00 ± 2.98%), 2C+3P (8.73 ± 2.75 %), 2(2C+P) (0.25 ± 0.30%) and 2(2C+3P) (0.42 ± 0.41%). In this study, a dominant 1.5C peak was only detected in root tips and axillary buds.

4.3.3. Flow cytometric analyses in *in vitro* cultured vanilla calli

In this study, prolonged callus culture duration has shown an increase in the aneuploid populations and average genome content in regenerated culture (**Table 4.4**, **Figure 4.7**). At 3M calli, majority of the populations were 1.5C, 2C and 2C+P. At 6M calli, there was a decrease in 1.5C populations and an increase in 2C, 2C+P and 4C (**Table 4.4**, **Figure 4.7**). Followed, at 12M calli, there was

no detectable 1.5C populations, but a dominated 2C population and sub-dominated 2C+P populations (**Table 4.4**, **Figure 4.7**). In overall, there is an increase in mean DNA content over prolonged callus culture duration (**Table 4.4**), where the highest mean DNA content was found in the populations from 12th month of culture.

Table 4.4. Percentage of nuclei and mean DNA content in 3, 6, 12 and 15 months of callus culture.

Callus Age	**Percentage of nuclei with given DNA content**						**Mean DNA content (pg)**
	1.5C	**2C**	**2C+P**	**4C**	**2C+3P**	**2(2C+P)**	
3M	49.62 ± 26.57	30.34 ± 13.77	27.56 ± 13.69	6.16 ± 3.00	ND	10.06 ± 17.13	825.52
6M	23.85 ± 15.04	55.78 ± 8.38	0.51 ± 0.67	28.84 ± 14.27	0.88 ± 0.58	0.08 ± 0.02	1264.24
12M	ND	55.08 ± 9.91	33.49 ± 4.42	ND	15.42 ± 7.46	0.34 ± 0.06	1309.50
15M	ND	58.98 ± 2.16	31.14 ± 7.71	8.46 ± 11.96	4.03 ± 5.33	0.16 ± 0.23	773.40

ND: not detectable. 3M, 6M, 12M, 15M: 3 months, 6 months, 12 months, 15 months.

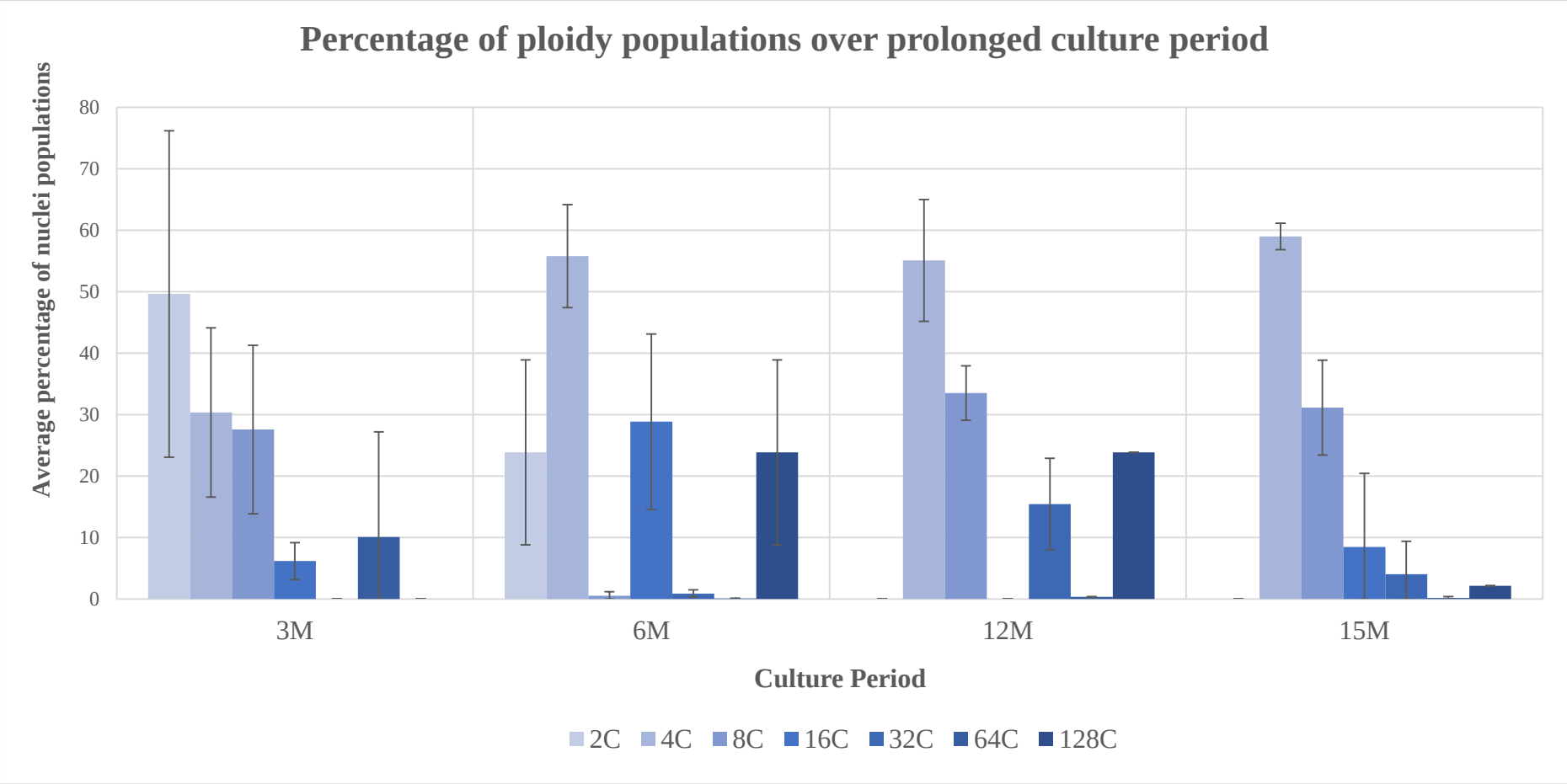


Figure 4.7. Ploidy changes in prolonged callus culture of vanilla.

4.4. Discussion

There is a range of various genome size estimation of *V. planifolia* from published literature. Different 2C DNA content has been reported for *V. planifolia* (2C = 6.3 pg, Kausch and Horner, 1984; 5.03 pg, Bory *et al.*, 2008; 5.22 pg, Lepers-Andrzejewski *et al.*, 2011) and atypical values of 15.5 pg (Zonneveld *et al.*, 2005) and 15.90 pg (Arumuganathan and Earle, 1991). Our estimation using *P. sativum* as internal control (2C = 5.5 ± 0.27 pg) (**Table 4.2**) was closer to the values from recent findings by Bory *et al.* (2008) and Lepers-Andrzejewski *et al.* (2011). Vanilla genome size could be easily over-estimated due to the choice of explants used (Lepers-Andrzejewski *et al.*, 2011). Leaf explants were the most commonly used tissue for plant FCM studies. However, in the case of vanilla, the 2C peak of leaf nuclei was almost not distinguishable in many cases (Bory *et al.*, 2008). This is due to the endopolyploidy phenomenon in vanilla, where 2C DNA content of leaf nuclei was low, causing the 4C peak of vanilla often mistaken as the 2C peak (Bory *et al.*, 2008; Lepers-Andrzejewski *et al.*, 2011).

4.4.1. Selection of suitable internal standards

It is generally agreed that flow cytometric estimation of nuclear DNA amount in absolute units should be performed using internal standard (the nuclei of a standard were isolated, stained and analysed simultaneously with the nuclei of a sample) (Dolezel *et al.*, 2005). The evaluation of a list of standard plants for their suitability as internal control for vanilla FCM was carried out using 7 standard plants (**Table 4.1**). The choice of internal standards could be dependent on the application of study. The optimal 2C genome size estimation of vanilla (2C = 5.5 ± 0.27 pg) was achieved with *P. sativum* as internal control as its 2C peak was the closest among the list, while not being overlapped by the peaks of vanilla plant. However, for the second objectives of this study, which was to evaluate the status of different ploidy levels between other explants, especially the leaf, root and callus tissues which contains multiple ploidy peaks, *R. sativus* L. Sava was a better option as internal standards for subsequent ploidy studies as

its 2C and 4C peaks did not overlap with any endopolyploidy peaks of vanilla (**Figure 4.4**).

In studies by Bory *et al.* (2008) and Lepers-Andrzejewski *et al.* (2011), wheat species *Triticum aestivum* cv. Chinese Spring (2C = 30.9 pg, GC 43.9%, Marie and Brown, 1993) was used as the internal standard, and reported a genome size of 5.03 pg and 5.22 pg, respectively. The differences in reading might be due to errors caused by non-linearity of the flow cytometry instruments (Dolezel *et al.*, 2005), and as the genome size of the wheat species is far from the vanilla, the error could be higher. This showed the importance of selecting a taxon whose DNA amount is not very different from the sample to be studied in order to decrease errors caused by non-linearity (Dolezel *et al.*, 2005). Besides, it is mandatory to check for peak overlapping using a sample without internal standard for each sample in endopolyploidy studies of vanilla.

4.4.2. Systemic differential progressively partial endopolyploidy (PPE) and its implications in PTC of vanilla

The results from flow cytometry demonstated that *V. planifolia* is a polysomatic plant, with systemic endopolyploid cells in different vegetative tissues. Barow and Meister (2003) showed that endopolyploidization is related to the taxonomic position of a species. Some taxonomic groups have many species with highly endopolyploidized tissues, and other do not. Our data are in line with findings from other orchids. In orchids, endoreduplication has been described in the raphid crystal idoblasts of *Vanilla* (Kausch and Horner, 1984) and in parenchyma cells of *Vanda* seedlings (Alvarez and Sagawa, 1965; Alvarez, 1968). Systemic endopolyploidy has been reported in orchids namely *Vanda* Miss Joaquim (Lim and Loh, 2003) and *Spathoglottis plicata* (Yang and Loh, 2004).

The pattern of endopolyploidization differed between the vegetative organs of vanilla. Our analysis of the DNA content in the different vegetative

organs of vanilla have exhibited the absence of endopolyploid cells in vanilla shoot tips and axillary buds (**Figure 4.5-6**), but the presence of partially progressive endoreduplication in vanilla leaf, root, callus and PLB.

Shoot meristems

The near absence of polysomaty in the shoot meristems (axillary buds and shoot tips) **(Figure 4.5A-B)** may be related to the prevention of genetically unstable germ line cells originating from shoot apical meristem tissues, as it has been observed in many plant species, such as *A. thaliana* and *Vanda* Miss Joaquim (Orchidaceae) (Galbraith *et al.*, 1991; Lim and Loh, 2003). This might explain the reason for genetic uniformity of direct nodal regenerants of vanilla as reported in this study (**Chapter 3**), as well as in other vanilla SV studies (Sreedhar *et al.*, 2007).

Roots

Vanilla root tips are highly endopolyploid. Endopolyploidy is the highest in roots of *Chenopodium quinoa* (Kolano *et al.*, 2008). The same was reported in quinoa roots which exhibited a high mean C-value, indicating the presence of endopolyploid cells (Kolano *et al.*¸ 2008). The finding of the highest endopolyploidy in root and hypocotyl might be explained by the presence of a large proportion of vascular tissue in these organs. Large cell size is required for the functioning of this tissue, and this can be achieved by endopolyploidization (Baluska, 1990; Sliwinska and Lukaszewska, 2005). In other orchids, such as *Dendrobium*, endopolyploidy was found in root tips and newly expanded young leaves (Jones and Kuehnle, 1998).

Leaves

Vanilla leaves are highly endopolyploid. Endoreduplication was found to occur in leaf tissues of *Spathoglottis plicata* (Yang and Loh, 2004) and *Dendrobium* species (Jones and Kuehnle, 1998). Furthermore, when the tip, middle and basal parts of the mature leaf were examined, it showed that all leaf tip, leaf base and mid of mature leaf exhibited endopolyploidy (**Figure 4.6**). In *S. plicata* (Orchidaceae), endopolyploidy was only found in the leaf base part of mature leaves, where less green meristem cells were present at the basal petiolar end of

the older, photosynthetically active cells at the leaf tip (Altamura *et al.*, 1987; Yang and Loh, 2004). This might possibly explain the reason why leaf explants of vanilla were more recalcitrant and less responsive to regeneration, as described in **Chapter 2** and Tan *et al.* (2011a). The marginal replication ratio calculated from this study (1.42 ± 0.02, 1.60 ± 0.01, 1.76 ± 0.01, 1.69 ± 0.03) corresponded to the findings by Bory *et al.* (2008), who found a ratio of 1.43 (± 0.03), 1.63 (± 0.04), 1.76 (± 0.04), 1.82 (± 0.02). This showed that the replication ratio of endopolyploidy peaks was consistent within genus (Trávníček *et al.*, 2015). The mechanism underlying PPE, whether they resulted from partial replication of the DNA or excision of DNA following whole genome replication, were not known, but this phenomenon occurred in many orchids (Brown *et al.*, 2007; Bory *et al.*¸2008). Hypothesis has it that if the partial replication is retained, each successive peak in *V. planifolia* should not be called 2C, 4C, 8C and 16C, but rather 2C, 2C+P, 2C+3P, 2C+7P, 2C+15P, where P is the DNA content of the replicated part of the 2*n* nucleus (Bory *et al.*, 2008). However, further analysis to confirm this hypothesis was required (Bory *et al.*, 2008). This finding also suggested that future studies of SV in leaf or root derived regenerants could take into consideration the possibility of partial replication during tissue culture process.

Systemic variation of cellular endopolyploidy in different tissues of *V. planifolia* was reported in this chapter. Pre-existing endopolyploidization in explant cells is probably one of the reasons for early polyploidization in callus cells (Fras and Mluszynska, 2004; Fras *et al.*¸ 2007). When organs with endopolyploid cells serve as explants, some regenerants can be polyploid (Chen *et al.*, 2009). Tissues expressing polysomaty should be avoided as explants, or else the regenerants should be monitored for ploidy (Kolano *et al.*, 2008).

The role of endoreduplication in plant development is still not well understood (Yang and Loh, 2004). It was proposed that the presence of endopolyploidy was associated with several factors, such as taxonomic position of a species, life cycle, genome size, and organ type (Barow and Meister, 2003). These patterns of endopolyploidy might be character reflection of the developmental and physiological properties of the tissue. This finding provides

useful information for understanding of the plant development and for industrial propagation of orchids.

4.4.3. Prolonged callus culture duration and SV

Prolonged callus culture duration has shown an increase in the aneuploidy populations in regenerated culture (**Table 4.4**). In this study, prolonged callus culture showed an increase in average genome content (**Table 4.4**) and ploidy patterns (**Figure 4.7**). Prolonged callus or PLB culture duration has also been shown to increase SV by chromosomal doubling or aberrations in multiple other plant species, namely *Arabidopsis thaliana* (Sedov *et al.*, 2014), *Asparagus officinalis* (Pontaroli and Camadro, 2005), *Elaeis guineensis* (Lucia *et al.*¸ 2011), *Phalaenopsis alliance* (Zhou and Sakaguchi, 2003), *Zea mays* (Molina and Garcia, 1998), *Dianthus caryophyllus* (Yaacob *et al.*, 2013b) and *Asparagus* spp. (Kunitake *et al.*, 1998; Raimondi *et al.*, 2001). Cytogenetic stability could be genotypic dependent. High cytogenetic stability in a callus culture and regenerants was reported for several species, namely *Crepis tectorum* (Sengupta *et al.*, 1988), *Pinus taeda* (Franklin *et al.*, 1989) *Ruscus hypophyllum* (Jha *et al.*, 1989) and *Larix* × *eurolepis* (Wyman *et al.*¸1992) and *Fagopyrum tataricum* (L.) Gaertn (Betekhtin *et al.*, 2017). The occurrence of endopolyploidy in callus-derived cells might explain why callus-derived plantlets were genetically distinct.

Besides, an increase in the ploidy level might lead to a decrease in or even a total loss of the regeneration potential of the polyploid plants (Ochatt *et al.*, 2000), for example, in *Arabidopsis thaliana* (Fras and Maluszynska 2003) and *Vitis vinifera* (Prado *et al.*, 2010). In this study, regeneration potential dropped after 6 months of callus culture, as the first to second passage (P1-P2) took 3 months, while the second to third passage (P2-P3) took 6 months for the callus the proliferate to a considerable size. Therefore, an increase in ploidy might be one explanation to the decrease in regeneration potential after 12 months of callus culture.

4.5. Conclusion and future prospects

The evaluation of a list of standard plants for their suitability as internal control for vanilla FCM was carried out. The optimal 2C genome size estimation of vanilla (2C = 5.5 ± 0.27 pg) was achieved with *P. sativum* as internal control as its 2C peak was the closest among list, while not being overlapped by the peaks of vanilla plant. The estimated 2C value for vanilla genome size was also close to the values from recent findings by Bory *et al.* (2008) and Lepers-Andrzejewski *et al.* (2011).

Endopolyploidy phenomenon was observed in vanilla. Irregular marginal replication ratios of 1.33 ± 0.02, 1.42 ± 0.02, 1.60 ± 0.01, 1.76 ± 0.01 and 1.69 ± 0.03 (instead of 2) between subsequent peaks were observed, which exhibit "progressively partial endoreduplications", a unique phenomenon in some orchid species (Hřibová *et al.*, 2016). The marginal replication ratio was constant between different tissues, but ploidy population levels varied among different tissues, with shoot tips having dominant 2C peaks and a small 4C peak, axillary buds having dominant 1.5C peak and relatively smaller populations of other endopolyploidy peaks, and leaf and root tissues having relatively balanced populations of different mixoploids, except that 1.5C population was also found in root tissues but not leaf tissues.

In callus culture and prolonged culture period, the pre-existing polyploidy populations showed amplification over time. The prolonged culture of vanilla correlated to an increase in polyploid populations and mean genome sizes over time, suggesting the role of prolonged culture stress in increasing SV levels of vanilla.

The knowledge of the basis underlying endopolyploidy is important for understanding the regulation of gene expression in differentiated tissues, and of the nature of the tissues used for plant regeneration and transformation. This may be useful in decreasing the occurrence of somaclonal variation and polyploidization *in vitro* (D'amato, 1985). The callus samples from prolonged culture durations (3 months, 6 months and 12 months) established in this chapter were thus used for further study at transcriptome level in **Chapter 5-6**.

CHAPTER V

DE NOVO TRANSCRIPTOME AND microRNA PROFILING OF LONG-TERM CULTURED VANILLA CALLUS

Chapter V

De novo transcriptome and microRNA profiling of long-term cultured vanilla callus

5.1. Introduction

The increasing occurrence of somaclonal variation (SV) in tissue culture of major crops has brought to light the importance of research on SV. A growing number of studies have started to focus on the investigation of SV to evaluate stability of the *in vitro* grown plants, where high frequencies of epigenetic variation have been revealed, in addition to genetic variation. Additionally, there is also an increasing use of these cultures as experimental model systems for uncovering dynamic epigenetic mechanisms acting during the dedifferentiation and subsequent developmental reprogramming of plant cells (William *et al.*, 2003; Avivi *et al.*, 2004; Koukaloa *et al.*, 2005; Berdasco *et al.*, 2008).

Previous studies have focused on the study of model species whereby genome sequences were available. The rapidly advancing technology of next-generation sequencing has now opened up the doors for the investigation of global transcriptional changes in non-model plants with high commercial values, like vanilla. Several studies have attempted using species such as *Arabidopsis thaliana* (Miguel and Marum, 2011), *Musa sp.* (Sahijram *et al.*, 2003; Nwauzoma and Jaja, 2013) and *Elaeis guinensis* (Jaligot *et al.*, 2009) as models to study SV. However, the occurrences of SV were also found to be species-dependent.

In vanilla, callus-mediated plant regeneration and prolonged *in vitro* culture duration were shown to cause abiotic stresses and increased incidence of SV. High SV rates were detected using genetic and epigenetic markers in callus-derived vanilla regenerants (**Chapter 3**), while prolonged culture duration has

been shown to increase polyploid populations and mean genome sizes in vanilla calli (**Chapter 4**). This makes the vanilla callus culture an interesting material to be studied for further understanding of somaclonal variations. In this chapter, total RNAs of different prolonged culture durations (3-month, 6-month and 12 month-old) of vanilla callus cultures were sequenced using the *de novo* NGS technology at mRNA and miRNA levels to study the *in vitro* gene expressions and regulations.

RNA-Seq has great sensitivity and a high base resolution capacity to capture the most complete picture of the transcriptome which can consist of both known and novel genes and small RNAs (Forestan *et al.*, 2016). It also allows the discrimination between splicing variants, alleles and sense/antisense isoforms. In addition, RNA-Seq enables the identification of the genome-wide mechanisms involved in plant stress response, adaptation and resilience (Agarwal *et al.*, 2014). RNA-Seq also allows the sequencing of transposable elements (TE) transcripts that move around the genome via RNA intermediates, called retrotransposons (Grandbastien, 2015), which is one of the proposed causes of SV.

The RT-qPCR method has been frequently used as a "gold standard" to corroborate data obtained using other techniques, such as microarrays and NGS sequencing (Bustin, 2010). It is a high sensitivity and high sequence-specific technique which allows the detection of a target sequence molecule of down to very low quantities with reproducible results (Begas *et al.*, 2014; Gotfred-Rasmussen *et al.*, 2016). This makes RT-qPCR a standard technique of choice for carrying out validation. Despite the extensive use of RT-qPCR, it is important to understand the basic principles, potential sources of errors and potential problems that might arise with the technique (Pffafl, 2010).

To date, epigenetic modifications were known to regulate the gene expressions of plants under stresses, which could lead to the occurrence of somaclonal variation (SV) at morphological and genetic levels (Krishna *et al.*, 2016). However, the mechanisms underlying SV were not well understood. It is proposed that a complicated mechanism involving activation and suppression of genes such retrotransposons is responsible for causing epigenetic variations *in*

vitro (Azman *et al.*, 2014). The expression of these genes could be regulated by post-transcriptional factors and the action of small RNAs, such as microRNA (miRNA) and piwi-interacting RNA (piRNA) (Rodriguez-Enriquez *et al.*, 2011). There is still little known about the mechanisms involved in this aspect of SV, and further knowledge in this field is important to contribute in future applications in vanilla, such as the development of effective pre-screening tools for clonal fidelity in commercial scale micropropagation, prevention of post-transcriptional gene silencing in genetic modifications, or *in vitro* crop improvements to increase gene pool for sustainable agriculture.

The main aim of this study was to study the global mRNA and miRNA expression profiles of vanilla calli maintained over prolonged culture duration using the next generation sequencing (NGS) and reverse transcription-quantitative real time PCR (RT-qPCR) technology. The specific objectives to be achieved for the main aim of the study were: (i) purification of high quality total RNA from 3-, 6- and 12-month-old vanilla calli samples for NGS sequencing and RT-qPCR analysis; (ii) NGS sequencing of mRNA and miRNA of the vanilla calli samples; (iii) global transcriptome profiling, relative gene expression analysis and functional annotation of mRNA and miRNA genes significantly expressed in vanilla calli over prolonged callus culture duration to identify potential genes related to SV; (iv) identification of potential miRNA target genes through target prediction analysis; (v) optimization of RT-qPCR protocols for mRNA and miRNA qPCR analysis: design of gene primers, evaluation of primer specificity and efficiency, selection of suitable and stable reference gene candidates for target gene analysis; and (vi) validation of NGS sequencing results by differential expression analysis of each mRNA or miRNA-of-interest using optimized RT-qPCR protocols.

5.2. Materials and methods

5.2.1. Chemicals reagents, kits and apparatus

Lists of chemicals and kits, and the apparatus and software used in this study are stated in **Appendix A** and **Appendix B**, respectively. Prior to any RNA-related work, the following procedures were conducted to eliminate any possible RNase that might be present and cause degradation to RNA samples: All working surfaces, micropipettes, tube holders, pipette tip boxes and gel electrophoresis tanks were wiped with RNaseZap® RNase Decontamination Solution (Thermo Fisher Scientific, USA). Other tubes and plastic wares were soaked in 0.5 M NaOH for at least 10 min, then rinsed three times with 0.1% (v/v) diethyl pyrocarbonate (DEPC)-treated ultrapure (18 MΩ) H_2O and dried in 80ºC oven overnight. UltraPure™ DNase/RNase-free distilled water (Invitrogen, Thermo Fisher Scientific, USA) was used as the universal diluent in all RNA extraction and RT-qPCR preparations.

5.2.2. Oligonucleotide Nucleotide Synthesis and Primer Source

The primer oligonucleotides used in this study were synthesized and de-salted by Integrated DNA Technologies (IDT), Singapore. The primer sequences used in RT-qPCR validation (listed in **Appendix I**) were designed from analysis output of the NGS sequencing (**Section 5.3.5**).

5.2.3. Source of plant materials

Vanilla calli materials were prepared as described in **Figure 4.1**. Calli were induced from nodal cultures of *in vitro* grown *V. planifolia* plantlets in callus inducing medium (CIM) (MS supplemented with 2.5 mg/L NAA and 1.0 mg/L BAP, 30 g/L sucrose and 3.5 g/L Phytagel™), and subcultured by transferring the cells onto fresh CIM medium every month. The primary calli masses were

macro-dissected at the 3rd, 6th and 12th month of culture (hence denoted as 3M, 6M and 12M) using sterile scalpels into two parts: the first half of the callus masses were excised into approximately 0.1 g each, and maintained on fresh CIM medium; while the other half of the callus was homogenized by grinding immediately with liquid nitrogen and stored at -80ºC in Trizol® Reagent to preserve the integrity of RNAs until use. For every 100 mg of ground tissue, 1.0 mL of Trizol® Reagent was added. Vanilla calli samples cultured under prolonged culture period were denoted as:

(i) 3M-1, 3M-2 and 3M-3 for biological replicates 1 to 3 of three-months-old calli,

(ii) 6M-1, 6M-2 and 6M-3 for biological replicates 1 to 3 of six-months-old calli, and

(iii) 12M-1, 12M-2 and 12M-3 for biological replicates 1 to 3 of twelve-months-old calli.

The mRNA and miRNA libraries used for RNA sequencing (RNA-seq) experiment were constructed from six vanilla calli samples. A total of three time points (3 months, 6 months and 12 months of culture), each with two biological replicates (biological replicates 1 and 2), were sequenced to obtain a more comprehensive profile of vanilla callus culture in response to prolonged culture duration stress. Genes-of-interest selected from the bioinformatics output were validated using RT-qPCR for all three biological replicates.

5.2.4. Overview of Research Methodology

An overview of research workflow applied in this part of the thesis study is illustrated in **Figure 5.1a** (RNA sequencing workflow) and **Figure 5.1b** (RT-qPCR workflow).

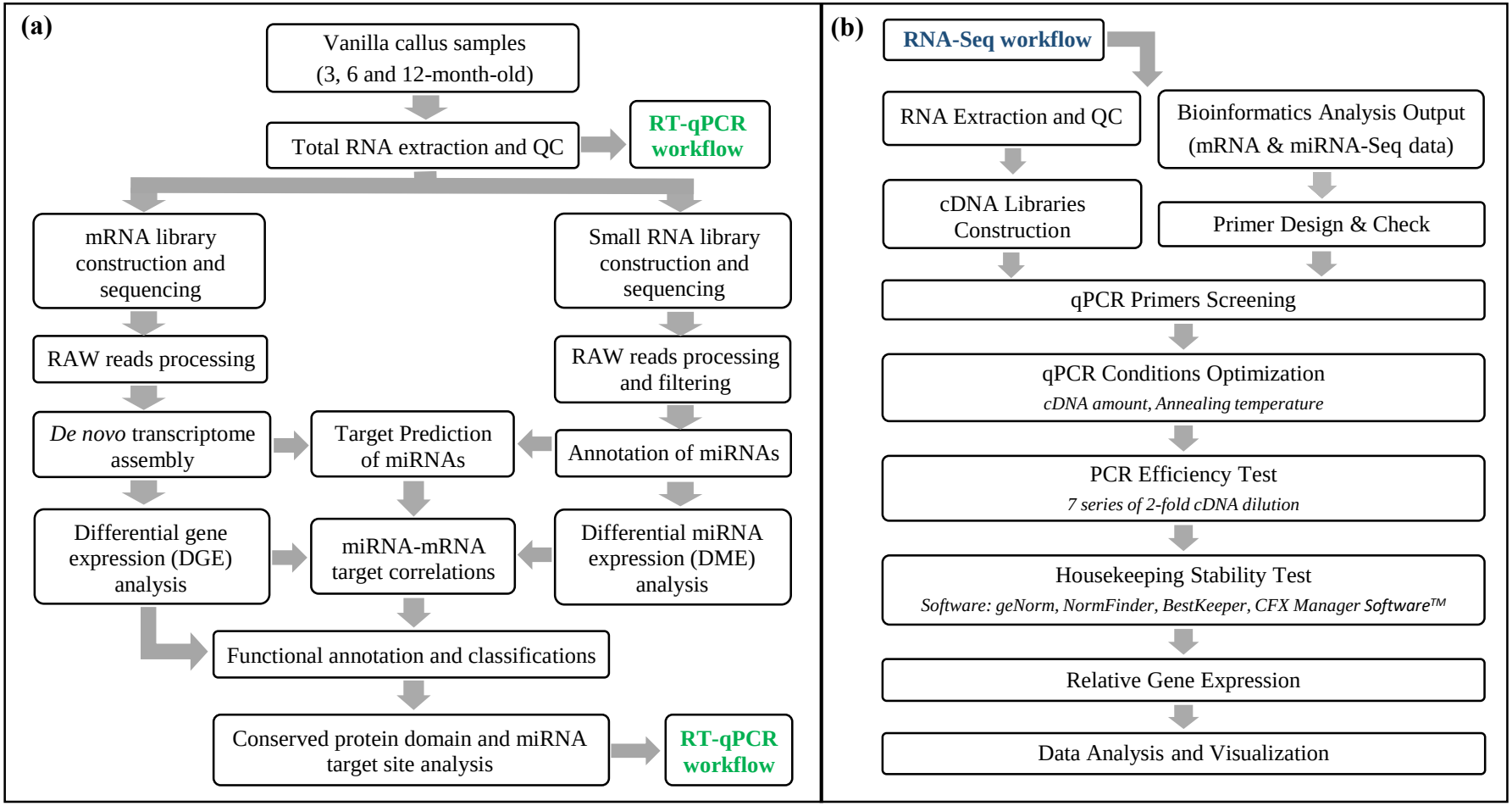


Figure 5.1. Overview of (a) RNA-Seq and (b) qRT-PCR workflow.

5.2.5. Total RNA Extraction, RNA Quantity and Quality Check

Total RNA was isolated from the ground tissues, which were preserved in Trizol® Reagent (**Section 5.2.3**), using Direct-Zol™ RNA MiniPrep kit (Zymo Research, USA) following the manufacturer's instruction. RNA quantity and quality check (QC) were performed prior to sequencing. Two per cent (2.0%) agarose gel stained with 3.0% (v/v) SybrSafe was used for the preliminary check of overall quality and yield of the RNA samples (Green and Sambrook, 2012). Intact total RNA should have sharp 25S and 18S rRNA bands showing 25S:18S ratio of 2:1, where partially degraded RNA will have a smeared appearance, will lack the sharp rRNA bands, or will not exhibit a 2:1 ratio. Completely degraded RNA will appear as a very low molecular weight smear. Before loading into an RNA gel, samples were mixed at equal volume with 2X RNA loading dye (Thermo Scientific, USA) and heated to 95°C for 4 min to denature the RNA prior to loading. RNA purities and concentrations were determined by Nanodrop ND-1000 (Thermo Scientific, USA) and Qubit RNA BR assay kit (Life technologies, USA; Q10210) measurement. Total RNA quality was assessed using the Agilent RNA 6000 Nano Kit (Agilent Technologies, USA; 5067-1511) and the Agilent 2100 Bioanalyzer (Agilent Technologies, USA). An absorbance ratio of A_{260}/A_{280} between 1.8 to 2.2 nm was accepted as pure, while a RIN value of > 7.0 (from the scale of 1, degraded, to 10, intact) was considered as good quality for qPCR analysis (Bustin *et al.*, 2009; Bustin *et al.*, 2010; Marx, 2013; Taylor *et al.*, 2010). Samples with A_{260}/A_{280} in the range of 1.8-2.2, minimum RNA Integrity Index (RIN) value of 7.6 and minimum concentration of 200 ng/µl in a total of 10 µg total RNA quantity were used for next generation sequencing (NGS) and RT-qPCR. High molecular weight and high integrity RNA samples were aliquoted for immediate shipping (in dry ice) to the DeepSeq department at University of Nottingham (UK campus) to perform RNA-sequencing, and the remaining sample aliquots were partially kept in RNA form and partially reverse transcribed into cDNA for RT-qPCR experiment, which were stored at -80ºC, to minimize the numbers of thawing cycle.

5.2.6. mRNA-sequencing and data analysis

5.2.6.1. mRNA poly-A enrichment, bar-coded sequencing libraries construction and sequencing

The mRNA sequencing wet lab was performed by the DeepSeq Department at the University of Nottingham, UK. An equal amount (1.0 µg) of total RNA was used for the enrichment of mRNA using NEBNext poly(A) mRNA Magnetic Isolation Module (NEB, E7490). Illumina stranded whole transcriptome sequencing libraries were prepared using NEBNext Ultra Directional RNA library prep kit for Illumina (NEB, E7420S) with NEBNext Multiplex Oligos for Illumina (Index Primers Sets 1 and 2) (NEB, E7335S and E7500S). The QC for library size distribution and concentration was performed using Bioanalyser HS kit (Agilent Biotechnologies, 5067-4626). Libraries were then quantified using qPCR (Kapa Biosystems, KK4824). Six barcoded libraries were pooled at desired concentrations, denatured and loaded for sequencing according to manufacturer's instructions. Sequencing was performed on the Illumina NextSeq500^{TM} sequencing platform (Illumina Inc., USA) to generate 2 x 150 bp reads at 30 million reads.

5.2.6.2. Raw sequence pre-processing and *de novo* transcriptome assembly

The processing of raw sequence data was conducted using CLCBio Genomic Workbench 8.0 (http://www.qiagenbioinformatics.com) (Qiagen, Germany). Contaminations of adaptor sequences, low quality tags with ambiguous sequences 'N' (Limit = 2) reads, low quality bases (lower than Q30) reads, short read length (reads less than 35 bp) and terminal nucleotides (15 nucleotides at the 5'end and 8 nucleotides at the 3' end) were removed in a trimming process. FastQC (Andrews, 2010) was run on reads before and after processing to check for trimming quality. Clean reads were then assembled with overlaps into longer fragments named contigs. Based on paired-end (PE) reads, contigs derived from the same transcript sequences and the distances between them were detected and

calculated respectively. These contigs were further combined into an "N" representing unknown sequences between two contigs and the scaffolds were repaired with PE reads to isolate sequences with the least Ns and no extension on either end. Such non-redundant transcripts are defined as unigenes. *De novo* assembled master contigs were submitted to the differential gene expression (DGE) analysis and annotation pipeline (**Section 5.2.6.3**).

5.2.6.3. Construction of a Digital Gene Expression Atlas and Annotation Pipeline

The resulting 145,550 contig sequences from the *de novo* combined assembly were used as a scaffold for transcripts mapping in differential gene expression (DGE) analysis. Gene expression levels were normalized and calculated using fragments per kilobase million (FPKM) metrics. Significance of gene expression differences were quantified in CLCBio Genomic Workbench software version 8.0 (http://www.qiagenbioinformatics.com) (Qiagen, Germany). False discovery rates (FDR) were used to determine the P-value threshold in multiple tests; while log two of fold change (log_2 FC) that was larger than 1 or smaller than -1 and FDR-corrected P-value that was smaller than 0.05, were used as criterion to judge the significance of DGEs under prolonged tissue culture stress condition. Time course expression analysis based on an R package, maSigPro (Nueda *et al.*, 2014), was performed in the Blast2GO PRO program (Götz *et al.*, 2008) to detect the genomic features with significant temporal expression changes using two steps regression strategy.

A total of 22,025 contig sequences with differentially expressed genes were annotated via local BLASTX alignment against several public databases using Basic Local Alignment Search Tool (BLAST) of the National Centre of Biotechnology Information (NCBI) at an e-value threshold of e < 1E-3 and sequence similarity > 30%. General protein databases (non-redundant protein (nr) – *Viridiplantae* (tax-id: 33090), Ref-Prot, Swissprot and pdbaa) were downloaded via the FTP server of NCBI BLAST databases (ftp://ftp.ncbi.nlm.nih.gov/blast/db/), while the specialized nucleotide databases

were downloaded from their respective web databases: Plant repetitive element resource (Nussbaumer *et al.*, 2013; Spannagi *et al.*, 2007, 2016) from the MIPS Repeat Element Database (mips-REdat) v_9.3p (ftp://ftpmips.helmholtz-muenchen.de/plants/REdat/); orchid transcriptomics resource (Su *et al.*, 2013, 2017; Chao *et al.*, 2014, 2017) from the Orchidstra v2.0 database (http://orchidstra2.abrc.sinica.edu.tw/orchidstra2/downloads.php) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Kanehisa, 2000; Kanehisa and Goto, 2000; Kanehisa *et al.*, 2016, 2017) using an E-value $<10^{-3}$.

The gene ontologies (GO) were then annotated according to the molecular functions, biological processes and cellular components ontologies with default parameter values using the Blast2GO PRO program (Götz *et al.*, 2008). After the Blast2GO mapping process, EC numbers from the KEGG pathway (http://www.genome.jp/kegg/) and GO terms were generated. Further annotation of GO functional classification, distribution and metabolic pathways at p-value ≤ 0.05 was performed.

5.2.7. Small RNA sequencing – microRNA analysis

5.2.7.1. Constructing and sequencing small RNA libraries

Small RNA sequencing wet lab was performed by the DeepSeq Department at the University of Nottingham, UK. RNA libraries were generated using the NEBNext® Small RNA Library Prep Set for Illumina® (Set 1). Library QC was performed using the Bioanalyzer HS kit (Agilent biotechnologies, 5067-4626) and quantified using qPCR (Kapa Biosystems, KK4824), pooled at desired concentrations, denatured and loaded for sequencing according to manufacturer's instructions. Sequencing of small RNAs was performed on the Illumina NextSeq500™ sequencing platform (Illumina Inc., USA) to generate 2 x 75 bp reads.

5.2.7.2. Sequence analysis of miRNA

An overall workflow of miRNA sequence analysis is summarized in **Figure 5.2**.

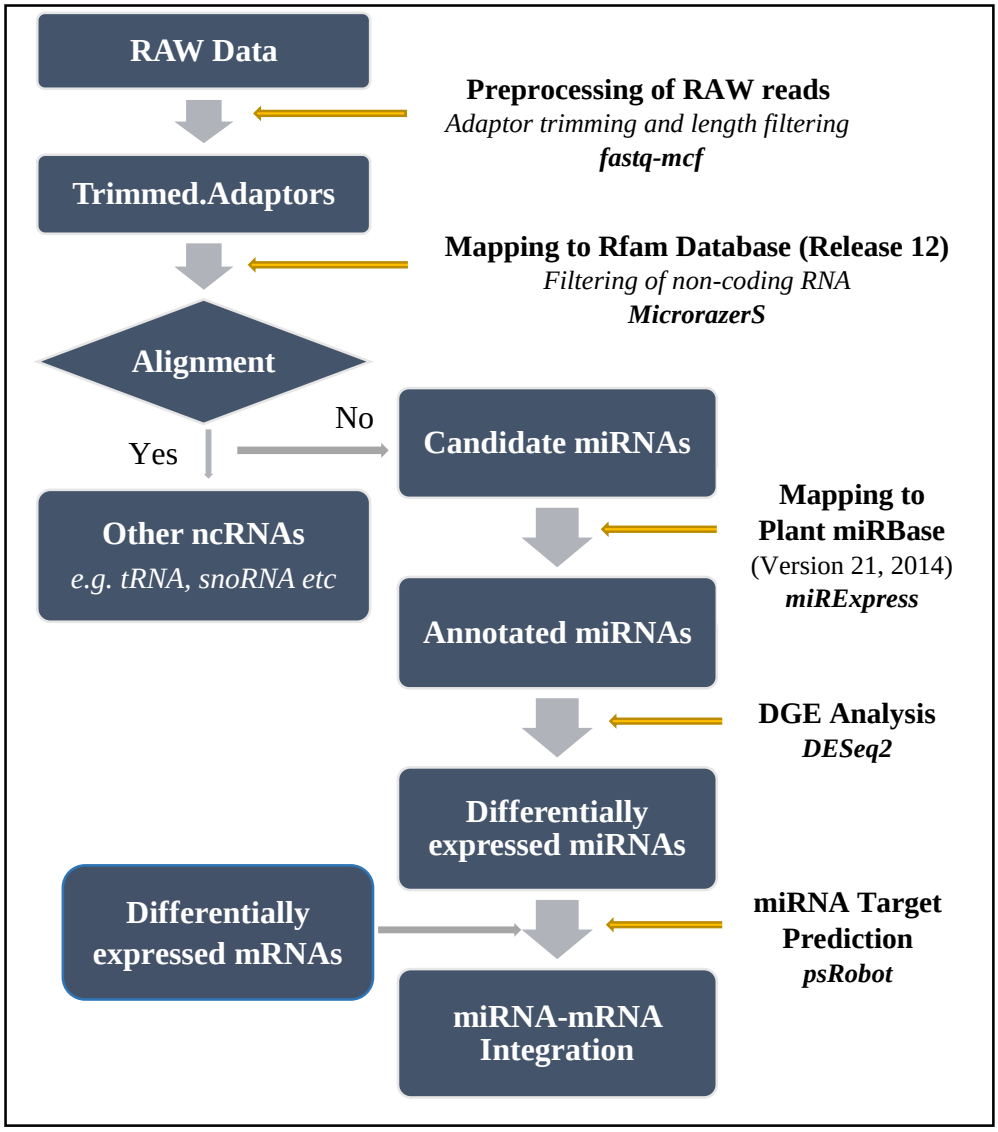


Figure 5.2. Schematic diagram of the miRNA bioinformatics pipeline.

The raw sequence files generated were subjected to quality control analyses using the FastQC software (Andrews, 2010). The fastq-q-mcf software version 1.04.636 (http://code.google.com/p/ea-utils/wiki/FastqMcf) was used to remove the Illumina small RNA adaptors, ambiguous nucleotides (no ambiguous nucleotide was allowed) and low-quality terminal ends, discard short sequences, and filter reads at the length from 15 to 30 bp. The total number of sequence tags was counted, while same sequences were merged and clustered.

Clean and processed reads were then mapped to the RFAM database (Release 12) to filter away non-coding RNAs (such as rRNAs, tRNAs, snRNAs and snoRNAs) using the MicrorazerS software (Emde *et al.*, 2010). The filtered miRNA candidates were annotated by mapping to the miRBase (Version 21, 2004) against known miRNAs of 73 plant species (**Appendix F**) using miRExpress (Wang *et al.*, 2009b). Differential gene expression analysis was carried out using DESeq2 software (Love *et al.*, 2014). The DESeq2 software provided methods to test for differential expression by use of negative binomial generalized linear models. The estimates of dispersion and logarithmic fold changes incorporate data-driven prior distributions (Love *et al.*, 2014). Differentially expressed genes were subjected to target prediction using psRobot module (Wu *et al.*, 2012). The miRNA candidates were aligned to the potential target transcripts from previous mRNA assembly (**Section 5.2.6.3**), by screening for perfect or near-perfect complementary match to their target-site. The penalty score parameters for complementary match have been set at 4 for a perfect target of transcript (Srivastava *et al.*, 2014).

Prediction of secondary structure

The secondary structure of candidate pre-miRNA sequences of potential miRNA homologs was predicted using the Zuker folding algorithm with the MFOLD-3.2 program (Zuker 2003). Two energy measurements, the adjusted minimal folding free energy (AMFE) and the minimum folding free energy index (MFEI) were calculated based on the formula below (Zhang *et al.*, 2006):

$$AMFE = \frac{MFE}{\text{Length of a potential pre-miRNA (LP)}}\ x\ 100$$

$$MFEI = \frac{AMFE}{(G+C)\%}$$

The ΔG(MFE) values (kcal/mol) of stem-loop structures generated by MFOLD program were applied to calculate their negative minimal free energies (MFEs), which is directly correlated with the sequence length, to normalize the potential effect of sequence length on MFE and to differentiate miRNAs from other RNAs. AMFE is defined as the MFE of a nucleotide length.

5.2.8. Integration of mRNA-seq and miRNA-seq data

miRNA-mRNA target relations and expression correlations were integrated for identifying reliable target (**Figure 5.3**). The results were represented in two formats of correlation heat map and network views.

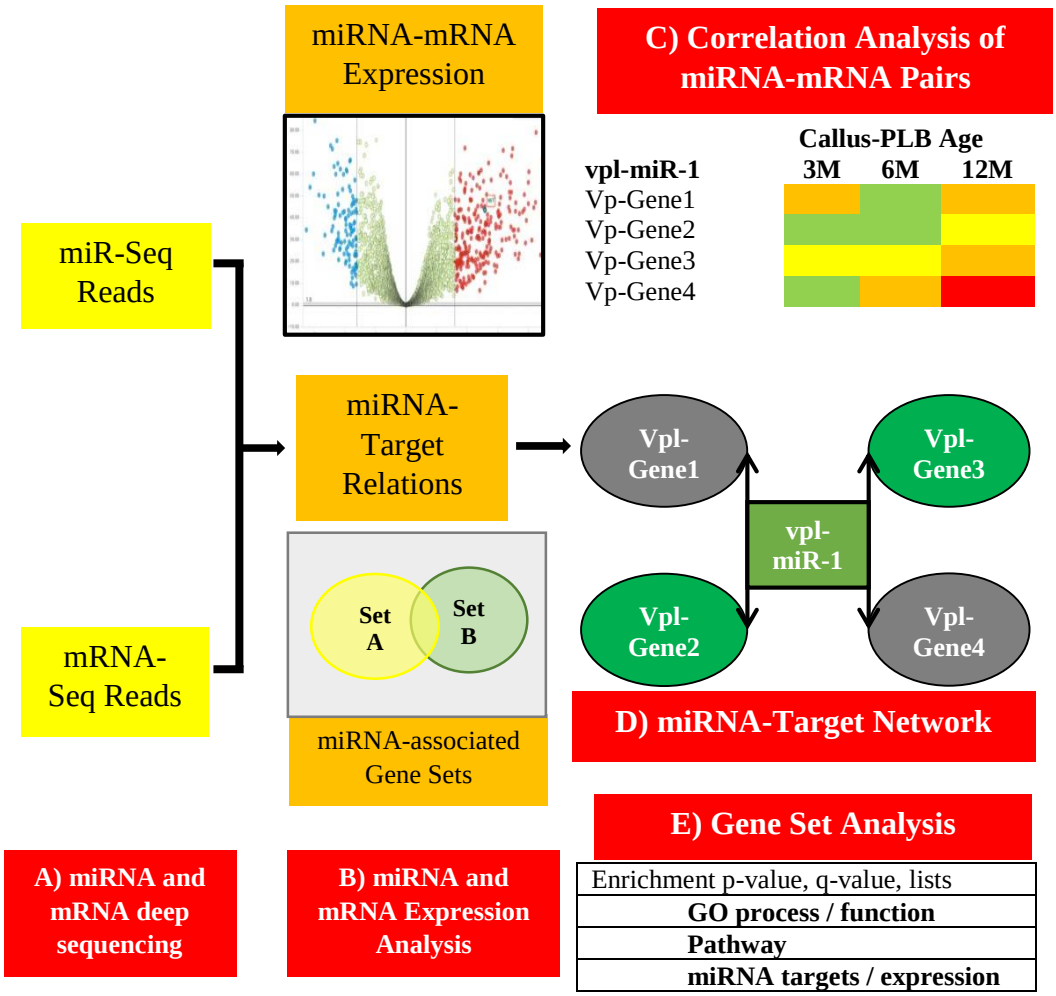


Figure 5.3. Integration pipeline for miRNA-mRNA target relations and expression correlations.

5.2.9. RT-qPCR validations of DEGs and DEMs in prolonged vanilla callus culture

5.2.9.1. RT-qPCR Primer Design

5.2.9.1.1. mRNA Primer Design

Primers were designed using the Primer3 software (Untergasser *et al.*, 2012; Koressaar and Remm, 2007), using the default parameters stated in **Table 5.1**, with the exception of contig vpl_032170 and contig vpl_024963, of which the GC content (%) were set at a wider range of 20-80% and 30-70% respectively. Long poly-X sequences were detected in both contigs, which might be the cause of irregular GC content. Conserved protein domains for each gene of interest were identified through gene annotation and were included in region for primer design. Housekeeping candidates were obtained from published literature (Rao *et al.*, 2014; Fock-Bastide *et al.*, 2014).

Table 5.1. Primer design parameters for qPCR.

Reaction Criteria			
Reaction Conditions	Min	Opt	Max
Divalent Salt (Mg++) (mM)			
Primer Criteria			
Primer Tm(C)	59	62	65
Primer GC (%)	35	50	65
Primer Size (nt)	17	22	30
Amplicon criteria			
Amplicon size (nt)	75	100	150
Custom Target Region		Yes	

Amplification specificities of primer candidates were then assessed using the Primer-BLAST (Ye *et al.*, 2012) software from NCBI, which uses BLAST and global alignment algorithm to screen primers against selected database. Apart from the original sequenced contigs, the primers were also blasted against the NCBI *nr* databases (Organism: *Viridiplantae*; taxid: 33090)

in order to avoid primer pairs (all combinations including forward-reverse primer pair, forward-forward and reverse-reverse pairs) that can cause non-specific amplifications. The selected primer sequences are listed in **Appendix H**.

The properties of designed primers were also checked using the IDT PrimerQuest (IDT, Sg). Only primers with no non-specific amplification in selected public databases, follows the mentioned primer parameters and with delta G higher than -8 was selected. Although there are slight differences in calculation algorithms between different primer design programs, the IDT PrimerQuest Tool (http://sg.idtdna.com/PrimerQuest/Home/Index?Display=AdvancedParams) were used as the main guideline for estimation of Tm and GC content, as the primers used in this experiment are synthesized by the same company.

5.2.9.1.2. miRNA Primer Design

Due to the chemistry of ligation-based reverse transcription method used (described later in **Section 5.2.9.2.2**), the reverse primers for miRNA RT-qPCR were a universal primer included in the miScript SYBR green PCR kit. The forward primers were designed with the miRPrimer 2.0 software (Busk, 2014). Sequences of housekeeping candidates were obtained from literature and stability screening of miRNA sequencing data using the GeNorm and NormFinder software (**Appendix J**). Primer sequences were listed in **Appendix H**.

5.2.9.2. cDNA library construction

The cDNA libraries for mRNA and miRNA were constructed using the QuantiNova™ Reverse Transcription Kit (Qiagen, Germany) and the miScript® Plant RT Kit (Qiagen, Germany) respectively, following the manufacturer's instruction. The total RNA amount of each sample was standardized to 2.0 µg (mRNA) and 250 ng (miRNA) prior to cDNA conversions. No-reverse-

transcription (NRT) reactions were included for each biological and each sample type with sample without RT enzymes.

5.2.9.2.1. mRNA

The library construction of mRNA consisted of two main steps, genomic DNA removal and reverse transcription. In a final volume of 15.0 µl, each RNA template was added with 2.0 µl of gDNA removal mix, followed by 2 min of incubation at 45ºC, for the removal of gDNA contamination. Then, the reverse transcription reaction was carried out by adding 1.0 µl of RT enzyme and 4.0 µl of RT mix to each template RNA, followed by incubation in a PCR thermocycler at 25ºC for 3 min for primer annealing, 45ºC for 10 min for reverse transcription and finally 85ºC for 5 min for inactivation of reaction. Each samples were then diluted 10X, aliquot and stored in -20ºC until use.

5.2.9.2.2. miRNA

The library construction of miRNA consisted of two main steps, ligation reaction and reverse-transcription reaction. Each RNA samples were firstly ligated to the miScript plant adaptors using miScript Plant RNA ligase mix (miScript ligation buffer, miScript ligation activator, miScript plant adaptor and miScript plant ligase) at 16ºC for an hour followed by enzyme deactivation at 65ºC for 20 min. Samples were allowed to equilibrate to room temperature for 2 min, then reverse-transcription reaction were carried out by adding 3.0 µl of ligation products to miScript plant RTase mix (miScript plant RT buffer, miScript plant RT nucleics, miScript plant RTase) at 37ºC for 2 hrs followed by enzyme deactivation at 95ºC for 5 min. All samples were diluted 10X, aliquot and stored in -20ºC until use.

5.2.9.3. Standard qPCR Conditions

5.2.9.3.1. mRNA

All qPCR samples were prepared using the master mix provided in the KAPA SYBR® FAST qPCR Kit (KAPA Biosystems, USA) according to the protocol handbook (**Table 5.2**). The KAPA SYBR® FAST qPCR Master Mix contained the KAPA SYBR® FAST DNA Polymerase, reaction buffer, dNTPs, SYBR® Green I dye, and $MgCl_2$ at a final concentration of 2.5 mM. For each qPCR sample, three technical replications were carried out. Each qPCR experiment is inclusive of an NTC (no template control, H_2O to replace cDNA sample) and an NRT (non-reverse transcribed sample, H_2O to replace RT enzyme).

Table 5.2. The components in an mRNA qPCR reaction.

PCR Component	Volume/reaction	Final concentration
DNase/RNase-free water	8.8 µl	-
2X KAPA SYBR FAST qPCR Master Mix Universal	10 µl	1X
10 µM forward primer	0.4 µl	200 nM
10 µM reverse primer	0.4 µl	200 nM
Template from RT reaction	0.4 µl	20 ng/reaction
Total volume:	20.0 µl	

The thermal cycling and fluorescence acquisition was done using the CFX Connect™ Real-Time PCR Detection System (Bio-Rad, USA) with the following conditions: (i) One cycle of initial denaturation at 95°C for 5 min (no acquisition, i.e. no fluorescence detection); (ii) 40 cycles of denaturation (95°C for 3 s, no acquisition), annealing (61.4°C for 20 s, single acquisition) and elongation (72°C for 12 s, single acquisition); (iii) one cycle of dissociation (melting) curve measurement — 95°C for 3s (no acquisition), 60°C for 5 s and 95°C at 0.5°C/s (continuous acquisition).

5.2.9.3.2. miRNA

All qPCR samples were prepared using the master mix provided in the miScript SYBR® Green PCR Kit (Qiagen, Germany) according to the protocol handbook (**Table 5.3**). The 2X QuantiTect® SYBR Green PCR Master Mix consists of HotStarTaq® DNA Polymerase, QuantiTect® SYBR Green PCR Buffer, dNTP mix (including dUTP), SYBR Green I dye, ROX™ passive reference dye and $MgCl_2$ at a final concentration of 5.0 mM. For each qPCR sample, three technical replications were carried out. Each qPCR experiment is inclusive of an NTC (H_2O to replace cDNA sample) and an NRT (non-reverse transcribed sample).

Table 5.3. The components in a miRNA qPCR reaction.

PCR Component	Volume / reaction	Final concentration
DNase/RNase-free water	7.75 µl	-
2X Quantitect SYBR Green PCR Master Mix	12.50 µl	1X
10X miScript Universal Primer	2.50 µl	1X
Template cDNA	1.00 µl	20 ng/reaction
10 µM Forward Primer	1.25 µl	0.5 µM
Total volume:	25.0 µl	

The thermal cycling and fluorescence acquisition was done using the CFX Connect™ Real-Time PCR Detection System (Bio-Rad, USA) with the following conditions: (i) One cycle of initial denaturation at 95°C for 15 min (no acquisition, i.e. no fluorescence detection); (ii) 40 cycles of denaturation (94°C for 15 s, no acquisition), annealing (43.9°C for 30 s, single acquisition) and elongation (70°C for 34 s, single acquisition); (iii) one cycle of dissociation (melting) curve measurement — 95°C for 3s (no acquisition), 60°C for 5 s and 95°C at 0.5°C/s (continuous acquisition).

5.2.9.4. Annealing Temperature Optimizations

5.2.9.4.1. mRNA

Optimization of annealing temperature were carried out with end-point PCR using the GoTaq™ DNA polymerase (Promega, USA) (1X PCR buffer, 2.5 mM $MgCl_2$, 1.6 mM dNTPs, 200 nM of each primer (forward and reverse), 1.25U of GoTaq™ DNA polymerase and 15 ng of template cDNA). For each primer sets, a gradient annealing temperature of 12 different temperatures ranging from 50-65˚C were tested using the Eppendorf Mastercycler (Eppendorf, USA). PCR conditions of annealing gradient PCR were carried out under the following PCR conditions: initial denaturation of 95ºC for 5 min, 40 cyclers of denaturation at 95ºC for 45s, annealing at 50-65ºC for 1 min and 72ºC for 30s, followed by final elongation at 72ºC for 1 min. PCR results were checked on 2.5% (w/v) agarose gel in 0.5X TBE stained in 3.0% SyBrSafe (Invitrogen, USA) and run at 90V for 1h.

5.2.9.4.2. miRNA

Optimization of annealing temperature were carried out with end-point PCR using the miScript SYBR® Green PCR Kit (Qiagen, Germany): 1X Quantitect SYBR Green PCR Master Mix, 1X miScript Universal Primer, 0.5 μM Forward Primer and 20 ng template cDNA. For each primer sets, a gradient annealing temperature of 8 different temperatures ranging from 50-60˚C were tested using the CFX Connect™ Real-Time PCR Detection System (Bio-Rad, USA) with the following conditions: (i) One cycle of initial denaturation at 95°C for 15 min (no acquisition, i.e. no fluorescence detection); (ii) 40 cycles of denaturation (94°C for 15 s, no acquisition), annealing (50-60°C for 30 s, single acquisition) and elongation (70°C for 34 s, single acquisition); (iii) one cycle of dissociation (melting) curve measurement — 95°C for 3s (no acquisition), 60°C for 5 s and 95°C at 0.5°C/s (continuous acquisition).

5.2.9.5. Validation of qPCR Amplification Efficiency

The qPCR amplification efficiency was validated using an equal pooled amount of each qPCR samples for each housekeeping and target gene primer (**Appendix H**). Seven dilution steps of two-fold (1:2) serial dilutions were done for each cDNA using DNase/RNase-free water. Each of the eight dilutions (the 0th to the 7th dilution) were amplified by qPCR in triplicates. A calibration curve was plotted for each target gene by the CFX Manager software (Bio-Rad, USA) and the qPCR amplification efficiency, *E*, was calculated from the slope of each concentration curve, as follows (Bustin *et al.*, 2009; Svec *et al.*, 2015):

$$E = (10^{-(1/\text{slope})} - 1) \times 100\%$$

5.2.9.6. Reference Stability Test

The stability of each reference gene candidate (**Table 5.4**) across different prolonged culture durations (3M, 6M and 12M) was evaluated and ranked using four algorithms, geNorm (Vandesompele *et al*., 2002; https://genorm.cmgg.be/), BestKeeper (Pfaffl *et al.*, 2004; https://www.gene-quantification.de/bestkeeper.html) and Normfinder (Andersen, 2004; https://moma.dk/normfinder-software), comparative ΔCT method (Silver *et al*., 2006) and RefFinder (Xie *et* al., 2012; http://leonxie.esy.es/RefFinder/?type=reference).

In geNorm, expression stability value (M) for each gene was calculated through pair-wise comparison of the genes; M value of less than 1.5 indicated stable expression; the lower the M value, the more stable the expression (Bustin *et al.*, 2010; De Spiegelaere *et al.*, 2015; Vandesompele *et al.*, 2002). In NormFinder, candidate reference genes were ranked according to the least of their estimated variations (Andersen, 2004). In comparative ΔCt method, relative transcription of pairs of genes were compared and ranked according to repeatability among all samples (Silver *et al*., 2006). BestKeeper algorithm determined the standard deviation and rated genes based upon variability (Pfaffl *et al.*, 2004). Finally, the RefFinder calculated the geometric mean based upon

rankings obtained from each of the four algorithm and provides the final comprehensive ranking (Xie *et* al., 2012).

Table 5.4. List of reference gene candidate for mRNA and miRNA RT-qPCR analysis.

mRNA	**miRNA**
Actin (*ACT)*	vpl_miR159n
Ubiquitin (*UBQ)*	vpl_miR168a-5p
Elongation Factor-1 alpha (*EF1*)	vpl_miR319g-3p
B-tubulin (*TUB*)	vpl_miR5179
18S-rRNA (*18S*)	vpl_SnoR6

5.2.9.7. Data Analysis for Relative Quantification of Gene Expression

The amount of mRNA or miRNA transcripts, which reflected the expression level, was determined relative to those of reference genes or reference miRNAs to normalise variations in individual reactions that were caused by initial quantities of cDNA templates. Therefore, the Cq values of the eight target mRNA genes (*vpl_contig_001034*, *vpl_contig_032170*, *vpl_contig_086799*, *vpl_contig_055903*, *vpl_contig_009944*, *vpl_contig_024863*, *vpl_contig_001827* and *vpl_contig_014640*) and three target miRNAs (*vpl_miR166r*, *vpl_miR395g* and *vpl_miR8155*) in three, six and twelve-month-old vanilla callus samples (3M, 6M and 12M) were normalised with the geometric Cq mean of selected reference genes (*18S, TUB* and *EF1* for the normalization of mRNA; *vpl_miR159n*, *vpl_miR168e-5p*, *vpl_miR319d-3p*) reference genes (Vandesompele *et al.*, 2002). The geometric mean was calculated as follows:

$$\left(\prod_{i=1}^{n} x\right)^{\frac{1}{n}} = \sqrt[n]{x_1 x_2 \dots x_n}$$

where a capital pi notation is used to show a series of multiplications, and x_n is the Cq value of the *n*th reference gene.

In the mRNA expression study, three reference genes were used (therefore, *n*=3). As such, the geometric mean of Cq values would be the cube root of the product of three Cq values:

$$\sqrt[3]{(Cq_{18s})(Cq_{TUB})(Cq_{EF1})}$$

In the miRNA expression study, two reference genes were used (therefore, *n*=2). As such, the geometric mean of Cq values would be the square root of the product of two Cq values:

$$\sqrt[2]{(Cq_{miR159})(Cq_{miR168})}$$

The relative expression of each target gene/miRNA in each sample type was calculated using the $2^{-\Delta\Delta Cq}$ method (Livak and Schmittgen, 2001; Pfaffl, 2001; Yuan *et al.*, 2006). For each target gene, the delta Cq of each sample of interest and control sample ($\Delta Cq_{\text{sample or control}}$) was first calculated by deducting the geometric Cq mean of the reference genes from the Cq mean of biological sets for the target gene ($\Delta Cq_{\text{sample or control}} = Cq_{\text{target}} - Cq_{\text{reference}}$). Then, the delta delta Cq of each target gene in the sample of interest ($\Delta\Delta Cq$) was calculated by deducting the $\Delta Cq_{\text{control}}$ from the $\Delta Cq_{\text{sample}}$ ($\Delta\Delta Cq = \Delta Cq_{\text{sample}} - \Delta Cq_{\text{control}}$).

For the comparison of relative expression, the normalized amount of target gene (a relative unit) was standardized to 6-month-old calli (denoted as *6M*) as an internal control (calibrator) and expressed as percent *3M* or percent *12M*. Finally, the relative expression, R, of a target gene/miRNA in a sample of interest was calculated as $R = 2^{-\Delta\Delta Cq}$. With that, the relative expression level of *6M* was 1.00, and the expression levels of other target genes/miRNAs in other samples (denoted as *target gene*[sample]) were compared against *Target[6M]*.

5.3. Results

To obtain a comprehensive transcriptome profile of vanilla callus culture in response to prolonged culture duration stress, a total of six mRNA libraries and the corresponding small RNA libraries of vanilla callus were constructed from three different time points (3 months, 6 months and 12 months of culture) for deep sequencing. The sequencing data was later validated using RT-qPCR.

5.3.1. Total RNA quality and quantity

The overall RNA quantity and quality measured with Nanodrop and Bioanalyzer was summarized in **Table 5.5**. Sample RIN values ranged from 7.60 to 8.80, while $OD_{260/280}$ purity ranged from 1.91 and 2.07. RNA concentrations, as estimated by Bioanalyzer, ranged from 365 to 1510 ng/µl. Electropherograms of RNA samples were shown in **Figure 5.4**. The arrow shows the 18S and 25S RNA peaks, which were used as the traditional evaluation of RNA integrity in the algorithm of RNA integrity number (RIN) value assignment.

Table 5.5. Nanodrop and Qubit readings.

Sample Name	Replicate Group	Concentration (ng/ul)		RIN value	$OD_{260/280}$	$OD_{260/230}$
		Nanodrop	Bioanalyzer			
3M-1	3M	618.15	1140	8.50	2.06	2.25
3M-2		774.00	1495	8.50	2.03	2.27
3M-3		761.60	1510	8.30	2.02	2.28
6M-1	6M	304.35	485	8.80	1.96	2.08
6M-2		392.85	585	8.30	1.91	1.73
6M-3		338.15	595	8.60	2.05	1.89
12M-1	12M	277.00	490	7.60	2.01	1.96
12M-2		241.25	365	7.80	1.92	1.91
12M-3		348.15	575	8.20	2.07	2.16

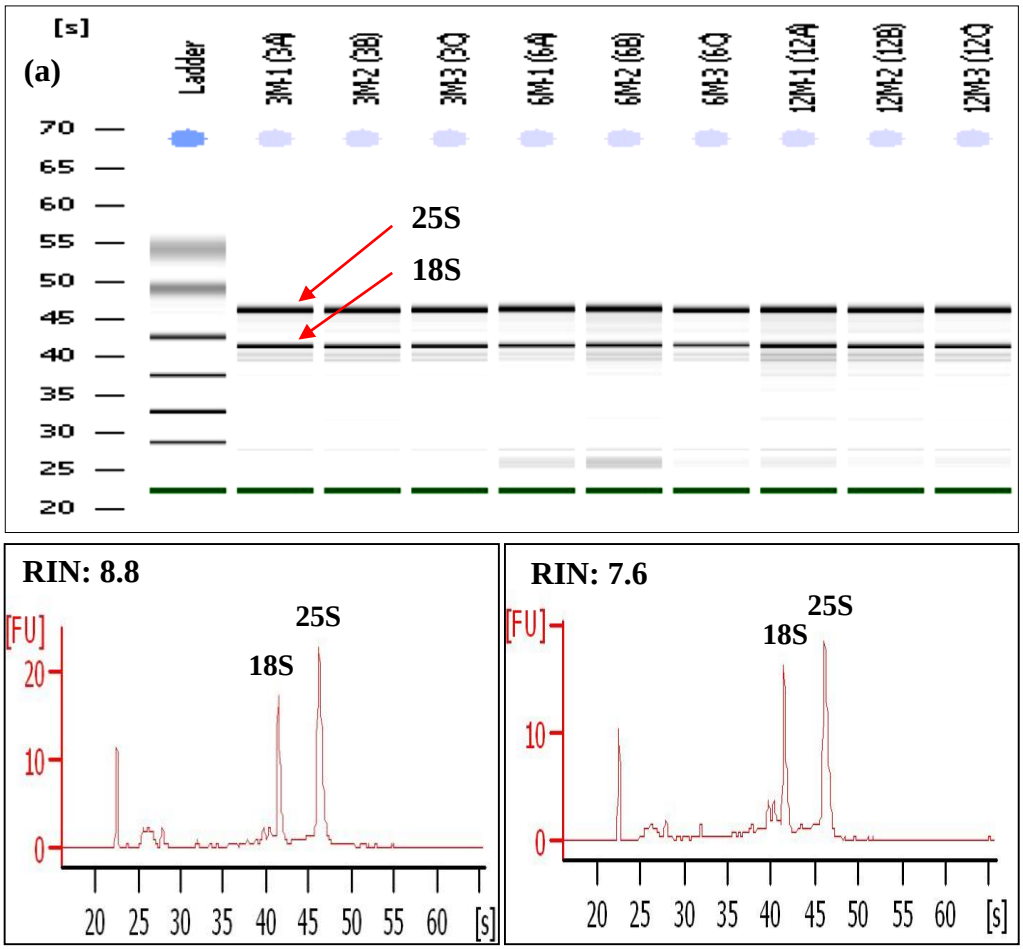


Figure 5.4. The RNA profiles from Agilent 2100 Bioanalyer analysis. The intact bands of 18S and 25S ribosomal RNAs (rRNAs) were shown in the virtual gel with arrows indications, and their distinct peaks were shown in electropherograms for the highest RIN (8.8) and the lowest RIN (7.6).

5.3.2. mRNA Sequence Profiles

5.3.2.1. Sequence processing and *de novo* assembly

More than 53 million raw reads were yielded for each RNA sample, with an average read length from 134.3 to 139.1 bp, which were then trimmed (for adaptor, ambiguity and low quality 3' terminals) and filtered (for quality at Q30 and short reads) (**Table 5.6**). After the removal of redundant sequences, more than 41 million reads per sample ranging from an average of 78.41 to 80.56 bp in length was retained (**Table 5.6**).

Table 5.6. Statistics of reads generated by mRNA sequencing of *V. planifolia* callus before and after pre-processing. Raw sequencing reads were trimmed for adaptor (NEBNext RNA directional kit adaptors), ambiguity, quality (Q30), short reads and low quality 3' terminals.

Sample Name	Paired RAW reads		Filtered RAW reads		
	Number of reads	Average length (bp)	Number of reads	Percentage trimmed	Average length (bp)
3M-1	60,498,688	138.2	48,739,066	80.56	81.2
3M-2	52,699,086	137.9	41,695,328	79.12	80.2
6M-1	56,772,196	134.3	44,517,474	78.41	77.4
6M-2	53,418,036	136.1	42,291,182	79.17	80.0
12M-1	53,415,184	138.2	42,206,260	79.02	80.9
12M-2	60,145,632	139.1	47,833,190	79.53	81.1

In *de novo* assembly, a total of 267,282,500 reads were assembled into 145,500 contigs with a GC content of 38.60% (**Table 5.7**). The assembled master contigs were used as reference for mapping in differential gene expression (DGE) analysis.

Table 5.7. Statistics for *de novo* transcriptome assembly.

Number of reads used in assembly	267,282,500
Number of contigs	145,550
Total size of contigs (Mbp)	82.01
Average contig length (bp)	563
Min contig length (bp)	33
Max contig length (bp)	14,520
N25 (bp)	1,402
N50 (bp)	679
N75 (bp)	374
Number of contigs > 1kb	16,562
Number of contigs > 5kb	202
Number of contigs > 10kb	4
GC content (%)	38.60%

5.3.2.2. Functional annotation and classifications of differentially expressed genes (DEGs)

Out of the 145,550 assembled master contigs, a total of 22,025 contigs which showed significant (FDR-p-value < 0.05) expression across the time points (3M, 6M and 12M), were subjected to functional annotation for a global transcriptome profiling. Time course expression analyses were conducted to study the expression patterns of these DEGs over prolonged callus culture duration.

5.3.2.2.1. BLAST, GO-mapping and Annotation Project Statistics

The primary protein database, *nr*, had the highest BLAST hit (47.90%) with the vanilla transcriptome, followed by the curated protein databases, Swiss-Prot (33.33%) and pdbaa (17.91%) (**Table 5.8**). InterproScan (Apweiler *et al.*, 2000, 2001; Fin *et al.*, 2017; Hunter *et al.*, 2011; Quevillon *et al.*, 2005) were carried out to functionally identify the new protein sequences, 40.34% of whole contigs showed IPS hits (**Table 5.8**). In primary protein database *nr*, vanilla has the most hits with protein database of the plant species, *Elaeis guineensis* (2278, 21.59%), *Phoenix dactylifera* (1960, 18.58%) and *Musa acuminata* (1116, 10.58%) (**Figure 5.5**). From the BLAST with Swiss-Prot, BLAST hits were limited to model species, such as *Arabidopsis thaliana, Oryza sativa* and *Zea mays*. BLAST search against 3D protein structure database, pdbaa, showed cross-kingdom top BLAST hits distribution to *Arabidopsis thaliana* (900, 22.81%), *Homo sapiens* (743, 18.84%) and *Saccharomyces cerevisiae* (112, 2.84%). The *nr* database was selected as the database used for functional protein screening and annotation as it gave the highest amount of annotatable genes as compared with other databases.

Table 5.8. BLAST search, GO mapping and annotation project statistics of vanilla transcriptome (a total of 22,025 sequences with DGEs) against the NCBI protein databases.

Database	Number (Percentage) of sequences		
	With BLAST hit	**GO Mapped**	**GO-Slim Annotated**
nr	10549 (47.90%)	8727 (39.62%)	6409 (29.10%)
Swiss-Prot	7340 (33.33%)	6946 (31.54%)	6145 (27.90%)
pdbaa	3944 (17.91%)	3366 (15.28%)	2366 (10.74%)
Interpro Scan	8,888 (40.34%)		

The definition of automatic annotation was assigned to assembled contigs based on the E-value (≤ 1E-03) and sequence identity from the BlastX result. A total of 22,025 assembled contigs with DGEs over prolonged callus culture duration were subjected to the BLASTX to protein databases, *nr* (organism limited to the taxa *Viridiplantae*, taxid: 33090), Ref-Prot, Swiss-Prot, pdbaa and interpro scan.

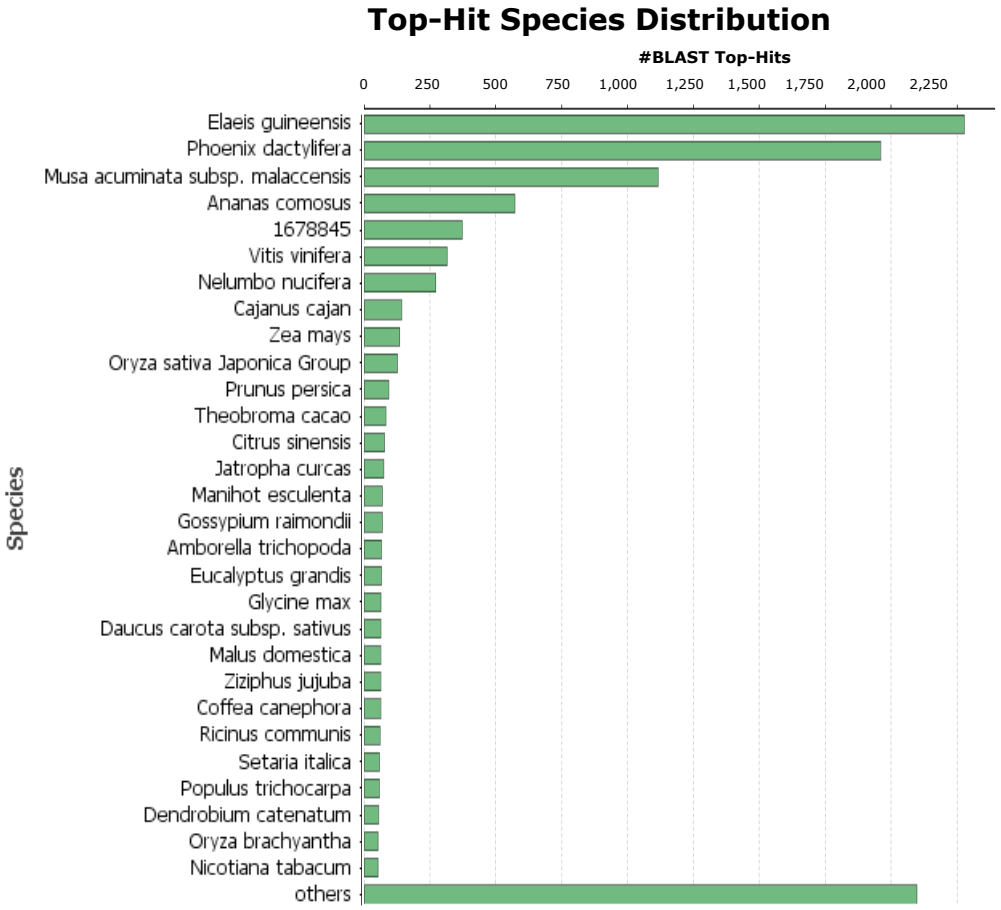


Figure 5.5. Top-hit species distribution of BLAST results for vanilla to *nr* (*Viridiplantae*) databases.

5.3.2.2.2 Global transcriptome dynamics

Gene ontology (GO) analysis of the global transcriptome profile of genes expressed in vanilla callus over prolonged callus cultures (3M, 6M and 12M) were classified into molecular function (MF), cellular component (CC) and biological process (BP) in **Figure 5.6**. In terms of cellular function of the DEG genes, most of gene expression occurred at the membrane, followed by nucleus, plastid and cytoplasm (**Figure 5.6b**). GO analysis of the statistics of genes has also given an overview of the molecular functions of these genes, including hydrolase activity, nucleotide binding, catalytic activity, binding, transferase activity, kinase activity, DNA binding, protein binding, transcription factor and RNA binding (**Figure 5.6a**). In the GO category of biological process (BP), it was observed that a high number of genes fell into the signal transduction and response to stress, in particularly, response to abiotic and endogenous stimulus (**Figure 5.6c**).

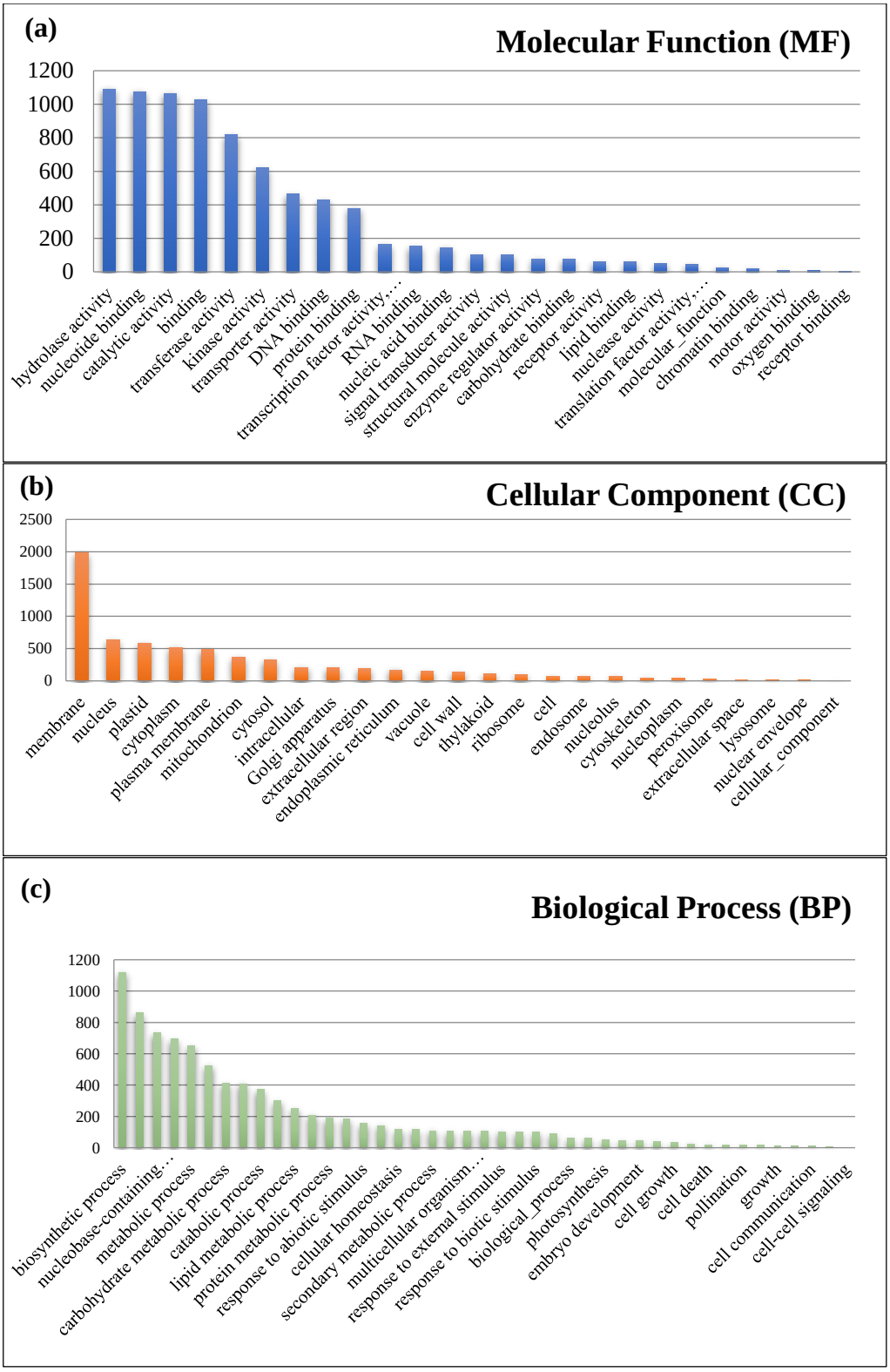


Figure 5.6. Gene ontology classification of assembled transcripts. GO analysis was performed for the three main categories (a: cellular component, b: molecular function, and c: biological process). The *X* axis represented different categories of all annotated genes differentially expressed in 3M, 6M and 12M vanilla callus. The right *Y* axis indicated the numbers of all annotated genes differentially expressed in 3M, 6M and 12M vanilla callus in a specific category.

5.3.2.2.3. DGE and time course expression analysis

A total of 3,789 contigs had significant (FDR-corrected *p* value ≤ 0.05) differences and a fold change of at least 2.0 in gene expressions between 3, 6 and 12 months old of callus cultures (**Table 5.9**).

Table 5.9. DGE summary of vanilla callus over 3, 6 and 12M prolonged culture durations.

DGE	Up-regulated		Down-regulated	
	Genes	**FC range**	**Genes**	**FC range**
3M vs 6M	1374	2.60 - 151.55	2415	2.63 – 12896
6M vs 12M	2359	2.69 - 22330.9	1430	2.51 – 11520
3M vs 12M	263	2.00 - 32.46	412	2.00 - 4526.6

DGE was analyzed at an FDR-corrected *p*-value ≤ 0.05 and fold change differences of at least 2.0.

From these genes, a total of seven clusters of gene expression patterns over prolonged callus culture duration were detected in the differentially expressed genes (DEGs) (**Table 5.10**). Amongst all expression patterns, cluster 3 had the highest number of DEGs, followed by Cluster 2.

Table 5.10. Time course expression analysis results.

Cluster	3M vs 6M	6M vs 12M	3M vs 12M	#contigs
1	Down-regulation	Down-regulation	Down-regulation	56
2	Up-regulation	Down-regulation	ND	1040
3	Down-regulation	Up-regulation	ND	2075
4	Up-regulation	Down-regulation	Up-regulation	43
5	Up-regulation	Down-regulation	Down-regulation	291
6	Down-regulation	Up-regulation	Up-regulation	217
7	Down-regulation	Up-regulation	Down-regulation	65
		Total:		3787

ND: No differences

5.3.2.2.4. Gene annotations and classifications

Among these genes, we identified a cascade of stress responsive genes and genes related to SV from literature (More detailed review in **Section 1.3.2.3**), namely the DNA methylation, activations of retrotransposable elements, histone modifications, DNA transcription factors and microRNAs gene regulations, which we then categorized into five gene families.

DNA methylation (MET) family

A number of genes encoding for key enzymes involved in the cysteine and methionine metabolism were expressed with significance (p-value ≤ 0.05) in prolonged culture of vanilla callus, as observed from the KEGG pathway (Kanehisa *et al.*, 2016, 2017) analysis (**Figure 5.7**), an important pathway for DNA methylation and gene regulation in plants (Kanehisa and Goto, 2000). The sequence conservation and expression of important gene in this family were further described with its potential interaction with microRNAs in the later part of this chapter (**Section 5.3.4.2.1**).

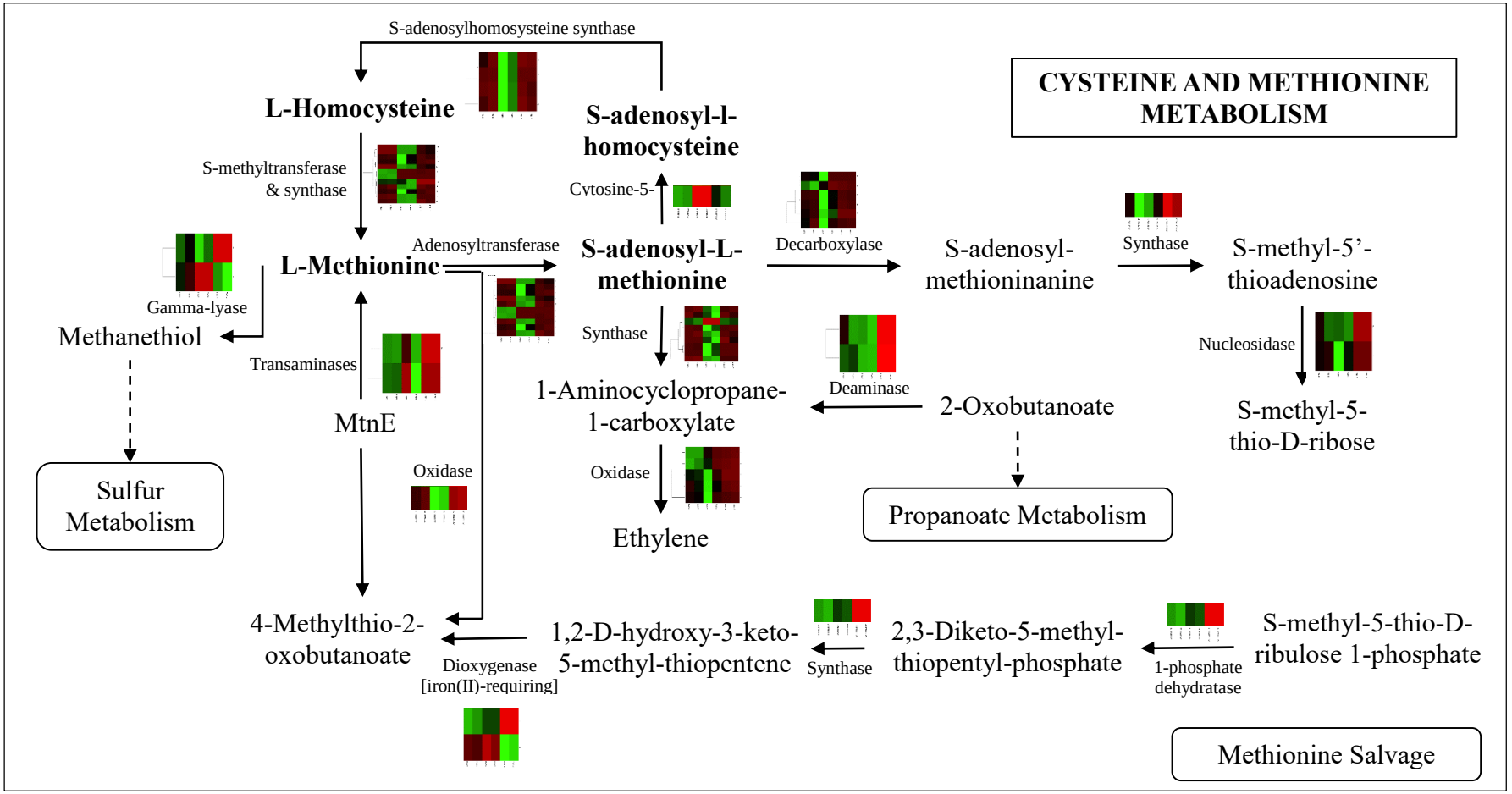


Figure 5.7. KEGG pathway for cysteine and methionine metabolism expressed in vanilla callus culture. This is an important pathway involved in DNA methylation (Kanehisa, 2000; Kanehisa and Goto, 2000). In prolonged callus culture of vanilla, enzymes involved in the

pathway process listed in this diagram were significantly expressed (Kanehisa, 2000; Kanehisa and Goto, 2000). The expression patterns of each gene were expressed in heat map next to each enzyme. Cysteine and methionine are sulfur-containing amino acid; in plants, cysteine is converted from serine (via acetylserine) by transfer of hydrogen sulphide and metabolized to pyruvate in multiple routes (Kanehisa, 2000; Kanehisa and Goto, 2000). S-Adenosylmethionine (SAM), synthesized from methionine and ATP, is a methyl group donor in many important transfer reactions including DNA methylation for regulation of gene expression, as well as used to regenerate methionine in the methionine salvage pathway (Kanehisa, 2000; Kanehisa and Goto, 2000).

Histone modification (HM) family

A total of 12 histone families were identified (**Table 5.11**). Despite having variants, the histone proteins were identifiable by their conserved protein domains. The histone lysine N-methyltransferase (HMT) family had the most number of members of variants and subunit forms, with a total of 11 contigs putatively encoding for HMT proteins, namely ASH2, ASHR1, ASHR2, ATXR7, ASHH2 and SUVR2 (**Figure 5.8**). The histone deacetylase (HDAC) family consisted of subunits SAP18, HDAC2, 6 and 19 (**Figure 5.9**), while histone acetyltransferase (HAT) family consisted of Elp3 and GCN5 (**Figure 5.10**). Besides, putative H1, H2A (5 variants), H2B, H3 (6 variants), H3K4 and H4 (6 variants) proteins were also identified.

Table 5.11. Histone protein families expressed in vanilla callus cultures.

Histone family	Subunits / Variants
HMT	ASH2, ASHH2, ASHR1, ASHR2, ATXR7 and SUVR2
HDAC	SAP18, HDAC2, 6 and 19
Histone H4	putative protein homologs
Histone H3	H3, H3.2, H3.3
Histone H2A	H2A, H2A variant 3, 6
HAT	Elp3, GCN5
LSD	putative protein homologs
Histone H1	-

Histone H2B	-
Histone H3K4	COMPASS-like H3K4 histone methylase component WDR5A
Histone chaperone	ASF1-like
Histone-binding	MSI1-like

HMT: LSD: Lysine-specific histone demethylase

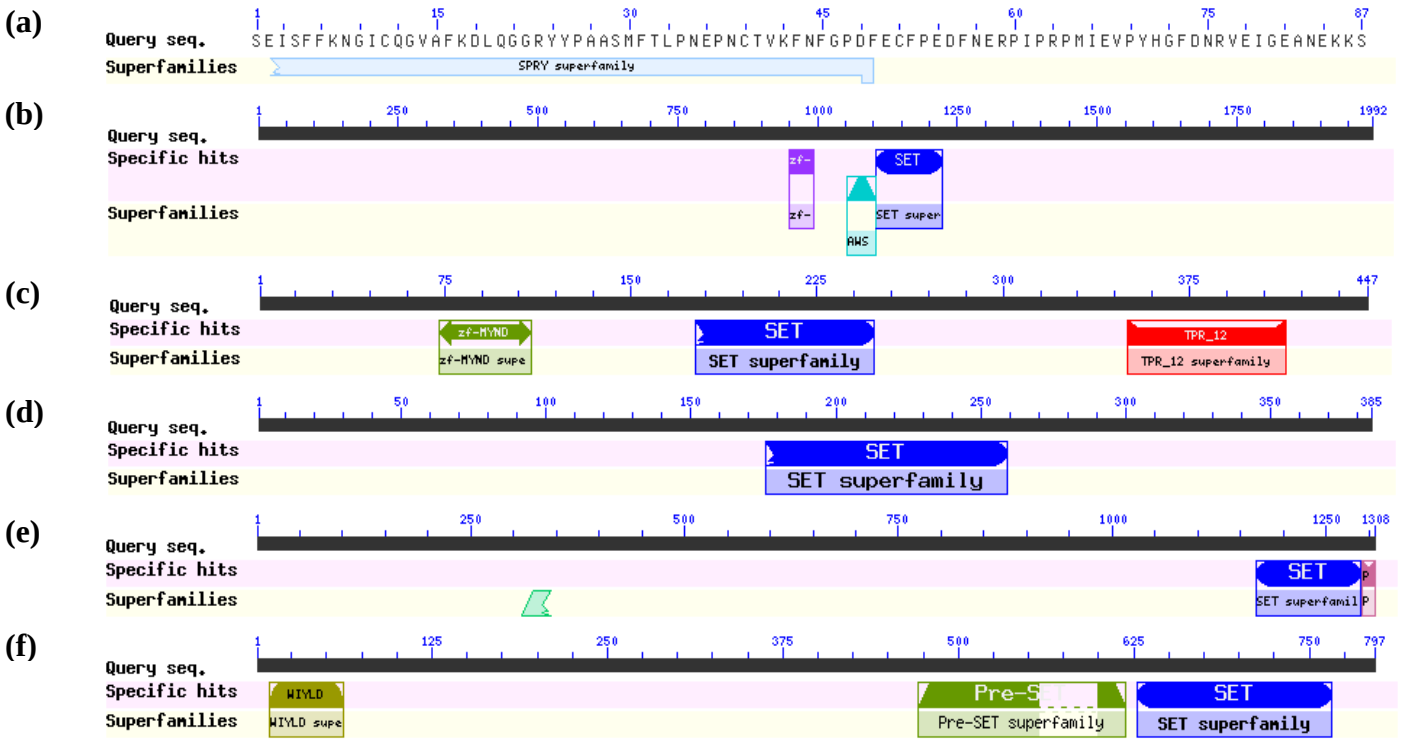


Figure 5.8. Conserved protein domain architectures for putative histone lysine N-methyltransferase (HMT) family members expressed in vanilla callus cultures: (a) ASH2 (SPRY superfamily protein; Accession: cl02614), (b) ASHH2 (AWS and SET domain-containing protein; ID: 10539360), (c) ASHR1 (zf-MYND and SET domain-containing protein; ID: 12026512), (d) ASHR2 (SET domain-containing protein; ID: 10645844), (e) ATXR7 (SET and PostSET-domain-containing protein, ID: 10209303) and (f) SUVR2 (Pre-SET and SET domain-containing protein; ID: 10564510). ID here referred to the NCBI protein architecture ID.

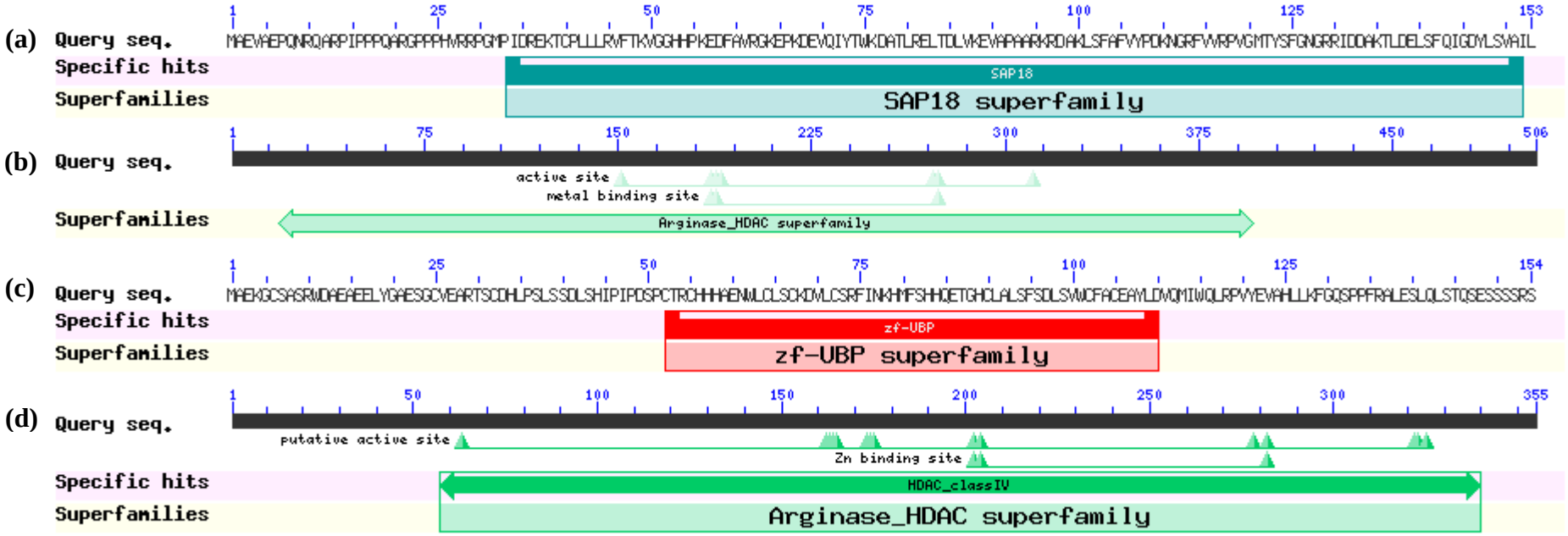


Figure 5.9. Conserved protein domain architectures for putative histone deacetylase (HDAC) family members expressed in vanilla callus cultures: (a) SAP18 (conserved SAP18 superfamily, pfam06487), (b) HDAC2 (conserved arginase_HDAC superfamily, cl17011), (c) HDAC6 (conserved sf-UBP superfamily, pfam02148) and (d) HDAC19 (conserved class IV HDAC protein domain, cd09993) proteins.

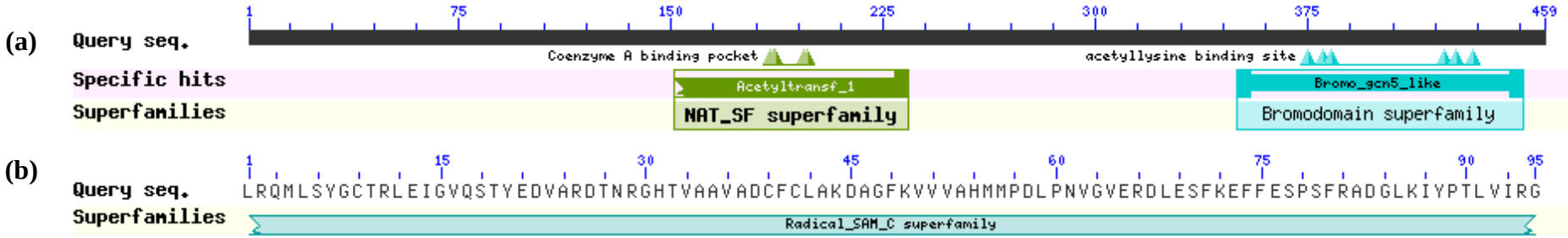


Figure 5.10. Conserved protein domain architectures for putative histone acetyltransferase (HAT) family members expressed in vanilla callus cultures: (a) GCN5 (conserved acetyltransferase (GNAT) family, pfam00583), and (b) Elp3 (conserved radical_SAM_C superfamily, cl25699) proteins.

In general, the histone proteins based on prediction of amino acid sequence expressed in vanilla callus cultures were mostly conserved at the predicted protein level, while some had different isoforms (indels or substitutions) and subunits (**Table 5.11**). Both transcript and protein sequence similarities were analyzed for each of the contig sequences encoding putative histone proteins. In general, the histone modification family proteins contained highly conserved protein domains, with protein isoforms and transcript variants. Multiple transcripts variants encoding for same histone protein were observed - some contigs had putative protein homologs which were 99% identical, such as histone deacetylase complex subunit SAP18 (XP_020685380.1) and lysine-specific histone demethylase 1 homolog 1 (XP_020675079.1), while some contigs had putative sequence homologs with base substitutions, such as the histone acetyltransferase of the MYST family (XP_020581186.1). In this study, these contigs were annotated with same putative gene function but different homologs, such as histone deacetylase complex subunit SAP18 putative protein homolog 1 (vpl_contig_5322) and histone deacetylase complex subunit SAP18 putative protein homolog 2 (vpl_contig_44861), which are both aligned to protein XP_020685380.1. Putative members of the protein family of histone H1-like protein (XP_020584714.1), histone H2A (4 transcript variants, **Figure 5.11a** and 2 protein variants, **Figure 5.11b**), histone H2B (EMS55288.1)**,** COMPASS-like H3K4 histone methylase component WDR5A protein (XP_020701228.1, **Figure 5.12**), histone H3 and histone H4 (4 protein variants with histone H4 conserved protein domain family PLN00035, **Figure 5.13**) were expressed in vanilla callus.

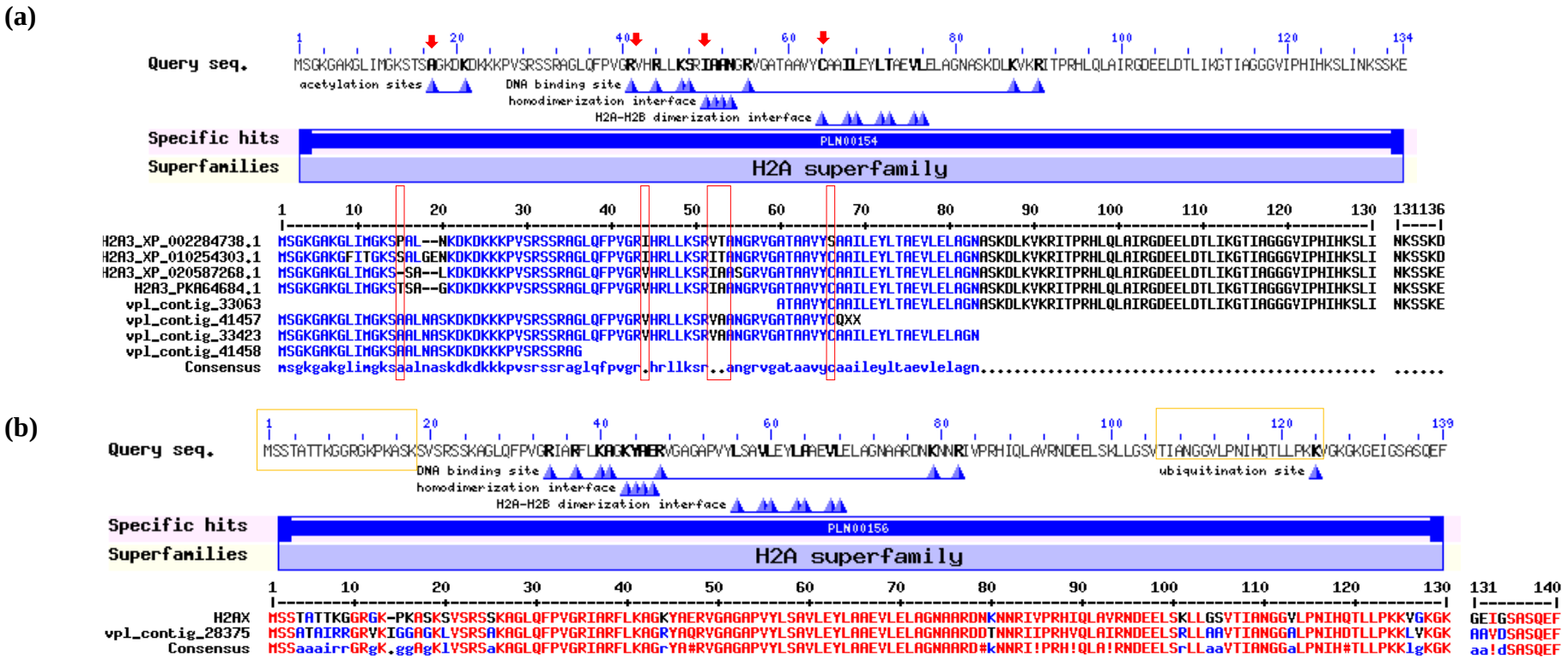


Figure 5.11. Conserved protein domains and multiple sequence alignment (MSA) of (a) vanilla putative H2A variant 3 (H2A3) proteins: The H2A3 proteins were conserved between plant species, except at the flanking regions of different conserved active sites (indicated by **red** arrows correlating to **red** boxes in sequential order). The vanilla H2A3 proteins were expressed in partial by a few transcript variants (vpl_contigs_33063, 41457, 22423 and 41458); **and (b) vanilla putative H2AX proteins**: As compared to the H2A3 protein, the H2AX protein lacked the acetylation site and had a ubiquitination site (indicated by **orange** boxes).

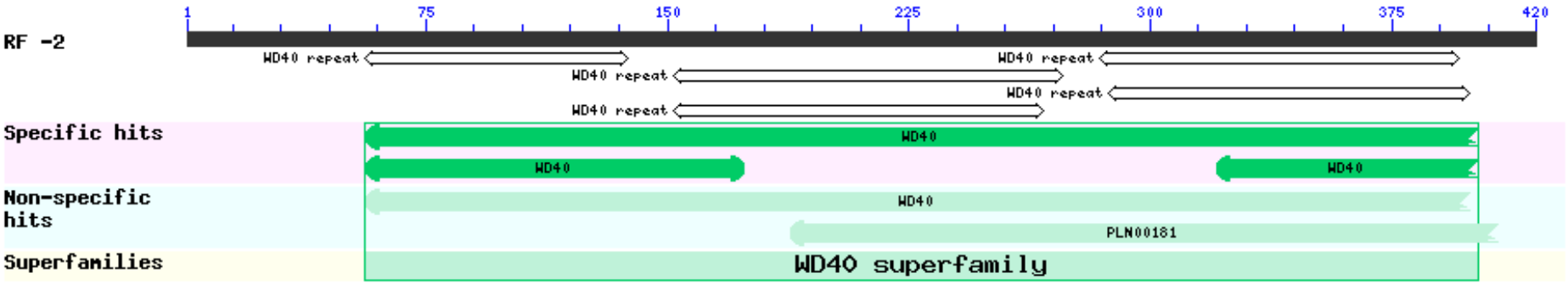


Figure 5.12. Conserved protein family domain of vpl_contig_60898 encoding for COMPASS-like H3K4 histone methylase component WDR5A protein (XP_020701228.1). In this sequence, the protein superfamily WD40 (COG2319), which consisted of G-beta repeats, were found.

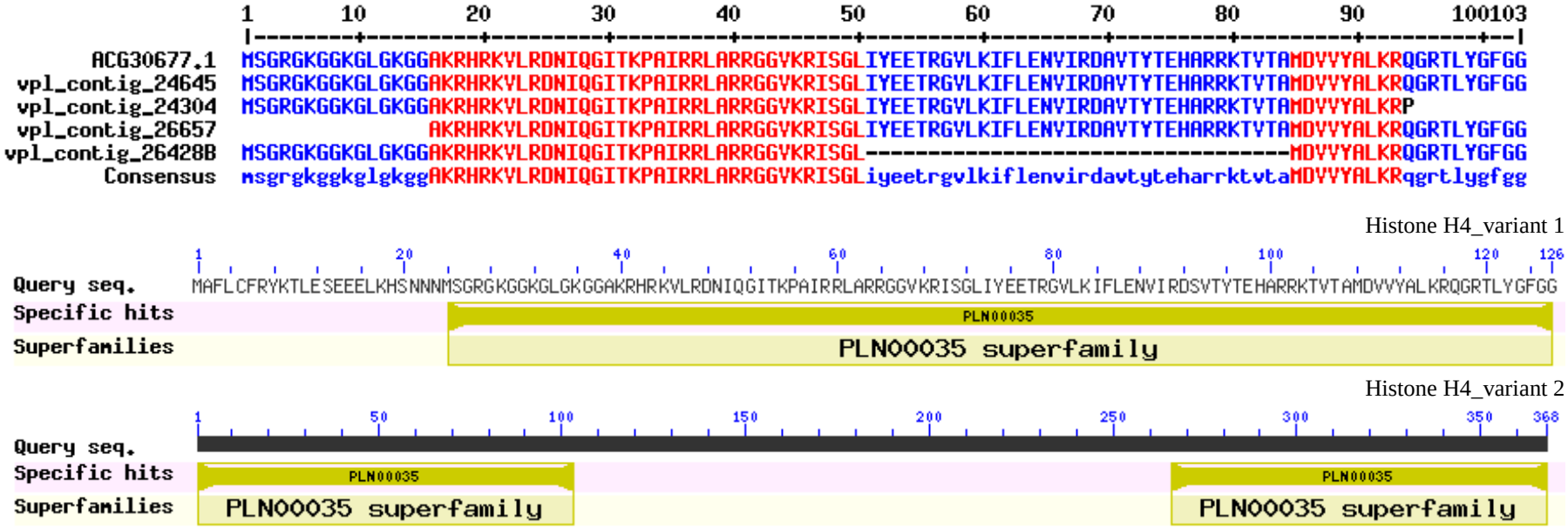


Figure 5.13. Conserved protein domain (PLN00035 superfamily, provisional histone H4) and multiple sequence alignment (MSA) of vanilla sequences encoding conserved histone H4. The conserved domain flanked at the interval of 24 to 126 aa of vpl_contig_24645 at an *e*-value of 7.50E-57. Vanilla contig sequences were aligned to protein sequences of **ACG30677.1** (histone H4 protein of *Zea mays*). Consensus amino acids found in all vpl contigs were highlighted **red**. Conserved amino acids were found in most of all sequences but were deleted in certain sequences (especially vpl_contig_26428) were highlighted **blue**.

Transposable element (TE) family

The vanilla contig sequences were queried to BLAST against the MIPS Repeat Element Database (mips-REdat) to screen for potential transposable element sequences. Out of 22032 vanilla contig sequences, 3200 (14.52%) showed BLAST match with the transposable element sequences in the MIPS Repeat Element Database. Majority of the BLAST hits were long terminal repeat (LTR) retroelement sequences, where the most dominant retro-element families were the *copia* LTR family (20.25%), followed by *gypsy* LTR (17.31%) and long interspersed nuclear elements (LINEs) (4.41%) (**Figure 5.14**). This sequence conservation and expressions of important genes in this gene family were further described with its interaction with microRNAs in the later part of the chapter (**Section 5.3.4.2.3 - 5.3.4.2.4**).

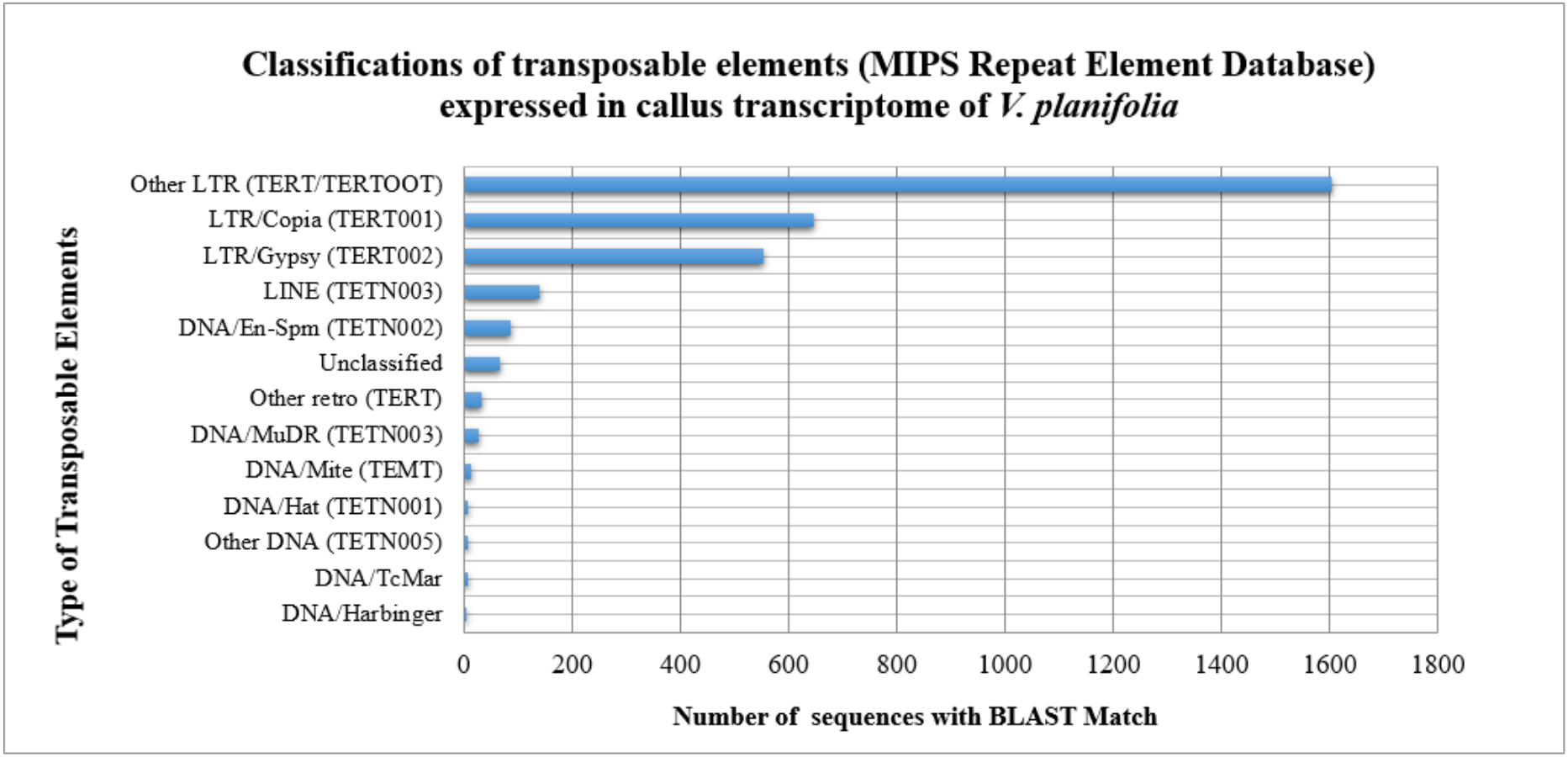


Figure 5.14. Bar chart showing classifications of vanilla callus transcriptome sequences with BLAST match with the transposable element sequences in MIPS Repeat Element Database.

DNA transcription factors (TFs) family

In vanilla calli, different transcription factors were expressed with significant changes over time (**Table 5.12**), which included the WRKY transcription factors, ethylene-responsive transcription factors (ERF), basic helix-loop-helix (bHLH), myelobastosis (Myb) family of transcripton factors, heat stress transcription factors, AP2-like ERFs, GATA, alfafin-like factor (ALF), core binding factor (CBF), DIVARICATA EMB1444, HY5, MYC4, NAC29, PAT1 and RAX3 transcription factor.

Table 5.12. List of transcription factor gene families and members expressed in vanilla callus cultures.

Transcription factor family	**Member**	**#contig**
WRKY transcription factor	WRKY 11, 17, 22, 33, 40, 48, 51, 53, 65, 70-72	17
ethylene-responsive transcription factor	ABR1, 2, 4, 14, 73, 96, 109, 1B	14
Transcription factor bHLH	bHLH18, 35, 87, 92, 130, 137	9
Myb family transcription factor	MYB1R1, 86, 108, APL	6
Heat stress transcription factor	HSF_A2c, A6b, B1, B2b	5
AP2-like ethylene-responsive transcription factor	X2, AIL1, AIL5-7, B3-RAV1	8
GATA transcription factor	GATA5, 16	3
Other	ALF, CBF, DIVARICATA, EMB1444, HY5, MYC4, NAC29, PAT1, RAX3	13

5.3.2.3. DGE analysis of vanilla calli over time

5.3.2.3.1. Expression of methylation and histone modification-related genes

Histone modifications during tissue culture could lead to disruption in the methylation pathway in a non-specific manner and hypo/hypermethylation patterns of DNA induced in tissue culture (Krishna *et al.*, 2016). In our study, among the differentially expressed histone protein-encoding genes, it was observed that most of the methylation genes, including S-Adenosylmethionine (SAM), S-adenosyl-L-homocysteine hydrolase (SAH), 5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (MET), 5 -methylthioadenosine S-adenosylhomocysteine nucleosidase 2-like (MTA-SAHN) and DNA-(cytosine-5)-methyltransferase (DNMT) genes, were up-regulated in 6M calli; while methylation genes such as S-acyltransferase (S-AT) genes were expressed in the opposite manner (**Figure 5.15**). As for histone family, the Myst-like histone acetyltransferase 1, histone chaperone ASF1-like and histone-lysine N-methyltransferase ASHH2-like isoform X1 were down-regulated in 6M; while histone H2A6, histone-lysine N-methyltransfearse SETD1B and histone superfamily protein-encoding genes were up-regulated in 6M (**Figure 5.16**).

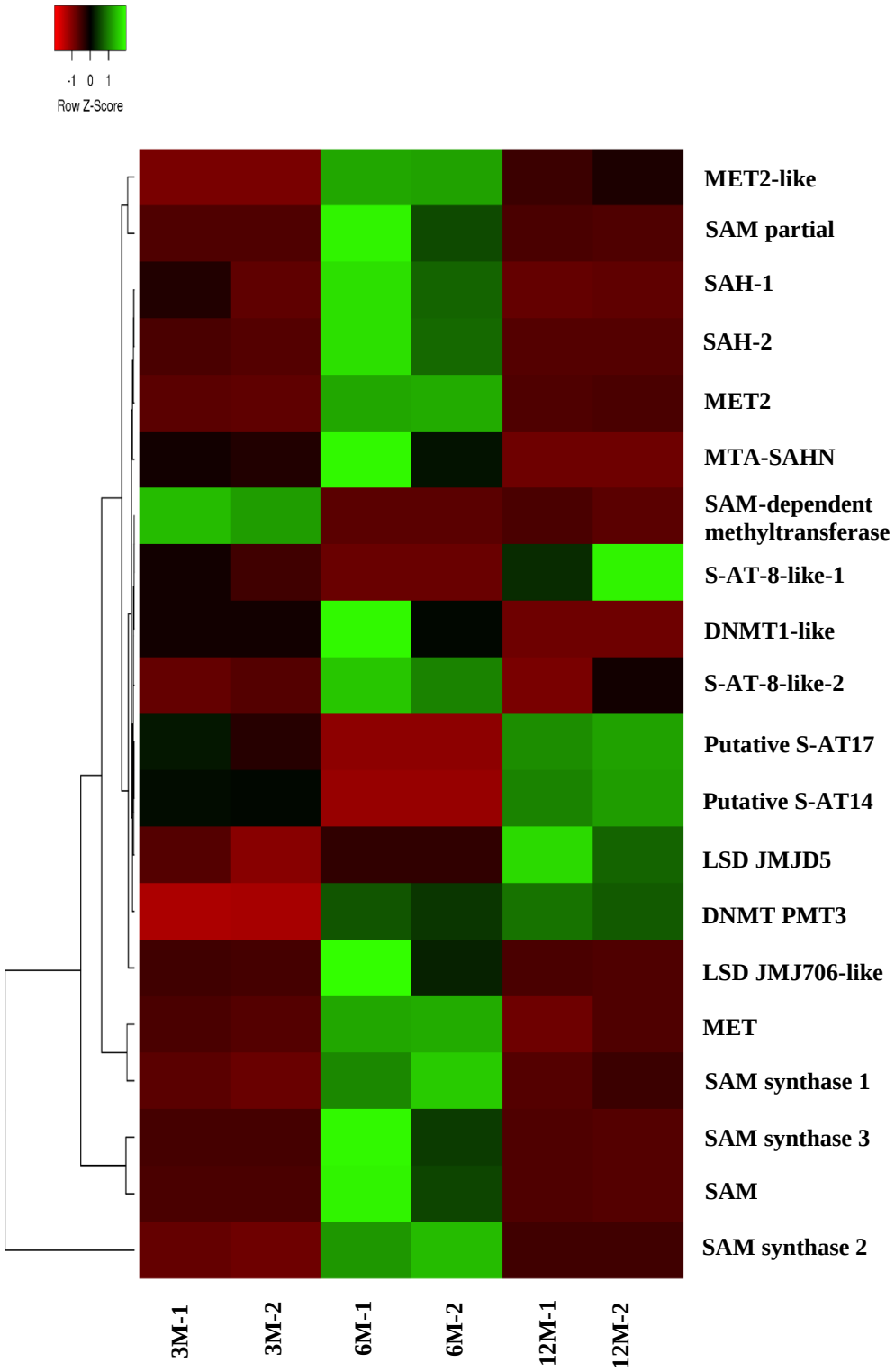


Figure 5.15. Heat maps of different methylation genes expressed in vanilla calli over time period 3M, 6M and 12M. **Red** indicates down-regulation, while **green** indicates up-regulation of genes. MET: 5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase; SAM: S-Adenosylmethionine; SAH: S-adenosyl-L-homocysteine hydrolase; MTA-SAHN: 5-methylthioadenosine S-adenosylhomocysteine nucleosidase 2-like; S-AT: S-acyltransferase; DMNT: DNA (cytosine-5)-methyltransferase; LSD: lysine-specific demethylase.

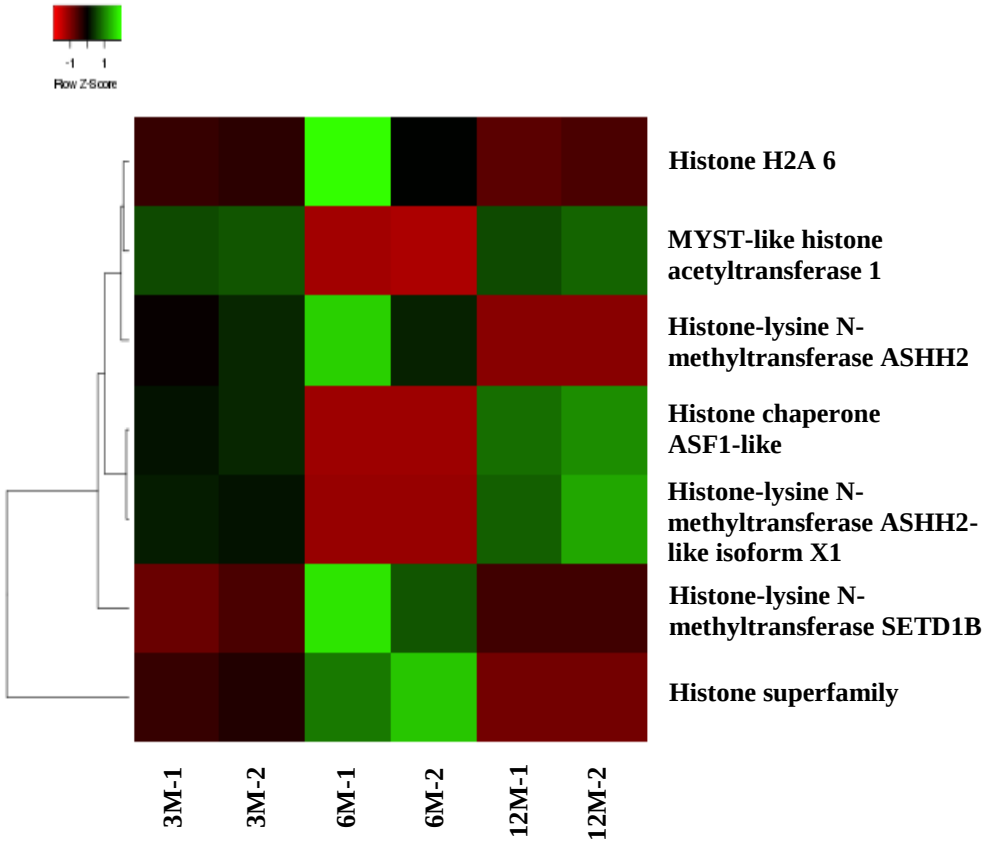


Figure 5.16. Heat maps of different histone protein-encoding genes expressed in vanilla calli over time period 3M, 6M and 12M. **Red** indicates down-regulation, while **green** indicates up-regulation of genes.

5.3.2.3.2. Expression of TFs, phytohormone- and stress-related genes

TFs play important roles in response to abiotic stresses in plants. In our study, a total of 75 contigs encoded for TFs (**Table 5.12**). Except for the DIVARATA (DIV), alfin-like and E2FB (**Figure 5.17**), most of the TFs were up-regulated at 6M, including WRKY, ERF, AP2-ERF, MYB, MYC, NAC, GATA, trihelix-ASIL, RAX3-like, CBX-like, LHW and heat-stress-responsive TF genes (**Figure 5.17**).

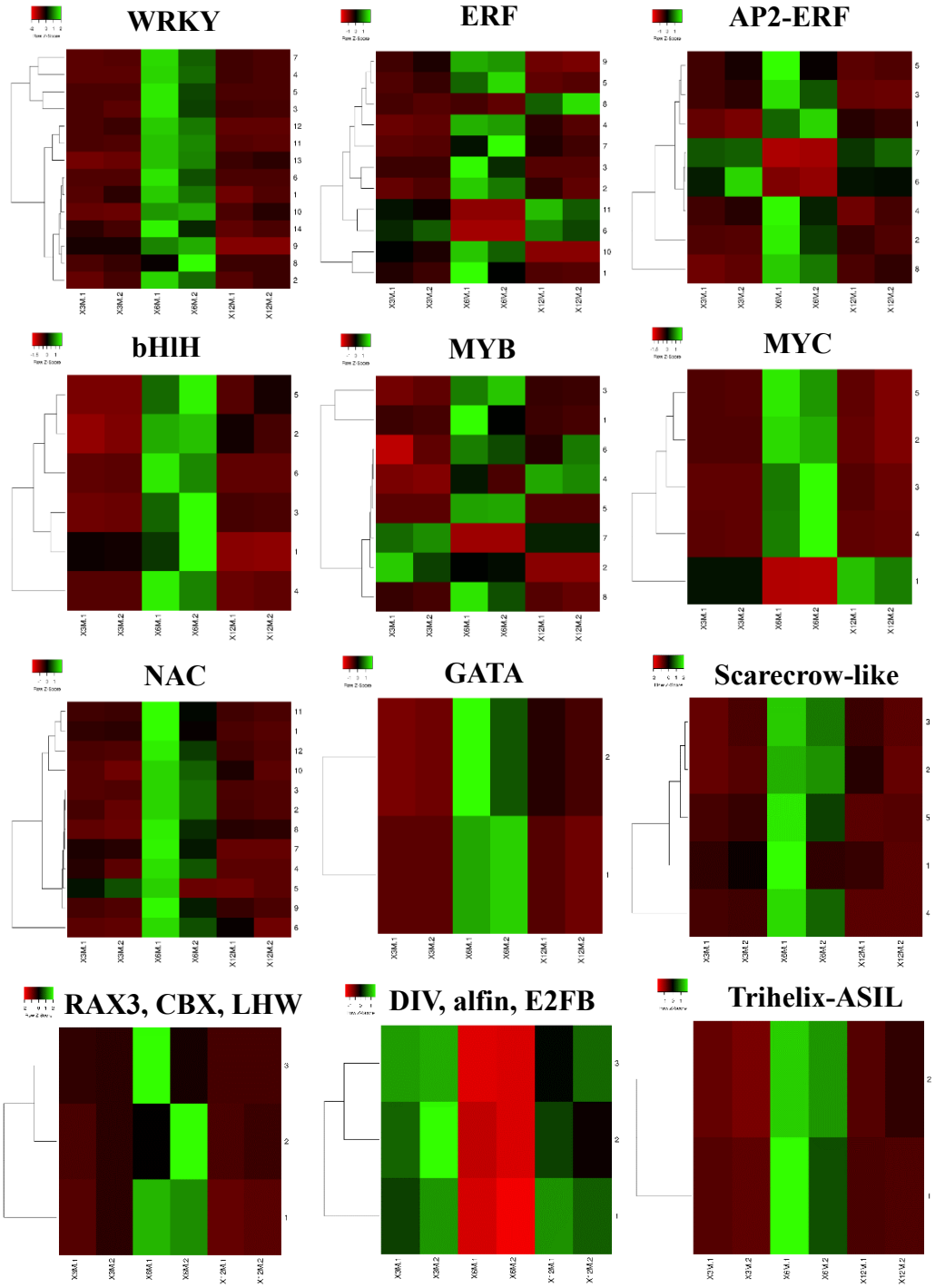


Figure 5.17. Heat maps of different transcription factors expressed in vanilla calli over time period 3M, 6M and 12M. **Red** indicates down-regulation, while **green** indicates up-regulation of genes.

5.3.2.3.3. Expression of TE-related genes

The expression patterns of TE-related genes, in particularly *gypsy*, *copia* and LINE, were different with the other gene families, where most of the genes were

down-regulated in 6M callus and up-regulated in 3M and 12M calli (**Figure 5.18**).

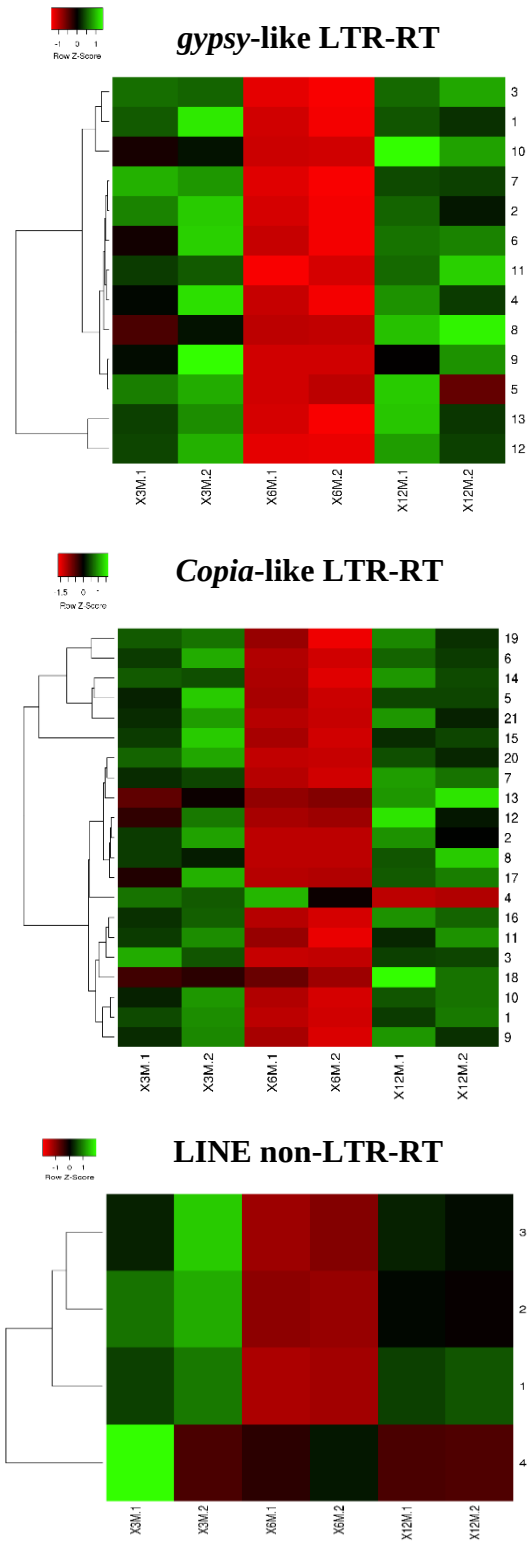


Figure 5.18. Heat maps of different retrotransposon genes expressed in vanilla calli over time period 3M, 6M and 12M. Red indicates down-regulation, while **green** indicates up-regulation of genes.

5.3.3. Small RNA Sequence Profiles

5.3.3.1. Sequence processing and filtering

More than 16 million raw reads were yielded for each RNA sample, which were then filtered for adaptor and contaminants (**Table 5.13**). After the removal of redundant sequences, more than 10 million unique sequences (**Table 5.13**) ranging from 15 to 30 nucleotides (nt) in length (**Figure 5.19**) was produced. Majority of the trimmed and filtered sequences were 21 to 24 nt in lengths, where a major peak in the abundance of small RNAs was observed at 24 nt (**Figure 5.19**).

Table 5.13. Summary of pre-analysis for miRNA sequencing of *V. planifolia* calli. Raw sequencing reads were trimmed for adaptor (Illumina small RNA adaptors), ambiguity (no ambiguous nucleotides were allowed) and filter with sequence length (15 to 30 nucleotides). Same sequences were clustered with the same sequence tags.

Sample name	Read length	Total number of reads	Number of reads after trim and filter	Discarded Reads	Total Number of sequence tag
3M-1	36	16,957,015	13,085,177	3,871,838	4,944,998
3M-2	36	16,985,597	13,400,761	3,584,836	5,384,491
6M-1	36	17,106,726	11,105,987	6,000,739	2,876,014
6M-2	36	17,071,501	11,367,729	5,703,772	3,145,714
12M-1	36	16,130,773	11,025,150	5,105,623	2,410,929
12M-2	36	16,661,118	10,950,240	5,710,878	2,174,664

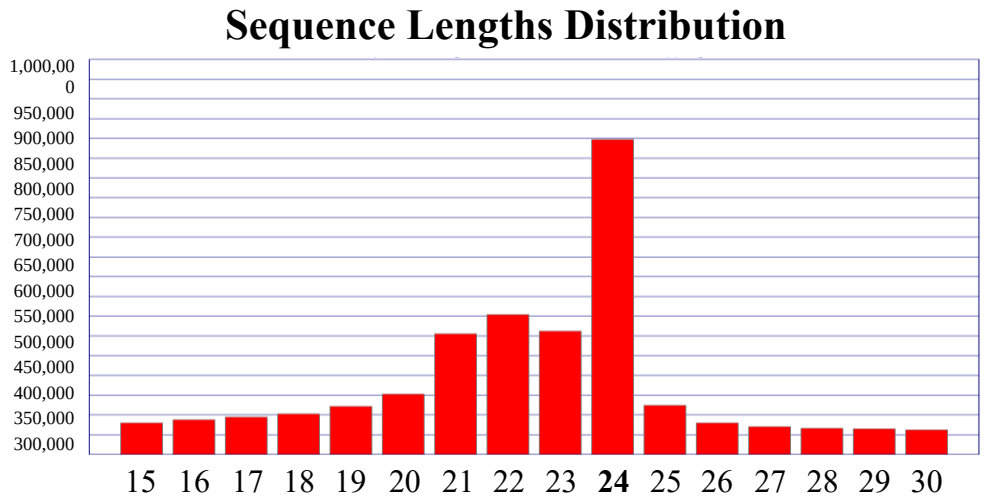


Figure 5.19. Length distribution of sequenced small RNAs of sample 6M-1 after pre-processing. Y axis represented the amount of reads. X axis represented the length of sequences in nucleotides (nt).

5.3.3.2. Small RNA Annotation

Sequences which were mapped against the non-coding RNAs (rRNA, tRNA, snRNA and snoRNA) in the Rfam databases comprised of only a small fraction (1.03%) of the overall sequences count, and were removed to avoid influencing RNA identification (**Table 5.14**). A total of 1.27 million reads which were mapped successfully to the pre-miRNAs sequences of 73 selected plant species (**Appendix F**) in miRBase version 21.0 were annotated, and mature miRNAs were grouped (**Table 5.14**).

Table 5.14. Summary of reads from raw data to annotated miRNA sequences of *V. planifolia.*

Sample name	Total number of sequence tag	Annotated (MirBase)		Unique sequences	Unknown unique sequences
		Number of reads	Mature miRNAs	(Annotated to RFAM)	
3M-1	4,994,998	295,011	207	21574	4,919,296
3M-2	5,384,491	364,461	192	18873	5,361,311
6M-1	2,876,014	163,644	178	62038	2,810,240
6M-2	3,145,714	168,763	188	47894	3,094,026
12M-1	2,410,929	131,977	173	48202	2,359,156
12M-2	2,174,664	145,610	168	17489	2,153,516

5.3.3.3. Global micronome dynamics

A total of 270 conserved unique sequences from 58 miRNA families (**Appendix G**), allowing for sequence mismatches of not more than 5%, were identified. Among these sequences, a number of miRNA families were conserved as we found orthologues of known miRNAs from all other plant species, such as *MIR156*, *MIR159*, *MIR160*, *MIR164*, *MIR166*, *MIR167*, *MIR168*, *MIR169*, *MIR171*, *MIR172*, *MIR319*, *MIR390*, *MIR395*, *MIR396*, *MIR397*, *MIR399* and *MIR408*; whereas some miRNA families were less conserved and were present only in some of the plant species, such as *MIR165*, *MIR5641* and *MIR8175* (conserved in *Arabidopsis*); *MIR2916* (conserved in *Populus euphratica* only); *MIR4995*, *MIR5368*, *MIR5770* and *MIR6300* (conserved in *Glycine max* only); and *MIR5072*, *MIR5077*, *MIR5083*, *MIR5523*, *MIR5538* and *MIR5539* (conserved in *Oryza sativa* only).

A majority of conserved miRNA families contained members that differed in not more than two nucleotides, which are regarded as miRNA isoforms. Nine families (*MIR156*, *MIR166*, *MIR167*, *MIR159*, *MIR319*, *MIR396*, *MIR399*, *MIR171* and *MIR164*) contained the highest number of multiple isoforms with 29, 25, 16, 15, 13, 13, 13, 12 and 11 members, respectively. Thirty-one miRNA families (such as *MIR528*, *MIR894*, *MIR2950*, *MIR3699* and *MIR4376*) contained only one member.

Interestingly, a notable divergence in expression frequency among miRNA unique sequences in vanilla callus cultures was observed (**Table 5.15**). For example, *MIR398*, *MIR319*, *MIR408* and *MIR159* were expressed in high frequencies, with an average of more than 30,000 counts across each callus culture; *MIR156*, *MIR162*, *MIR166*, *MIR168* and *MIR396* were expressed in moderately high frequencies (1,000 to 10,000 average counts); while the remaining miRNA populations (10.04%, 18.28% ad 65.95%) were expressed in moderate (100 to 1,000 average counts), moderately low (10 to 1000 average counts) and extremely low (less than 10 average counts) frequencies, respectively (**Table 5.15**) across the callus cultures.

Table 5.15. Expression frequencies of miRNA unique sequences.

Expression frequency	Percentage (%)	Average count range
High	1.43	>30,000
Moderately high	4.30	1,000-10,000
Moderate	10.04	100-1,000
Moderately low	18.28	10-100
Low	65.95	<10

5.3.3.4. Differential miRNA expressions (DME)

In total, 45 of the annotated miRNA candidates were significant during pair-wise comparisons of 3M versus 6M, of which 18 of them were significantly up-regulated and 27 of them were significantly down-regulated; while 7 of the annotated miRNA candidates were significant during comparisons of 6M versus 12M, of which 2 of them were significantly up-regulated and 5 of them were significantly down-regulated (**Table 5.16**, **Figure 5.20**); at FDR-corrected *p*-value ≤ 0.05 and fold-change differences of at least 2.0.

Table 5.16. DME summary of vanilla callus over 3, 6 and 12M prolonged culture durations.

	Significant	Up-regulation	Down-regulation
3M VS 6M	45	18	27
6M VS 12M	7	2	5

The differential expression by project which have been filter at FDR *p*-value ≤ 0.05 and FC ≥ 2

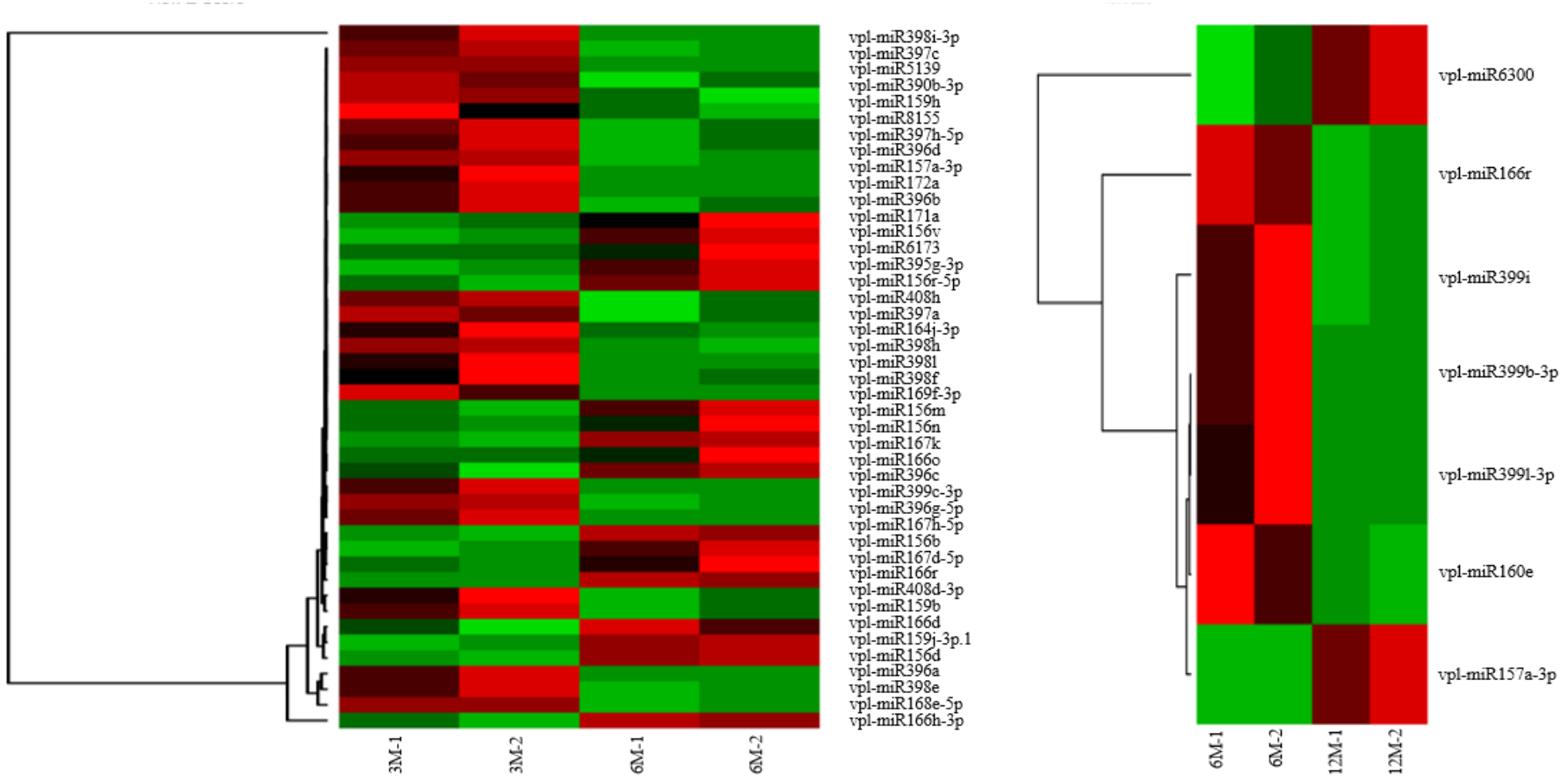


Figure 5.20. Heat map graphs of significantly expressed miRNAs during pairwise comparisons between (a) 3M versus 6M, and (b) 6M vs 12M. Red bar indicated genes down-regulated, while **green** bars indicated genes up-regulated.

5.3.4. miRNA-mRNA Target Correlations

5.3.4.1. Target prediction and correlation analysis

The putative targets of the identified miRNAs were predicted from the mRNA assembled contigs (**Section 5.3.2.1**) using psRobot, on the basis of complete complementarity between the miRNA and target mRNA (Kidner and Martienssen, 2005) with a total ComiR score not higher than 4.0 (**Figure 5.21**).

```
>vpl-miR395g-3p_34  Score: 4.0  vpl_contig_14640

Query:           1 CTGAAGTGTTTGGGGGAACTC 21
                   ||*||*|||:|||:|||*|*|
Sbjct:         387 GAGTTGACAGACCTCCT-GTG 368
```

Figure 5.21. Sequence alignment during target prediction of vpl-miR395g and one of its potential mRNA target, vpl_contig_14640 at a ComiR score of 4.0. The symbols, | represents reverse complement match, * represents identical sequence match, while : represents mismatch.

Subsequently, the list of DGEs for miRNAs and mRNAs were combined and analyzed to check for potential miRNA-mRNA target correlations between the two sequence profiles. Potential mRNA targets of the miRNAs were filtered based on the negative correlations of expressions between mRNA and miRNA. In total, there are 54 miRNAs with significant DGEs in vanilla callus culture, at negative correlations with significant DGEs of predicted mRNA targets (**Figure 5.22**). Amongst all, 3 of the miRNAs, namely vpl_miR166r, vpl_miR395g and vpl_miR8155, were significantly expressed throughout the prolonged callus culture of 3M, 6M and 12M: vpl_miR395g and vpl_miR166r were significantly up-regulated in 6M callus, while vpl_miR8155 was significantly down-regulated in 6M callus. These three miRNAs, vpl_miR395g, vpl_miR166r and vpl_miR8155 were found to have 57, 4 and 11 potential target mRNAs respectively. These targets were further analyzed to validate their functions and target sites.

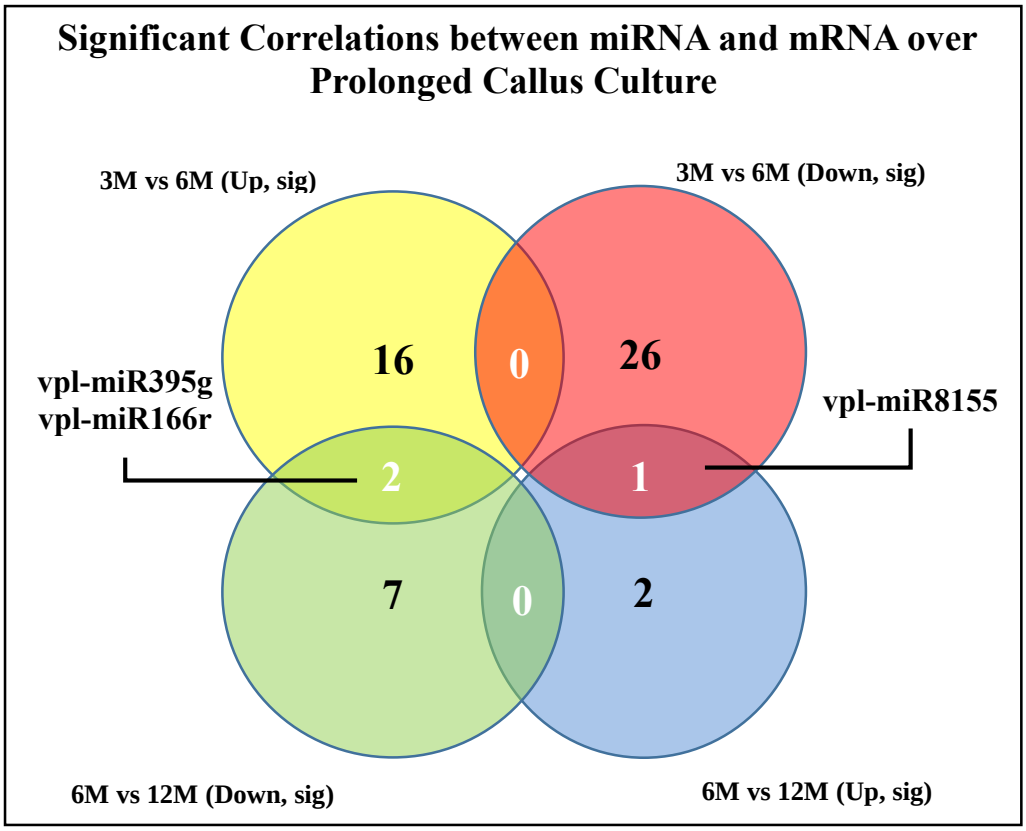


Figure 5.22. Venn diagram showing significant (FDR-corrected *p*-value ≤ 0.05) correlations between miRNA and mRNA expressions. Arrow showed the common miRNA between timepoints, miR8155, miR395g and miR166r.

5.3.4.2. Conserved Protein Domains and miRNAs Target Sites

After an overall screening, the 3 miRNAs, vpl_miR395g, vpl_miR166r and vpl_miR8155 were found to potentially correlate to the gene expression values of eight annotatable genes, which fell into the gene families of DNA methylation and transposable elements. These 8 genes were further characterized through more in-depth BLAST analysis. Multiple gene targets were detected for vpl-miR395g-3p and vpl-miR166r, and could be categorized by their gene ontologies, DNA integration (GO:0015074), DNA binding (GO:0003676) and cofactor-binding (GO:0048037), which was related to retrotransposable activities. As for vpl_miR8155, one potential predicted target gene which had both binding (GO:0005488) and transferase activity (GO:0016740), which was related to DNA methylation were studied.

5.3.4.2.1. DNA methyltransferase and miRNA

Target gene of *vpl_miR8155* was annotated to the 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase (MET) 1 isoform X2 gene with 99% query cover, 91% identity and an e-value of 5E-143 to the gene sequence of *Phalaenopsis equestris* (XP_020589210.1) (**Figure 5.23**). This gene appeared conserved across species, with high alignment scores (above 200) with other species such as *Dendrobium catenatum* (XP_020687014.1) and *Apostasia shenzhenica* (PKA56998.1) and conserved protein domains (**Figure 5.23**, **Table 5.17**).

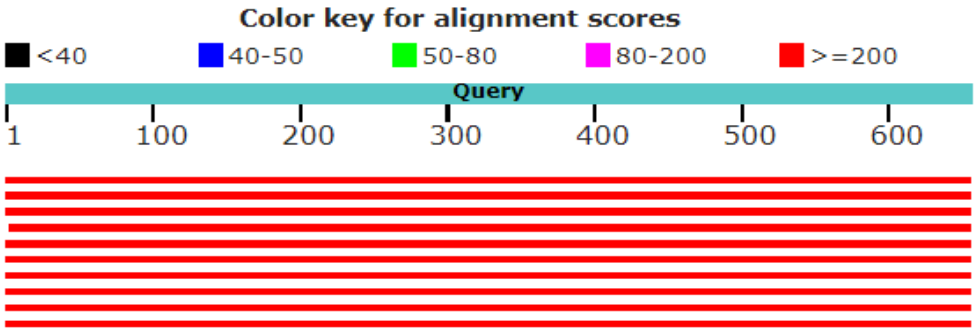


Figure 5.23. BLAST Hit Alignment score of top ten hits to vpl_contig_9944 to *nr* database. The **red** bars showed the alignment scores of higher than 200, to the 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase （MET） 1 isoform X2 gene sequence of *Phalaenopsis equestris* (XP_020589210.1). The other hits were also conserved in other species, namely *Dendrobium catenatum* (XP_020687014.1), *Apostasia shenzhenica* (PKA56998.1) .

Table 5.17. Conserved domains from BLAST results of vpl_contig_9944 to putative methione synthase gene with gene ontology of transferase activity and binding (GO: 0016740 and GO:0005488).

Symbol	Superfamily	Domain
PLN02475 or met_syn_B12ind	PRK05222 (cl28374)	5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase (PLN02475 or TIGR01371)
Meth_synt_1 or CIMS_N_terminal_like	URO-D_CIMS_like (cl00464)	Cobalamine-independent methionine synthase (pfam08267), or MetE, N-terminal domain-like (cd03312)

Interestingly, in a global view of transcriptome expressed in vanilla callus culture, apart from vpl_contig_9944, a series of genes encoding for enzymes involved in the cysteine and methionine metabolism were found to be co-expressed in the callus culture of vanilla (**Figure 5.7**).

5.3.4.2.2. Co-factor-binding and miRNA

Another *vpl_miR395g-3p* potential target contig region, vpl_contig_14640, encoded for a different protein family, the C-terminal binding protein AN (PKA60438.1) (**Figure 5.24**). This gene contained a conserved nicotinamide adenine dinucleotide (NAD)-binding domain (IPR006140) from "FDH_GDH_like superfamily" (CDD: 330194), which consisted of multiple NAD binding sites (cd12154).

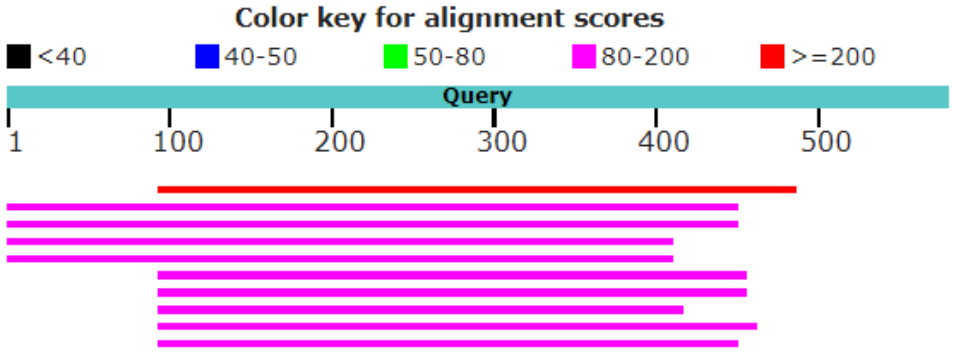


Figure 5.24. BLAST Hit Alignment score of top ten hits to vpl_contig_14640 to *nr* database. The top hit (in **red** bar) showed an alignment score higher than 200, to the C-terminal binding protein AN sequence of *Apostasia shenzhenica* (PKA60438.1). Other hits with alignment scores between 80 and 200 (in **pink** bars) were also specific to C-terminal binding protein, but in other species, namely *Elaeis guineensis* (XP_019705218.1), *Phoenix dactylifera* (XP_008805571.1) and *Dendrobium catenatum* (XP_020685610.1).

5.3.4.2.3. DNA integration and miRNA

Analysis of the potential targets of vpl-miR395g-3p and miR166r with gene ontology of DNA integration (P:GO:0015074), RNA binding (F:GO:0003723), RNA-directed DNA polymerase activity (F:GO:003964) and RNA-dependent DNA replication (P:GO:0006278), such as vpl_contig_24863, showed interesting pattern of conserved proteins related to long terminal repeats (LTR) retroviral polyproteins (**Table 5.18**), where local alignments of the conserved regions showed resemblance to putative *ty1/copia*-type polyprotein model (**Figure 5.25**). The miRNA target binding site to these groups of genes is mostly bound at the upstream of RNAse H-like domain (**Figure 5.25**). The other contigs with similar *ty1/copia*-like patterns included vpl_contig_86799 (potentially bound by vpl-miR166r) and vpl_contig_32170 (potentially bound by vpl-miR395-3p). For vpl_contig_55903 (potentially bound by vpl_miR395g-3p), the polyprotein shared conserved protein domains to the ty3/gypsy-like family instead.

Table 5.18. Conserved domains from BLAST results of vpl_contig_24683 to putative retroviral polyproteins with gene ontology of DNA integration (GO: 15074).

Symbol	Superfamily	Domain
Rve	Rve (cl21549)	Integrase core domain (pfam00665)
RVT_2	RVT_2 (cl06662)	Reverse transcriptase RNA-dependent DNA polymerase (pfam00078)
RNase_HI_RT_Ty1	RNAse_H_like (cd06222)	Ty1/Copia family of RNase HI in long-term repeat retroelements (cd09272)

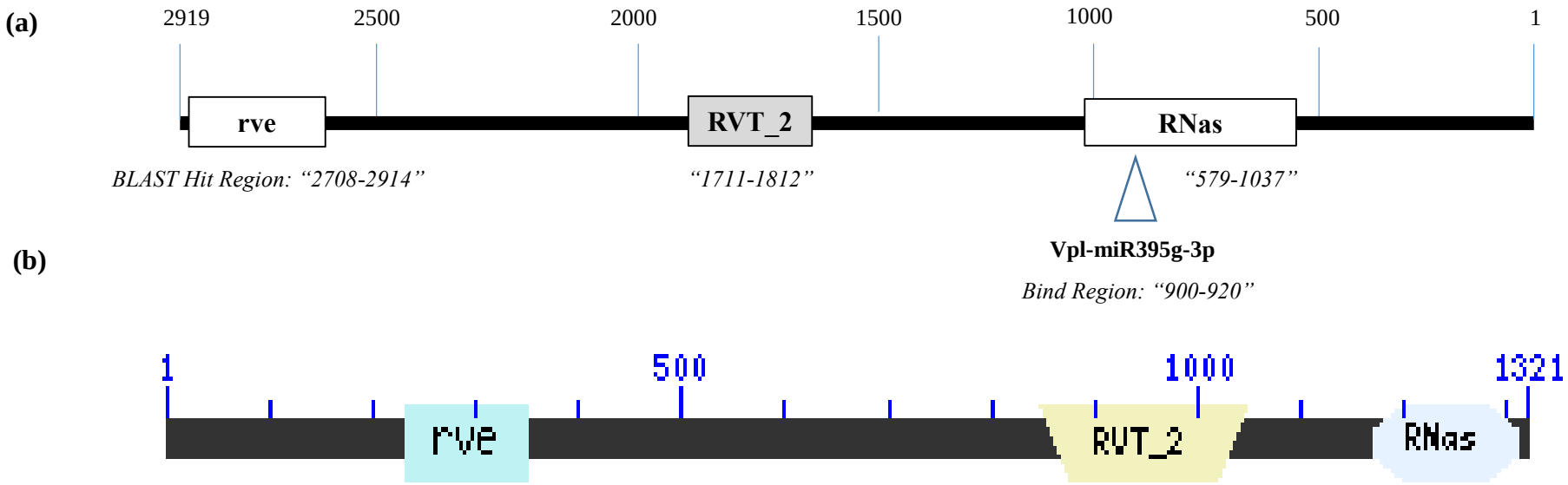


Figure 5.25. Illustration of BLAST sequence alignment and positions of conserved regions of vpl_contig_24683: (a) Conserved regions in vpl_contig_24863 and (b) similar protein architecture to vpl_contig_24683 found in 644 other species in *nr* databases (taxonomy spanned: Eukaryota) through CDART search (Geer *et al.*, 2002) of NCBI, which was closest to this contig based on the conserved regions. The "rve" represented the "integrase core domain" (pfam: 00605); "RVT" represented the "reverse transcriptase RNA-dependent DNA polymerase" (pfam00078) and "RNAs" represented the "RNase_HI_RT_Ty1" or "Ty1/Copia family of RNase HI in long-term repeat retroelements" (cd09272). The triangle pointed to the target region of vpl_miR395g-3p predicted using *psRobot*, which flanked the "RNase_HI_RT_Ty1" region.

5.3.4.2.4. DNA binding and miRNA

The analysis of miRNA targets with gene ontology of DNA binding (GO: 003676), such as vpl_contig_18270, showed a different pattern of conserved proteins which resembled non-LTR retrotransposon polyproteins (**Table 5.19**), where local alignments of the conserved regions showed five conserved domains that closely resembled one of the non-LTR retrotransposon polyprotein models found in 42 other angiosperms sequences in NCBI *nr* database (**Figure 5.26**). The miRNA target binding site of this gene flanked at a region closely before the "zf-CCHC_4" domain (**Figure 5.26**). Other contigs with similar patterns included vpl_contig_1034 (bound by vpl-miR166r).

Table 5.19. Conserved domains from BLAST results of vpl_contig_18270 to putative retroviral polyproteins with gene ontology of DNA binding (GO: 003676).

Symbol	Superfamily	Domain
zf_RVT	zf-RVT (cl16506)	Zinc-binding in reverse transcriptase (DNA-dependent DNA polymerase) (pfam13966)
RT_nLTR_like	RT_like (cl02808)	Non-LTR (long terminal repeat) retrotransposon and non-LTR retrovirus transcriptase (RT) (cd01650)
RVT_1	RVT_1 (cl26764)	Reverse transcriptase (RNA-dependent DNA polymerase) (pfam00078)
zf-CCHC_4	zf-CCHC_4 (cl18787)	Zinc knuckle (pfam14392)
DUF4283	DUF4283 (cl16623)	Domain of unknown function (DUF4283) (pfam14111)

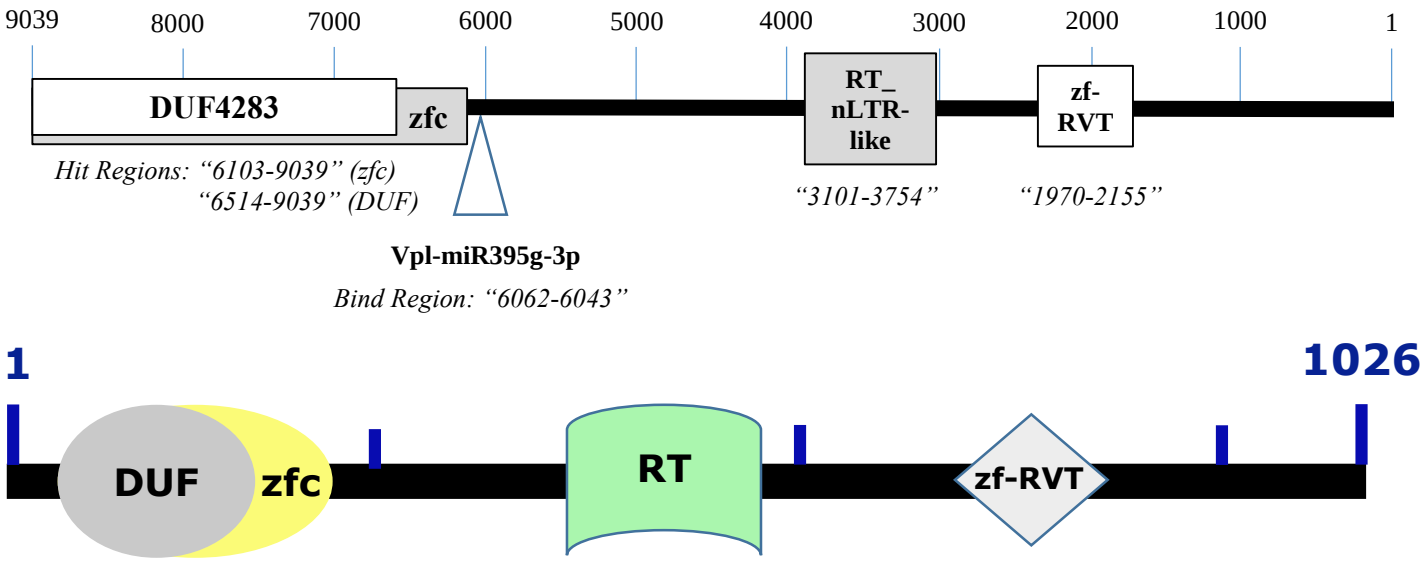


Figure 5.26. Illustration of BLAST sequence alignment and positions of conserved regions of vpl_contig_14640: (a) Conserved regions in vpl_contig_14640 and (b) similar protein architecture to vpl_contig_14640 found in 42 other species in *nr* databases (taxonomy spanned: Mesangiospermae) through CDART search (Geer *et al.*, 2002) of NCBI, which was closest to this contig based on the conserved regions. The "zf-RVT" represented the "Zinc-binding in reverse transcriptase (DNA-dependent DNA polymerase)" (pfam13966)"; "RT_nLTR_like" represented the "non-LTR (long terminal repeat) retrotransposon and non-LTR retrovirus transcriptase (RT) (cd01650)"; the "RVT_1" represented the "reverse transcriptase (RNA-dependent DNA polymerase)" (pfam00078); the "zf-CCHC_4" represented "zinc knuckle (pfam14392)"and "DUF4283" represented the "domain of unknown function (DUF4283) (pfam14111)". The triangle (△) pointed to the target region of vpl_miR395g-3p predicted with *psRobot*, which flanked closely before the "zfc" region.

5.3.5. Primer Design for RT-qPCR

A total of eight mRNAs and three miRNAs were selected for RT-qPCR validations (**Table 5.20**). Validation of specificity of Primer3-designed primers using PRIMER-BLAST against the National Center for Biotechnology Information (NCBI) databases (division of *nr viridiplantae*) showed that the primers were specific **(Figure 5.27).**

Table 5.20. Summary of genes to be validated using qRT-PCR.

Gene_ID	Relative gene expression			
	3M vs 6M		6M vs 12M	
	Log_2 FC[1]	*p*-value	Log_2 FC	*p*-value
mRNA sequencing				
vpl_contig_001034	-3.29	5.26E-07	2.07	3.12E-03
vpl_contig_032170	-5.48	1.35E-07	4.74	9.97E-06
vpl_contig_086799	-3.94	1.42E-02	3.54	3.30E-02
vpl_contig_055903	-3.00	1.73E-02	3.27	3.82E-03
vpl_contig_009944	2.44	2.78E-05	-1.57	1.08E-02
vpl_contig_024863	-1.89	2.05E-02	1.99	9.35E-03
vpl_contig_001827	-2.20	1.82E-03	1.67	2.15E-02
vpl_contig_014640	-9.00	7.40E-10	10.32	2.47E-15
miRNA-sequencing				
vpl_miR166r	2.84	2.61E-33	-1.01	2.34E-02
vpl_miR395g	1.75	1.73E-05	-1.18	5.38E-02
vpl_miR8155	-1.97	6.05E-03	1.32	5.39E-02

[1]FC: fold change

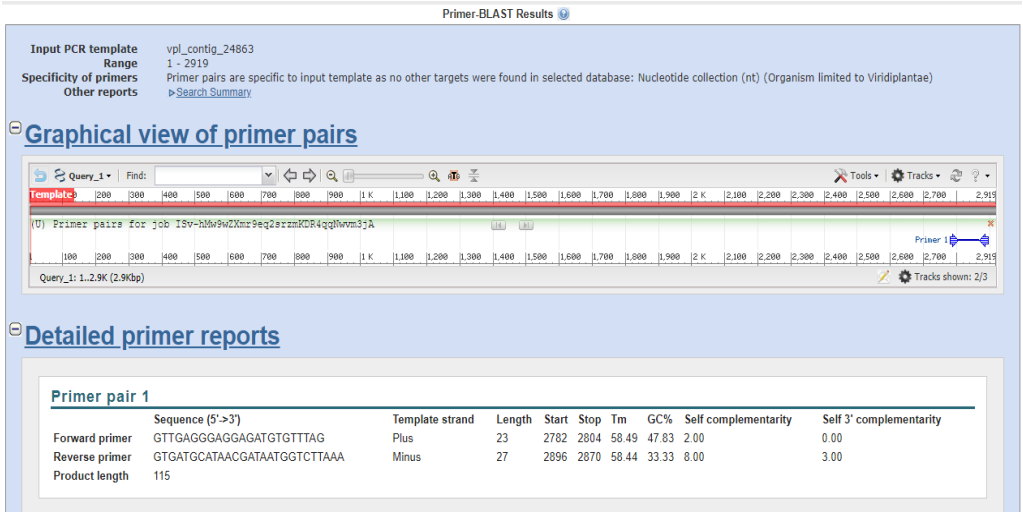


Figure 5.27. Preliminary results of Primer-BLAST report using Vpl_contig_24863 and its forward and reverse primer pair.

5.3.6. Annealing Temperature Optimizations

5.3.6.1. mRNA

PCR amplification reactions with gradient annealing temperatures (50.0, 50.4, 51.3, 52.7, 54.7, 56.6, 58.4, 60.3, 60.3, 62.3, 63.7, 64.6 and 65.0ºC) were performed to verify the optimum annealing temperature for each mRNA primer pairs (**Table 5.21**). For all the five mRNA reference gene candidates, amplifications were stable, with similar band intensity and single band observed across all temperature (50-65ºC) (**Figure 5.28**), except that EF-1α had a fainter PCR amplicon at 64.6ºC (**Lanes 11**, **Figure 5.28**) and above. Among the seven mRNA target genes, unspecific banding patterns were observed in Vpl_18270 and Vpl_24863 at temperatures below 52.7ºC (**Lane 4**, **Figure 5.28**) and 63.7ºC (**Lane 8**, **Figure 5.29**) respectively. Besides, for genes Vpl_32170 and Vpl_55903, amplifications were stable but bands were fainter at 63.7ºC (**Lane 10**, **Figure 5.29**) and above. Therefore, the annealing temperature of 60.3ºC (**Lane 9**) was chosen as the optimal annealing temperature to be used across all genes (both reference and target genes) in further RT-qPCR analyses**.**

Table 5.21. Size ranges of expected amplicon size and PCR-amplified amplicon.

Genes	Expected amplicon (bp)	PCR amplicon (bp)
References Genes		
Actin	112	100 - 200
Ubiquitin	60	< 100
EF-1α	96	~100
β-tubulin	121	100 - 200
18S rRNA	178	100 - 200
Target Genes		
Vpl_contig_001034	126	100 - 200
Vpl_contig_014640	100	~100
Vpl_contig_018270	122	100 - 200
Vpl_contig_024863	115	100 - 200
Vpl_contig_032170	116	100 - 200
Vpl_contig_055903	88	< 100
Vpl_contig_086799	120	100 – 200

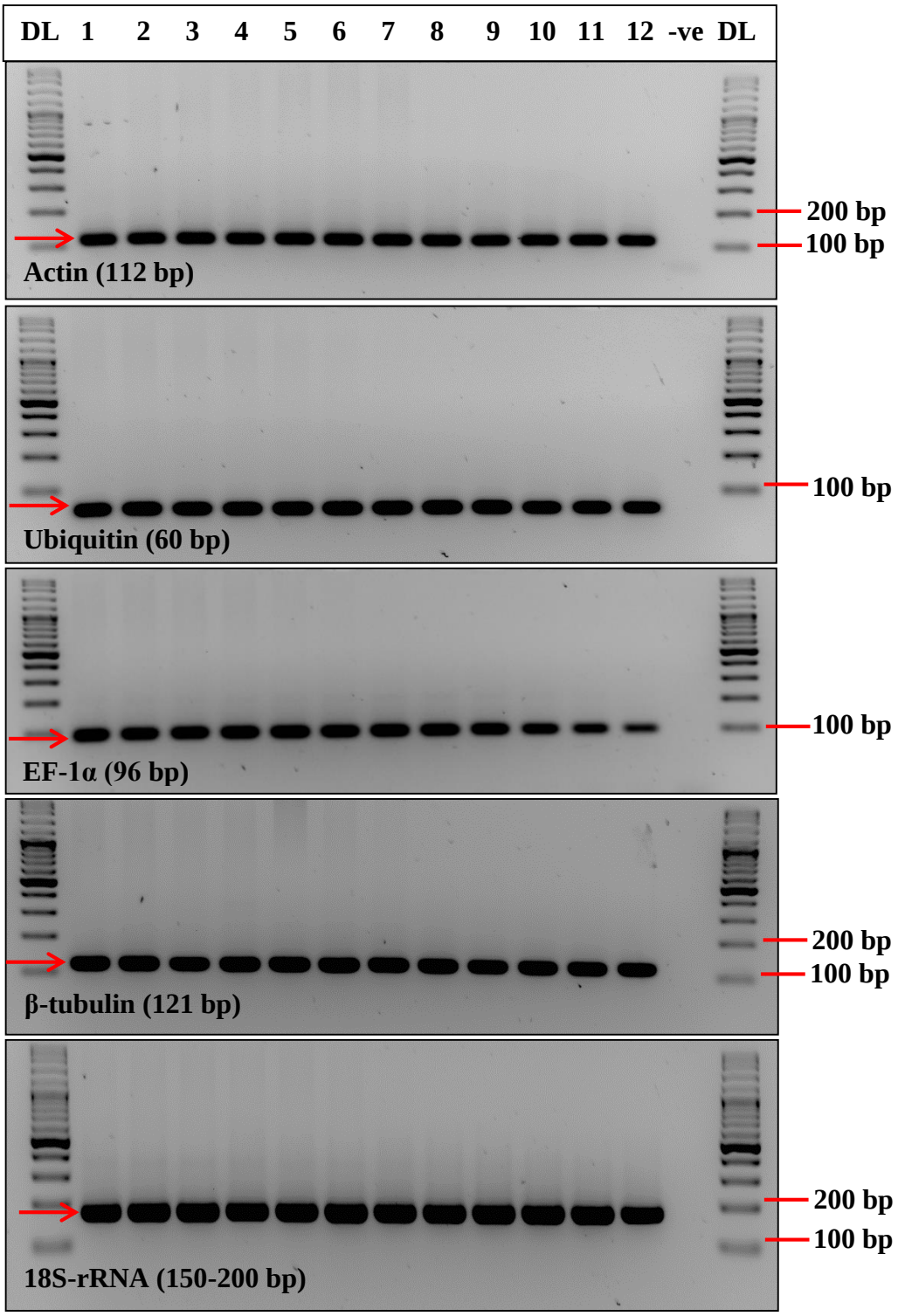


Figure 5.28. Agarose gel electrophoresis (2.5% w/v) of *V. planifolia* cDNA amplified by PCR to determine the optimum annealing temperature of each reference genes primer pair. Lane DL: 100 bp plus DNA Ladder (Thermo Scientific, USA); **Lane 1-12:** Gradient annealing temperature (50.0, 50.4, 51.3, 52.7, 54.7, 56.6, 58.4, 60.3, 60.3, 62.3, 63.7, 64.6 and 65.0ºC); **Lane -ve:** PCR Negative Control. **Arrow** indicated PCR amplicons.

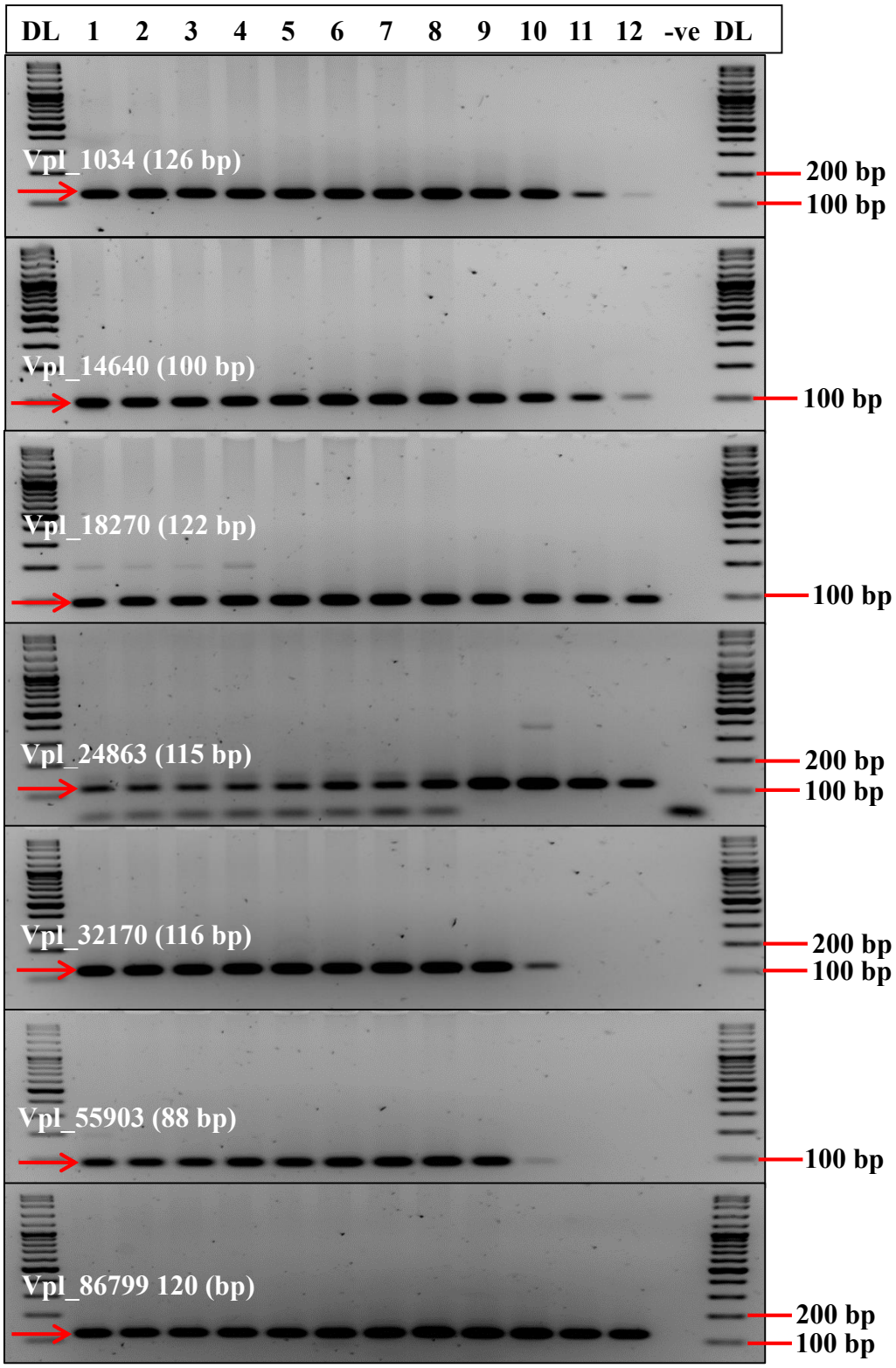


Figure 5.29. Gradient Annealing Temperature PCR Profiles of *V. planifolia* cDNA amplified by mRNA target genes primer pairs on 2.5% (w/v) agarose gel. Lane DL: 100 bp plus DNA Ladder (Thermo Scientific, USA); **Lane 1-12:** Gradient annealing temperature (50.0, 50.4, 51.3, 52.7, 54.7, 56.6, 58.4, 60.3, 60.3, 62.3, 63.7, 64.6 and 65.0ºC); **Lane -ve:** PCR Negative Control. **Arrows** indicated PCR amplicons.

5.3.6.2. miRNA

PCR amplification with gradient annealing temperatures (40.0, 41.2, 43.9, 47.7, 52.6, 56.5, 58.8 and 60.0ºC) were performed to verify the optimum annealing temperature for each miRNA primer pair (**Table 5.21**). Optimal amplification for miRNA reference and target gene candidates was analyzed from the melt peaks of gradient amplifications (**Figure 5.30**). Among the five reference and three target miRNA candidates, amplifications were stable across 40.0 to 60.0ºC, with single melting peak across all temperatures, except for *vpl-miR5179* (**Figure 5.30d**) which showed no amplification above 58.8ºC, *vpl-miR166r* which showed low amplification peak at 40.0C (**Figure 5.30e**), and *vpl-miR395g-3p* which showed a potential dimer (as indicated by **blue arrow** in **Figure 5.30g**). Therefore, 43.9ºC was selected as the optimal annealing temperature that works for all miRNAs in a single experiment for subsequent RT-qPCR analyses.

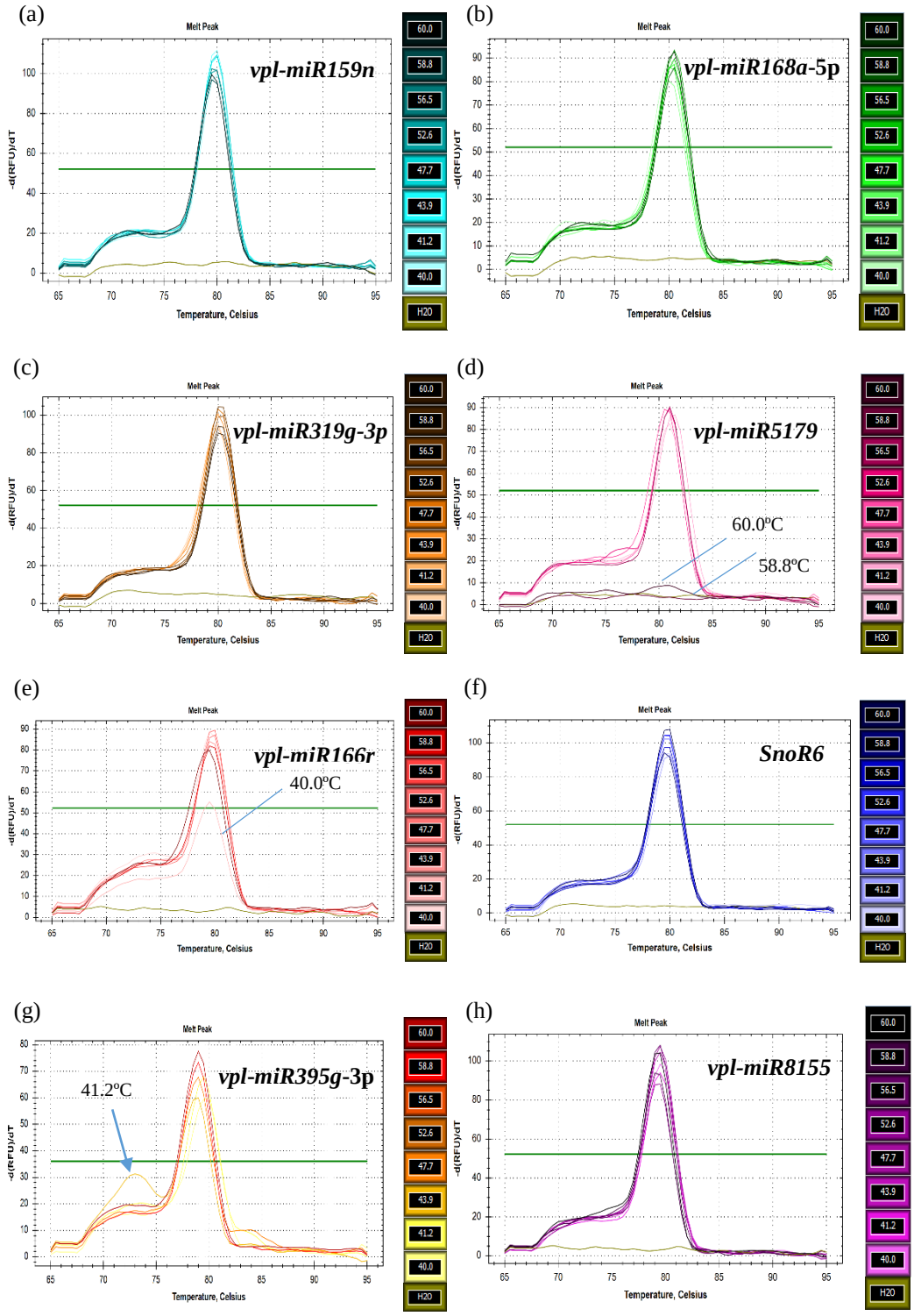


Figure 5.30. Melt peak in gradient PCR of reference and target miRNAs: (a) *vpl-miR159n*, (b) *vpl-miR168a-5p*, (c) *vpl-miR319g-3p*, (d) *vpl-miR5179*, (e) *vpl-miR166r*, (f) *SnoR6*, (g) *vpl-miR395g-3p* and (h) *vpl-miR8155*.

5.3.7. Estimation of Reference Genes Stability

The references genes were subjected to qPCR analysis using pooled cDNA samples of three biological samples (3M, 6M and 12M), and three technical replicates for each biological set, to estimate the stability of each reference gene (β-actin (*ACT*), α-tubulin (*TUB*), ubiquitin (*UBQ*), 18S ribosomal RNA (*18S*) and elongation factor-1 alpha (*EF1*)) under the stress of prolonged culture durations. The stability of each reference genes was analyzed using five algorithms, NormFinder, Delta CT method, BestKeeper, GeNorm and RefFinder.

In GeNorm, the M value of reference gene candidates ranged from 0.027 to 0.049; while M value of reference miRNA candidates ranged from 0.057 to 0.152 (**Table 5.22**). As all M values fell below 1.5, these reference genes are considered stable (Bustin *et al.*, 2010; De Spiegelaere *et al.*, 2015; Vandesompele *et al.*, 2002). From the geomean ranking value, the overall estimated stability of the five reference gene candidates ranked from *18S* being the most stable reference gene (average ranking value of 1.732), followed by *EF1* (2.000), *TUB* (2.449), *ACT* (2.991) and *UBQ* (3.976) (**Table 5.22**); and thus *EF1*, *18S* and *TUB* were chosen as the estimated top three stable reference genes to be used for subsequent mRNA RT-qPCR experiments. For miRNA reference candidates, *vpl-miR159r* was the estimated most stable reference miRNA candidate (average ranking value of 1.189), followed by *vpl-miR168a-5p* (1.565), *vpl-miR319g-3p* (2.711), *SnoR6* (4.000) and *vpl-miR5179* (5.000) (**Table 5.22**), and thus *vpl-miR159r* and *miR169a-5p* were selected as the estimated most stable reference miRNA candidates to be used in subsequent miRNA RT-qPCR experiments.

Table 5.22. Estimation of expression stability of reference gene / miRNA candidates across different prolonged culture duration of vanilla callus computed by five algorithms (RefFinder, Delta CT, BestKeeper, NormFinder and GeNorm).

Gene / miRNA	Geomean of ranking values (RefFinder)	Average of STDEV (Delta CT method)	Coeff. Of Corr. [r] (BestKeeper)	Stability value (NormFinder)	M value (GeNorm)
mRNA					
Actin (*ACT*)	2.991	0.537	0.987	0.456	0.034
Ubiquitin (*UBQ*)	3.976+	0.801+	-0.997+	0.773+	0.049+
E 1	2.000	0.432	0.908*	0.068*	0.027*
β-tubulin (*TUB*)	2.449	0.516	0.942	0.399	0.029
18S-rRNA (*18S*)	1.732*	0.429*	0.928	0.068*	0.028
MiRNA					
vpl-miR159r	1.189*	1.008*	0.977	0.008*	0.057*
vpl-miR319g-3p	2.711	1.129	0.792	0.009	0.063
vpl-miR168a-5p	1.565	1.128	0.298	0.024	0.064
SnoR6	4.000	1.426	0.214*	0.039	0.076
vpl-miR5179	5.000+	2.605+	0.999+	0.104+	0.152+

*indicates the most stably expressed gene(s)/miRNA(s) and + the least stably expressed gene as determined by each software.

5.3.8. qPCR Amplification Efficiency of Genes

5.3.8.1. mRNA

The efficiencies of PCR amplification of mRNA primers was determined by the construction of concentration curves using pooled cDNAs of vanilla callus through seven series of two-fold serial dilutions (**Figure 5.31**; **Table 5.23**). The results showed PCR efficiency at the range of 85.3 to 107.9%.

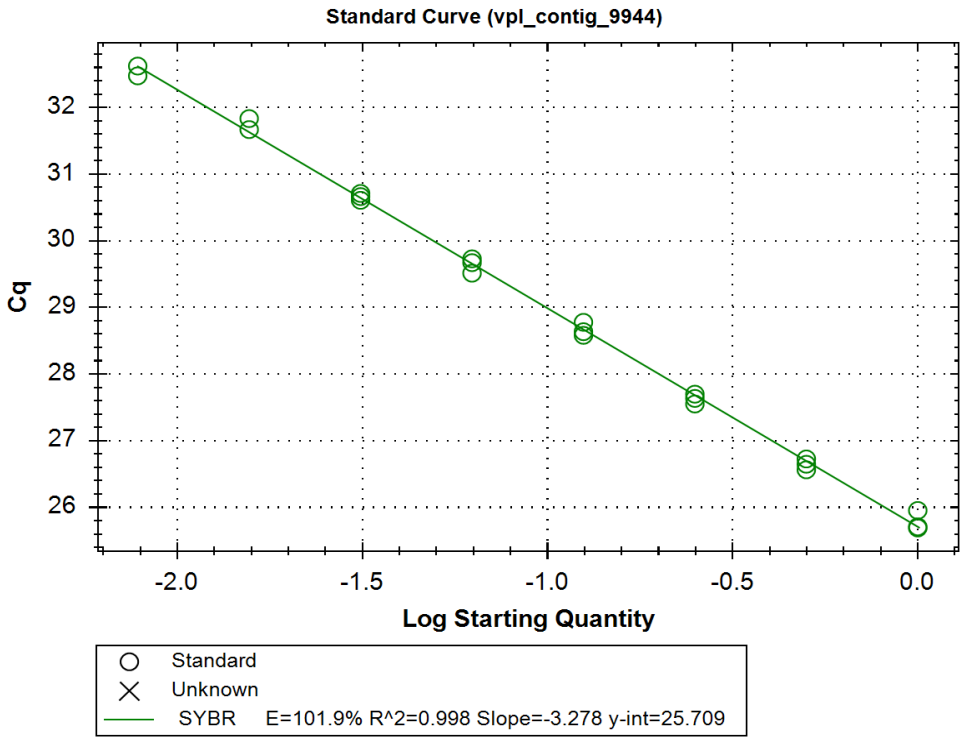


Figure 5.31. Standard curve of *vpl_contig_9944*. Standard curve was constructed from seven series of two-fold dilution.

Table 5.23. PCR efficiency results of mRNA reference and target genes.

mRNA	PCR Efficiency, E (%)	R^2	Slope	y-int
EF1	91.2	0.995	-3.553	26.50
18S	95.4	0.998	-3.437	19.75
vpl_contig_001034	117.8	0.964	-2.958	30.22
vpl_contig_009944	101.9	0.998	-3.278	25.71
vpl_contig_014640	92.8	0.988	-3.509	31.63
vpl_contig_018270	101.7	0.958	-3.283	30.79
vpl_contig_024863	123.0	0.867	-2.871	31.82
vpl_contig_032170	85.3	0.989	-3.733	31.40
vpl_contig_055903	107.9	0.952	-3.145	32.62
vpl_contig_086799	99.5	0.906	-3.334	32.31

5.3.8.2. miRNA

The efficiencies of PCR amplification of miRNA primers showed efficiency at the range of 82.4 to 99.6% (**Figure 5.32**; **Table 5.24**).

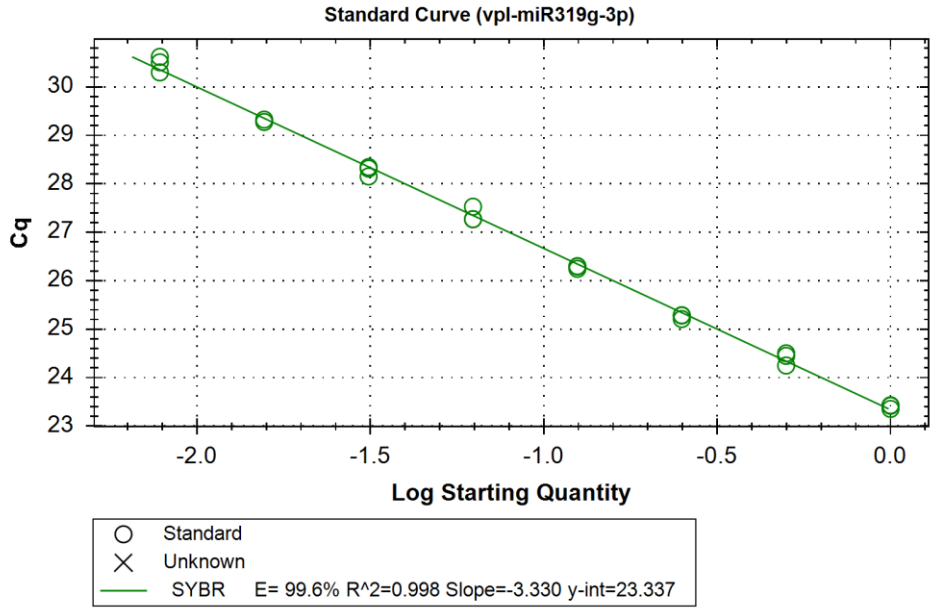


Figure 5.32. Standard curve of *vpl-miR395g-3p*. Standard curve was constructed from seven series of two-fold dilution.

Table 5.24. PCR efficiency results of miRNA reference and target genes.

miRNA	PCR Efficiency, E (%)	R^2	Slope	y-int
vpl_miR166r	93.4	0.997	-3.491	25.164
vpl_miR395g	82.4	0.969	-3.830	35.212
vpl_miR8155	91.6	0.987	-3.542	27.096
vpl_miR159n	98.9	0.990	-3.349	25.489
vpl_miR319g-3p	99.6	0.998	-3.330	23.337

5.3.9. Melt Curve Analysis

Besides, a melt curve analysis was conducted at the end of each qPCR to check whether only a single and specific PCR product was produced, based on the dissociation characteristics of the PCR amplicons. All the individual reactions showed single melt peaks for each reference or target genes (**Figure 5.33**).

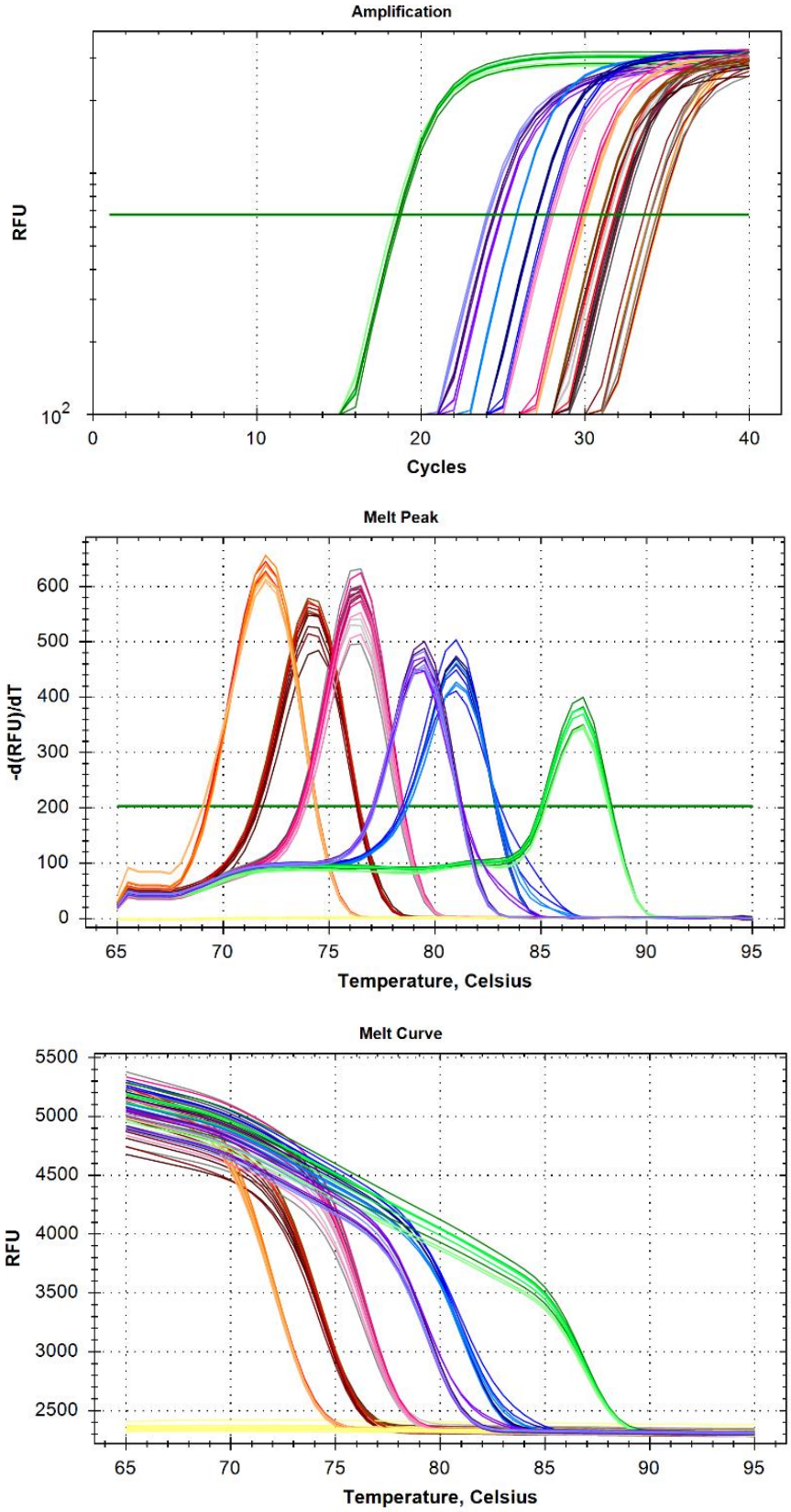


Figure 5.33. Log amplification plot (a), melt curve (b) and melt peak (c) of one of the qPCR amplification experiments. Each color indicated an mRNA sample being amplified. Distinct melt peaks showed the specificity of RT-qPCR products.

5.3.10. Validation of RNA-Seq data by RT-qPCR

To validate the expression profiles of differentially expressed genes (DEGs) and differentially expressed miRNAs (DEMs) from RNA-seq data (**Figure 5.34)**, an RT-qPCR was performed. A total of 3 DEMs and 8 DEGs, which are their potential corresponding targets, were selected to validate the expression profiles from RNA-Seq over prolonged callus culture duration (**Figure 5.35-5.37**). The results showed that, most of the miRNAs were negatively correlated with the expression level of their respective targets (**Figure 5.35-5.37**), as predicted by RNA-Seq (**Figure 5.34**). In particular, *vpl-miR395g-3p* (**Figure 5.35a**) and its corresponding potential target genes (**Figure 5.35b-f**) have negative correlation in their expression over time. When *vpl-miR395g* was up-regulated at 6M, its corresponding gene was down-regulated, and opposite was observed at 12M. It corresponded to the RNA-seq results (**Figure 5.34**). Similar was observed for *vpl-miR166r* (**Figure 5.36a**) and its corresponding potential target genes *vpl_contig1034* and *vpl_contig86799* (**Figure 5.36b and c**). The only exception was with *vpl-miR8155* and its target. In contrast to the RNA-seq results, the RT-qPCR results of miRNA *vpl-miR8155* showed up-regulation at 6M, instead of down-regulation (**Figure 5.37a**).

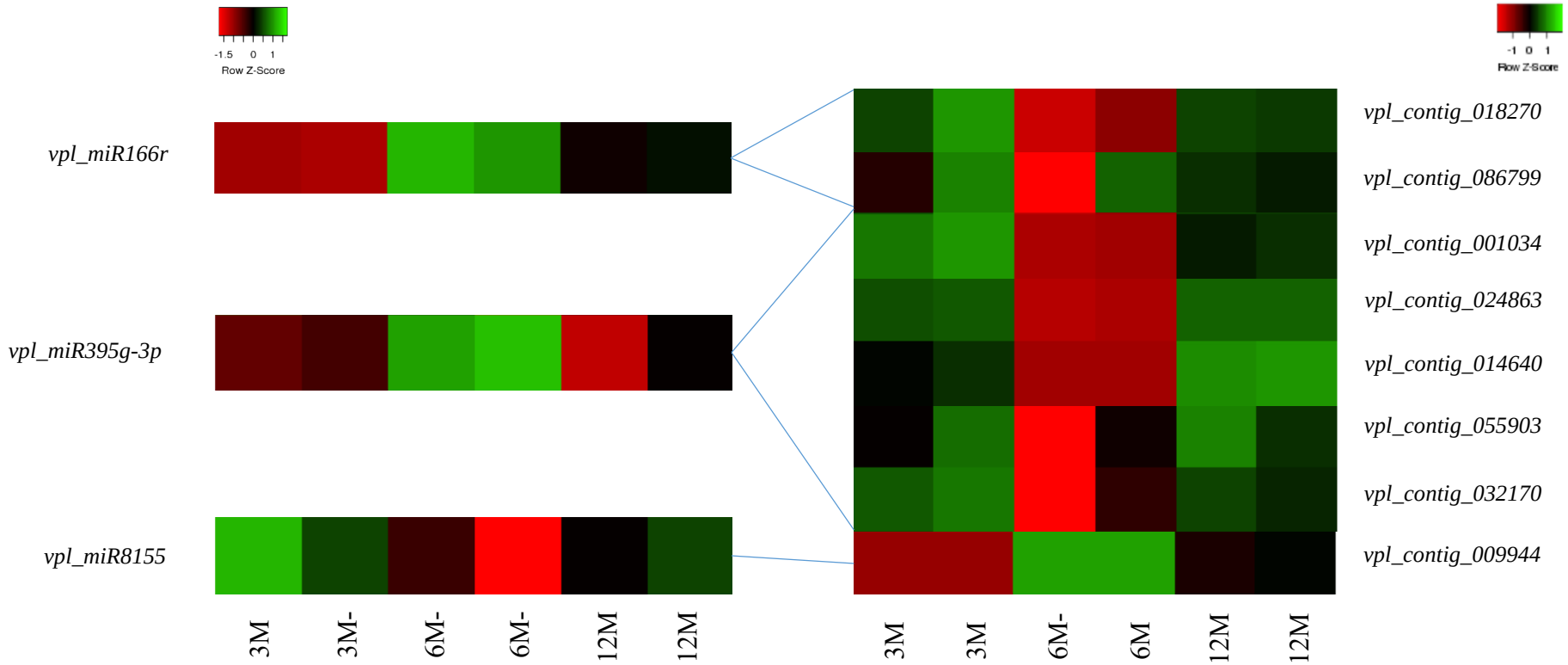


Figure 5.34. The expression levels of predicted miRNAs and mRNAs. Hierarchical heat map displaying transformed expression values (log_2FPKM) for 3 and 8 significant differentially expressed miRNAs (**left**) and mRNAs (**right**) respectively. **Red** shows up-regulation while **green** shows down-regulation of miRNAs or mRNAs. **Blue** lines between miRNA and mRNA heat maps indicates respective potential target gene(s) of each miRNA.

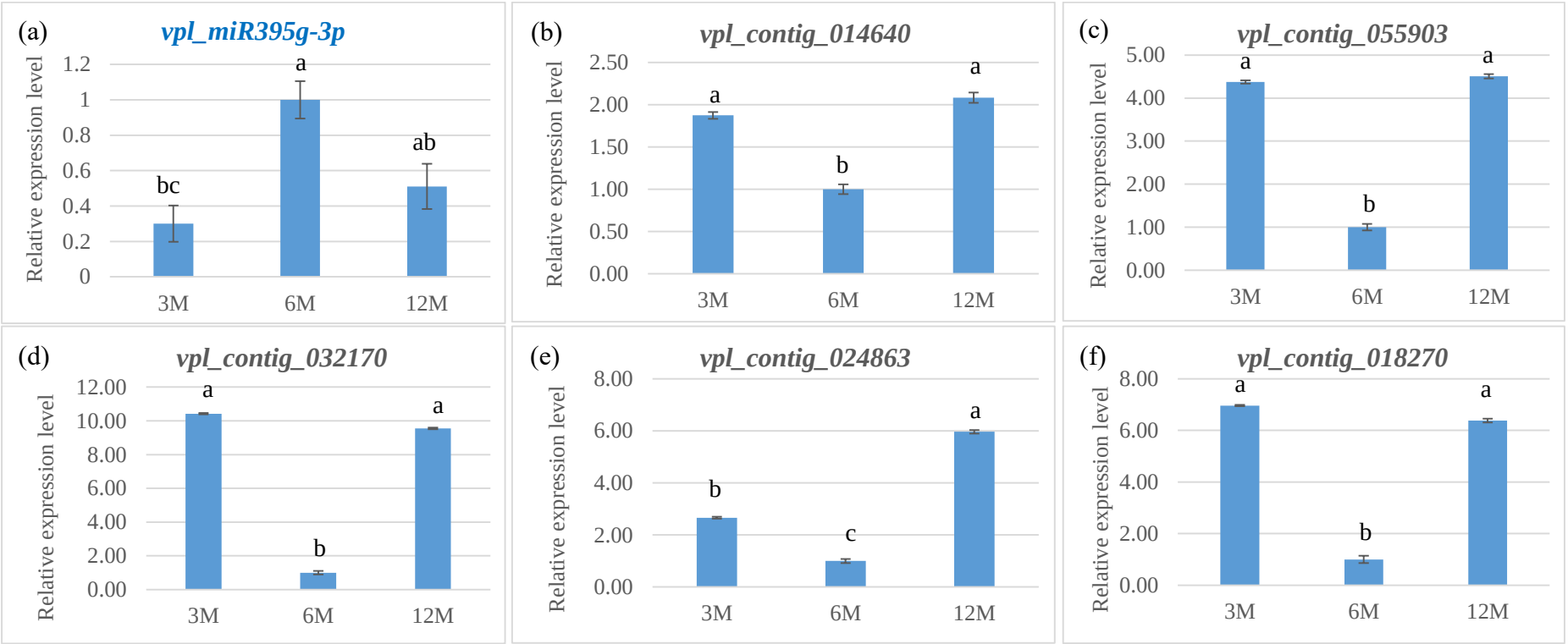


Figure 5.35. RT-qPCR results of differentially expressed miRNAs and mRNAs of vanilla callus cultures under prolonged culture durations of 3M, 6M and 12M: Expression levels of *vpl_miR395g-3p* (a) and its potential targets, *vpl_contig_014640* (b), *vpl_contig_055903* (c), *vpl_contig_032170* (d), *vpl_contig_023863* (e) and *vpl_contig_018270* (f). Bar represents the mean of three biological replicates. Expression levels for miRNAs were normalized to *vpl_miR159n* and *vpl_miR319g-3p*. Expression levels for mRNAs were normalized to *EF1*, *TUB* and *18S*. Error bars indicate standard deviations (n=3). Bars with the same letters are not significantly different at p=0.05 (n=3).

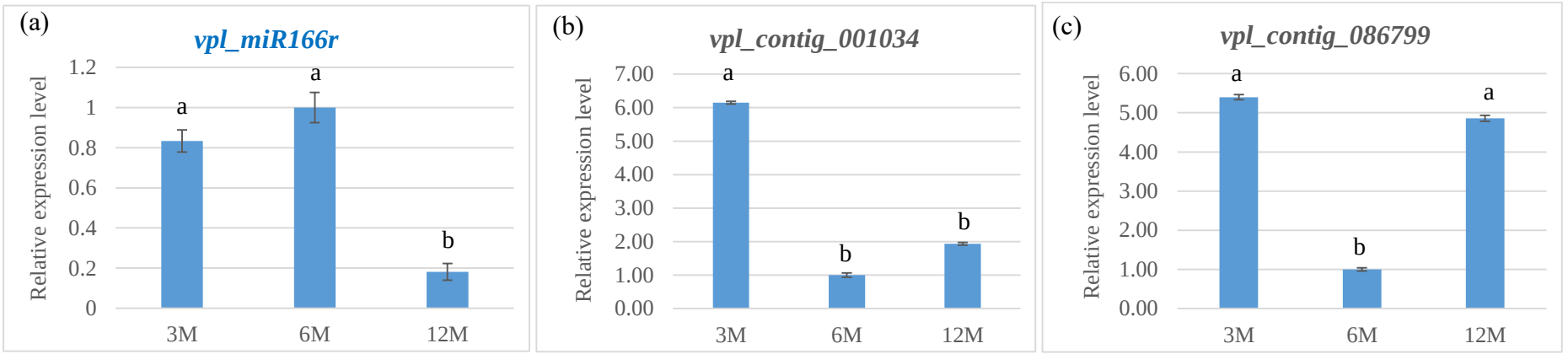


Figure 5.36. RT-qPCR results of differentially expressed miRNAs and mRNAs of vanilla callus cultures under prolonged culture durations of 3M, 6M and 12M: Expression levels of *vpl_miR166r* (a) and its potential targets, *vpl_contig_001034* (b), *vpl_contig_086799* (c). Bar represents the mean of three biological replicates. Expression levels for miRNAs were normalized to *vpl_miR159n* and *vpl_miR319g-3p*. Expression levels for mRNAs were normalized to *EF1*, *TUB* and *18S*. Error bars indicate standard deviations (n=3). Bars with the same letters are not significantly different at p=0.05 (n=3).

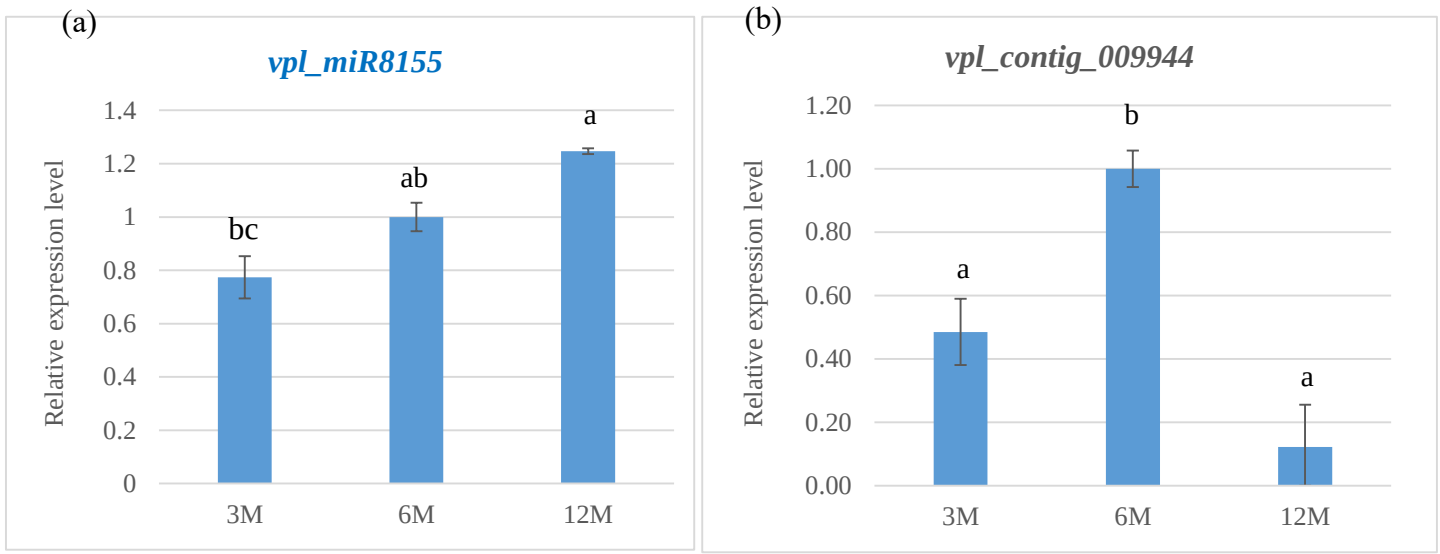


Figure 5.37. RT-qPCR results of differentially expressed miRNAs and mRNAs of vanilla callus cultures under prolonged culture durations of 3M, 6M and 12M: Expression levels of *vpl_miR8155* (a) and its potential targets, *vpl_contig_009944* (b). Bar represents the mean of three biological replicates. Expression levels for miRNAs were normalized to *vpl_miR159n* and *vpl_miR319g-3p*. Expression levels for mRNAs were normalized to *EF1*, *TUB* and *18S*. Error bars indicate standard deviations (n=3). Bars with the same letters are not significantly different at p=0.05 (n=3).

The qPCR values (**Table 5.25**), measured in normalized CT values, were in negative correlation with expression level; while the RNA-seq values (**Table 5.25**), measured in normalized expression values, were in positive correlation with expression level. The closer the correlation score to 1, the more positive the qPCR data correlates to the RNA-seq output. In overall, with the exception of *vpl-miR8155*, the RT-qPCR results were in positive correlation with RNA-Seq data (Pearson correlation coefficient ranged from 0.7122 to 0.9996) (**Table 5.25**).

Table 5.25. Correlation analysis of RNA-seq and RT-qPCR results.

Contig_ID		3M[2]	6M[2]	12M[2]	Correlation score[1]
vpl_contig_001034	qPCR	5.37	7.99	7.03	0.9996
	RNASeq	21.15	1.63	9.14	
vpl_contig_009944	qPCR	3.22	2.18	3.03	0.9912
	RNASeq	44.20	183.53	82.11	
vpl_contig_014640	qPCR	7.97	8.88	7.82	0.7603
	RNASeq	4.55	0.00	11.48	
vpl_contig_018270	qPCR	6.93	9.73	7.06	0.8999
	RNASeq	32.31	5.29	22.59	
vpl_contig_024863	qPCR	7.83	9.24	6.67	0.8637
	RNASeq	4.95	1.01	5.38	
vpl_contig_032170	qPCR	7.38	10.76	7.51	0.8649
	RNASeq	4.22	0.06	2.55	
vpl_contig_055903	qPCR	8.70	10.83	8.65	0.9746
	RNASeq	1.45	0.13	1.76	
vpl_contig_086799	qPCR	8.82	11.25	8.97	0.9573
	RNASeq	0.81	0.03	0.60	
miRNA_ID		**3M**	**6M**	**12M**	**Correlation score[1]**
vpl-miR166r	qPCR	2.67	2.40	2.46	0.7122
	RNASeq	161.50	1180.00	440.00	
vpl-miR395g-3p	qPCR	10.75	9.01	9.98	0.8123
	RNASeq	37.00	131.00	38.00	
vpl-miR8155	qPCR	3.13	2.76	2.44	0.3884
	RNASeq	144.00	23.50	72.50	

[1]Pearson correlation coefficients were calculated by GraphPad Prism 5.0 software.
[2]The qPCR values were measured in normalized CT value, while the RNASeq values were measured in normalized expression values.

5.4. Discussion

Somaclonal variation (SV) is an important, yet not well explained phenomenon in *in vitro* culture systems of plants. In *V. planifolia* tissue culture system, the indirect (callus-mediated) nodal regeneration generates somaclonal variants at genetic and epigenetic levels, and the prolonged callus culture duration resulted in increased polyploidy populations in vanilla calli. It was hypothesized that *in vitro* culture stress were regulated by epigenetic modifications of gene expression, which could lead to the occurrence of SV at morphological and genetic levels. In this study, we performed an integrated mRNA and miRNA profiling of callus cultures at prolonged culture durations of 3, 6 and 12 months. Our results provided some new insights into the response of *V. planifolia* callus culture at transcriptomic levels.

5.4.1. BLAST and functional gene annotation

According to the top-hit species statistics of BLAST (**Figure 5.5**) against the primary database in NCBI, contigs with DEGs in vanilla prolonged callus cultures had the highest annotation to the databases of species such as *Elaeis guineensis, Phoenix dactylifera, Musa acuminate*, *Ananas comosus* and *Vitis vinifera*. According to literature, these monocot plant species were also commercially important crop plants with the high research reports of SV (Jaya *et al.*, 2002; Pérez *et al.*, 2009, 2012; Kuksova *et al.*, 1997; Rival *et al.*, 2013; Sanputawong and Te-chato, 2011, Lucia *et al.*, 2011; Tang *et al.*, 2000; Hwang, 2002, Ghag *et al.*, 2014; Giménez *et al.*, 2001; Hwang and Ko, 1992, 2004; Martin *et al.*, 2006; Lee *et al.*, 2011; Krishna *et al.*, 2016; Hadrami *et al.*, 2011). Besides, these plant species had common genome features of high whole genome duplication (WGD) rate in their evolutionary history (Jiao *et al.*, 2014). This suggested that the occurrence of SV in a plant species could be related to their genome evolution history, for example, the whole genome duplication (WGD) rate and retain of WGD through the evolutionary history.

This phenomenon could be related to the domestication aspect of agriculture, of which the genome consequences of polyploidy (Pecinka *et al.*,

2011; Renny-Byfield and Wendel, 2014; Paterson, 2005) might have increased the adaptive plasticity of those plants, enabling successful domestication (Salman-Minkov *et al.*, 2016); while the stressful condition of tissue culture resembled such selection process by triggering abiotic stress responses. The GO and KEGG pathway analysis has revealed that a cascade of stress response genes was enriched in vanilla prolonged callus cultures (**Figure 5.6**-**5.8**).

Among these genes, we further classified and focused on discussing genes in five different genes families potentially related to SV, namely the DNA methylation gene families, retrotransposon gene families, histone modification gene families, transcription factor gene families and finally the microRNA families. Global mRNA (**Section 5.3.2.2.2**) and miRNA (**Section 5.3.3.3**) profiles showed all significantly expressed genes and miRNAs expressed in the 3M, 6M and 12M vanilla calli; while time course DGE and DME analysis of the mRNA (**Section 5.3.2.2.4**) and miRNA (**Section 5.3.3.4**) discussed the expression patterns of most significant DEGs and DEMs expressed in each gene family over the time course.

5.4.2. DNA methylation profile

From KEGG pathway analysis, a number of genes expressed in vanilla callus were inter-connected into a network involved in the sulfur metabolism, propanoate metabolism and methionine salvage parts of the cysteine and methionine pathway (**Figure 5.7**). This agreed with other studies which has highlighted the importance of DNA methylation in callus development and somaclonal variation, as reported in the case of flower mantling defect in oil palm (Jaligot, 2000) and variation in DNA methylation patterns in grapevine somaclones (*Vitis vinifera* L.) (Schellenbaum *et al.*, 2008). This is probably because DNA methylation in tissue culture is involved in the regulation of various genes transcription, such as the activation of retrotransposon and microRNAs (Saze *et al.*, 2012), such as found in oil palm and wheat (Cantu *et al.*, 2010).

A number of methylation genes, namely SAM, SAH, MET, MTH/SAHN and DNMT were up-regulated in 6M calli (**Figure 5.16**). These

genes were involved in the maintenance of methylation homeostasis in eukaryotes (Mull *et al.*, 2006). Loss of SAH function were reported to cause pleiotropic effects including developmental abnormalities and reduced cytosine methylation (Mull *et al.*, 2006). This might explain why SAH and DNMT were up-regulated at the same time.

5.4.3. Retrotransposon profile

A global preview of potential genes encoding for retrotransposable elements that were expressed in callus transcriptome of vanilla (**Figure 5.14**) has shown that majority of them mostly belonged to the LTR retrotransposons family, in particularly the *copia*-like (20.25%) and *gypsy*-like (17.31%) LTR retrotransposable elements. This agreed with the findings from other studies that LTR retrotransposons (LTR-RTs) are activated in callus cultures, contributing to various mutations at gene level in various plants species (Grandbastien, 2015). For example, Hirochika (1993) reported high level of transposition of *Tto1* in tobacco cultured cells. Activation of retrotransposons (*Tos17* and *Tos19*) was involved in causing tissue culture induced mutations (Hirochika *et al.*, 1996; Picault *et al.*, 2009). In barley, *BARE-1* retrotransposon plays a role in mutation in tissue culture (Evrensel *et al.*, 2011; Jääskeläinen *et al.*¸ 2013). Other similar examples included callus culture-active LTR-RT *Athila* in *Arabidopsis thaliana* (Pischke *et al.*, 2006), *Rtsp-1* in *Ipomoeae batatas* (180), *Hopscotch-like* LTR-RT in *Saccharum offinarum* (Araujo *et al.*, 2005), *CCR* in *Oryza sativa* (Neumann *et al.* 2007), *MERE1* from *Medicago truncatula* (Rakocevic *et al.*, 2009), *ONSEN/Atcopia78* in *Arabidopsis thaliana* (Matsunaga *et al.*, 2012), *Ac* in *Zea mays* (Brettell and Dennis, 1991), *Lullaby* in *Oryza sativa* (Picault *et al.*, 2009), *Tnt1* in *Medicago truncatula* (d'Erfuth *et al.*¸ 2004) and *Rtsp-1* in *Ipomeda batatas* (Tahara *et al.*, 2004). Besides, in a preliminary study of vanilla callus transcriptome using X-species microarray (Chai, 2014), it was shown that *matR*, an RT gene usually indicative of a mobile element such as retrotransposon or retrovirus, were up-regulated in vanilla callus as compared with shoot. Interestingly, our study also shown that these RTs, which are inclusive of *gypsy*-like LTR, *copia*-like LTR and LINE non-LTR, were highly

expressed in 3M and 12M vanilla calli (**Figure 5.18**). This suggests that the activation of these RTs might occur twice independently, instead of consecutively throughout the culture duration. First activation might have occurred during early callus (3M) phase and second activation might have happened during prolonged callus culture duration phase (12M). This might also explain why most of the DEGs were expressed in cluster 2 and 3 (**Table 5.10**). We observed that transposition of retrotransposable elements seemed to play a major role in SV of *in vitro* cultured calli, which reflects a survival strategy and stress-induced generators of genomic diversity as proposed by McClintock (Wessler, 1996).

5.4.4. Histone profile

Thirdly, members of histone modification family were also found to be expressed in vanilla callus cultures (**Table 5.11**). In this study, the histone encoding genes where classified into gene families of three biological function groups, histone deacetylases (HDACs), histone cores and histone acetyltransferases (HATs). This agreed with the evidence from other studies that differential expression of chromatin remodeling genes and histone methylation genes happens during tissue culture, which leads to disruption in the methylation pathway in a non-specific manner and hypo/hypermethylation patterns of DNA induced in tissue culture (Krishna *et al.*, 2016).

5.4.4.1. HDACs: HDAC2, HDAC19, SAP18

Among putative genes coding for different histone modification proteins (**Table 5.11**), the putative genes encoding for histone deacetylases (HDACs) had been found to be related to the occurrence of SV in numerous crop species like the oil palm (*Elaeis guineensis*) and pineapples (*Ananas comosus*). In oil palm, HDACs suppressed the gene expression and contributed to the development of somaclonal variants - higher HDACs was found in the tissue culture-derived mantled leaf samples than in phenotypically normal leaf samples (*Yaacob et al.*, 2013a). In pineapple (*Ananas comosus* var. MD2) somaclonal variants, the

histone deacetylase (HDAC) enzyme activities were also higher in dwarf phenotypes than in non-dwarf phenotypes (Halim *et al.*, 2017). In vanilla, three putative genes encoding for the HDAC enzymes, namely HDAC2, HDAC19 and SAP18 were found to be expressed in vanilla calli and in higher levels during the early callus phases (3 and 6-month-old calli).

5.4.1.2. Histone Cores: H2A, H3, H4 and H3K4

The second group of histone protein-encoding genes expressed in vanilla callus were the putative genes encoding for histone cores H2A, H3, H4 and H3K4. This group of histone protein-encoding genes could be involved in epigenetic processes by powering transcription and growth induction in vanilla callus cells. In potato cell culture, DNA cytosine methylation and core histone multi-acetylation were observed after subculture (Law and Suttle, 2005).

5.4.1.3. HATs: ELP and GCN5

The third group of histone protein-encoding genes were the genes encoding for histone acetyltransferases (HATs). In our study, we identified two HATs of interest, which are the ELP and GCN5. The involvement of these genes in SV could be explained by the role of a fine-tuned acetylation-deactylation balance during transcription *in vivo* in order for the development of normal phenotypes (Wittschieben *et al.*, 2000). Elp3 are involved in growth adaptation, gene activation and temperature sensitivity (Wittschieben *et al.*, 1999, 2000). More recently, Kornet and Scheres (2009) reported that the histone acetyltransferase GCN5 is also essential for Arabidopsis root stem cell niche maintenance, playing an important role in shaping a developmentally instructive gradient in the root which is formed by the PLETHORA (PLT) stem cell transcription factors.

5.4.5. Stress regulating TFs might be key regulators of the RTs expression

Activation of transcription factors may imply the activation of their target genes, and hence, identifying master regulators of key processes and analysing the underlying interactions and regulatory network (Neelakandan and Wang, 2012; Zheng *et al.*, 2009) may help us to understand and predict *in vitro* development phenomena (Neelakandan and Wang, 2012). In our studies, we identified a few important transcription factors, including the WRKY transcription factor, ethylene-responsive transcription factor, transcription factor bHLH, Myb family transcription factor, heat stress transcription factor, AP2-like ethylene-responsive transcription factor, GATA transcription factor and others (**Table 5.12**). These transcription factor are regulators for plant biotic and abiotic stresses immunity response (Pandey and Somssich, 2009; Ambawat *et al.*, 2013; Naika *et al.*, 2013). Similar were also reported in *Arabidopsis*, of which the *ONSEN* LTR-copia type RT acquired activation through heat stress transcription factor, HSFA2 under heat stress (Cavrak *et al.*, 2014). In our study, the expression patterns of most of the TFs (**Figure 5.17**) were totally opposite of the expression patterns of RTs (**Figure 5.18**). This further supported the hypothesis that activation of stress regulating TFs might be key regulators of the RTs expression.

5.4.6. miR166r, miR395g and miR8155 as potential suppressor for DNA methylation and RT genes

There is still little known about the response of microRNAs to tissue culture conditions. It has been proposed that *in vitro* callus cells generate a unique set of microRNAs (Rodriguez-Enriquez *et al.*, 2011). In vanilla callus cultures, three microRNAs, namely vpl-miR166r, vpl_miR395g and vpl_miR8155, were also identified through its potential binding site (**Section 5.2.4.2**) and correlation in gene expressions (**Figure 5.34**), that potentially

regulates eight annotatable genes, which fell into the gene families of DNA methylation and retrotransposable elements. The potential target binding sites of miRNA to RT-encoding genes is found to mostly bound at the upstream of RNAase H-like domain (**Figure 5.25**). This has supported one of the SV hypothesis that the epigenetic variations could involve activation and suppressions of genes such as retrotransposons (Azman *et al.*, 2014), as caused by post-transcription factors and affected by the action of small RNAs such as microRNA and piwiRNAs (Rodriguez-Enriquez *et al.*, 2011).

5.4.7. Selection of suitable reference candidates for RT-qPCR

Proper set-up of qPCR experiments is of critical importance for the success of the study. A proper qPCR experiment requires an in-depth understanding of the qPCR technique, objectives of study, and careful considerations of factors such as sample size, replicates and the role of normalization. In this study, three biological replicates per time course (3M, 6M and 12M) with two technical replicates each were carried out for each DGE or DEM experiment, with NTC and NRT included in each experiments as negative controls. One of the important pre-requisites to DGE or DGM analysis is the selection of optimal reference gene or combination of references for normalization of DGE or DGM analysis.

Normalization with reference genes has been the guideline for reliable RT-qPCR experiment (Bustin *et al.* 2009). It is important that the stability of reference genes is empirically validated for each study prior to any RT-qPCR analysis (Bustin *et al.* 2009). No single reference gene can be recommended for use across all tissue and cell types, even among the most commonly used housekeeping genes (De Kok *et al.*, 2005). In present work, the expression levels of five frequently used housekeeping genes, *ACT, EF1, UBQ, 18S* and *TUB* were estimated using RT-qPCR in our experimental samples. Based on the five statistical algorithms used (**Table 5.22**), *18S, EF1* and *TUB* were estimated to be the most stably expressed reference genes, while the least stable reference

genes were *UBQ* followed by *ACT*. In other recent transcriptomic studies in *V. planifolia*, *ACT* and *UBQ* were selected as the most stably expressed reference gene in RT-qPCR validation of RNA-seq data during early seed development (Rao *et al.*, 2014); while *ACT* and *EF1* were selected as the most stable reference genes during pod development of vanilla (Fock-Bastide *et al.*, 2014); which differed from our results. This outcome has highlighted the importance of evaluating the expression stability of classical housekeeping genes before selecting them as an internal control for RT-qPCR.

One of the potential reasons that *UBQ* was the most unstable reference gene candidate could be due to the influence of hormonal treatments and abiotic stresses on gene expressions during vanilla callus culture. Jain and co-workers (2006) reported that both *TUB* and *ACT* were unstable in *Oryza sativa* under hormone treatment, salt and drought stress. Similarly, unstable gene expression of *UBQ* under the influence of abscisic acid, methyl jasmonate and salicylic acid in *Saccharum sp.* (Ling *et al.* 2014); and the unstable gene expression of *ACT* in *Nicotiana tabacum* under salt and drought stresses (Guo *et al.*, 2014) were also reported. This could be related to the primary role of ubiquitin in hormone synthesis, hormone signalling cascades and plant defence mechanisms (Dreher and Callis, 2007; Santner and Estelle, 2011), in particularly auxin. Magyar and co-workers (2005) reported that in *Arabidopsis* cell culture, ubiquitination of related proteins played a role in hormone signalling of exogenous auxin, 1-naphthalene-acetic-acid (1-NAA), which was also the main auxin used in the callus culture of vanilla in this study.

There has not been published work on miRNA reference candidate to date. As a new approach in *V. planifolia*, we have taken advantage of miRNA-seq data in previous chapter to design potential miRNA reference candidate. Therefore, miRNA reference candidates were designed from RNA-seq data in previous chapter. A total of 86 miRNA candidates from RNA-seq data with moderate to high expression (**Table 5.15**) in vanilla callus cultures were subjected to statistical algorithm screening (**Appendix J**) to select few potential stable candidates for RT-qPCR reference stability test. An additional SnoRNA candidate, *SnoR6* were included from published literature. Our results estimated

that *miR159*, *miR168* and *miR319* outperformed *SnoR*6 in terms of expression stability (**Table 5.22**). These miRNAs are estimated to be better options of reference miRNAs for this experiment than *SnoR6*, as although many experiments have used SnoRNAs as reference for miRNA studies (Ferdous *et al.*, 2015), it may not be the optimal option for our study. In cases where neither of the expressed miRNAs can be used as stable miRNAs, non-coding RNAs with less than 100 nucleotides long (which in contrast to the ~18-25 nucleotides for miRNAs), such as *SnoR6* or *SnoR14* could be used as reference (Luo *et al.*, 2018). However, the endogenous references are more recommended to be stably expressed miRNAs, rather than longer RNA species such as snoRNAs and snRNAs (Luo *et al.*, 2018). This was due to the potential different behaviors of shorter miRNAs during the RNA extraction and RT reaction, as compared to the longer transcripts (Luo *et al.*, 2018; Vandesompele *et al.*, 2002; Pantaleo *et al.*¸ 2016), which could introduce bias to the results (Gee *et al.*, 2011; Kou *et al.*, 2012).

The miR159, miR168 and miR319 families were found to be involved in the regulation of various genes and were expressed under different abiotic stress conditions (**Table 5.26**). However, there were also different sequence and length variants of miRNAs (isomiRs), such as different isomiRs of miR159 (*vpl-miR159l*-3p), miR168a-3p (*vpl-miR168d*-3p) and miR319 (*vpl-miR319a* and *vpl-miR319h*-3p) (**Appendix J**), which had different expressions than our reference miRNA candidates, *vpl-miR159n*, *vpl-miR168a*-3p and *vpl-miR319g-3p*. This finding has also once again emphasized the importance of pre-selecting references in each miRNA expression analysis. This is especially for the case of *in vitro* culture miRNA studies, where the expressions of miRNAs were often influenced by the condition such as abiotic stresses and different isomiR could have altered target capacity or specificity (Cammaerts *et al.*, 2015) and thus different expressions. Our results thus strongly support the importance of taking advantage of high-throughput transcriptomic data for the selection of proper reference in RT-qPCR experiments (Pombo *et al.*, 2017).

Table 5.26. A summary of role of miRNA candidates (miR159, miR168 and miR319) in plants, compiled from literature review (Koroban *et al.*, 2016; Sunkar, 2010).

MiR Family	Target regulation	Abiotic stress	Species	References
miR159	MYB transcription factors (MYB33 and MYB101)	ABA treatment, drought, salt and cold stress	*A. thaliana*	Zhang *et al.* (2014); Reyes and Chua (2007); Liu *et al.* (2008); Zhou *et al.* (2008)
		ABA treatment, drought stress	*Phaseolus vulgaris*	Arenas-Huertero *et al.* (2009)
		Hypoxia stress (long term exposure)	Maize (seedlings)	Zhang *et al.* (2008)
miR168	ARGONAUTE and MAPK	Salt and drought stress	*A. thaliana*	Giacomelli *et al.* (2012); Liu *et al.* (2008)
		Cold stress (*miR*168a, b)	*P. trichocarpa*	Lu *et al.* (2008)
		UV-B radiation	*P. tremula*	Jia *et al.* (2009)
miR319	TCP transcription factor	Nitrogen deficiency, drought, salt and cold stress	*A. thaliana*	Paul *et al.* (2015); Sunkar and Zhu (2004); Liu *et al.* (2008); Zhou *et al.* (2008)
		Drought stress	*Oryza sativa*	Jagadeeswaran *et al.* (2014)

5.4.8. RT-qPCR validation of RNA-seq data

Transcriptomic study is an approach that allows quick overview of the genes expressed that could be leading to epigenetic events that occur in the *in vitro*-cultured vanilla cells. However, the robustness of a tool needs to be further validated with a second tool, therefore, several genes that were found to be expressed were further validated using RT-qPCR. Overall, the current study found that RT-qPCR was sensitive in validating the results of RNA-seq, as shown by Pearson correlation coefficients ranged from 0.7122 to 0.9996 (**Table 5.25**). Besides, the expression of miRNA and mRNA showed negative correlations as predicted from the output of bioinformatics analysis, i.e. when miRNA was up-regulated, the potential mRNA target was down-regulated. An exception was observed with one of the miRNAs, *vpl-miR8155*, which belonged to the low to moderate expression frequency miRNAs (**Table 5.15**). This could be one of the technical limitations to RT-qPCR of miRNAs, which was the less accurate quantifications of miRNAs with low expressions. On the other hand, it was observed that miRNA with expression frequencies above moderately low (**Table 5.15**) had higher correlation score (**Table 5.25**), such as *vpl-miR395g* (moderately low to moderate; Pearson correlation coefficient 0.8123) and *vpl-miR166r* (moderate to moderately high, Pearson correlation coefficient 0.7122). This suggested that the accuracy of RT-qPCR analysis of miRNAs could be affected by the low expression frequency of the miRNAs. Similarly, in other studies, a lower limit of detection of 10 copies in qPCR of miRNA were reported (Schmittgen and co-workers, 2008).

5.5. Conclusion and future prospects

An overall view of global transcriptome expressed in vanilla prolonged callus culture has been studied using *de novo* RNA-seq and bioinformatics analysis. Activation of complex abiotic stress responsive genes and epigenetic reprogramming involving DNA methylation, retrotransposons activation, histone modifications, transcription factors and microRNA gene regulation, were found to be expressed in vanilla calli over prolonged culture duration.

Selection of candidate mRNA and miRNA references was established. Based on the analysis of five different statistical algorithms (geNorm, NormFinder, Bestkeeper, delta delta CT method and RefFinder), *18S*, *EF1* and *TUB* were recommended as the most appropriate reference genes while *vpl-miR159*, *vpl-miR168* and *vpl-miR319* were recommended as the most stable reference miRNAs, for the accurate normalization of gene and miRNA expression data from RT-qPCR of vanilla callus samples. Correlation of expression data between two approaches (RNA-Seq and RT-qPCR) has also strengthened the confidence of the data output from RNA-Seq. The outcome of this study hoped to provide a reference to facilitate the future mRNA and miRNA profile analysis of genes associated with SV and *in vitro* culture of vanilla. RT-qPCR workflow in this study could also serve as a workflow for other validation analysis in vanilla miRNA or miRNA profile studies.

CHAPTER VI

GENERAL DISCUSSION AND PERSPECTIVES

Chapter VI

General discussion and perspectives

6.1. Summary of findings

The main aim of this study was to investigate the occurrence of somaclonal variations in tissue culture smples of *Vanilla planifolia* (*V. planifolia*) at the morphological, genetic, epigenetic, cytological and transcriptomic levels. Several specific objectives of this study was achieved: (i) Aseptic culture of vanilla was successfully established using modified surface sterilization protocol. Two types of regeneration pathways were generated for vanilla using the nodes as explants, which included the direct and indirect regeneration systems (Chapter 2). The indirect regeneration system produced significantly higher number of regenerated shoots (at an average of 10.5 shoots per node in 6 months), as compared to the direct regeneration systems (which produced an average of 8.0 shoots per node in 6 months). (ii) Morphological, genetic and epigenetic marker analysis using ISSR and MSAP marker systems (Chapter 3) have shown that the direct nodal regeneration retained genetic fidelity of vanilla plants, and thus is more suitable for clonal mass propagation of elite vanilla materials. In contrast to the direct regeneration, the indirect nodal regeneration resulted in genetic heterogeneity of vanilla regenerants. (iii) Ploidy analysis of different vanilla vegetative tissues using flow cytometry (Chapter 4) demonstrated that *V. planifolia* is a polysomatic plant, with systemic endopolyploidy cells in different vegetative tissues, which could be the source of tissue culture variations. In addition, poidy analysis of callus culture and prolonged culture period (Chapter 4) showed that the pre-existing polyploid populations were amplified over time. The prolonged culture of vanilla correlated to an increase in polyploid populations and mean genome sizes over time, suggesting the role of prolonged culture stress in increasing SV levels of vanilla. (v) Transcriptomic and micronomic analysis of prolonged vanilla callus culture using RNA-seq technology has revealed complex expression patterns and cross-talks between

several epigenetic gene players, including microRNAs, DNA methylation and retrotransposable elements (Chapter 5). Amongst these genes, 3 miRNAs and 8 mRNAs have been validated by RT-qPCR.

6.2. Implication of findings

In the rapid development of biological science, plant tissue culture (PTC) has taken lead as one of the most promising areas of biotechnological application tools for today and tomorrow agriculture (Suman, 2017). PTC offers tremendous potential for fundamental research and for commercial applications such as clonal multiplication, genetic engineering, plant breeding and crop improvement, production of valuable metabolites or pharmaceutically important compounds and germplasm conservation of rare species (Neelakandan and Wang, 2012; Suman, 2017). PTC relies on the fact that many plant cells have the ability to regenerate into a whole plant, termed totipotency (Hussain *et al.*, 2012). Depending on the plant species, an explant may be taken from many parts of the plant, namely the shoot tip, leaf, axillary bud, stem or root tissues of the source mother plant; and a small piece of plant (the explant) could undergo rapid multiplications in relatively short periods of time under controlled aseptic environment (Hussain *et al.*, 2012). Establishment of *de novo* developmental cell fate *in vitro* is governed by various factors such as the type or concentrations of plant growth regulators (PGRs), temperature, quality and intensity of light, size and source of explants, genetic make-up and stress (Smith, 2000). A good balance of all these factors defines the success for PTC of a particular plant species. Yet, problems could also arise from the manipulation of these factors in PTC, which can cause the occurrence of variations in culture, a phenomenon termed as somaclonal variations (SV) (Neelakandan and Wang, 2012).

Theoretically, PTC involves only mitotic cell division which is an asexual process that should produce clonal, genetically uniform plants (Larkin and Scowcroft, 1981). Nevertheless, in practical situations, the occurrence of uncontrolled and random spontaneous SV could occur (Bairu and Kane, 2011). Some 50 years ago, SV was first reported by Braun (1959). Followed, it has been

a continuous report of SV cases in PTC of many plant species (Bairu *et al.*, 2011), including vanilla. In vanilla, the observation of SV occurrence has been briefly reported but not well investigated (Minoo *et al.*, 2002). Palama and co-workers (2010) suggested the potential occurrence of SV in vanilla through callus culture, however, there has been only a few studies reported on micropropagation via indirect organogenesis in *V. planifolia,* where callusing rate was reported to be poor (Gu *et al.* 1987, Kononowicz and Janick 1984; Philip and Nainar 1986). Besides, Tan *et al.* (2011a) reported that 35% and 10% of callus was recorded in nodal and leaf explants respectively when cultured on MS medium containing 2.0 mg/L 1-naphthaleneacetic acid (NAA) and 1.0 mg/L 6-Benzylaminopurine (BAP). Unfortunately, evaluation on SV has not been carried out on its regenerants. Our study investigated the SV aspect of vanilla through the morphological, genetic, epigenetic and cytological screening of vanilla, through the analysis of vanilla regenerants from both direct and indirect nodal regenerations.

Somaclonal variation (SV) has been a topic of commercial interest and agricultural importance, playing major roles in mass propagation of elite plant materials as well as *in vitro* crop improvement. In mass propagation of elite plant materials, the occurrence of SV could cause great economic losses; while on the other hand, it could be a source of genetic variations to be utilized as an effective tool for improving crop variety *in vitro*. Vanilla being a crop with poor genetic variability, efforts should be made to induce more variation to identify new traits such as disease resistance or tolerance (Minoo, 2002). SV is actually a double-edged sword whereby its presence in micropropagation programmes is inimical, while it can be gainfully exploited as effective crop improvement tool to create stable variations such as disease resistance, where other methods failed or are cumbersome to achieve (Sahijram, 2015). Commercially useful crop varieties have been successfully generated from *in vitro* selection of somaclonal variants in numerous plants species, such as apple (*Malus sp.*), banana (*Musa sp.*), chilli (*Capsicum sp.*) and many others for superior horticultural traits such as disease resistance and increased crop yield (Krishna *et al.*, 2016). In both applications, effective screenings and basic understanding of SV is of importance. In our study, our utilized callus as a study material for SV mechanisms.

With the advances in molecular technology and NGS, there is a great potential for researchers to unravel the reasons and mechanisms underlying these abnormalities. In the past, molecular screenings and studies on SV have majorly focused on the use of random fingerprinting DNA markers, which has the advantage of simplicity and do not need the prior knowledge of genome sequences (Bairu *et al.*, 2011). However, these methods are based on random screening and had the limitations of false negative in the detection of SV. And often, they do not reveal the cause and mechanisms behind SV, especially at epigenetic level (Jaligot *et al.*, 2000). Recent advances in next generation technologies, such as the next generation transcriptome (mRNA-seq), smRNA transcriptome (smRNA-seq) and methylome sequencing (methylC-seq), have allowed the high-throughput global profiling of samples that could correlate expressions between these epigenetic players. By directly sequencing the transcriptome, methylome and smRNA transcriptome, Lister *et al.* (2008) have reported a direct relationship between the location of smRNAs and DNA methylation, perturbation of smRNA biogenesis upon loss of CpG DNA methylation, and a tendency for smRNAs to direct strand-specific DNA methylation in regions of RNA–DNA homology in *Arabidopsis* cell culture. In addition, the authors stated that strand-specific mRNA-seq detected variation in transcript abundance of hundreds of genes, transposons, and unannotated intergenic transcripts upon modification of the DNA methylation state.

To date, high throughput research on SV has only been focusing on model plant species and a few commercial crop like oil palm, due to the technical limitations caused by lack of genomic sequences in non-model plants and the challenges in bioinformatics analysis. Knowledge gained from model species have provided a few possible causes of the epigenetic regulations of SV (Miguel and Marum, 2011), while the advances in current next generation sequencing and bioinformatics tools in gene profiling and novel gene discovery has allowed unprecedented opportunities for more plant biologist to increase the understanding of the functions and dynamics of plant cells and populations in non-model plants (Lister *et al.*, 2009). At gene level, our studies agreed with several current hypotheses of SV occurrence, which consists of a few important gene families including DNA methylation, histone modification, activation of

retro-elements, transcription factors and miRNA gene regulation. These genes were found to be differentially expressed over time in prolonged callus culture of vanilla.

On the other hand, in the case of miRNAs, it would be thinkable that in the future, knowledge about miRNA mobility in tissue culture could reach its full potential as a strategy for propagating recalcitrant species such as orchids or trees to promote rejuvenation (Wang *et al.*, 2010). Furthermore, the secretion of small RNA into the culture media has been found in animal cells (Wang *et al.*, 2010). Therefore, it is tempting to speculate that plants can also secrete miRNAs into the media. If so, it would be interesting to explore the effects that the addition of certain small RNAs to culture media could have in recalcitrant plants. The current limitation would be the lack of genomic sequences for *V. planifolia*, in order for future characterization of novel miRNAs found in the sequencing data. This could be a potential future work to be incorporated into SV studies as well as to increase the regeneration rate of vanilla tissue culture.

6.3. Future recommendations

There are a few important considerations and screenings that could be done prior to tissue culture of vanilla. Firstly, the ploidy level of explants should be determined for regeneration involving cellular dedifferentiation, such as the callus-mediated regenerations. As observed in our study, pre-existing endoreduplication in an organ could lead to the multiplication of polyploidy populations in an *in vitro* culture system in the callus-mediated regeneration system of vanilla. The polyploid cell populations were multiplied over time in the prolonged nodal-derived callus culture. For long term culture, shoot tips which had dominant 2N population should be a better option. Secondly, although callus-mediated plant regeneration gave rise to higher number of regenerants, it should be avoided for clonal multiplication as a higher level of somaclonal variations were observed. Thirdly, we suggest that callus-mediated plant regeneration is a potential source for genetic variations in order to be used for selection for *in vitro* crop improvements of vanilla. In addition, we propose that

root tip regeneration using our optimized protocol could be another viable explants for callus-mediated regeneration for *in vitro* crop improvement. This is because of its shorter regeneration time and higher endopolyploidy levels pre-existed in root tips, which could potentially be multiplied in callus culture over prolonged time period. Fourthly, the output from RNA-seq has provided a list of retrotransposable element sequences, which could be further incorporated into the marker systems in future work. Characterization of microRNAs and integration of data with methylome could also be carried out with the availability of genomic sequences.

Plant tissue culture, although is a controlled system, is still a complex system which is affected by a network of stress factors, such as temperature, exogenous PGR (types and concentrations), wounding and other stress responses, which are involved in triggering morphogenesis through cell-cycle disturbance and inducing variability (Peschke and Phillips, 1992). The factors that contribute to SV incidence remains a subject for debate and warrants further stringent and larger scales of experiments, which is a great challenge in real laboratory condition. This limited application of SV study due to complexity of biological systems has been acknowledged in other studies, such as observed in study by Côte *et al.* (2001) to develop a statistical model for SV prediction (Bairu *et al.*, 2011). Besides, a long term future study could include the characterization of vanilla somaclones at field evaluation level to examine the effects of SV on the flowering and yield potential of vanilla plants when it reaches maturity, which usually will take minimum 3 years (Castro-Bobadilla and García-Franco, 2007; Diez *et al.*, 2016).

It is hoped that a more precise and comprehensive knowledge of the molecular basis of these variations would help to devise strategies to improve the totipotency and embryogenic capability in recalcitrant species and genotypes, and to address bottlenecks associated with clonal propagation, so that more effective tools could be developed and applied in routine screening for early detection of SV in TC industry or for the selection of SV for crop improvements of vanilla.

APPENDICES

Appendix A. The list of standard solutions, chemicals, enzymes and kits used in the study and their respective manufacturers.

Standard solutions / Chemicals	Manufacturer	Country
Plant tissue culture: sterilization agents, culture media, PGRs and vitamin (Chapter 2)		
1-Naphthaleneacetic acid (NAA)	Duchefa Biochemie B.V.	The Netherlands
2,4-Dichlorophenoxyacetic acid (2,4-D)	Sigma Aldrich	USA
6-Benzyaminopurine (BAP)	Duchefa Biochemie B.V.	The Netherlands
Agargel™	Sigma Aldrich	USA
Carbendazim	Mei Zim	Malaysia
Clorox®	Clorox	USA
Ethyl Alcohol	R&M chemical	USA
Gamborg B5 medium	Duchefa Biochemie B.V.	The Netherlands
Knudson C orchid medium	Duchefa Biochemie B.V.	The Netherlands
Murashige and Skoog basal salt with Vitamins	*Phyto*technology Laboratory® / Duchefa Biochemie B.V.	USA / The Netherlands
Myo-inositol	Nacalai Tesque	Korea
Phytagel™	Sigma Aldrich	USA
Sucrose	Fisher Scientific	USA
Thiamine-HCl		USA
Thidiazuron (TDZ)	Sigma Aldrich	USA
Tween 20	Sigma Aldrich	USA
Cleaning Agent Decon® 90	Decon Laboratories Limited	UK
Plant cell histology and microscopy (Chapter 2, 4)		
0.2 M phosphate buffer (pH 7.2)		
10% (v/v) Paraformaldehyde, Ultra pure	Merck (4005)	USA
25% (v/v) Glutaraldehyde (aqueous solution,	Sigma (G-6257)	USA
Butanol		
Caffeine	Sigma (C-0750)	USA
Ethanol		
Mineral oil		
Technovit® 7100 slicable cold-curing resin - for infiltration	Heraceus Kulzer GmbH	Germany
Molecular analysis: DNA and RNA extractions, RE digestions, PCR and gel electrophoresis reagents (Chapter 3, 5-6)		
10X Tris-borate-EDTA (TBE) buffer	Thermo Fisher Scientific	USA
2X RNA loading dye	Thermo Fisher Scientific	USA
50X Tris-acetate-EDTA (TAE) buffer	Thermo Fisher Scientific	USA
6X DNA loading dye	Thermo Fisher Scientific	USA
Agarose	Vivantis	Malaysia
Ascorbic acid	Sigma-Aldrich	USA
Bovine Serum Albumin (BSA)	Promega	USA
Direct-Zol™ RNA MiniPrep kit	Zymo Research	USA
DNase I (RNase-free)	Thermo Fisher Scientific	USA
dNTPs	Promega	USA
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich	USA
Fastdigest® *Eco*RI	Thermo Fisher Scientific	USA

Fastdigest® *Hpa*II	Thermo Fisher Scientific	USA
Fastdigest® *Msp*I	Thermo Fisher Scientific	USA
GeneRuler 1 kb Plus DNA ladder (0.5 µg/µl, ready-to-use)	Thermo Fisher Scientific	USA
GeneRuler 100 bp Plus DNA ladder (0.5 µg/µl, ready-to-use)	Thermo Fisher Scientific	USA
GoTaq™ Flexi Taq Polymerase, buffer and $MgCl_2$	Promega	USA
Hexadecyltrimethylammonium bromide (CTAB)	Promega	USA
Phenol: Chloroform Isoamyl alcohol (PCI) (25:24:1)	Sigma Aldrich	USA
Polyvinylpyrrolidone (PVP-40)	Sigma-Aldrich	USA
Potassium acetate	Sigma Aldrich	USA
RiboRuler High Range RNA Ladder, ready-to-use	Thermo Fisher Scientific	USA
RNAse A	Thermo Fisher Scientific	USA
RNeasy® MinElute cleanup kit	Qiagen	Germany
Sodium acetate	Sigma Aldrich	USA
Sodium chloride (NaCl)	Sigma-Aldrich	USA
Sodium dodecyl sulfate (SDS)	Sigma-Aldrich	USA
SybrSafe® DNA gel stain, 10,000X concentrate in dimethyl sulfoxide (DMSO)	Invitrogen	USA
T4 DNA ligase	Thermo Fisher Scientific	USA
Tris base	Promega	USA
Trizol® Reagent	Invitrogen	USA
UltraPure™ DNase/RNase-free distilled water	Invitrogen	USA
β-mercaptoethanol	Sigma-Aldrich	USA

Plant ploidy experiment: Plant nuclei extraction staining and flow cytometry analysis (Chapter 4)

Citric acid		
HCl (1N, 5N, concentrated)		
$MgCl_2.6H_2O$ / $MgCl_2$		
MOPS		
Murashige and Skoog basal salt with Vitamins	*Phyto*technology Laboratory®	
Na_2EDTA		
$Na_2PO_4.12H_2O$		
NaOH (1N, 5N, pellets)		
Potassium chloride (KCl)		
Propidium iodide (PI) (Cat. No.: P4170)	Sigma Aldrich, USA	
PVP-10		
RNAse A stock solution		
Sodium chloride (NaCl)		
Sodium citrate	Sigma Aldrich	USA
Sodium metabisulfite		
Sorbitol		
Spermine tetrahydrochloride (Spermine.4HCl)	Sigma Aldrich	USA
Tris	1st base	Singapore

Triton X-100		
Tween 20		
β-mercaptoethanol	Sigma Aldrich	USA
RNA sequencing: RNA handling, qubit and bioanalyzer analysis, mRNA and miRNA sequencing, cDNA libraries preparation and quantitative real time PCR analysis (Chapter 5)		
Agilent RNA 6000 Nano Kit (5067-1511)	Agilent Technologies	USA
Bioanalyser HS kit (5067-4626)	Agilent Biotechnologies	USA
Direct-Zol™ RNA MiniPrep kit	Zymo Research	USA
KAPA SYBR® FAST qPCR Master Mix (2X) Kit (KR0389)	Kapa Biosystems	USA
miScript® Plant RT Kit	Qiagen	Germany
miScript® SYBR Green PCR Kit	Qiagen	Germany
NEBNext Multiplex Oligos for Illumina (Index Primers Sets 1 and 2) (E7335S and E7500S)	NEB	UK
NEBNext poly(A) mRNA Magnetic Isolation Module (E7490)	NEB	UK
NEBNext Ultra Directional RNA library prep kit for Illumina (E7420S)	NEB	UK
NEBNext® Small RNA Library Prep Set for Illumina® (Set 1)	NEB	UK
QuantiNova™ Reverse Transcription Kit	Qiagen	Germany
Qubit RNA BR assay kit (Q10210)	Life technologies	USA
RNAse Away	Sigma Aldrich,	USA
Trizol® Reagent	Invitrogen	USA
Universal qPCR Master Mix (KK4824)	Kapa Biosystems	USA

Appendix B. Equipment, apparatus and software used in this study.

Apparatus	Brand, Country (Product Code)
Plant tissue culture: media preparation, surface sterilization, cell culture (Chapter 2)	
Esco Airstream Horizontal Laminar Flow Cabinet	ESCO®, USA
HOBO Wireless Data logger system	OnSet, USA
Steri 350 Glass beads sterilizer	Innotech Biosystems international Inc., USA
Plant cell histology and microscopy (Chapter 2, 4)	
Nikon Eclipse 80i Light microscope	Nikon, Japan
Nikon SMZ1500 stereomicroscope	Nikon, Japan
NIS-Element 3.0	Nikon, Japan
Molecular analysis: DNA and RNA extractions, RE digestions, PCR and gel electrophoresis reagents (Chapter 3, 5-6)	
Eppendorf 5810R Refrigerated Centrifuge	Eppendorf, Germany
FluorChem® HD 2 Gel Imaging systems	Alpha Innotech, USA
Nanodrop ND-1000	Thermo Scientific, USA
Nexus Gradient thermocycler	Eppendorf, Germany
Thermomixer comfort	Eppendorf, Germany
Vacuum dessicator (Concentrator plus)	Eppendorf, Germany
Water bath	Memmert Germany (WWB 29)
Plant ploidy experiment: Plant nuclei extraction staining and flow cytometry analysis (Chapter 4)	
BD Accuri™ C6 Flow cytometer	BD Biosciences, USA
BD Accuri™ C6 Software	BD Biosciences, USA
Cell strainers, 40 µm pore size	BD Falcon, USA (Cat. No. LAB352340)
Petri dishes, Triple Vent, Sterile, 94 x 15 mm	Greiner BIO-ONE, Austria (Cat. No. 633185)
Razor blades and holders, double edged, sharp-ended	Xinda, China (BS-2982; #11)
R-BTM sample tubes with cell strainer cap, 5 ml, 12 X 75 mm	BD Falcon, USA (Cat. No. LAB352235)
RNA sequencing and RT-qPCR analysis: RNA handling, qubit and bioanalyzer analysis, mRNA and miRNA sequencing, cDNA libraries preparation and quantitative real time PCR analysis (Chapter 5)	
Agilent 2100 Bioanalyzer	Agilent Technologies, USA
Bioinformatics Workstation	Dell, USA
BioRad CFX Connect™ / CFX Manager	BioRad, USA
CFX Connect™ Real-Time PCR Detection System	Bio-Rad, USA
CLCBio Genomic Workbench 8.0	Qiagen, Germany
Eppendorf™ real-time PCR Tube Strips and Masterclear™ Cap Strips	Eppendorf, USA
FastQC software	Babraham Bioinformatics, UK

Illumina NextSeq500 Sequencer System	Illumina, USA
Partek Genomic Suite (Version 6.6)	Partek, Sg
PCR Workstation	
Statistical software (Chapter 2-7)	
Genstat (Version 18)	VSN International, UK

Appendix C. Chemical buffers used in this study and their buffer components.

Buffer	Composition	Reference
CTAB DNA Extraction buffers		
Extraction Buffer A	2% (w/v) Hexadecyltrimethylammonium bromide (CTAB) (Sigma-Aldrich, USA), 100 mM Tris-HCl (pH 8.0) (Promega, USA), 20 mM Ethylenediaminetetraacetic acid (EDTA), 1.4 M Sodium chloride (NaCl) (Sigma Aldrich, USA), 4% (w/v) polyvinylpyrrolidone (PVP-40) (Sigma-Aldrich, USA), 0.1% (w/v) ascorbic acid (Sigma-Aldrich, USA) and 10 mM β-mercaptoethanol (Sigma-Aldrich, USA)	Keb-Llanes (2002)
Extraction Buffer B	100 mM Tris-HCl (pH 8), 50 mM EDTA (Promega, USA), 100 mM NaCl and 10 mM β-mercaptoethanol (Sigma-Aldrich, USA)	
SDS	10% Sodium dodecyl sulfate (SDS) (Sigma Aldrich, USA)	
Plant histological analysis buffers		
Glutaldehyde-Paraformaldehyde-Caffeine (GPC) fixative solution	50% (v/v) of 0.2 M phosphate buffer (pH 7.2), 20% (v/v) of 10% (v/v) Paraformaldehyde (Merk 4005, ultra pure), 4% (v/v) of 25% (v/v) Glutaraldehyde (aqueous solution, Sigma G-6257), 1% (w/v) Caffeine (Sigma C-0750) (predissolved in 2 ml distilled water with little heating), 26 ml distilled water to a total volume of 100ml	MPOB Histology Lab Manual
[1]Nuclei isolation buffers (used in plant flow cytometry)		
LB01	15 mM Tris, 2 mM Na_2EDTA, 0.5 mM spermine.4HCl, 80 mM KCl, 20 mM NaCl, 15 mM mercaptoethanol, 0.1% (v/v) Triton X-100, pH 8.0	Doležel *et al.* (1989)
Tris.$MgCl_2$	200 mM Tris, 4 mM $MgCl_2.6H_2O$, 0.5% (v/v) Triton X-100, pH 7.5	Pfosser *et al.* (1995)
Galbraith	45 mM $MgCl_2$, 30 mM sodium citrate, 20 mM MOPS, 0.1% (v/v) Triton X-100, pH 7.0	Galbraith *et al.* (1983)
Otto	Otto I: 100 mM citric acid, 0.5% (v/v) Tween 20[*2] (pH 2–3) Otto II: 400 mM $Na_2PO_4.12H_2O$ (pH 8–9)	Otto (1992) Doležel and Göhde (1995)
Bory	1X MS macroelements (Murashige and Skoog, 1962), 600 mmol/L sorbitol, 1% Triton X-100, 10 mmol/L sodium metabisulfite	Bory *et al.* (2008)

WPB	0.2M Tris.HCl, 4 mM MgCl2, 2 mM Na_2EDTA, 86 mM NaCl, 1.0% (w/v) PVP, 1.0% (v/v) Triton X-100, 0.01 M sodium metabisulfite (added fresh)	Loureiro *et al.* (2007)
GPB	0.5 mM spermine tetrahydrochloride, 30 mM sodium citrate, 20 mM MOPS, 80 mM KCl, 20 mM NaCl, 1.0% (v/v) Triton x-100, pH 7.0	Loureiro *et al.* (2007)
Partec buffer A	10 mM Tris, 2 mM MgCl2, 50 mM sodium citrate, 1% PVP (PVP K-30 -> PVP-10), 0.1% (v/v) Triton-X, pH 7.5	Silva and Fukai (2003)

*[1] All buffers were filtered through 0.22 µm filter to remove small particles. All buffers were stored at -20°C (except Otto buffers I and II, which were prepared fresh) in 5 ml aliquots and never refreeze after thawing.

*[2] Cell culture-tested grade (Sigma-Aldrich, cat no. P2287). This is essential to prevent high sample background in flow cytometry analysis.

Appendix D. Formulation of media used in this study.

Murashige & Skoog medium (1962) with vitamins

Micro Elements	**mg/l**	**µM**
$CoCl_2.6H_2O$	0.025	0.11
$CuSO_4.5H_2O$	0.025	0.10
FeNaEDTA	36.70	100.00
H_3BO_3	6.20	100.27
KI	0.83	5.00
$MnSO_4.H_2O$	16.90	100.00
$Na_2MoO_4.2H_2O$	0.25	1.03
$ZnSO_4.7H_2O$	8.60	29.91
Macro Elements	**mg/l**	**mM**
$CaCl_2$	332.02	2.99
KH_2PO_4	170.00	1.25
KNO_3	1900.00	18.79
$MgSO_4$	180.54	1.50
NH_4NO_3	1650.00	20.61
Vitamins	**mg/l**	**Vitamins**
Glycine	2.00	Glycine
myo-Inositol	100.00	myo-Inositol
Nicotinic acid	0.50	Nicotinic acid
Pyridoxine HCl	0.50	Pyridoxine HCl
Thiamine HCl	0.10	Thiamine HCl

Knudson C Orchid Medium (Morel 1965)

Micro Elements	**mg/l**	**µM**
$FeSO_4.7H_2O$	25.00	89.92
$MnSO_4.H_2O$	5.68	33.61
Macro Elements	**mg/l**	**mM**
$Ca(NO_3)2$	241.30	1.43
KCl	250.00	3.35
KH_2PO_4	250.00	1.84
$MgSO_4$	122.15	1.02
NH_4NO_3	500.00	6.25
$(NH_4)2SO_4$	500.00	3.78

Gamborg B5 medium (Gamborg *et al.* 1968)

Micro Elements	**mg/l**	**µM**
$CoCl_2.6H_2O$	0.025	0.11
$CuSO_4.5H_2O$	0.025	0.10
FeNaEDTA	36.70	100.00
H_3BO_3	3.00	48.52
KI	0.75	4.52
$MnSO_4.H_2O$	10.00	59.16
$Na_2MoO_4.2H_2O$	0.25	1.03
$ZnSO_4.7H_2O$	2.00	6.96
Macro Elements	**mg/l**	**mM**
$CaCl_2$	113.23	1.02
KNO_3	2500.00	24.73
$MgSO_4$	121.56	1.01
NaH_2PO_4	130.44	1.09
$(NH_4)2SO_4$	134.00	1.01

Appendix E. List of adapter and primer sequences used in ISSR and MSAP analysis (**Chapter 3**).

Name	Enzyme	Type	Sequence (5'-3')
ISSR Analysis (Chapter 3)			
UBC811			$(GA)_8C$
UBC824			$(TC)_8G$
UBC826			$(AC)_8C$
VP-ISSR2			$(CA)_8RY$
VP-ISSR4			$(GA)_8YC$
VP-ISSR6			$(AG)_8YT$
VP-ISSR7			$(AG)_7YC$
VP-ISSR10			$(GT)_8YG$
VP-ISSR23			$(CT)_8RC$
VP-ISSR30			$HVH(TG)_7$
VP-ISSR32			$C_3(GT)_7$
MSAP Analysis (Chapter 3)			
Name	**Enzyme**	**Type**	**Sequence (5'-3')**
*Eco*A+	*Eco*RI	Adapter(+)	CTCGTAGACTGCGTACC
*Eco*A-	*Eco*RI	Adapter(-)	AATTGGTACGCAGTC
Hpa/MspA+	*Hpa*II/*Msp*I	Adapter(+)	GACGATGAGTCCTGAT
Hpa/MspA -	*Hpa*II/*Msp*I	Adapter(-)	CGATCAGGACTCAT
E01	*Eco*RI	Primer +1	CTGCGTACCAATTCa
E02	*Eco*RI	Primer +1	CTGCGTACCAATTCt
E03	*Eco*RI	Primer +1	CTGCGTACCAATTCc
E04	*Eco*RI	Primer +1	CTGCGTACCAATTCg
MH01	*Hpa*II/*Msp*I	Primer +1	GATGAGTCCTGATCGGa
MH02	*Hpa*II/*Msp*I	Primer +1	GATGAGTCCTGATCGGt
MH03	*Hpa*II/*Msp*I	Primer +1	GATGAGTCCTGATCGGc
MH04	*Hpa*II/*Msp*I	Primer +1	GATGAGTCCTGATCGGg
E01-1	*Eco*RI	Primer +2	CTGCGTACCAATTCaa
E01-2	*Eco*RI	Primer +2	CTGCGTACCAATTCat
E01-3	*Eco*RI	Primer +2	CTGCGTACCAATTCac
E01-4	*Eco*RI	Primer +2	CTGCGTACCAATTCag
E02-1	*Eco*RI	Primer +2	CTGCGTACCAATTCta
E02-2	*Eco*RI	Primer +2	CTGCGTACCAATTCtt
E02-3	*Eco*RI	Primer +2	CTGCGTACCAATTCtc
E02-4	*Eco*RI	Primer +2	CTGCGTACCAATTCtg
MH01-1	*Hpa*II/*Msp*I	Primer +2	GATGAGTCCTGATCGGaa
MH01-2	*Hpa*II/*Msp*I	Primer +2	GATGAGTCCTGATCGGat
MH01-3	*Hpa*II/*Msp*I	Primer +2	GATGAGTCCTGATCGGac
MH01-4	*Hpa*II/*Msp*I	Primer +2	GATGAGTCCTGATCGGag
MH04-1	*Hpa*II/*Msp*I	Primer +2	GATGAGTCCTGATCGGga
MH04-2	*Hpa*II/*Msp*I	Primer +2	GATGAGTCCTGATCGGgt
MH04-3	*Hpa*II/*Msp*I	Primer +2	GATGAGTCCTGATCGGgc
MH04-4	*Hpa*II/*Msp*I	Primer +2	GATGAGTCCTGATCGGgg

Nucleotide ambiguity codes: **Y** represented C/T, **R** represented A/G, **H** represented A or C or T, **V** represented A or C or G;

Appendix F. List of plant (*Viridiplantae*) with available miRNA sequences deposited in mirBase version 21.0, which was used for miRNA annotation in this study.

- Viridiplantae
  - Chlorophyta
    - *Chlamydomonas reinhardtii* (50 precursors, 86 mature) [CHLRE3.1]

    Coniferophyta
    - *Cunninghamia lanceolata* (4 precursors, 4 mature)
    - *Picea abies* (40 precursors, 41 mature)
    - *Pinus densata* (29 precursors, 29 mature)
    - *Pinus taeda* (35 precursors, 36 mature)
  - Embryophyta
    - *Physcomitrella patens* (229 precursors, 280 mature) [JGI1.1]
    - *Selaginella moellendorffii* (58 precursors, 64 mature)
  - Magnoliophyta
    - *Amborella trichopoda* (124 precursors, 129 mature) [AmTr_v1.0]
    - eudicotyledons

      Araliaceae
      - *Panax ginseng* (29 precursors, 32 mature)

      Asteraceae
      - *Cynara cardunculus* (48 precursors, 57 mature)
      - *Helianthus annuus* (6 precursors, 7 mature)
      - *Helianthus argophyllus* (3 precursors, 3 mature)
      - *Helianthus ciliaris* (3 precursors, 3 mature)
      - *Helianthus exilis* (2 precursors, 2 mature)
      - *Helianthus paradoxus* (3 precursors, 3 mature)
      - *Helianthus petiolaris* (3 precursors, 3 mature)
      - *Helianthus tuberosus* (16 precursors, 16 mature)

      Brassicaceae
      - *Arabidopsis lyrata* (205 precursors, 384 mature) [v1.0]
      - *Arabidopsis thaliana* (325 precursors, 427 mature) [TAIR10]
      - *Brassica napus* (90 precursors, 92 mature) [BrassicaDB-20080411]
      - *Brassica oleracea* (10 precursors, 11 mature)
      - *Brassica rapa* (96 precursors, 157 mature) [Brapa_1.1]

      Caricaceae
      - *Carica papaya* (79 precursors, 81 mature)

      Cucurbitaceae
      - *Cucumis melo (*120 precursors, 120 mature)

      Euphorbiaceae
      - *Hevea brasiliensis* (31 precursors, 31 mature) [Hevbra_1.0]
      - *Manihot esculenta* (153 precursors, 153 mature) [Cassava4]
      - *Ricinus communis* (63 precursors, 63 mature) [TIGR0.1]

Fabaceae

- *Acacia auriculiformis* (7 precursors, 7 mature)
- *Arachis hypogaea* (23 precursors, 32 mature)
- *Acacia mangium* (3 precursors, 3 mature)
- *Glycine max* (573 precursors, 639 mature) [v1.0]
- *Glycine soja* (13 precursors, 13 mature)
- *Lotus japonicus* (62 precursors, 67 mature) [lj2.5]
- *Medicago truncatula* (672 precursors, 756 mature) [JCVI_Mt3.5.2]
- *Phaseolus vulgaris* (8 precursors, 10 mature)
- *Vigna unguiculata* (18 precursors, 18 mature)

Lamiales

- *Avicennia marina* (2 precursors, 3 mature)
- *Digitalis purpurea* (13 precursors, 13 mature)
- *Rehmannia glutinosa* (32 precursors, 37 mature)
- *Salvia sclarea* (18 precursors, 18 mature)

Linaceae

- *Linum usitatissimum* (124 precursors, 124 mature)

Malvaceae

- *Gossypium arboreum* (1 precursors, 1 mature)
- *Gossypium herbaceum* (1 precursors, 1 mature)
- *Gossypium hirsutum* (78 precursors, 80 mature)
- *Gossypium raimondii* (296 precursors, 296 mature) [Graimondii2_0]
- *Theobroma cacao* (82 precursors, 82 mature)

Ranunculaceae

- *Aquilegia caerulea* (45 precursors, 45 mature)

Rhizophoraceae

- *Bruguiera cylindrica* (4 precursors, 4 mature)
- *Bruguiera gymnorhiza* (4 precursors, 4 mature)

Rosaceae

- *Malus domestica* (206 precursors, 207 mature) [maldom-pseudo1.0]
- *Prunus persica* (180 precursors, 214 mature) [Prupe1_0]

Rutaceae

- *Citrus clementina* (5 precursors, 5 mature)
- *Citrus reticulata* (4 precursors, 4 mature)
- *Citrus sinensis* (60 precursors, 64 mature)
- *Citrus trifoliata* (6 precursors, 6 mature)

Salicaceae

- *Populus euphratica* (4 precursors, 4 mature)
- *Populus trichocarpa* (352 precursors, 401 mature) [JGI_Poptr2.0]

Solanaceae

- *Nicotiana tabacum* (162 precursors, 164 mature)
- *Solanum lycopersicu*m (77 precursors, 110 mature) [SL2.40]
- *Solanum tuberosum* (224 precursors, 343 mature) [SolTub3.0]

Vitaceae

- *Vitis vinifera* (163 precursors, 186 mature) [Genoscope-20100122]

- monocotyledons
  - *Aegilops tauschii* (88 precursors, 173 mature) [ASM34733v1]
  - *Brachypodium distachyon* (317 precursors, 525 mature) [Bd21]
  - *Elaeis guineensis* (6 precursors, 6 mature)
  - *Festuca arundinacea* (15 precursors, 15 mature)
  - *Hordeum vulgare* (69 precursors, 71 mature)
  - *Oryza sativa* (592 precursors, 713 mature) [MSU7]
  - *Sorghum bicolor* (205 precursors, 241 mature) [Sorbi1]
  - *Saccharum officinarum* (16 precursors, 16 mature)
  - *Saccharum sp.* (19 precursors, 20 mature)
  - *Triticum aestivum* (116 precursors, 119 mature)
  - *Triticum turgidum* (1 precursors, 1 mature)
  - *Zea mays* (172 precursors, 321 mature) [RefGen_v2]

Appendix G. Identification of conserved miRNA families in *Vanilla planifolia* with corresponding isoforms.

miRNA family	Gene name	Length (bp)	Sequence	Reference miRNA	Reference species
MIR156	vpl-miR156a	21	CUGACAGAAGAGAGUGAGCAC	ama-miR156	*Avicennia marina*
	vpl-miR156b	20	UGACAGAAGAGAGUGAGUAC	hci-miR156a	*Helianthus ciliaris*
	vpl-miR156c	20	UGACAGAAGAAAGUGAGCAC	mdm-miR156ad	*Malus domestica*
	vpl-miR156d	21	UGACAGAAGAGAGUGAGCAUA	cca-miR156b	*Cynara cardunculus*
	vpl-miR156e-5p	21	UUGACAGAAGAUAGAGAGCAC	ahy-miR156b-5p	*Arachis hypogaea*
	vpl-miR156e-3p	21	GCUCUCUAAGCUUCUGUCAUC	ahy-miR156b-3p	*Arachis hypogaea*
	vpl-miR156f-3p	22	GCUCACCUCUCUUUCUGUCAGU	aly-miR156b-3p	*Arabidopsis lyrata*
	vpl-miR156g-3p	21	GCUCACUCUCUAUCUGUCAGC	osa-miR156b-3p	*Oryza sativa*
	vpl-miR156h-3p	21	GCUCACCCUCUAUCUGUCAGU	zma-miR156b-3p	*Zea mays*
	vpl-miR156i	21	UUGACAGAAGAGAGAGAGCAC	ahy-miR156c	*Arachis hypogaea*
	vpl-miR156j	20	UGUCAGAAGAGAGUGAGCAC	ghr-miR156c	*Gossypium hirsutum*
	vpl-miR156k	19	UGACAGAAGAGAGGGAGCA	har-miR156c	*Helianthus argophyllus*
	vpl-miR156l	21	UUGACAGAAGAUAGAGGGCAC	cme-miR156e	*Cucumis melo*
	vpl-miR156m	20	UGACAGAAGAGAGCGAGCAC	sbi-miR156e	*Sorghum bicolor*
	vpl-miR156n	20	UGACAGAGGAGAGUGAGCAC	vvi-miR156e	*Vitis vinifera*
	vpl-miR156o-5p	20	UGAUAGAAGAGAGUGAGCAC	sly-miR156e-5p	*Solanum lycopersicum*
	vpl-miR156p	20	UGACAGAAGAGAAUGAGCAC	nta-miR156f	*Nicotiana tabacum*
	vpl-miR156q-3p	21	GCUCACUCUCUAUCCGUCACC	ath-miR156f-3p	*Arabidopsis thaliana*

	vpl-miR156r-5p	20	CGACAGAAGAGAGUGAGCAC	aly-miR156g-5p	*Arabidopsis lyrata*
	vpl-miR156s	20	UGACAGAAGAGAGGGAGCAC	cme-miR156h	*Cucumis melo*
	vpl-miR156t	20	UGACAGAAGAGAGAGAGCAU	vvi-miR156h	*Vitis vinifera*
	vpl-miR156u-5p	20	UGACAGAAGAAAGAGAGCAC	aly-miR156h-5p	*Arabidopsis lyrata*
	vpl-miR156v	20	UGACAGAAGAGGGUGAGCAC	mtr-miR156j	*Medicago truncatula*
	vpl-miR156w	21	UUGACAGAAGAUGGAGAGCAC	ptc-miR156l	*Populus trichocarpa*
	vpl-miR156x-5p	21	CGACAGAAGAGAGUGAGCAUA	osa-miR156l-5p	*Oryza sativa*
	vpl-miR156y	21	UUGACAGAAGAGAGUGAGCAC	gma-miR156o	*Glycine max*
	vpl-miR156z	21	UGACAGAAGAGAGUGAGCACU	gma-miR156q	*Glycine max*
	vpl-miR156aa	21	CUGACAGAAGAUAGAGAGCAU	gma-miR156r	*Glycine max*
	vpl-miR156ab	18	AUUGGAGUGAAGGGAGCU	gma-miR156z	*Glycine max*
MIR157	vpl-miR157a-3p	21	GCUCUCUAUACUUCUGUCACC	aly-miR157c-3p	*Arabidopsis lyrata*
	vpl-miR157b-3p	21	GCUCUCUAUGCUUCUGUCAUC	aly-miR157d-3p	*Arabidopsis lyrata*
	vpl-miR157b-5p	20	UGACAGAAGAUAGAGAGCAC	aly-miR157d-5p	*Arabidopsis lyrata*
MIR159	vpl-miR159a	21	UUUGGAUUGAAGGGAGCUCUA	ahy-miR159	*Arachis hypogaea*
	vpl-miR159b	21	UUUGGACUGAAGGGAGCUCUA	aqc-miR159	*Aquilegia caerulea*
	vpl-miR159c	21	UUUGGUUUGAAGGGAGCUCUA	pde-miR159	*Pinus densata*
	vpl-miR159d	21	CUUGGAGUGAAGGGAGCUCUC	vvi-miR159a	*Vitis vinifera*
	vpl-miR159e-3p	20	CUUGGAUUGAAGGGAGCUCU	bdi-miR159a-3p	*Brachypodium distachyon*
	vpl-miR159f	21	AUUGGAUUGAAGGGAGCUCCU	cme-miR159b	*Cucumis melo*
	vpl-miR159g	20	CUUGGAUUGAAGGGAGCUCC	cpa-miR159b	*Carica papaya*
	vpl-miR159h	21	UUUGGAUUGAAGGGAGCUCUC	lus-miR159b	*Linum usitatissimum*

	vpl-miR159i-3p	21	UUUGGAUUGAAGGGAGCUCUU	aly-miR159b-3p	*Arabidopsis lyrata*
	vpl-miR159j-3p.1	21	UUUGGAUUGAAGGGAGCUCUG	bdi-miR159b-3p.1	*Brachypodium distachyon*
	vpl-miR159k	21	AUUGGAUUGAAGGGAGCUCCA	osa-miR159c	*Oryza sativa*
	vpl-miR159l-3p	21	UUUGGAUUGAAGGGAGCUCCU	aly-miR159c-3p	*Arabidopsis lyrata*
	vpl-miR159m	21	UUUGGAUUGAAAGGAGCUCUU	sof-miR159e	*Saccharum officinarum*
	vpl-miR159n	21	CUUGGAUUGAAGGGAGCUCUA	osa-miR159f	*Oryza sativa*
	vpl-miR159o-3p	21	UUUGGAGUGAAGGGAGCUCUG	zma-miR159i-3p	*Zea mays*
MIR160	vpl-miR160a	21	UGCCUGGCUCCUUGUAUGCCA	dpr-miR160	*Digitalis purpurea*
	vpl-miR160b-5p	21	UGCCUGGCUCCCUGAAUGCCA	ahy-miR160-5p	*Arachis hypogaea*
	vpl-miR160c	21	UGCCUGGCUCCCUGGAUGCCA	aqc-miR160a	*Aquilegia caerulea*
	vpl-miR160d	21	CGCCUGGCUCCCUGUAUGCCA	ppt-miR160b	*Physcomitrella patens*
	vpl-miR160e	20	UGCCUGGCUCCCUGUAUGCC	bdi-miR160f	*Brachypodium distachyon*
MIR162	vpl-miR162a	21	UCGAUAAACCUCUGCAUCCAG	aau-miR162	*Acacia auriculiformis*
	vpl-miR162b	21	UCGAUAAACCUCUGCAUCCGG	bdi-miR162	*Brachypodium distachyon*
	vpl-miR162c	21	UCGAUAAGCCUCUGCAUCCAG	osa-miR162b	*Oryza sativa*
MIR164	vpl-miR164a	21	UGGAGAAGCAGGGUACGUGCA	cca-miR164	*Cynara cardunculus*
	vpl-miR164b	21	UGGAGAAGCAGGGCACUUGCU	far-miR164a	*Festuca arundinacea*
	vpl-miR164c	21	UGGAGAAGCAGGGCACGUGAA	hci-miR164a	*Helianthus ciliaris*
	vpl-miR164d-3p	21	CAUGUGCCCUUCUUCUCCACC	bdi-miR164a-3p	*Brachypodium distachyon*
	vpl-miR164e-5p	21	UGGAGAAGCAGGGCACGUGCA	aly-miR164a-5p	*Arabidopsis lyrata*
	vpl-miR164f-5p	21	UGGAGAAGCAGGGCACGUGCU	ata-miR164a-5p	*Aegilops tauschii*
	vpl-miR164g	20	UGGAGAAGCAGGGCACGUGC	gma-miR164b	*Glycine max*

	vpl-miR164h	21	UGGAGAAGCAGGACACGUGAG	sbi-miR164c	*Sorghum bicolor*
	vpl-miR164i-5p	21	UGGAGAAGCAGGGCACGUGCG	aly-miR164c-5p	*Arabidopsis lyrata*
	vpl-miR164j-3p	21	CAUGUGCCCUUCUUCUCCAUC	zma-miR164h-3p	*Zea mays*
	vpl-miR164j-5p	21	UGGAGAAGCAGGGCACGUGUG	zma-miR164h-5p	*Zea mays*
MIR165	vpl-miR165a-3p	20	UCGGACCAGGCUUCAUCCCC	aly-miR165a-3p	*Arabidopsis lyrata*
	vpl-miR165a-5p	20	GAAUGUUGUCUGGAUCGAGG	aly-miR165a-5p	*Arabidopsis lyrata*
MIR166	vpl-miR166a	21	CCGGACCAGGCUUCAUCCCAG	cln-miR166	*Cunninghamia lanceolata*
	vpl-miR166b	21	UCGGACCAGGCUUCAUUCCUC	aqc-miR166a	*Aquilegia caerulea*
	vpl-miR166c	21	UCUCGGACCAGGCUUCAUUCC	atr-miR166b	*Arabidopsis thaliana*
	vpl-miR166d	22	UCGGACCAGGCUUCAUUCCCUU	crt-miR166b	*Citrus reticulata*
	vpl-miR166e	21	UCGGACCAGGCUUCAUUCCUA	mtr-miR166b	*Medicago truncatula*
	vpl-miR166f-5p	21	GGACUGUUGUCUGGCUCGAGG	aly-miR166b-5p	*Arabidopsis lyrata*
	vpl-miR166g	20	UCGGACCAGGCUUCAUUCCU	aqc-miR166c	*Aquilegia caerulea*
	vpl-miR166h-3p	21	UCGGACCAGGCUUCAUUCCCC	aly-miR166c-3p	*Arabidopsis lyrata*
	vpl-miR166i-3p	21	UCGGACCAGGCUUCAUUCCUU	ata-miR166c-3p	*Aegilops tauschii*
	vpl-miR166h-5p	21	GGAAUGUUGUCUGGCUCGAGG	aly-miR166c-5p	*Arabidopsis lyrata*
	vpl-miR166j-5p	21	GGGAUGUUGUCUGGCUCGACA	sly-miR166c-5p	*Solanum lycopersicum*
	vpl-miR166k	21	UCGGGCCAGGCUUCAUCCCCC	mtr-miR166d	*Medicago truncatula*
	vpl-miR166l-5p	21	GGAAUGUUGUCUGGCUCGGGG	ata-miR166d-5p	*Aegilops tauschii*
	vpl-miR166m-3p	21	UCGAACCAGGCUUCAUUCCCC	osa-miR166e-3p	*Oryza sativa*
	vpl-miR166n-5p	21	GGAAUGUUGUCUGGCACGAGG	aly-miR166e-5p	*Arabidopsis lyrata*
	vpl-miR166o	20	UCGGACCAGGCUUCAUUCCC	cme-miR166g	*Cucumis melo*

	vpl-miR166p-5p	21	GGAAUGUUGUUUGGCUCGAGG	aly-miR166g-5p	*Arabidopsis lyrata*
	vpl-miR166q	20	UCGGACCAGGCUUCAUUCUC	cme-miR166i	*Cucumis melo*
	vpl-miR166r	21	UUGGACCAGGCUUCAUUCCCC	mes-miR166i	*Manihot esculenta*
	vpl-miR166s-3p	21	UCGGAUCAGGCUUCAUUCCUC	osa-miR166i-3p	*Oryza sativa*
	vpl-miR166t	21	UCCGGACCAGGCUUCAUUCCC	ppt-miR166j	*Physcomitrella patens*
	vpl-miR166u	20	CGGACCAGGCUUCAUUCCCC	gma-miR166m	*Glycine max*
	vpl-miR166v	21	UCGGACCAGGCUUCAUUCCCU	osa-miR166m	*Oryza sativa*
	vpl-miR166w	21	UCGGACCAGGCAUCAUUCCUU	ppt-miR166m	*Physcomitrella patens*
	vpl-miR166x	21	UCGGACCAGGCUCCAUUCCUU	ptc-miR166p	*Populus trichocarpa*
MIR167	vpl-miR167a	21	UCAAGCUGCCAGCAUGAUCUA	aqc-miR167	*Aquilegia caerulea*
	vpl-miR167b	21	UGAAGCUGCCAGCAUGAUCUG	atr-miR167	*Arabidopsis thaliana*
	vpl-miR167c	21	GGAAGCUGCCAGCAUGAUCCU	ppt-miR167	*Physcomitrella patens*
	vpl-miR167d-3p	21	AGAUCAUGUGGCAGUUUCACC	ahy-miR167-3p	*Arachis hypogaea*
	vpl-miR167d-5p	21	UGAAGCUGCCAGCAUGAUCUU	ahy-miR167-5p	*Arachis hypogaea*
	vpl-miR167e	21	UGAAGCUGCCAGCAUGAUCUC	lus-miR167a	*Linum usitatissimum*
	vpl-miR167f	21	UCAGAUCAUCUUGCAGCUUCA	gra-miR167b	*Gossypium raimondii*
	vpl-miR167g	21	UGAAGCUGACAGCAUGAUCUA	tae-miR167b	*Triticum aestivum*
	vpl-miR167h-5p	21	UGAAGCUGCCAGCAUGAUCUA	aly-miR167b-5p	*Arabidopsis lyrata*
	vpl-miR167i-5p	22	UAAAGCUGCCAGCAUGAUCUGG	sly-miR167b-5p	*Solanum lycopersicum*
	vpl-miR167j-5p	21	UAAGCUGCCAGCAUGAUCUUG	aly-miR167c-5p	*Arabidopsis lyrata*
	vpl-miR167k	20	UGAAGCUGCCAGCAUGAUCU	bna-miR167d	*Brassica napus*
	vpl-miR167l-5p	22	UGAAGCUGCCAGCAUGAUCUGG	aly-miR167d-5p	*Arabidopsis lyrata*
	vpl-miR167m-5p	22	UGAAGCUGCCAGCAUGAUCUGC	ata-miR167f-5p	*Aegilops tauschii*

	vpl-miR167n-5p	21	UGAAGCUGCCAACAUGAUCUG	ptc-miR167h-5p	*Populus trichocarpa*
	vpl-miR167o	22	UGAAGCUGCCAGCCUGAUCUUA	gma-miR167k	*Glycine max*
MIR168	vpl-miR168a	20	UCGCUUGGUGCAGGUCGGGA	cme-miR168	*Cucumis melo*
	vpl-miR168b-5p	21	UCGCUUGGUGCAGAUCGGGAC	ata-miR168-5p	*Aegilops tauschii*
	vpl-miR168c-3p	21	CCCGCCUUGCAUCAACUGAAU	aly-miR168a-3p	*Arabidopsis lyrata*
	vpl-miR168d-3p	21	CCUGCCUUGCAUCAACUGAAU	sly-miR168a-3p	*Solanum lycopersicum*
	vpl-miR168e-5p	21	UCGCUUGGUGCAGGUCGGGAA	aly-miR168a-5p	*Arabidopsis lyrata*
	vpl-miR168f	21	UCGCUUGGUGCAGGUCGAGAA	bna-miR168b	*Brassica napus*
	vpl-miR168g-3p	20	CCCGCCUUGCAUCAAGUGAA	zma-miR168b-3p	*Zea mays*
	vpl-miR168h-5p	21	UCGCUUGGUGCAGGUCGGGAC	bra-miR168c-5p	*Brassica rapa*
MIR169	vpl-miR169a-3p	21	GGCAAGUUGUCCUUGGCUACA	aly-miR169a-3p	*Arabidopsis lyrata*
	vpl-miR169a-5p	21	CAGCCAAGGAUGACUUGCCGA	aly-miR169a-5p	*Arabidopsis lyrata*
	vpl-miR169b-5p	23	UAGCCAAGGAUGAUUUGCCUGUG	ata-miR169b-5p	*Aegilops tauschii*
	vpl-miR169c-5p	21	CAGCCAAGGAUGACUUGCCGG	aly-miR169c-5p	*Arabidopsis lyrata*
	vpl-miR169d-3p	19	GGCAGUCUCCUUGGCUAUU	aly-miR169h-3p	*Arabidopsis lyrata*
	vpl-miR169e-3p	19	GGCAGUCUCCUUGGCUAGC	ata-miR169i-3p	*Aegilops tauschii*
	vpl-miR169f-3p	19	GGCAGUCUCCUUGGCUAUC	aly-miR169l-3p	*Arabidopsis lyrata*
	vpl-miR169g-5p	21	UAGCCAAAGAUGACUUGCCUG	aly-miR169n-5p	*Arabidopsis lyrata*
	vpl-miR169h	21	AAGCCAAGGAUGACUUGCCUG	ptc-miR169o	*Populus trichocarpa*
MIR171	vpl-miR171a	21	UUGAGCCGCGUCAAUAUCUCC	ctr-miR171	*Citrus trifoliata*
	vpl-miR171b-3p	21	UGAUUGAGCCGCGCCAAUAUC	aly-miR171a-3p	*Arabidopsis lyrata*
	vpl-miR171c	21	UGAUUGAGCCGUGCCAAUAUU	mtr-miR171c	*Medicago truncatula*

	vpl-miR171d-3p	21	UUGAGCCGUGCCAAUAUCACG	aly-miR171c-3p	*Arabidopsis lyrata*
	vpl-miR171d-5p	21	AGAUAUUGGUGCGGUUCAAUC	aly-miR171c-5p	*Arabidopsis lyrata*
	vpl-miR171e	21	UGAUUGAGCCGUGCCAAUAUC	aqc-miR171d	*Aquilegia caerulea*
	vpl-miR171f	21	UGAUUGAGUCGUGUCAAUAUC	gma-miR171d	*Glycine max*
	vpl-miR171g-5p	21	UGUUGGCUCGACUCACUCAGA	ata-miR171d-5p	*Aegilops tauschii*
	vpl-miR171h-5p	21	UGUUGGCUCGGCUCACUCAGA	osa-miR171e-5p	*Oryza sativa*
	vpl-miR171i	21	UGAUUGAGCCGCGUCAAUAUC	cme-miR171h	*Cucumis melo*
	vpl-miR171j	21	UUGAGCCGCGUCAAUAUCUUA	gma-miR171p	*Glycine max*
	vpl-miR171k	21	UUGAGCCGCGUCAAUAUCUCA	gma-miR171t	*Glycine max*
MIR172	vpl-miR172a	23	UGAGAAUCUUGAUGAUGCUGCAU	aau-miR172	*Acacia auriculiformis*
	vpl-miR172b-3p	21	AGAAUCUUGAUGAUGCUGCGU	ata-miR172a-3p	*Aegilops tauschii*
	vpl-miR172c-3p	20	AGAAUCUUGAUGAUGCUGCA	csi-miR172a-3p	*Citrus sinensis*
	vpl-miR172d-5p	21	GUGGCAUCAUCAAGAUUCACA	aly-miR172a-5p	*Arabidopsis lyrata*
	vpl-miR172e-5p	21	GUAGCAUCAUCAAGAUUCACG	ppe-miR172a-5p	*Prunus persica*
	vpl-miR172f	21	AGAAUCUUGAUGAUGCGGCAA	csi-miR172b	*Citrus sinensis*
	vpl-miR172g-3p	21	AGAAUCUUGAUGAUGCUGCAG	aly-miR172f-3p	*Arabidopsis lyrata*
MIR319	vpl-miR319a	21	UUGGACUGAAGGGAGCUCCCU	aau-miR319	*Acacia auriculiformis*
	vpl-miR319b	21	UUUGGACUGAAGGGAGCUCCU	atr-miR319b	*Arabidopsis thaliana*
	vpl-miR319c	20	CUUGGACUGAAGGGAGCUCC	ppt-miR319b	*Physcomitrella patens*
	vpl-miR319d	20	UUGGACUGAAGGGAGCUCCC	atr-miR319c	*Arabidopsis thaliana*
	vpl-miR319e	20	UUGGACUGAAAGGAGCUCCU	gma-miR319c	*Glycine max*
	vpl-miR319f	20	UUGGACUGAAGGGAGCUCCU	cme-miR319d	*Cucumis melo*

	vpl-miR319g-3p	21	UUGGACUGAAGGGAGCUCCUU	aly-miR319d-3p	*Arabidopsis lyrata*
	vpl-miR319h-3p	21	CUUGGACUGAAGGGAGCUCCC	ppt-miR319d-3p	*Physcomitrella patens*
	vpl-miR319i	23	UUGGACUGAAGGGGAGCUCCUUC	gma-miR319i	*Glycine max*
	vpl-miR319j	20	UUGGGCUGAAGGGAGCUCCC	ptc-miR319i	*Populus trichocarpa*
	vpl-miR319k	21	UUUGGACCGAAGGGAGCCCCU	gma-miR319n	*Glycine max*
	vpl-miR319l	22	UGGACUGAAGGGGAGCUCCUUC	gma-miR319o	*Glycine max*
	vpl-miR319m	21	UGGACUGAAGGGAGCUCCUUC	gma-miR319q	*Glycine max*
MIR390	vpl-miR390a	21	AAGCUCAGGAGGGAUAGCGUC	cca-miR390	*Cynara cardunculus*
	vpl-miR390b-3p	20	CGCUAUCUAUCCUGAGCUCC	ata-miR390-3p	*Aegilops tauschii*
	vpl-miR390c-3p	21	CGCUAUCCAUCUUGAGUUUUA	stu-miR390-3p	*Solanum tuberosum*
	vpl-miR390d-3p	21	CGCUAUCCAUCCUGAGUUUCA	aly-miR390a-3p	*Arabidopsis lyrata*
	vpl-miR390e-3p	21	CGCUAUCCAUCCUGAGUUCCA	aly-miR390b-3p	*Arabidopsis lyrata*
	vpl-miR390f-5p	21	GAGCUCAGGAGGGAUAGCGCC	ppt-miR390c-5p	*Physcomitrella patens*
	vpl-miR390g	21	AAGCUCAGGAGGGAUAGCACC	gma-miR390d	*Glycine max*
	vpl-miR390h-3p	21	CGCUAUCCAUCCUGAGUUUUA	ptc-miR390d-3p	*Populus trichocarpa*
MIR393	vpl-miR393a-5p	22	UCCAAAGGGAUCGCAUUGAUCC	aly-miR393a-5p	*Arabidopsis lyrata*
MIR394	vpl-miR394a	20	UUGGCAUUCUGUCCACCUCC	ahy-miR394	*Arachis hypogaea*
	vpl-miR394b	22	UUGGCAUUCUGUCCACCUCCAU	mes-miR394c	*Manihot esculenta*
MIR395	vpl-miR395a	21	GUGAAGUGCUUGGGGGAACUC	osa-miR395a	*Oryza sativa*
	vpl-miR395b-3p	18	AAGUGUUUGGGGGAACUC	ata-miR395b-3p	*Aegilops tauschii*
	vpl-miR395c-3p	21	CUGAAGUGUUUGGGGGGACUU	aly-miR395c-3p	*Arabidopsis lyrata*

	vpl-miR395d-3p	20	UGAAGUGUUUGGGGGAACUC	ata-miR395c-3p	*Aegilops tauschii*
	vpl-miR395e-3p	21	CUGAAGUGUUUGGGGGGACUC	aly-miR395f-3p	*Arabidopsis lyrata*
	vpl-miR395f	21	UGAAGUGUUUGGGGGAACUUU	gma-miR395g	*Glycine max*
	vpl-miR395g-3p	21	CUGAAGUGUUUGGGGGAACUC	aly-miR395g-3p	*Arabidopsis lyrata*
MIR396	vpl-miR396a	21	UUCCACAGCUUUCUUGAACUG	aau-miR396	*Acacia auriculiformis*
	vpl-miR396b	21	UUCCACAGCUUUCUUGAACAG	mdm-miR396a	*Malus domestica*
	vpl-miR396c	21	UUCCACAGCUUUCUUGAACUA	pab-miR396a	*Picea abies*
	vpl-miR396d	21	UUCCACAGCUUUCUUGAACGU	ppe-miR396a	*Prunus persica*
	vpl-miR396e-3p	21	GUUCAAGAAAGCUGUGGGAAA	cca-miR396a-3p	*Cynara cardunculus*
	vpl-miR396f	21	UCCCACAGCUUUAUUGAACUG	cca-miR396b	*Cynara cardunculus*
	vpl-miR396g-5p	21	UUCCACAGCUUUCUUGAACUU	aly-miR396b-5p	*Arabidopsis lyrata*
	vpl-miR396h	21	UUCAAGAAAGCUGUGGGAAAA	cca-miR396c	*Cynara cardunculus*
	vpl-miR396i	21	UUCCACGGCUUUCUUGAACUU	pab-miR396c	*Picea abies*
	vpl-miR396j	21	UUCCACGGCUUUCUUGAACUG	cme-miR396e	*Cucumis melo*
	vpl-miR396k	22	UUCCACAGCUUUCUUGAACUGU	gma-miR396e	*Glycine max*
	vpl-miR396l-3p	21	GUUCAAUAAAGCUGUGGGAAA	ata-miR396e-3p	*Aegilops tauschii*
	vpl-miR396m	20	UCCACAGCUUUCUUGAACUG	gma-miR396h	*Glycine max*
MIR397	vpl-miR397a	20	UUUGAGUGCAGCGUUGAUGA	mes-miR397	*Manihot esculenta*
	vpl-miR397b	21	UCAUUGAGUGCAGCGUUGACG	pab-miR397	*Picea abies*
	vpl-miR397c	21	UCCUUGAGUGCAGCGUUGAUG	atr-miR397a	*Arabidopsis thaliana*
	vpl-miR397d	21	CCGUUGAGUGCAGCGUUGAUG	hvu-miR397a	*Hordeum vulgare*
	vpl-miR397e-5p	21	UCAUUGAGCGCAGCGUUGAUG	zma-miR397a-5p	*Zea mays*

	vpl-miR397f	21	UCAUUGAGUGCAGCGUUGGUG	atr-miR397b	*Arabidopsis thaliana*
	vpl-miR397g-3p	21	AUCAACGCUGCACUCAACGGC	hvu-miR397b-3p	*Hordeum vulgare*
	vpl-miR397h-5p	21	UCAUUGAGUGCAGCGUUGAUG	aly-miR397b-5p	*Arabidopsis lyrata*
MIR398	vpl-miR398a	20	UGUGUUCUCAGGUCACCCCU	ahy-miR398	*Arachis hypogaea*
	vpl-miR398b	21	UGUGUUCCCAGGUCGCCCCUG	atr-miR398	*Arabidopsis thaliana*
	vpl-miR398c-3p	21	UGUGUUCUCAGGUCACCCCUU	aly-miR398a-3p	*Arabidopsis lyrata*
	vpl-miR398d-3p	21	UAUGUUCUCAGGUCGCCCCUG	stu-miR398a-3p	*Solanum tuberosum*
	vpl-miR398e	21	UGUGUUCUCAGGUCGCCCCUG	aqc-miR398b	*Aquilegia caerulea*
	vpl-miR398f	21	CGUGUUCUCAGGUCGCCCCUG	ppe-miR398b	*Prunus persica*
	vpl-miR398g-3p	21	UGUGUUCUCAGGUCACCCCUG	aly-miR398b-3p	*Arabidopsis lyrata*
	vpl-miR398h	21	GGUGUUCUCAGGUCGCCCCUG	lus-miR398f	*Linum usitatissimum*
	vpl-miR398i-3p	21	UGUGUUCUCAGGUCGCCCCCG	ata-miR398g-3p	*Aegilops tauschii*
MIR399	vpl-miR399a	21	UGCCAAAGGAGAAUUGCCCGG	ssl-miR399	*Salvia sclarea*
	vpl-miR399b-3p	21	UGCCAAAGGAGAUUUGCCCGG	aly-miR399a-3p	*Arabidopsis lyrata*
	vpl-miR399c-3p	21	UGCCAAAGGAGAAUUGCCCUG	ata-miR399a-3p	*Aegilops tauschii*
	vpl-miR399d-3p	21	UGCCAAAGGAGAGUUGCCCUG	aly-miR399b-3p	*Arabidopsis lyrata*
	vpl-miR399e	21	UGCCAAAGGAGAUUUGUCCGG	bna-miR399c	*Brassica napus*
	vpl-miR399f	21	CGCCAAAGGAGAAUUGCCCUG	tcc-miR399e	*Theobroma cacao*
	vpl-miR399g-3p	21	UGCCAAAGGAGAUUUGCCUCG	aly-miR399e-3p	*Arabidopsis lyrata*
	vpl-miR399h	19	UGCCAAAGGAGAAUUGCAC	cme-miR399f	*Cucumis melo*
	vpl-miR399i	21	UGCCAAAGGAGAUUUGCCCAG	gma-miR399g	*Glycine max*
	vpl-miR399j	21	UGCCAAAGGAGAUUUGCUCGG	mdm-miR399g	*Malus domestica*

	vpl-miR399k	21	UGCCAAAGGAGACUUGCCCAG	osa-miR399h	*Oryza sativa*
	vpl-miR399l-3p	21	UGCCAAAGGAGAUUUGCCCCG	aly-miR399i-3p	*Arabidopsis lyrata*
	vpl-miR399m	21	UGCCAAAGGGGAUUUGCCCGG	sbi-miR399k	*Sorghum bicolor*
MIR408	vpl-miR408a	21	UGCACUGCCUCUUCCCUGGCU	cca-miR408	*Cynara cardunculus*
	vpl-miR408b	22	UGCACUGCCUCUUCCCUGGCUG	smo-miR408	*Selaginella moellendorffii*
	vpl-miR408c-3p	21	AUGCACUGCCUCUUCCCUGGC	ahy-miR408-3p	*Arachis hypogaea*
	vpl-miR408d-3p	20	UGCACUGCCUCUUCCCUGCC	ata-miR408-3p	*Aegilops tauschii*
	vpl-miR408e-3p	21	CUGCACUGCCUCUUCCCUGGC	bdi-miR408-3p	*Brachypodium distachyon*
	vpl-miR408f	22	CUGCACUGCAUCUUCCCUGUGC	ppt-miR408a	*Physcomitrella patens*
	vpl-miR408g-3p	21	UGCACAGCCUCUUCCCUGGUU	stu-miR408a-3p	*Solanum tuberosum*
	vpl-miR408h	21	CUGCACUGACUCUUCCCUGGC	sof-miR408e	*Saccharum officinarum*
MIR482	vpl-miR482a-3p	22	UCUUCCCUACACCUCCCAUACC	gma-miR482b-3p	*Glycine max*
MIR528	vpl-miR528a-5p	21	UGGAAGGGGCAUGCAGAGGAG	ata-miR528-5p	*Aegilops tauschii*
MIR535	vpl-miR535a	22	UGACAACGAGAGAGAGCACGCG	aqc-miR535	*Aquilegia caerulea*
	vpl-miR535b	21	UGACAACGAGAGAGAGCACGU	mes-miR535a	*Manihot esculenta*
	vpl-miR535c	21	UGACAACGAGAGAGAGCACGG	mes-miR535b	*Manihot esculenta*
	vpl-miR535d	21	UGACGACGAGAGAGAGCACGC	mdm-miR535d	*Malus domestica*
MIR845	vpl-miR845a	21	UGCUCUGAUACCAAUUGUUGG	bdi-miR845	*Brachypodium distachyon*
	vpl-miR845b-5p	21	UCAAUUGGUAUCAGAGCAACG	aly-miR845b-5p	*Arabidopsis lyrata*
	vpl-miR845c	21	UGGCUCUGAUACCAAUUGAUG	vvi-miR845d	*Vitis vinifera*

MIR894	vpl-miR894a	20	CGUUUCACGUCGGGUUCACC	ppt-miR894	*Physcomitrella patens*
MIR1507	vpl-miR1507a	22	UCUCAUUCCAUACAUCGUCUGA	gma-miR1507a	*Glycine max*
	vpl-miR1507b	21	UCUCAUUCCAUACAUCGUCUG	gma-miR1507b	*Glycine max*
MIR1511	vpl-miR1511a	20	AACCAGGCUCUGAUACCAUG	gma-miR1511	*Glycine max*
	vpl-miR1511b	21	ACCUAGCUCUGAUACCAUGAA	mdm-miR1511	*Malus domestica*
MIR2118	vpl-miR2118a-3p	21	UUGCCGAUUCCACCCAUUCCU	gma-miR2118a-3p	*Glycine max*
MIR2916	vpl-miR2916a	23	UGGGGACUCGAAGACGAUCAUAU	peu-miR2916	*Populus euphratica*
MIR2950	vpl-miR2950a	21	UUCCAUCUCUUGCACACUGGA	atr-miR2950	*Arabidopsis thaliana*
MIR3630	vpl-miR3630a-3p	22	UGUGGGAAUCUCUCUGAUGCUU	han-miR3630-3p	*Helianthus annuus*
	vpl-miR3630b-3p	22	UUUGGGAAUCUCUCUGAUGCAC	vvi-miR3630-3p	*Vitis vinifera*
MIR3699	vpl-miR3699	21	GACAGAAGAUAGACUUUGGUC	pab-miR3699	*Picea abies*
MIR4376	vpl-miR4376	22	ACGCAGGAGAGAUGAUGCUGGA	sly-miR4376	*Solanum lycopersicum*
MIR4995	vpl-miR4995	21	AGGCAGUGGCUUGGUUAAGGG	gma-miR4995	*Glycine max*
MIR5054	vpl-miR5054	18	UCCCCACGGUCGGCGCCA	bdi-miR5054	*Brachypodium distachyon*
MIR5072	vpl-miR5072	22	CGAUUCCCCAGCGGAGUCGCCA	osa-miR5072	*Oryza sativa*

MIR5077	vpl-miR5077	19	GUUCGCGUCGGGUUCACCA	osa-miR5077	*Oryza sativa*
MIR5083	vpl-miR5083	20	AGACUACAAUUAUCUGAUCA	osa-miR5083	*Oryza sativa*
MIR5139	vpl-miR5139	19	AAACCUGGCUCUGAUACCA	rgl-miR5139	*Rehmannia glutinosa*
MIR5141	vpl-miR5141	24	AGACCCGACGCGACUGACAGAUAA	rgl-miR5141	*Rehmannia glutinosa*
MIR5168	vpl-miR5168-3p	21	UCGGACCAGGCUUCAAUCCCU	ata-miR5168-3p	*Aegilops tauschii*
MIR5179	vpl-miR5179	21	UUUUGCUCAAGACCGCGCAAC	bdi-miR5179	*Brachypodium distachyon*
MIR5368	vpl-miR5368	19	GGACAGUCUCAGGUAGACA	gma-miR5368	*Glycine max*
MIR5523	vpl-miR5523	22	UGAGGAGGAACAUAUUUACUAG	osa-miR5523	*Oryza sativa*
MIR5538	vpl-miR5538	22	ACUGAACUCAAUCACUUGCUGC	osa-miR5538	*Oryza sativa*
MIR5539	vpl-miR5539b	22	AAGAAAACGGAUGCGCGUGCUA	osa-miR5539b	*Oryza sativa*
MIR5641	vpl-miR5641	21	UGGAAGAAGAUGAUAGAAUUA	ath-miR5641	*Arabidopsis thaliana*
MIR5770	vpl-miR5770a	21	UUAGGACUAUGGUUUGGACGA	gma-miR5770a	*Glycine max*
MIR6173	vpl-miR6173	20	AGCCGUAAACGAUGGAUACU	hbr-miR6173	*Hevea brasiliensis*
MIR6300	vpl-miR6300	18	GUCGUUGUAGUAUAGUGG	gma-miR6300	*Glycine max*

MIR6478	vpl-miR6478	21	CCGACCUUAGCUCAGUUGGUG	ptc-miR6478	*Populus trichocarpa*
MIR6485	vpl-miR6485	20	UAGGAUGUAGAAGAGCAUAA	hbr-miR6485	*Hevea brasiliensis*
MIR8155	vpl-miR8155	19	UAACCUGGCUCUGAUACCA	cpa-miR8155	*Carica papaya*
MIR8175	vpl-miR8175	20	GAUCCCCGGCAACGGCGCCA	ath-miR8175	*Arabidopsis thaliana*
MIR8741	vpl-miR8741	21	UAGCACUGAAGAUGAUGAUGG	gra-miR8741	*Gossypium raimondii*

Only 95% matched reads were considered to be known miRNA matched with the 73 plant species in miRBase v21.0.

Appendix H. Contig sequences of mRNA used in qRT-PCR analysis, in FASTA format.

>vpl_contig_86799

AAAGGGAACATCTTATTTTCCTTGAAATTAAATAGTAAATTTGATGACCTTTATATAGAC
ACCAATAGGAGAGAACCTCTGGAATAATCTATTAACTATTCCACCTATTATTTATATCCA
ATATTTATAGATTTGCCTAATATTCTATAACTTGTAACAGTTTCTCTAATGAAGTGACAA
TCAACTTCAATGTAATTGATTTGCTCATAAAACATTAGATAGCTTGCAATGTCAATGGTG
GCCTGATTGTCACGTCACTTATGCAATGCTTTGGATACAGAAAACCAAGTTTTTCACATA
ATTAAAAACAATAGCATTTCATAAGTTGTATAAGCCATGACTCGATACTCAACTTCAATA
GTAGACCATGCTATCATCATTTATTTTTGACTTTTAAAGAGATGTGAAGCACTATTAAAG
GATTATAGACAGACAATGAGCGTGGAAGGAATCAAAGGGCTAGAGGTTTGCAAAGGGCTA
TGGAGATGGTTACAAAAGCCAGTGTAGGTCTTCAAAGAGAGTGTTGAAGTGCTATGCAAC
TTTTCAAGAGGATGGGATCGAGTCTTAGGTCTCATTAGGAGAATAGAGAGGAACATGAAC
CAACACCAGCCAGAAAGAGAGGGAAGGGTTAGGGTTATGGCATGTGAGATAAGGTTGGGG
GAGGATCATGGGTTCAAGCCTTATACATAAGGGAAAAGTTGGGGCTACAACTTGTGAGAT
GAGCAAGGATTAGGGCTTTTGGTTGGTGTTGGGAAAGGCTTAGTAGTGAAAAATGATGTT
AAGATTGGGTGGGTCATTGTAGAAGGCAAAGTAAGATAAAAATGCTCTTATATCATTTAG
AAATGAAAATGTCTTGTTTTCTTTGATATTAGAGAGGAAATTATATAAAGAATAGACCTT
TGAAAAAAAATGACCACAACAAGGAACGAACTATAGAGTCACTAAACGAACACAACATAT
GGATGAACACAAGATAGTTTTACAGAACAAACAAAACACATTAGAGAGATCAACCAATTA
AAACATGTCCGTGACAATCAACCAAATCCATCACCTAGGATTGATTCATTAGACCAAGAG
GCAATGGAGGTTGTTCCAAAATCATGATGCGACTGATGACAATGTT

>vpl_contig_24863

GAACTGAAACTGGGGTAGGAAACCTGCATTACAATATTATTTCTCAGGATGGCCTTAGAA
GCAAAATCTTTGTGCATGCAATAAGAGAGAATTGAGGAAAGTAAACGATAACAGCCACTA
ATATCAAATGACCTTGCATGAAACTATTGAAAAAGATAGTATTAAATTTGTAGTAATTAA
GAACCACAAGAAAGGAGTCTCCAGTTCCAATCTTATCATTTTATACACAAACCCTTGATG
CTTAGTCCTCCCATGATTCTGTACAAGCAGTAGCAGTATCCCTTTTGAATAACAAGAATT
GAAATAAGTGCTCAAGGTGAAATTAATTACCTCAATTCAAATTTTTTTTCATAAAATATA
AAGGATTTATATAGAATATTGTTTTCTAATCATCCTAAACATATAATAAAAAACCTATCA
TATATAATTCAATTATTACAATCATTCAAAATTATTTGCTTTTAAGAAAGTAGCCTTTCT
ATTTCCTTTGTTTAAGGATAAAGATCCTCAAGGATCCTACCAAATCTCCAATACTCTCCT
TCAAGCTAATAAATCAATCATTCTCAACTTGCTTATGATAGATTCAAACTATTTTTTTTC
ATTGTCTTGAATAGAATATCTACTTTTTGTCGTACCAATAAAATATAGGTCAAGTTTATT
CCACCTCTCTCTCTCTTAATGAAGTTCCTGTCAATTTTCACATGTTTCATACAATTATGA
ACAAAATTTTTTACAGCACTAATTGCAAGCTTAAAAGAATAATTTAGTAACTAAAAGCTA
TAGATATTTAAAAAAATCTTTCTGGTTATATTAACTAACAAATTTATTGGACATTATCTT
ATGCTCAACTTTAGCATTGCTTTGTACAACACTTACTGTTTCTTACTCCGCTAGGTCACT
AGATTAACCAAATACTTCAGCATGAAAACCGACTTGCCATCTTCATATAAGCTGGCCCAA
GCTGCATCCATAAAGCCAATAATTCTTTTATCTTCATTTTTCCTAAATATTAGACCCTTA
TATGAAATCTTTTTAAGTATGTCATAATGTACATAATTCCTTTAGGATAATTTTGTGTAG
GTGAGTTCATGAATTGATTTACAACAATTACTATGCATAAGACATCAGGCTTTGCAAGGG
ATAAATAGATTAATTTTTCTACTAATCTTTTACATATTCCTTTATCAACTGTGGAATTAT
CGCTAATATTGTTTCCTTTCCAGGCATGCTCCAAGGGGTGTTTGTTAGTTTGATACAATC
AAATGGGTTTCTTTTAGAAGATCTAGGAATCTATTTCTTTTAAAAGATGTATATACCTTG
CTTTCTCATAACCACCAAATGCCAAAAAAATATTTCTTTTGACCTAGATTTTTCACCTCA
AAGTCTAATTTGAGCTTTTGTTCTAGGTTCACAATTTTAACAGTCATCGTCGATAATGAT
AATATCAACTATATAAACAATCAAAGTAGATATTTTGCCATTAAAATTGTTCTTTATGAA
CAAGGTATGATCAATTGATCTTTGCTCGTATCATCACTTGTTTGAGGGTTGTATTAAATC
TTATAAACCATGCACTTGATGATTACTTGAGCCCACATATTGGCTTCTTCAATCATCTTT
TTATTTTTCATTTTTTTTTCAAATTTTGGTGACATTCTCATATTTATCTCTTCAAAATTT
CCATTTAAGAAGGCTTTTTTTATAAAAAGCTGATAATAAGATCCTAATAGAATTTAACTT
TGATACTAGTGCAAATATTTCTTCATAATCAATCACATAAATTTGTGTAAAACCCTGGGC
TACTAAGTGGCTTCCTATCTTTCAAGGCTCTCGTCACCCTTGCATTTAATAATGTATACC
CATTTATGACTATCCATTTTCTTATTTGTAAGTAAATCAACTACATCCTAGGTATCATTC
TTTGTCAAAGCTCTCGTTTCTCTATTAATTGTAAGTCTTCTTCGATACTATTCGAAATAT
TGACTCCATCTACTTTCAACATGTAAGCTTCGAAATTGGTATATAATTTTGAGTATGTTA
CCAAGTTAGAGAGTGAATATTTGAAGCAAACAATAACACTATTTCAAAGTGAAATAGTGA

TATCTAAATCACTTGCGATCATTAGAACACTTTTAAGTCCTTGTTCAACTAGGAAATATT
GATGGCACTTGAAGGGATGCCTTGGTTTTCTACAACAAACATAACAAGCTTCTTGAATCT
TCTCTCTCCCTTGATTTTTCTTAATTATGGGTTCACATAGATGGTTAGAAATTAGTATGT
TAGGTGGATATAGGTTTGGCTTGAAGAACAATTTTCTATTAATAGTTGGAGGATATTATT
ATGAACAAACTAGTTCAAGTAGTAAGAGTAAAATGGGTAAGAAAATAATGGGATTCTAAT
GTTAATTTTTACCTAATCTCTTCTCCTGAAAAGATCCATAAAGAAAAATGAGGCTGAATC
TCAAATAATGTAACATTTGATTTAATGAAGAGCTTTTTAGCAAAAGAACAATAACATTTA
TATTCTTATTGGCTAGGTTATTATCCCATAAAGTCACAACTTTAAAGCTTCTGGTTAGAG
CTCATGATATTGAAAGTCATAATTATAAACATAGTAAAATTGGACACCTTTACAAGAATG
GAATTTAGTACATAAATAAGAAGGTATTGATTCAACAAGATGCTAAATGGTATTTGAAAC
AAAGTAGAACATGATTAGTACATATAATGGAACAAAAAACTTCAAGTGGATAACAATTCT
ATCTTTTAACAATATCTTTTTGTTGAGGGAGGAGATGTGTTTAGGCACAAGCACTTTGAT
GCAAGATACCATAGGATACCAAGTATTGAATAAAATCATTTGAGACATATTTAAGACCAT
TATCGTTATGCATCACAAAGATTAGAGTTTAAAATATAT

>vpl_contig_55903
AAAACATGCTATTTTTGACAAAAGAACTATCATGAATATTACTAAAACATTTCTTTATTG
AAGTTTAGATAAACCTTATACAAAATCAATACTAATATCCTCTTATGGTTTATGAAGTAT
AGGTAACGACATATATAAAGAAATATTTTCCTCTCTTTATAGTCTAGCAAATTTGACATT
AACAAACAACAAAAGTTAAATCTTTCCTTAATTACAGCCAATAAAATCGAACTTTTACAT
TAAAAACAACTTATTTCTACCAAAGTGTCTACCCATTTCATCTCAATATAATTTATCCAT
AATAAATTTAGAAAAAGGATGAGTCAAGAATATATAATTTTAACAACTTAAATAGATAGT
CATCTTTCATTATTCACAATCAAGAATACGTAGGTTATCTATTAATTCTTGAAAATTTTC
ATCAAAATCTTTGTAGTCAGCATGCCAAAATTCAACAACCTCTATGCTTATTTTATCAAG
CAAAAGCACCATGCAACTCCAAGCATCCATTAAAAATATAATTTTGCAAATACTTCATCC
ATTTTATATACCTTACATCGAATTTCTTTTCAGATAGTAAATATCTGAGTGCTTAATAAT
ACAAGAGCAGAATATATTTCTTTGCTAAAAGGTTATGATACCAGTGCTTTTGTGTTTGAA
TTACTACATCAAATTCTTTATTACAATTTGAGTATTTTTTTCTTGATTTCATTTGATTTT
TCCCTATAAATCCTTCCTAACTAAGAATAGCAACTACATTAACACCAGATCCATTACAAG
TAACCTCTAAACTTTTATTTAAATTCAATGTCCTAAAACATGGTGTTTGCATCAAAGATT
CTTTAATTCTTTCAAATACTGCAGCTCCCTTTTCCATTTGAATTCATTTTTTCTTAAAAT
CAATTACTGGAGTTATAATCATATTAAATTCTTACAAAAAAACTATAGTGAAAAGTAACC
AAACATTTACAGAACAATAATAAACATATAATTAATTCCTAAATGGATCATTTTACATCA
ATTTTTTCCTCCAATGAAATTCTCAACTATAGGATGAAGAACTTTAGAATACAAAGTATG
CCAAAATGCCAAAAAACATTTCTCATGCAAATATGAAAATAGCATTCGTCTTTTTGGTTT
TGATAGAGGTATTAAAGAAATTATGAACTTAAAACAAGTTTCGGATGAAGCATACTACCA
CAGTCATAAGGTACGAACACTTAAGGAAGAAAAGAAAAAATCAAAATGCTCCAAAATTAA
GACCTTAATTAAGGTTGCAGTTTCAATTCAAAGTGCTGCATTGTTCGTGAGACATGTCTG
AACAAAAGTGCAAATCATTATCAACATGCTAACTGTTTGTAATAAAGGAAAATATGCTGC
CATGAGTGTGATGCAGAGAACTGCCATTGACAT

>vpl_contig_32170
GTTCTTCTAGCATTGTCACTAAGTCTAAACAGTCTCATCTCCCACTTATCTAGTCCATTA
GTACTAGCCACTTCCCCTTCAACTATTTTATGCGAATGTATGTCAAACTCCTCAAACTTC
TTATAATGGCTTTAAATATTATGTTTTCTTTTTAAATGATTTCTTCAAACATCTTAGGTT
TTAACCCATGAATAATAAATTTGACACTACATATATATTCCTTTAATTTCATAAAAATGT
CAAAAACCAATATTATAAAGAATTTATTAATTTTCTAACCAATAACAACGGTGAATTTCA
ATCCCTACTTTCCTATCTTTATATAGATGAAATCTCAAATCACTTGTCCTCAAATTCAAG
GGCAAAGTGGCATAGTTGAAAAGCATCATCGCCAAATTGTTGAATTGGGACTTACCATGC
TTGATGATCCAACATTACCTCTCTATTTATGGAGTTCCGCTTTTGAAAAAATTGTTTTTA
TCATTAAAAAGCTCCCATTTAAATCCCTCAATAGGGCATCTCCTCAAAAATGTTTTTTGA
AAATGTATTTGATTACCAATTTTTGGCGTTCATTAGATGCTCATGCTTTCCATACCTCAA
GCTTTATAATAAGCACAAACTTGAATATAAGTTTTCCTTATATGTTTTCTTGGATAAACC
AAATAGATATGGCAACAAATTTTTATATTTGAGTTCCTATTGTGTTTACTTATGCTATAT
GAATGTTTCAATGAATAGAAGTTTCCATTTAAGACCTCTTCACTTGCTCTACCAAGTCCA
TGCCCTCCTAAAATTCAATAAGACCAAATCTTT

>vpl_contig_18270
ACTTTTATGATCGATGGTATCTGTAAGTTTCTATCGCTGGAGTTTATTTGTTCTTTTTCG
TTGTTGAATTTGTTGCATGGATCTGTTTATGCTATTCAATCTCTATTTACTTCTTGTAAA
ATGATGTCTGAATTTACTTTTTCTTGACATGGTTTATGCTGATCCTGGCGTGGTGCCGTG

TAGCATTGGAAATGTTGAAAATCAATTTTGTATTTTGAGGTGAAGCAAGAAAAAAGGGGA
AGTGAGGCTTTTAGATGGGTGGAAAACCATCACAAGAAATTTGTTAGGTTCTGGACAGTT
ATCTCAAATTGAGATCACTTTTGTAATCCCTGAATGCGGGATTCCTTTATTTGATTAGTG
TAGGCCAAAGATGATAAGGTTTTGTGCCGTGTGAAGATAAGGTTGTTTTGTGCGTGCTAC
TGATAGCAATTTTTTAGGGGGAGAGTTTGATTTGCATTTTGTGTGCTGGCCTGCCATAGT
TGATTGCGAATTGACAAGACATTTTGGATATGTTATGATTGATGCTGTGAGCCATGTGAC
AACAGCTTCTTAGTAGTGTCTGTAAAGGCTAGGTTGTGGCATATTGGTTAGTGGATTACT
CTATGTTGATAAAAGGATGCTTATAACATTTGTTGTGAATAATGTAAAGTAAAATGGGTT
GTTATGTTTCATGGCTCCTTTACACATTAACATGTTTGCCATTCCTTTTTTTTGTACATG
ATGATAAGACCTGTAAATTACTGTCATGAAAGAAATTGAGTTTAAAACATTAAAATTTCT
AATGATGTTGGAGCTAAACAAGAAAGATATTTAAATTCATTTCATTAACTTGTCTTGTTA
GGGGACCCTTATGGTATCTTTAGTAGCTTAAAAGCTTGATTTGGGCTAATGTTTGCATTT
GTATTTTCAAAGACTAGTTATAGGATAATTTTAAAGCCTTGTGGCTCCCCAATTCAATTT
TTTGGTTATGGTAACTTTTTTCATATTAATTGACATGATAACTTTTTGATATTCAAAGTT
CAGTATTTAGGATAGTGTTTTCATGATTTTTTTTACTGTAATGCTTCTTTACTTTTTTTT
TTGATGAAATTATGATATTATCGAAAGAGGCATAAGCCATACAAGAAATCCACATGTGGG
ATCCTCAAGGAGGATCAAACATGAAATCAGACAGCATAAGTTTGAGTAGGGGCTTGGGAA
TGGATTCCCCGGGATCTTATACGAACTCATTAAAGCATGCAGCGCGAGCCCCCCAGTCAG
TAGGTGCATTCCCATCTCTATGGATATGGGAAAACCGGACTGTCTCCAGACACCTGGCAT
ACCTAAGAGTCTCAGCAAAGTGGCCATCCACAGGGGTTGGGATATCATCTCACGGACCAC
CTGGAGCGCCGTCATCGAGTCAGTCTGAATCTCCATGGCCCGGTGAAGGGAAAGGGAACC
CCTTAGCATGTTGAGGGCCTCGATCATAGCCTGGACCTCTATCATATTAGCAGCCTCGAA
GAAAACCCTCTTGACAGCCATCAAGAGCAGCTCCCCATGCTGTCTCTGACAACCAGGCCA
AGACTGACAATGCAATCATCTACCACAGCATTGCAATTCACTTTTATCCATCATGTTGGG
GGGGAATCCAACCATGGGTTACAAGCAAGTTCGAAGAGTTGTAGGAATTCCCATTGGATA
GGGATGAATAGAAGTAATGTTGCCTGATTGCTGTAATCACCTCGATCTTGATCGTCCGTA
GCAGTGAGCACGGCATGGCCATCGGATTATTATGCCTCATCGCATTCCGCCTTACCCAAA
TGTAATACATAGTGGTGGCCAACGTGGCAGCATGCCTAGCACCATCACCACGGTGCGCTG
TCGGCGAACCCACAAACCTCAAGAGCCGGTCAATGAATTGAAGCCCAGCTCCCGCACCTG
TGGACCGACCCAGCGTAGCCAATGCCCCGCGAGCCTGGCAGTCAGCTCAGAGTTCCTGTG
TGCGTGGGCAGTCCCAAAACACATGTTGCAAGGATTCATCCACCGCACAGCCCCATTGGC
AGTTGCTGTTCTGGAGCATTCCCCTGCGGGCCAGCCACGCCGCCGTAGGAAGGAAATCAC
CGACGACCCGCCACATGAACATGCGCACCCTCGGCATTGCTTTAATCTTCCAAATGCCCC
TCATCTCGTCTGCATGATGCCGGCCACCCCGCCTGACGGTAGCATTAGCACCGCACACCA
TGTGCGTGTAGATTCTTCGGCTATCGACCCTCGCCGTCTCGATGTGCAGCTGGAGGAGCA
CGTCCTGGGCTGACGTCCGAGCTCTCGAAAGATGCAGAATTCTGTCAATTAGCTCCTCGC
CCGCCCAAGCATGGATGATATCCACATCCTAGTCCCTAGCAGCCGTAATGCAGTCAGCAA
CGCTCGCCGGCCAATTGATATTCATGTTAATGAAGGTCGGCCATTAGCAGAGTGGAAGGA
TGCCCATCCAATTGTCCTTCGCGAATCTCACCGTCCTCCCATCACCGATGCGCCAGACGA
AGATTGATATTGCCACTGGCATTGCCTCCTGCACCATTCTGCACACAGGAGATGCGTTTC
GGGGAGCTTTCATAAGAGGATTCATCAGCATGTACTTCTGCCTCATAATTTCAACCCAAA
GGCTCGCCTCGCGACTGAGAATGCACTTGCAATTGTTCATTTGTAGCACGCGCTGCCATA
CGCGGAAGTCCCGGAGTCCAAGCCCGCCAACGGCCTTCGGCAAGGTGATACAGTCCCATG
GCGCGTAATGCATCGTGTGCGCGCTGCTCGTCCCGCTCCATAGGAACCTTCTAATTCTCT
TTTCCATGTCATTAAGTACTCCTGCAGGCACCGCACAATGCCGCAAAACGTGATTTGGGA
TGGTAGTAAGAATAGAGGAGGCCAAGGAGGCTCTGCCCGCAAGCGACGGCGGACAGATAC
GCTAATTGACCACTGCAGCGTCAATCCTTTGGACCAAAAAATTAAGGTCGGCCTTGTTCA
GGGGCCTCGCAGCAATGGGGACCCCGAGGTACGTGAGCGTCCGGCCTTGCTCCATGCGGA
GTTTCCTCAAGACTAGCTTGGCTTTCTTTTTGTTGAACTTTGTGCTGAAGTATACTTTGG
ATTTCATAAGGTTAATGCGCAACCCTGCAGCCTCACCGAATGTCCGCATGATACTACGCA
CAGCGCGCGCGGCCTCGCGCGTGGCGGGAGCAAACAGCAGCACATCGTCTGCGTATAGGA
GGTGCAAAATCCGCGGCGCGTGTTGCCATGGTCTGAATCCAAGCAGCTGTTTCGTCTCGT
AGTCCCTCACCAGCCATGAAAGGAATTCAGTAGCCATGACGTACAGGTAGGGCGAAAGCG
GACATCCCTGCCGGAGTCCCCGCTGCGCCACCACCCAAGGGAGATGAGAACCATTGACGA
GAACACCAAATCTCGGACAAGTTATGCACACAGCGATCCATTCACACCAGACAGGGTGGA
ATCCCATGATCTTCATCATCTCGAACAGGAAGTCCCAACGCATGCTGTCGTAGGCCTGCT
CCATGTCGATCTTCACAGCCATGAGGCCCGAAGCGGAAGTGCAAGCATGCATCTTTGATG
CCATCTCCTGTGCGATATAAACATTATCTGATATACTACGCCCAGGGATAAATGCTGCTT
GCTTGCGTAAGATGAAACTTGACATCAGCGGTTTCAGCCTCTCGACCAGAATGCACGGCA
CCACCTTGTAGATGGTGTTGCAGAGGTGGATTGGCCAGAAGTGCTCCGGCCTGCACGCCC
CCTATTTCTTCAGGATAAGGACAATGAGCGTGTCATTCCATTCCTCTGGCAACCTCCATT

CCTCAAAGAAAGCCTGCATCGCGCTAATGACATCTTAGCACGTGATGTCCCAATAGCCCT
GCAGAATCTCCGCCTGATCCCGTCGATGCCCGGGGATTTGTTCCTCCTGAGTCCCCGCAT
GGCCGCAATAATCTCACTTTTGGTGACCTTGCATTCAAGTTTGGATGCCACATCATTGCA
GATGATGCCAGATAGCATAGAGATCGGCAGCAAGGCCACCCGGCCCCCGCTCCGCCATTT
TCCTTCAAAGTAAGCCAAAAGAACCCTCCTAATGCTTTTAGGTGAATCCGCCACCTTGTC
TTTTACGTCCCTTAGCTGAGTAATTTGGCTTGTTCGAGCACGAGCCGTGGCAGCAGCATG
GAAAAACCTGGTGTTGCTATCCCCGAGCTGCGCCTGATGCACTTTAGTGCACTGCCACCA
CCATTGTTCCTCATGTGCACAACCGCGGGGGAAGGCCGCGACCATTTGACCAAGTCGCAT
ATGCTCCTTTTCTGTTAGTCCCTCCCGAGATTCCTTTGATTGAAGTTCATTGATTCCTCT
TTCCAGCTCCCCGTTCCTATTAACAATGTTTCCCACCTCCTTCCGGCTCCATCTCCGCAG
TGCCATCAGCGAGCACTTGAATTTAGTCATAATCACCTCAATGCAATTTCCCCCCACCCG
TCTCTGCCATGCACGATTCACCACCCGGAACGATTGTGGAAAGTCTAACCACGCACTCTC
AAATTGCGGCCTGCCATTGTTCAGCCGTGGATTCCTGCATCCCAGGTCAACCAACAGCGG
GCAATGATCAGACCCGAGCCGGCAGAGATGTTGCACCGCAAAGTTCATGAGGACATGGTC
TAATTGCTCCGGGATGCATGTGTCACCCTGCTTCCCGTTGCACCAGGTAAAGCGATTGTC
GGTGAAGCCCATGTCCACCAGCTGCAACTCCTGCATAAAGTCAGCAAGTTCTTGCTGTCC
CCTGTCGTAGATGAACTTCCGCCCGCCCCGCTTCTCCTTCTCTGCCATCACAATGTTGAG
GTCCCCTGCGACAATGAGTGGGTCATCCCCACTGGACGTGTGCCTGATGGCGTCCCAGAC
CTCCCAGCACTTATGGTGGTCTTTAGGGTGCTAAACAATGAAGCAATTGATCTTGCAGTC
ACCACGTAGCCTCACCTCGCACGCCACAACATGCGTTGAGGCCGCCATGATCGTGATATC
AAACCAGTTTTTCCATAATAGAATGATACCTGCCGACATGCCGGCAGAAGGGTGAGCAAC
GAAGCCCGTCCTTCCCAAGGAGCCAATTCGCATCCTGTCGTGAGAAACAGTGCGCTTTCG
TCTCCATCAAGGCCACGACTGCAGCCCCTCGCGTACCAATCTTTGCAGATATCTTCTTGA
CTTCTTGTCCCTAACACCTCTGATATTCCAAACCACGAACCGCATCTTCACTTGGAACTT
GTCATTGGTACCCACATGCAGCCTCGCAATTCTCCCGCATGAAGCCTCCTCCATCTTCCT
CCTCGATGACCCCTACGCCTCACGAAGCGCAGCCCAGGAGAAATTCTCCTATCTTGTTTT
GGTGGAGTGTCCATCTGGGTTTATAATTATGTCCACCTCCCATTCCTCCTGTAGCAAAGG
CAATTTCCCTTTGTTCCCTCCTACCTGCATTGCCTGTCCCGTTCCCACATCCTTCATGGA
AGCATCTGCCTGTTTTATCCGCACCCAATCCTCGCGGGCCCTTGGGTGCGAGGCCTGCCT
TGCCATACCTCTACCACGGTCCATGGATGGGCTTCTCATGCCGGGATACATGCTAGCGGC
CCCCTGGTCATGAATCCTGCGATTGACGGTCTTCCGAGGATTCTGGCTCACCTTGCTTGC
TATGCTTCTGTGGTCAGCCACGCCCTTTCTCCTACGTCGGGGTTGGACAAGCATCCATGG
GCCCAACTGTTGTGCCTTTTCCTTTGGAGCATACATGTTCTCCTCCCGCTTCCGAGGACT
TCTGGCGCATTCTGCCTTAGTGGTCTTCCCGTCACCATTTTACGTATCCTCCACCTCCAT
CAGTCCTTCTCGAACGACATTATCTTCTGGAGTCACAGAGGAAGCTGGACATGCGTGCAC
CGCTGCCATGGGCTTATGCGTGTTGCCCTGGCCCGCGTGGTGACCAACGTGCCCGCATGA
GAAGCAAATCCGCGACAATCCCTCGTACTCGAAAACCTGGTAGGTTCTTCCAAATTCTTC
GTTGAGCCATATACCCAGTTTCAGAGGCTTCGAGAGGTCCATCTGGACGCACACCTGAGC
ATAGGAGGATCGTTCCATCTCAGCAGTTCAGTCGTTAAGGAGAAAATGCTCCCCAATAGC
CTTGGCAATTCTTAAAATGTTCTCTACGTCCCAATATATGAGTGGGAGGCACGGAAAGCG
CACCTACACCGGTGCGCACATCGCTGTTGCGGCGTTTGGTCCGAATTCGTGTGTCCAGAT
ATCAATCCCCATGGCTTGGCCAGCGACCTGCCACAGGCCGCCAAGCCAGACTTCATTCCT
GTCATTGATGGAATCGAAGTGCAAGCAAAAATAGCCGGATGACAGGTCTAGAATCCTAAA
ATTCTTGTAGTGACCCCATCTCTTTAGCAACGCGCGTGCAATGACGGGGAATGGGAGTGT
GTTCCCGTATACTCGCCCAACCAAATCAAATTTGTGAAATTTGATACGCCTATGGCAAGC
TTCCGATGACATAAACACATCACCACCCTCAACGTGATCAGGAACACCCGCAGCACGATC
CTTGACATCGATATTGGCCGGACACACATAGGAGGTTGCACGACTATTTGGGGCAATGTT
GTTTGCAATTAGGGCTTCTCTGAAACTAATCCTTGTTGCGGGAACCTCCTGTCTCCCAGC
AGCCCCATGTTCGGGGTTGATCGTTTCTCCCGAGAGCTTGCTCACTGCAGGCTCCGCAAC
CATATGCGCTGTGTCCTCTAGCTCCATCATGCTAGTGATCGCGGTGTGACCTCCATTTGC
TTCCCCACCTCCGTCCTCCTGGCTCTAACAGTCAGTTAGCTCAGCTGGAGCTTCCACATG
TTCCACCTTGCTTGCCGGGTAATCCCTGAAGTCATTCGTGCCAATAGTTTCGTTTAATAC
GTCCGCCGCTGCATCTGAATTTTGAGAGAAGAGGGGATTTCCTTCGGCGATCATCGTTTC
CCCTCCAGAAGTGGAAGAGAGGCATTCTCTCACTGGCAAGTTGCTTCAGTGGTCCTCACC
TTCCTGTTGCTCCTCCTAACCTTGCCTCCCGCCATCTTTTGTCCCGTTAGGGCTGCAGGT
GATCCAGAACCATGCAAAGGGACGATCCGGTGTAGTGCGCCGTAGAGTAACTTTGCAGTT
TAAAGCTGTCTCCACGCCCCCGCTTCAGCCAAATATTCAGAATTTCAAATTCAAAAACTC
TCGGATCCACGTGACCCGACAAATTTCCACAACGGTCTAGATACCTCACGTGCCCACTTG
CACGTGCGACCGTTAGCCGTTACCAGCTGCCTGTCACGTGCCCTCCTGCTCCTCCGTCCT
GCCGTGAGTTCCCCTGATCCAGTCTCCGACGGTGGCCCGGTATCTCACCCTCCGTTTACA
CCTCCTCACCAGCTGTAGATGCCAACTACTCCTCACCTGCTATCCACGACTCGCGTTCCG

TCCCGTGTGCTTCGCCACCAGCAATCCGTCTCCAGCCCTCAGTCACCGCTCTCTGTCCCC
TGCTGCCAGCTCCCCGTCGCGAACTCCTTCCTGTCGCTGCCTTCCGATCGCTGCCTTCCG
TCGCTGGCTCCCGGTCGCTACCTTCCCGTGACTCTTAAGTTTGTTATGCTTCCAACTTTT
AAAAGATTGTTATGGCATGCCCCATTATGATGTTCATTATGTAGTACTTGTTTTTTGTGT
AGAAGCTTATCACTAGTTATAACAAGGAACAAGATAAAACAAATTCTTATTTTAATGGCC
TTAATGATTGTCCATTGCTAATTAGGTGTTGTAGAAGTAGATCTACTGCAAGGAAGACAC
AGTTGGTCATTTACCTCTAGTTATTACAATGGACAATATAAAACACATTCTCATTTTAAT
GGCGTTAATGCTTGTCCATTTCTAATCAGGTATTGTGGTGGTAGATCTACTACAAAGAAG
ACAAAATCAGTCATTGTCTTCACTTTGTAGCATTCTTTTACAGCTACTATTTTAGACTGG
TTTAGTGATCCATATTTCGTTGTTTGTTCTTGTTGGATGCTAATTAGATGTTCTAGTTTT
AAACTTTTTTTTTGAAATCTTCCTAGTCATATTTTCTTTGTCTTTCCATGTATAATTAAT
TACCTTGAATTTTGCAAAAAATAAAATAAAATAAAAATAACAAATGGTACAACTAAATTG
AAATGACCAGTACAACCAAATAAACAATCTTGTCAAAACATTAAAATCTCATAATTGAGA
ACGTATTAAGTTTTGGTTTTTATTATCCTCGACTGAGCATCCTTGTTTGATTATTTGAGA
GGAGAGGAGAGCAGAGGGGAGGGAGGGTCTCCTATGAGATTAACAAGAGCATCGACTCAC
AACATGAGCTTTGTTGAGACCTTGAGATCAGTTCGTGATTTCTTCGAGATGGCCTTTCGG
TGTTCAGGGTTTTTGGTGGGTATTATTTTTAGGAATAACGTCTCTGGAGCATAATAACCA
ATTTCAATTGGTTAATACTAACTTGTTTGTGTAATAGAATTTAGATTGGGCATTAGCAAG
CAGTCGCCTTTTGTTGTTTTTATGGCATTGTAAAAAAGAATTTGGGTTTGTGACTTTGGT
ATTTCTCTTATATGCATGCTATTTTACTGACTATTTTGTTGGTGATGAAATATTAGCTAA
GCTAATGTAAGTATTTTACTTTTAGAGGGCTTACTAACATTATTTATGAAGATGCATTGT
TGTTAAAAAAGATATTAGTGTTATGTAAATTTTAGGAATTTGATGATATCAAGAAAATTA
AAGGAATTAGTATTGGATTGATAGTGAAGGTTGAGTACAGGTCCAATAATTGACTTTAAA
ATTAGGAATTTTAGATTTTTTGATATTCACATAAAACTTATTATGAAAGGACTTGATATT
ACTTTGTATTTTAATCATTATGATTGTTCATAACATGAATCTTATACCTTGTTGCTTTAT
CAAATGTGTTTTTAGAACTTATTTATTAGTTTATGAGGA

>vpl_contig_14640
GATGGAGCCCAGCCAGCCCGAGGCAGCAGAAGACGAGGATGAATGCCGCGCCAAGAGGTG
TGTCCTGCGAAGTAAAGCAAGGAATAGGGCCGTCACAGTGTCAGCAATCTCCTCGGCACG
GCTTACATCGACGTGGACAAGGCGGAGACCGAGATCAGTAGCGAGGGCAGAGTCCACTGC
ACGGTCAGCTGAGCCGAGGCACAGAATAAGCTGCCATGGCTGTAGCCTGCGCTGGGCGGC
GCGGGGCAAGAAGGCGAGGGAGTGGAGAAGCACGACAGCGGCAGACTCGATCCGACCATC
CGGGAGGCGGGAAAGGCCAACGTGCTCCACGGAGGCGATCCCGGCTAGAGACTCCTGCTC
AAAGGAGGTGTCCTCCAGACAGTTGAGCGTCACGACCAGTGGGAATGACGCCATCGGAGT
TGGATCGACAGGAGATCGGCCGCCGGCAGACCGCTGGATCTTGCTGCCGTCGCCTCTTTC
CCTCATCAACGTTTGAGCTTTTTATCTTTGCGCTGTCTCCTTAGTGACGCGACGGCTTTG
CTGCTTTTCCTTTCCGTGACGTGGGGAAGTGTGGGATACT

>vpl_contig_contig_27274
TGAAATAGAAATCACTGCAAACCCGGCTGCCACTATTTCATGCTATGACTTCGTTTGCAT
TGCCATCCAACAGGTTCTAAATTGTACAAACTTCACTAATTTAAAATACAAGGTATAAAG
AGATGCCTTTTGAGTTATTTTATTTAATCATTTAAAATGACCAGTTGAAAACAAAAACAA
ACCGCATAGCAGTATGCACTTGGGGTTAAGGTGGTTTGCTATTCATAATACTCATAAGAT
ATATGCACAAAGGTTTCTACGTGGTATCAGATTTTTGGTTTTTTTTTCTCTTCTTTTTGA
AATAGTGGACAAAATCTATATGAATAAAAATGGGTCAGGTAAAATACTAACGAACCTTGG
CCCACATAAATCAAAAAATACTCGACTCAAGAGAACAAATAAATCAACCATAATACTATT
CACAAAAATAAATACAATAGATCTATGTGATGTGAAAAATAACTATATGATAAACTCAAT
CTCAACTAGAGCAACAATAATATGGTAGGCAAAAAAGGTGCCATGAACATTTGTGAACAT
GAACTCTCATCATGAATATCAACATGAGCTCTACGGCAATCTCCATGAATCTTTCATCGA
CAACCATTACATTACCAATAACCTTCCATACATCACATCATTTTCAATCTCCCATGGAAC
TAAATTTATTAGCATTCTCAGTTAATGGATATTCTAATGATTCATGATTTTCAACGTATT
TTGCTTTAGATTTTGTAGCCTCAGGTGACAATCTTGCTTCTTGACTGCAACAAGATCATT
TAATATTGAGAAGGATTAAAGCAATTATGTCTTTAATCAAAACTATTTTGATTTCCTGCA
ATTTGATCTTGCAAGCATGTATTCTTCTTAAAAAACATCTATCACCTATACGCCAAAACC
CACATTTGAAACATTCAAGATCAAATAAAATCACTTTACAAAACCTTCAAAATCTCCTTG
CCTAATTTTCTCCTCCAAATGAAATCTCTATTTGATAATCTCATTTTTGTTGGACACTAT
TTCCGATTTCAATCTCATCAATTTTATAGTTGATGAACTTAGTTCTAAGCACCATTGTTA
TAATTCTATCCGAATATCTCATTTGATAATCTATACAACCTCTGGCTTGAGGAAGAAAAG
CTAATCAAATCACTTACTACTCCATCCATCCCTTGTGCTCGTGCTCCAACAAACCCATCC
AATACCTTTGTTTTCAGTAGTTGTAATTGAGGCTGTGGCAATGCTGCAAATCATCTTGCA

ATCCTCTTGAAGAATCATTGCTTTGTATTTCTCCAATAGTTATTATTAACCACCACTTCA
TCCCACTCTTAGAATCATTATTGGAATGCTCATCAAATTTTGTAACAAATATGGTTATAA
GGTTTGTGTTTGCCTAGATCTTGAAAATTGCATATGTAACTACAATGATCTTTGCTTAAA
TTTAACAACCAGCTTGAATGATTTGTTTGATTCACATTGGTGCCTCAACGTAGAAGTAAC
GCATCACATAACCCCTTCAAACACTTCTCTTCGAAACCATGACCATTATCATGATGTTGA
TACAATCACAATCAACAACATAACTCTTTACCCATGACACACTTTGGGCCTCAATGGTAA
ATTTCTACTAAATGAAATTCTTTATGTTCCCAAGTTCTTCAAATATTTCAAAAAAATATT
TATTGGTTTAATTTTTTCCACCTTTGTGGAAGGTTCACATAAACTTCCTTTGAACACTTC
GCATCTAGATATTTCGGGTCACTCTCCAATCCCTTCAATTTGAAGTTATAGGTTTTATAG
ATTCATTGATGATAATAAATCCCAATATATATGGTTTTCTCACATGTTTTATAAATCTTT
CTTCGTTCTTTTGTATCTTCAATGCTTTCAAAATTGGTTGCAAAAATCTTAAATTTTTAC
TCAATTGTATATGATTTTAACAATTTTCTATTGACACATTCAACACAAAAGCAATTTGTT
TGTCCCTATATTGCTCAACAAAATAGTGTTGTCAAATCTCAATATAGTCACATTGTTGAG
ATGCATCATTGCTTGCTGTCATAATCCGTAACACACCACATTATTATTGGACTATTATTT
TTGTTATAATTGTTTTCTCGATTAATAATCTTACTACACCTAACAAAAAAAAACACAAAT
CATCTTTTGAAGTAAATTTTTTCCATCCATCCAGTCACACACACCTAATATGTCCATAAA
CTCATGCCTAGATCCTTCCCATGTGTGCTCGACTATAACAATGCCTACAAACTCTATCAC
TAACTAGAGTTCAACAAAACAAAAATTTAGATATCCAAATCATAAAGTTTGATGAAATGA
TTTTTTACTTCAAAAACTTCACCACAAAAAACAAAAAACCAATAATATCTTTGATTTTCC
CATTCACCCCAACCATTAACTATAACACTCCTACAACCTTAACTATAAACATAATTCACT
ACACTAGCCTTAAAAATACCAATGGTCCTTTGTAATAGCCGATCTTTAATACATCCAATA
ATTACTATATTATAGTTCCTGACAATCAACTCACTATTCAAAATTCCAACTTGGTTGACA
GAATTCCAATTCCATATATATTGTCATACAAATCCCCATTCTCTTCCTCAAGTCCCTTAC
TCCTCCATAATTGATCGTTTGCACTTTGAAAACACATTCAACCAAATAACAAAACATCCC
CATTGGTTGTGAGCTTCGCAAAATGCGTACAATGCTCCCATCAAAAATTCCACTTGCCAA
TTGATTCTAGCCTCACACGATATGAATATCATTGGGAAACAAAATGGGTGATTAACATAA
AATGTCATGCAAATAACACCATAAAGTGGTATAAAGTTCGAACCATTGCAAAGACTACAA
TCAATAAGGACTTGACAGTGATGACACTTTAGTCCTATTGTCAAGGTCATAAACTTTCAA
ACTGTTTTTGTTATTGTTAGCTTCAGCAGTTGCCTTTTACATCAAATTGATGTGCAAAAT
ATCTTCTAAATGGATTTCTTTTTGAAGCAATGTTTATGTGCAAAGCAAGCTTCCATGACA
AAAAAACCTCACCCTCATCTCGTATCATAAATTTGTAAGACAATATATGAAATCAAACAA
ACACTATGATCATGACACCAATGTCAGCATTCACAACTCTTATACAGCTTGCTTTTGCAA
GTCTTTTGAATTCTTATACTCTTACACAATAATGGTTTATCTTATATTTTTATTTTCGTT
TATTTACATGACTACAATGACTAGGAAAAACCTTATGAAAATTGATAATATGATACTAGC
TTTCTACAATACATTTCAATGTGAGGATCTTGGCTCATGTTTTTTTTTGTTGGTATGAAA
ACATCAAATGTTGCAAATGGCATGATGATTTCCCAAGCAAGAGAAGTGTTCAACCTCTTG
TAAAAATTTAATATGGCAAACTATATACCTT

>vpl_contig_1034
CATTAAGTGCTTCCATAGCCACTGGAACTCATCAGGTATTGCACCGGTGCAAGTTTGGTG
GCACAAAAAAGTACAGAAACAATATAAACAGTTCTCCGACAACAACCTTCAAATACATTT
ATTAACTAAATCACAATATACTTTGTCAACTTGTGAACCTATTTGTCAATAATGACAATC
AAAACCTGCCCTTCTTGTGATATTGTTCAAATTATAGTATCTCACCTACAAGCTTTGATT
AAGTTATCCCCTGCTTTCAGAATGTAATATACGTTCACATTTTATGCTCAATCACTACTA
CTTAGCCTTATGATTAATAAAAGACATTCCAAAGGCACCACGGCTTCATCATTACTCACT
TCATTTTCCTTCTTTTAAAGCACATAAAACTTTACTTGCTCAGGTACATTACACCAAACT
TGCCACAATGTTTTTACAACTAGGTGCATGCATTCTATTGGTACGAATTGACCGCTGGAC
AATTTTGGTTATCAAAAGTTATAAATTGAATCAAGCTAAACTCTAACAAAAAGAACATAA
CTGCTCAAATTTGGTTTATTATCAAATCCTAAACCAAAATTCTAAGCTGAGCTAAAGTCA
CAAACTAAAATGATCAATTCAGTTTTGTAAATACAAAAAACAAGCTGAAAAATAAAAAAT
CAAGAAAAAAAATCTTGTTGAAAAAGATCCAAAATTACTAAACTTTATGGAAACCACTAA
AAGATATCAGTTTAGTCTAGTGTTTTAATTTAAATTAGCTAATGCCCACTTCTGTTTCAA
AATTTTACCTATAAGTTATGGTACGAAAAATGTGAATAAAAAGCCAAATGAGATTCTATC
TAATAAACACATCGTAACGGTAAATAAAATCAATTCTTCGAGAAAAGTCTGAAATAAAGA
AAATAAGGTTGGGTTGGTGGACTTGAGAGTAACAAGATCAACTGAGGTATTAGTGATTGT
AATCTGATCAATTGAGATATCTCAATAGTGATAGGAACTGGAACAGTGATACCCTGATCA
GTAGAAATTTAAATAGGAGCAATATTAAAAGAGCTTAGCTATAAGTCTGTGTTTTGCCTA
TGATTGATTAGGTTGATACTTCTGTTTCCAGATTTTCTGATTTTTCATCTTGCTATTCAC
AGTATAATGTTGATTCCCATGGACATGCTCAGGGTTTAGTCGGTAAGTTGATTCCTTGTG
GTTGAGCTTGCAAAAGAGAGAGCAAAAATCAAGGAGATTTTCATATTTCACTATCTACCA
ATATACAATTTAACAATTATCTATCCATATTTTATTTGGTAATGCATTATGGAGATTGAT

TTTCACACATATCATTACAACAATGGATCGGAAAAAAGAAACAGTTACAGCATCAATCTT
AAGAGGTTCCCCAATAAGAGATAGTTTAGAGAAAATGAGTCCCTTGCTGAAAAAGTGCAC
TAATAAATGAAGTTTGATCCAAATATGCATAATAAAGGATTCGAACTTAGCGAAGTAAAA
ATGAGACTACTTAAATGTTCTCATTTAGAAGTCGTGGAAGATCTAATTTAACGCAACAAG
CACATGAGGAAATTCTCTTAACATTAAAATTGAATTAAAATGTGCCAAGAGTCCAATGGC
CCAATCGAGAAGAAACCGCCAAAACCAATTGCCAAAAGGCCTCTTCGCACGACTTTATAT
TAAGGCAATCTGAAAAGAATTTTCTATCAAGAGCAAACCTAAAAGGGTTCAAAAAGCAAA
TTCGCATCCTCGTCGATAAAAATAACTATCGACATGCCTTTATACGTGATGGCATATCTC
ACCATAACCAACTAATCCATAATGAAGGTGCAATAAATATCAACATTATCCATCCACAAA
GGACTATTTTGAGTGCACTGAGAACACCCCAGCATGGTTGTGAAAGGCTACAGTATGAGA
AGAAATGGAGGCCAAGGCACCCTTAACTACCTAAGAAAAGATAACAGCTAAAGATGGAGC
CAGAGAAGACCTAAGTAGAGGAATTGAAACTATAGAAAAAATGGGAGCAAAAGTAAACTG
AACAAAGACCAAAGGGAATGGAATGGAGACCAAAGAGGTAGGATCAACCATCGCAAATTG
TGTCCTGCAGCAAAAAGAGACCTTTACGGTAGCGAAACGCCAGCAAAAAGGTACGACCAG
GATAGAAGACTACAAGCCAAAGCAATTAGCCGAAGGTTGATTATCAACAGCCCCAATATC
AACGACGATCCAATACGAGAT

>contig_9944
AGCAAGGTCATTGGCCCAAATGTTACGTCCATCCACAATTCCAGCAAAAAGATACTTTCC
AGATGGAAAACCAGCACTTTTGATCACATCAAGGGTCTTGGTTCCACGGACAAGATCGAA
TCCAAAACCAGTGATTGTACTCAAGGCAGTGATGGTCTTGTATGCCTCAGCAGGGACATC
AGCAAAGTAGGTCTCAATAACCACATTAAGGCCAGACAAAGATGATTCGAGTTCAGTATA
TGCCTTGGTAAATGTTTCCAACTGGTGAGAATCAAGATCCTTGACAAGGGTGGGTTCATC
AAACTGAACCCATGAAACACCAGCTGCTTTCAACTCAGCAATAACTTGCTTGTAGATGGG
TAGAATATGACCAAGAAGAGAAAGAATATTGAAGGTTTTTTCAACACCCTTGGCGGGCTT
GGAGAGCAGCAAGTATGTCACAGGTCCAACAAGCACGGGAACAGTGTCCACTCCAAGTGC
TTTTGCCTCCTTATACTCATTGACTGCCTTGTGAGAAGCATAGGAGAACTTGGTATTGGG
GCCTAGTTCAGGGACAATGAAATGGTAGTTCGTGTCAAACCATTTTGTCATCTCCATAGC
AGGTACAGAGGCATTTCCTCTGGCCATGGAGAAATAGGTGTCAAAACCAATCTCA

Appendix I. List of primer sequences used in qRT-PCR analysis (**Chapter 6**).

Name	Type	Sequence (5'-3')
Vpl-Actin-F	Housekeeping genes (mRNA)	GGGTTACTCCTTTACGACCACA
Vpl-Actin-R		GCTGCTCTTCGCTGTCTCAA
Vpl-βtub-F		TGGGCGAAAGGACACTACAC
Vpl-βtub-R		CACCCAAAGAATGGCACACT
Vpl-EF1a-F		TTGCTTGCCTTCACTCTTGG
Vpl-EF1a-R		TTCATCGTACCTTGCCTTGG
Vpl-Ubi-F		AGAGTGCCCTAACGCTGAGT
Vpl-Ubi-R		TGACGATCAAAGTGATTCGC
Vpl-GAPDH-F		GCTCCGATGTTTGTTGTTGG
Vpl-GAPDH-R		CTTCTGTGTTGCTGTGGTTGC
Vpl-18S-F		GCCTGGAATCTTTCGGTTTG
Vpl-18S-R		GCTCGTCGCACAACAACATC
Vpl_9944-F	Target Genes (mRNA)	GACTGCCTTGTGAGAAGCATAG
Vpl_9944R-		CCTCTGTACCTGCTATGGAGAT
Vpl_24863-F		GTTGAGGGAGGAGATGTGTTTAG
Vpl_24863-R		GTGATGCATAACGATAATGGTCTTAAA
Vpl_18270-F		TCGTGTGTCCAGATATCAATCC
Vpl_18270-R		GACCTGTCATCCGGCTATTT
Vpl_14640F		CTGCGAAGTAAAGCAAGGAATAG
Vpl_14640R		TCGCTACTGATCTCGGTCT
Vpl_1034F		GCATCAATCTTAAGAGGTTCCC
Vpl_1034R		CTTCGCTAAGTTCGAATCCTTT
Vpl_32170F		ACCCATGAATAATAAATTTGACACTACA
Vpl_32170R		GAAATTCACCGTTGTTATTGGTTAGA
Vpl_86799F		GCCATGACTCGATACTCAACTTC
Vpl_86799R		GATTCCTTCCACGCTCATTGT
Vpl_55903F		GCAACTACATTAACACCAGATCC
Vpl_55903R		CTTTGATGCAAACACCATGT
Vpl_miR168a-5p-F	Housekeeping genes (miRNA)	TCGCTTGGTGCAGGT
Vpl_miR162-F		CGCAGTCGATAAACCTCT
Vpl_miR5179-F		GCGCAGTTTTGCTCAAGA
Vpl_miR319d-3p-F		GCAGTTGGACTGAAGGGA
Vpl_miR159-F		GCGCAGTTTGGATTGAAG
Vpl_SnoR6-F		AATTGGAACGATACAGAGAAGATTAGC
vpl_miR166i-F	Target genes (miRNA)	GCAGTTGGACCAGGCT
vpl_miR166i-R		CAGTTTTTTTTTTTTTTTGGGGAATG
vpl_miR395g-3p-F		GCTGAAGTGTTTGGGGGAA
vpl_miR395g-3p-R		GGTCCAGTTTTTTTTTTTTTTTGAGT
vpl_miR8155-F		CGCAGTAACCTGGCTCT
vpl_miR8155-R		GTCCAGTTTTTTTTTTTTTTTGGTATC

Appendix J. Reference stability pre-screening results for 86 potential miRNA candidates from RNA-seq data, computed by five algorithms (RefFinder, Delta CT, BestKeeper, NormFinder and GeNorm).

Method ---------- Ranking	Delta CT	BestKeeper	Normfinder	Genorm	Recommended comprehensive ranking
1	vpl-miR160e	vpl-miR167e	vpl-miR408a	vpl-miR166h-5p	**vpl-miR894a | vpl-miR164j-5p**
2	vpl-miR319m	vpl-miR159e-3p	vpl-miR408d-3p	vpl-miR168e-5p	
3	vpl-miR5072	vpl-miR319j	vpl-miR319g-3p	vpl-miR319g-3p	**vpl-miR167e**
4	vpl-miR167e	vpl-miR164j-5p	vpl-miR168e-5p	vpl-miR166h-3p	**vpl-miR164j-5p**
5	vpl-miR164j-5p	vpl-miR319m	vpl-miR162b	vpl-miR408c-3p	**vpl-miR319m**
6	vpl-miR159e-3p	vpl-miR166t	vpl-miR894a	vpl-miR319a	**vpl-miR159e-3p**
7	vpl-miR162a	vpl-miR168f	vpl-miR159b	vpl-miR396a	**vpl-miR160e**
8	vpl-miR3630a-3p	vpl-miR535a	vpl-miR396g-5p	vpl-miR159a	**vpl-miR319j**
9	vpl-miR319j	vpl-miR171i	vpl-miR157b-3p	vpl-miR398e	**vpl-miR5072**
10	vpl-miR164j-3p	vpl-miR5072	vpl-miR399c-3p	vpl-miR398i-3p	**vpl-miR319g-3p**
11	vpl-miR166t	vpl-miR156f-3p	vpl-miR396c	vpl-miR6300	**vpl-miR168e-5p**
12	vpl-miR390e-3p	vpl-miR390e-3p	vpl-miR6478	vpl-miR166r	**vpl-miR166t**
13	vpl-miR166g	vpl-miR156m	vpl-miR319h-3p		**vpl-miR408a**
14	vpl-miR171i	vpl-miR156n	vpl-miR398b	vpl-miR894a | vpl-miR164j-5p	**vpl-miR171i**
15	vpl-miR535a	vpl-miR159n	vpl-miR397h-5p	vpl-miR166d	**vpl-miR535a**
16	vpl-miR390d-3p	vpl-miR160e	vpl-miR168d-3p	vpl-miR408a	**vpl-miR390e-3p**

17	vpl-miR159n	vpl-miR8175	vpl-miR397c	vpl-miR408d-3p	**vpl-miR168f**
18	vpl-miR168f	vpl-miR164j-3p	vpl-miR396d	vpl-miR159j-3p.1	**vpl-miR408d-3p**
19	vpl-miR164i-5p	vpl-miR390d-3p	vpl-miR164e-5p	vpl-miR156d	**vpl-miR162a**
20	vpl-miR156m	vpl-miR167m-5p	vpl-miR396b	vpl-miR159e-3p	**vpl-miR164j-3p**
21	vpl-miR156n	vpl-miR166g	vpl-miR172a	vpl-miR167e	**vpl-miR3630a-3p**
22	vpl-miR156f-3p	vpl-miR162a	vpl-miR390b-3p	vpl-miR319m	**vpl-miR156f-3p**
23	vpl-miR168h-5p	vpl-miR168h-5p	vpl-miR5139	vpl-miR319j	**vpl-miR390d-3p**
24	vpl-miR167m-5p	vpl-miR164i-5p	vpl-miR5179	vpl-miR166t	**vpl-miR159n**
25	vpl-miR159h	vpl-miR156v	vpl-miR159h	vpl-miR171i	**vpl-miR156m**
26	vpl-miR172a	vpl-miR165a-3p	vpl-miR8155	vpl-miR535a	**vpl-miR166g**
27	vpl-miR156v	vpl-miR3630a-3p	vpl-miR399i	vpl-miR168f	**vpl-miR166h-5p**
28	vpl-miR390b-3p	vpl-miR166b	vpl-miR166q	vpl-miR5072	**vpl-miR156n**
29	vpl-miR8175	vpl-miR171a	vpl-miR2950a	vpl-miR156m	**vpl-miR164i-5p**
30	vpl-miR396b	vpl-miR5083	vpl-miR3630a-3p	vpl-miR156n	**vpl-miR894a**
31	vpl-miR390h-3p	vpl-miR390h-3p	vpl-miR164i-5p	vpl-miR390e-3p	**vpl-miR172a**
32	vpl-miR166b	vpl-miR156r-5p	vpl-miR164j-3p	vpl-miR156f-3p	**vpl-miR167m-5p**
33	vpl-miR5139	vpl-miR159h	vpl-miR390h-3p	vpl-miR159n	**vpl-miR168h-5p**
34	vpl-miR5179	vpl-miR172a	vpl-miR160e	vpl-miR160e	**vpl-miR8175**
35	vpl-miR396d	vpl-miR396b	vpl-miR390d-3p	vpl-miR390d-3p	**vpl-miR159h**
36	vpl-miR397c	vpl-miR390b-3p	vpl-miR162a	vpl-miR164j-3p	**vpl-miR396b**
37	vpl-miR165a-3p	vpl-miR5139	vpl-miR390e-3p	vpl-miR162a	**vpl-miR390b-3p**
38	vpl-miR5083	vpl-miR5179	vpl-miR5072	vpl-miR166g	**vpl-miR5139**
39	vpl-miR395g-3p	vpl-miR395g-3p	vpl-miR156f-3p	vpl-miR167m-5p	**vpl-miR390h-3p**
40	vpl-miR397h-5p	vpl-miR6173	vpl-miR319m	vpl-miR8175	**vpl-miR396d**

41	vpl-miR171a	vpl-miR8155	vpl-miR167e	vpl-miR156v	**vpl-miR397c**
42	vpl-miR156r-5p	vpl-miR166o	vpl-miR394a	vpl-miR168h-5p	**vpl-miR156v**
43	vpl-miR8155	vpl-miR396d	vpl-miR164j-5p	vpl-miR3630a-3p	**vpl-miR397h-5p**
44	vpl-miR6173	vpl-miR397c	vpl-miR159e-3p	vpl-miR164i-5p	**vpl-miR5179**
45	vpl-miR398b	vpl-miR167k	vpl-miR166g	vpl-miR165a-3p	**vpl-miR162b**
46	vpl-miR166o	vpl-miR397h-5p	vpl-miR319j	vpl-miR166b	**vpl-miR398b**
47	vpl-miR2950a	vpl-miR159i-3p	vpl-miR168h-5p	vpl-miR5083	**vpl-miR159b**
48	vpl-miR164e-5p	vpl-miR164e-5p	vpl-miR171i	vpl-miR171a	**vpl-miR396g-5p**
49	vpl-miR399i	vpl-miR2950a	vpl-miR159n	vpl-miR156r-5p	**vpl-miR166b**
50	vpl-miR159i-3p	vpl-miR398b	vpl-miR166t	vpl-miR390h-3p	**vpl-miR399c-3p**
51	vpl-miR167k	vpl-miR399i	vpl-miR535a	vpl-miR159h	**vpl-miR396c**
52	vpl-miR166q	vpl-miR167d-5p	vpl-miR168f	vpl-miR172a	**vpl-miR165a-3p**
53	vpl-miR168d-3p	vpl-miR166q	vpl-miR8175	vpl-miR396b	**vpl-miR166h-3p**
54	vpl-miR319h-3p	vpl-miR168d-3p	vpl-miR398e	vpl-miR5139	**vpl-miR6478**
55	vpl-miR167d-5p	vpl-miR6478	vpl-miR167m-5p	vpl-miR390b-3p	**vpl-miR8155**
56	vpl-miR396c	vpl-miR159l-3p	vpl-miR156m	vpl-miR5179	**vpl-miR164e-5p**
57	vpl-miR6478	vpl-miR156b	vpl-miR156n	vpl-miR395g-3p	**vpl-miR319h-3p**
58	vpl-miR399c-3p	vpl-miR319h-3p	vpl-miR159i-3p	vpl-miR6173	**vpl-miR157b-3p**
59	vpl-miR159l-3p	vpl-miR396c	vpl-miR156v	vpl-miR396d	**vpl-miR168d-3p**
60	vpl-miR156b	vpl-miR399c-3p	vpl-miR166b	vpl-miR397c	**vpl-miR5083**
61	vpl-miR396g-5p	vpl-miR167h-5p	vpl-miR395g-3p	vpl-miR166o	**vpl-miR398e**
62	vpl-miR159b	vpl-miR394a	vpl-miR5083	vpl-miR397h-5p	**vpl-miR171a**
63	vpl-miR167h-5p	vpl-miR396g-5p	vpl-miR165a-3p	vpl-miR8155	**vpl-miR408c-3p**
64	vpl-miR394a	vpl-miR159b	vpl-miR171a	vpl-miR398b	**vpl-miR319a**

65	vpl-miR157b-3p	vpl-miR167l-5p	vpl-miR156r-5p	vpl-miR167k	**vpl-miR396a**
66	vpl-miR162b	vpl-miR157b-3p	vpl-miR6173	vpl-miR164e-5p	**vpl-miR156r-5p**
67	vpl-miR167l-5p	vpl-miR159j-3p.1	vpl-miR167k	vpl-miR2950a	**vpl-miR2950a**
68	vpl-miR159j-3p.1	vpl-miR162b	vpl-miR166o	vpl-miR159i-3p	**vpl-miR399i**
69	vpl-miR156e-5p	vpl-miR156e-5p	vpl-miR167d-5p	vpl-miR399i	**vpl-miR166r**
70	vpl-miR894a	vpl-miR166d	vpl-miR159l-3p	vpl-miR166q	**vpl-miR159a**
71	vpl-miR166r	vpl-miR894a	vpl-miR159j-3p.1	vpl-miR168d-3p	**vpl-miR395g-3p**
72	vpl-miR166d	vpl-miR166r	vpl-miR156e-5p	vpl-miR167d-5p	**vpl-miR6300**
73	vpl-miR408a	vpl-miR408a	vpl-miR167l-5p	vpl-miR319h-3p	**vpl-miR166q**
74	vpl-miR408d-3p	vpl-miR408d-3p	vpl-miR167h-5p	vpl-miR396c	**vpl-miR166d**
75	vpl-miR156d	vpl-miR6300	vpl-miR156b	vpl-miR6478	**vpl-miR159j-3p.1**
76	vpl-miR6300	vpl-miR156d	vpl-miR166r	vpl-miR399c-3p	**vpl-miR6173**
77	vpl-miR319g-3p	vpl-miR319g-3p	vpl-miR166d	vpl-miR159l-3p	**vpl-miR398i-3p**
78	vpl-miR166h-5p	vpl-miR166h-5p	vpl-miR6300	vpl-miR156b	**vpl-miR166o**
79	vpl-miR168e-5p	vpl-miR168e-5p	vpl-miR156d	vpl-miR396g-5p	**vpl-miR156d**
80	vpl-miR398e	vpl-miR398e	vpl-miR396a	vpl-miR167h-5p	**vpl-miR159i-3p**
81	vpl-miR396a	vpl-miR396a	vpl-miR166h-5p	vpl-miR159b	**vpl-miR167k**
82	vpl-miR166h-3p	vpl-miR166h-3p	vpl-miR166h-3p	vpl-miR394a	**vpl-miR394a**
83	vpl-miR159a	vpl-miR159a	vpl-miR159a	vpl-miR157b-3p	**vpl-miR167d-5p**
84	vpl-miR319a	vpl-miR319a	vpl-miR319a	vpl-miR162b	**vpl-miR159l-3p**
85	vpl-miR408c-3p	vpl-miR408c-3p	vpl-miR408c-3p	vpl-miR167l-5p	**vpl-miR156b**
86	vpl-miR398i-3p	vpl-miR398i-3p	vpl-miR398i-3p	vpl-miR156e-5p	**vpl-miR167h-5p**

REFERENCES

Abd-Elsalam KA (2003) Non-gel based techniques for plant pathogen genotyping. *Acta Microbiologica Polonica* **52(4)**: 329-341.

Adamczuk A, Siegień I and Ciereszko (2012) Morphogenesis of plants *in vitro* under stress condition. In: G. Łaska (ed.) *Biological diversity - from cell to ecosystem*, pp. 25-40. Polish Botanical Society - Branch in Białystok. ISBN 978-83-62069-28-6.

Agarwal P, Parida SK, Mahto A, Das S, Mathew IE, Malik N and Tyagi AK (2014) Expanding frontiers in plant transcriptomics in aid of functional genomics and molecular breeding. *Biotechnology Journal* **9**: 1480-1492.

Ahloowalia BS (1992) *In vitro* radiation induced mutants in Chrysanthemum. *Mutation Breeding Newsletter* **39**: 6.

Ahmed EU, Hayashi T and Yazawa S (2004) Auxins increase the occurrence of leaf-colour variants in *Caladium* regenerated from leaf explants. *Scientia Horticulturae* **100**: 153–159.

Akagi H, Yokozeki Y, Inagaki A, Nakamura A and Fujimura T (1996) A co-dominant DNA marker closely linked to the rice nuclear restorer gene, Rf-1, identified with inter-SSR fingerprinting. *Genome* **39**: 1205–1209.

Alatzas A and Foundouli A (2006) Distribution of ubiquitinated histone H2A during plant cell differentiation in maize root and dedifferentiation in callus culture. *Plant Science* **171**: 481–487.

Aldape K, Ginzinger DG and Godfrey TE (2002) Real-time quantitative chain reaction: a potential tool for genetic analysis in neuropathology. *Brain pathology* **12**: 54-66.

Aloni R., Aloni E, Langhans M, Ullrich, CI (2006). Role of cytokinin and auxin in shaping root architecture: Regulating vascular differentiation, lateral root initiation, root apical dominance and root gravitropism. *Annals of Botany* **97**: 883-893.

Alonso A and Fernándex OG (2007) Real time quantitative PCR in forensic science. In: R Rapley (Ed.) *Molecular Forensics*, pp. 59-70. New York, USA: John Wiley & Sons Ltd.

Altamura MM, Bassi P, Cavallini A, Gionini G, Cremonini R, Monacelli B, Pasqua G, Sassoli O, Tran TVK and Cionini PG (1987) Nuclear DNA changes during plant development and the morphogenetic response *in vitro* of *Nicotiana tobacum* tissues. *Plant Science* **53**: 73–79. doi: 10.1016/0168-9452(87)90180-4

Altschul SF, Madden TL, Schäffer AA, Zhang JH, Zhang Z, Miller W and Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**: 3389-3402.

Alvarez MR (1968) Quantitative changes in nuclear DNA accompanying postgermination embryonic development in *Vanda* (*Orchidaceae*). *American Journal of Botany* **55:** 1036-1041.

Alvarez MR and Sagawa Y (1965) A histochemical study of embryo development in *Vanda* (*Orchidaceae*). *Caryologia* **18**: 251-261.

Alves E, Ballesteros I, Linacero R and Vázquez AM (2005) *RYS1*, a foldback transposon, is activated by tissue culture and shows preferential insertion points into the rye genome. *Theor Appl Genet* **111**: 431–436.

Amberger LA, Palmer RG and Shoemaker RC (1992) Analysis of culture-induced variation in soybean. *Crop Science* **32**: 1103-1108.

Anuradha K, Shyamala BN and Naidu MM (2013) Vanilla – Its Science of Cultivation, Curing, Chemistry, and Nutraceutical Properties. *Critical Reviews in Food Science and Nutrition* **53(12)**: 1250-1276. doi: 10.1080/10408398.2011.563879.

Anwar A, Kikuchi A and Watanabe KN (2010) Assessment of somaclonal variation for salinity tolerance in sweet potato regenerated plants. *Afr J Biotechnol* **9**: 7256–7265.

Anzola JM, Sieberer T, Ortbauer M, Butt H, Korbei B, Weinhofer I, Mullner AE and Luschnig C (2010) Putative *Arabidopsis* transcriptional adaptor protein (PROPORZ1) is required to modulate histone acetylation in response to auxin. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* **107**: 10308–10313.

Apweiler R, Attwood TK, Bairoch A, Bateman A, Birney E, Biswas M, Bucher P, Cerutti L, Corpet F, Croning MDR, Durbin R, Falquet L, Fleischmann W, Gouzy J, Hermjakob H, Hulo N, Jonassen I, Kahn D, Kanapin A, Karavidopoulou Y, Lopez R, Marx B, Mulder NJ, Oinn TM, Pagni M, Servant F, Sigrist CJA, Zdobnov EM and Interpro C (2000) InterPro -- an integrated documentation resource for protein families, domains and functional sites. *Bioinformatics* **16 (12):** 1145-1150. doi:10.1093/bioinformatics/16.12.1145.

Apweiler R, Attwood TK, Bairoch A, Bateman A, Birney E, Biswas M, Bucher P, Cerutti L, Corpet F, Croning MD, Durbin R, Falquet L, Fleischmann W, Gouzy J, Hermjakob H, Hulo N, Jonassen I, Kahn D, Kanapin A, Karavidopoulou Y, Lopez R, Marx B, Mulder NJ, Oinn TM, Pagni M, Servant F, Sigrist CJ and Zdobnov EM (2001). The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Research* **29** (1): 37–40. doi:10.1093/nar/29.1.37.

Araki H, Shimazaki H, Hirata Y, Oridate T, Harada T and Yakuwa T (1992) Chromosome number variation of callus cells and regenerated plants in *Asparagus officinalis* L. *Plant Tis Cult Lett* **9**: 169-175.

Araújo LG, Prabhu AS, Filippi MC and Chaves LJ (2001) RAPD analysis of blast resistant somaclones from upland rice cultivar IAC 47 for genetic divergence. *Plant Cell, Tissue and Organ Culture (PCTOC)* **67**: 165–172.

Arditti J (2009) "Vanilla". in Micropropagation of Orchids, 2nd Edition, Volume II (ed J. Arditti), pp. 1239-1291. Wiley-Blackwell, Oxford, UK. ISBN: 978-1-4443-0040-6.

Arditti J, Ghani AKA (2000) Numerical and physical properties of orchid seeds and their biological implications. *New Phytologist* **145**: 367-421.
Arenas-Huertero C, Perez B, Rabanal F, Blanco-Melo D, De la Rosa C, Estrada-Navarrete G, *et al.* (2009) Conserved and novel miRNAs in the legume *Phaseolus vulgaris* in response to stress. *Plant Molecular Biol*ogy **70**: 385–401.
Arif IA, Bakir MA, Khan HA, Al Farhan AH, Al Homaidan AA, Bahkali AH, Al Sadoon M and Shobrak M (2010) A Brief Review of Molecular Techniques to Assess Plant Diversity. *Interational Journal of Molecular Sciences* **11**: 2079–2096.
Arihara A, Kita T, Igarashi S, Goto M and Irikura Y (1995) White Baron: a non-browning somaclonal variant of *Danshakuimo* (Irish cobbler). *Am Potato J* **72(11)**: 701–705.
Arumuganathan K and Earle ED (1991) Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* **9(3)**: 208-218.
Asif M, Eudes F, Randhawa H, Amundsen E (2013) Cefotaxime prevents microbial contamination and improves microspore embryogenesis in wheat and triticale. *Plant Cell Rep* **32**L 1637-1646.
Azofeifa-Bolaňos JB, Gigant LR, Nicolás-García M, Pignal M, Tavares-González FB, Hágsater E, Salazar-Chávez GA, Reyes-López D, Archila-Morales FL, García-García JA, Silva D, Allibert A, Solano-Campos F, Rodríguez-Jimenes GC, Paniagua-Vásquez A, Besse P, Pérez-Silva A and Grisoni M (2017) A new vanilla species from Costa Rica closely related to *V. planifolia* (*Orchidaceae*). *European Journal of Taxonomy* **284**: 1-26.
Babu SCV and Gassmann M (2016) *Assessing Integrity of Plant RNA with the Agilent 2100 Bioanalyzer System*. Application Note. Available at: http://www.agilent.com/cs/library/applications/5990-8850EN.pdf (Accessed: 7th June 2017).
Bai D and Knott DR (1993) The effects of level of 2,4-D and time in culture on regeneration rate and chromosome number of regenerants from calli of the hybrid *Triticum aestivum* cv. Chinese Spring ph1b x *Thinopyrum ponticum* (2n =10x = 70). *Genome* **36**: 166-172.
Bairu MW and Kane ME (2011) Physiological and developmental problems encountered by *in vitro* cultured plants. *Plant Growth Regulation* **63**: 101-103.
Bairu MW, Aremu AO and Staden JV (2011) Somaclonal variation in plants: causes and detection methods. *Journal of Plant Growth Regulation* **63**: 147-173.
Bairu MW, Fennell CW and van Staden J (2006) The effect of plant growth regulators on somaclonal variation in Cavendish banana (*Musa* AAA cv. 'Zelig'). *Scientia Horticulturae* **108**: 347–351.
Balilashaki K, Gantait S, Naderi R and Vahedi M (2015) Capsule formation and asymbiotic seed germination in some hybrids of *Phalaenopsis*, influenced by pollination season and capsule maturity. *Physiology and Molecular Biology of Plants* **21(3)**: 341-347. doi:10.1007/s12298-015-.
Baluska F (1990) Nuclear size, DNA content, and chromatin condensation are different in individual tissues of the maize root apex. *Protoplasma* **158:** 45-52.

Baránek M, Křižan B, Ondrušíková E, Pidra M (2010) DNA methylation changes in grapevine somaclones following *in vitro* culture and thermotherapy. *Plant Cell, Tissue and Organ Culture (PCTOC)* **101**: 11–22.
Barghouthi SA (2011) A universal method for the identification of bacteria based on general PCR primers. *Indian Journal of Microbiology* **51(4)**: 430-444.
Barow M (2006). Endopolyploidy in seed plants. *BioEssays* **28**: 271-281.
Barow M and Meister A (2003). Endopolyploidy in seed plants is differently correlated to systematics, organ, life strategy and genome size. *Plant Cell Environ* **26**: 571-584.
Barret P, Brinkman M and Beckert M (2006) A sequence related to rice *Pong* transposable element displays transcriptional activation by *in vitro* culture and reveals somaclonal variations in maize. *Genome* **49**: 1399–1407.
Battistini C and Rosati P (1991) In vitro evaluation of somaclonal strawberry (*Fragaria x ananassa* 'Brighton') variants for susceptibility to *Phytophthora cactorum.* In: Dale A, Lubby D (eds) *The strawberry into the 21st century*. Timber Press, Portland, pp 121–123.
Bauer MJ and Birchler JA (2006) Organization of endoreduplicated chromosomes in the endosperm of *Zea mays* L. *Chromosoma* **115**: 383-394.
Baurens FC, Nicolleau J, Legavre T, Verdeil JL and Monteuuis O (2004) Genomic DNA methylation of juvenile and mature *Acacia mangium* micropropagated *in vitro* with reference to leaf morphology as a phase change marker. *Tree Physiology* **24**: 401–407.
Bayliss MW (1973) Origin of chromosome number variation in cultured plant cells. *Nature* **246**: 21-28.
Bayliss MW (1980) Chromosomal variation in tissue culture. In: IK Vasil (Ed.) *Perspectives in plant cell and tissue culture*, pp. 113–144. New York, USA: Academic Press.
BD Biosciences (2012) BD Accuri™ C6 Flow Cytometer Instrument Manual (7820018-Rev-2).
Becker C, Hammerle-Fickinger A, Riedmaier I and Pfaffl MW (2010) mRNA and microRNA quality control for RT-qPCR analysis. *Methods* **50**: 237-243.
Bednarek PT, Orlowska R and Niedziela A (2017) A relative quantitative methylation-sensitive amplified polymorphism (MSAP) method for the analysis of abiotic stress. *BMC Plant Biology* **17**: 79. doi: 10.1186/s12870-017-1028-0.
Beemster GT, Baskin TI (2000) Stunted plant 1 mediates effects of cytokinin, but not on auxin, on cell division and expansion in the root of *Arabidopsis. Plant Physiology* **124**: 1718-1727.
Benes V and Castoldi M (2010) Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available. *Methods* **50**: 244-249.
Ben-Hayyim G and Goffer Y (1989) Plantlet regeneration from NaCl selected salt-tolerant callus culture of Shamouti oranges (*Citrus sinesis* L. Osbeck). *Plant Cell Rep* **7**: 680–683.
Benková E, Hejátko J (2009) Hormone interactions at the root apical meristem. *Plant Molecular Biology* **69**: 383-396.

Bennett MD and Leitch IJ (2005) Genome size evolution in plants. *In:* Gregory TR (eds.), *The evolution of the genome, pp.* 89-162. Elsevier Academic Press, Burlington, Massachusetts, USA.

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, *et al.* (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456(7218)**: 53-59.

Berdasco M, Alcázar R, García-Ortiz MV, *et al.* (2008) Promoter DNA hypermethylation and gene repression in undifferentiated *Arabidopsis* cells. *PLoS ONE* **3**: e3306.

Berger SL (2007) The complex language of chromatin regulation during transcription. *Nature* **447**: 407–412.

Bergervoet JHW, Verhoeven HA, Gilissen LJW and Bino RJ (1996) High amount of nuclear DNA in tomato (*Lycopersicon esculentum* Mill.) pericarp. *Plant Sci* **116**: 141-145.

Bergkvist A, Rusnakova V, Sindelka R, Garda JMA, Sjögreen B, Lindh D, Forootan A and Kubista M (2010) Gene expression profiling – Clusters of possibilities. *Methods* **50**: 323-335.

Bernard PS and Wittwer CT (2002) Real-time PCR technology for cancer diagnostics. *Clinical Chemistry* **48(8)**: 1178-1185.

Betekhtin A, Rojek M, Jaskowlak J, Milewska-Hendel A, Kwasniewska J, Kostyukova Y, Kurczynska E, Rumyantseva N and Hasterok (2017) Nuclear genome stability in long-term cultivated callus lines of *Fagoyrum tataricum* (L.) Gaert. *PLOS One* **12(3)**: e0173537. doi: 10.1371/journal.pone.0173537.

Bhardwaj SV, Thakur T, Sharma R and Sharma P (2012) *In vitro* selection of resistant mutants of ginger (*Zingiber officinale* Rosc.) against wilt pathogen (*Fusarium oxysporum* f. sp. zingiberi Trujillo). *Plant Dis Res* **27**: 194–199.

Bhatia R, Singh KP, Jhang T and Sharma TR (2009) Assessment of clonal fidelity of micropropagated gerbera plants by ISSR markers. *Scientia Horticulturae* **119**: 208–211.

Bhatia S (2015) Application of Plant Biotechnology. In: Modern Applications of Plant Biotechnology in Pharmaceutical Sciences, pp. 157-207.

Bhatia S, Sharma K, Dahiya R and Bera T (2015) Plant Tissue Culture. In: S. Bhatia (Ed.) Modern Applications of Plant Biotechnology in Pharmaceutical Sciences. Academic Press, Cambridge, US. pp. 31-107. ISBN: 978-0-12-802221-4.

Bi CW, Qiu JB, Zhou MG, Chen CJ, Wang JX (2009) Effects of carbendazim on conidial germination and mitosis in germlings of *Fursasium graminearum* and *Botrytis cinerea*. *International Journal of Pest Management* **55(2)**: 157-163.

Bio-Rad Laboratories Inc. (2013) *CFX96 Touch™, CFX96 Touch Deep Well™, CFX Connect™, and CFX384 Touch™ Real-Time PCR Detection Systems - Instruction Manual*. Available at: http://www.bio-rad.com/webroot/web/pdf/lsr/literature/10021337.pdf (Accessed: 21th June 2017).

Biswas MK, Dutt M, Roy UK, Islam R and Hossain M (2009) Development and evaluation of *in vitro* somaclonal variation in strawberry for improved horticultural traits. *Sci Hortic* **122**: 409–416.

Bloom (2017) Vanilla. In: *Natural and Artificial Flavour: What's the difference?*, pp. 13-14. New York, USA: American Council on Science and Health (ACSH). ISBN: 978-0-9910055-9-8.

Bohra P, Wamam AA, Sathyanarayana BN, Umesha K, Anu SR, Swetha HG, Gourish RK (2013) Aseptic culture establishment using antibiotics with reference to their efficiency and phytotoxicity in difficult-to-establish native Ney Poovan Banana (*Musa,*AB). *Proc Natl Acad Sci India Sect B Biol Sci* **84**: 257-263.

Boland CR and Ricciardiello (1999) How many mutations does it take to make a tumor? *PNAS* **96(26)**: 14675-14677.

Bonetta L (2005) Prime time for real-time PCR. *Nature Methods* **2**: 305-312.

Bonnet E, Peer YV and Rouzé (2006) The small RNA world of plants. *New Phytologist* **171**: 451-468.

Bonnet E, Wuyts J, Rouzé and Peer YV (2005) Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identified important target genes. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* **101(31)**: 11511-11516.

Bory SO, Catrice S, Brown IJ, Leitch R, Gigant F, Chiroleu M, Grisoni M, Duval MF and Besse P (2008) Natural polyploidy in *Vanilla planifolia* (*Orchidaceae*). *Genome* **51**: 816 – 826. doi: 10.1139/G08-068

Botstein D, White RL, Skolnick M and Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *The American Journal of Human Genetics* **32(3)**: 314-331.

Bouman H and De Klerk GJ (2001) Measurement of the extent of somaclonal variation in begonia plants regenerated under various conditions. Comparison of three assays. *Theoretical and Applied Genetics* **102**: 111–117.

Bouriquet G (1954) *Le vanillier et la vanille dans le monde. Encyclopédié Bíologique XLVI*. Paul Lechevaliet, Paris VI.

Braun AC (1959) A demonstration of the recovery of the crown-gall tumor cell with the use of complex tumors of single-cell origin. *Proceedings of the National Academy of Sciences of the United States of America* **45**: 932–938.

Brettell RIS and Dennis ES (1991) Reactivation of a silent *Ac* following tissue culture is associated with heritable alterations in its methylation pattern. *Mol Gen Genet* **3**: 365–372.

Brown ML (2009) Madagascar's Cyclone Vulnerability and the Global Vanilla Economy. In: Jones EC and Murphy AD (Eds.) *The Political Economy of Hazards and Disasters*, pp. 241-264. Lanham, USA: Altamira Press.

Brummer EC, Cazcarro PM and Luth D (1999) Ploidy determination of alfalfa germplasm accessions using flow cytometry. *Crop Science* **39**: 1202–1207.

Bulpitt CH, Li Y, Bulpitt PF and Wang JG (2007) The use of orchids in Chinese medicine. *Journal of the Royal Society of Medicine* **100**: 558-563.

Bulpitt CJ (2005) The uses and misuses of orchids in medicine. *The Quarterly Journal of Medicine* **98**: 625–631.

Busconi M, Lucini L, Soffritti G, Bernardi J, Bernardo L, Brunschwig C, Lepers-Andrzejewski S, Raharivelomanana P, Fernandez JA (2017) Phenolic profiling of traceability of *Vanilla xtahitensis*. *Frontiers in Plant Science* **8**: 1746. doi: 10.3389/fpls.2017.01746.

Busk PK (2014) A tool for design of primers for microRNA-specific quantitative RT-qPCR. *BMC Bioinformatics* **15**: 29.

Bustin SA (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology* **25**: 169-193.

Bustin SA (2010) Why the need for qPCR publication guidelines? – The case for MIQE. *Methods* **50**: 217-226.

Bustin SA, Beaulieu JF, Huggett J, Jaggi R, Kibenge FS, Olsvik PA, Penning LC and Toegel S (2010) MIQE précis: Practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. *BMC Molecular Biology* **11**: 74.

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J and Wittwer CT (2009) The MIQE guidelines: *M*inimum *I*nformation for Publication of *Q*uantitative Real-Time PCR *E*xperiments. *Clinical Chemistry* **55(4)**: 611-622.

Bythrow JD (2005) Vanilla as Medicinal Plant. *Seminars in Integrative Medicine* **3**: 129-131.

Caboni E, Anselmi S, Donato E and Manes F (2003) *In vitro* selection of *Actinidia deliciosa* clones tolerance to NaCl and their molecular and *in vivo* ecophysiological characterisation. *Acta Hortic* **618**: 77–83.

Caetano-Anollés G, Bassam BJ and Gresshoff PM (1991) DNA amplification using very short arbitary oligonucleotide primers. *Bio/Technology* **9**: 553-557.

Cameron KM (2004) Utility of plastid psaB gene sequences for investigating intrafamilial relationships within *Orchidaceae*. *Molecular Phylogenetics and Evolution* **31(3)**: 1157-1176.

Cameron KM (2009) On the value of nuclear and mitochondrial gene sequences for reconstructing the phylogeny of vanilloid orchids (Vanilloideae, Orchidaceae). *Annals of Botany* **104(3)**: 377-385.

Cameron KM and Molina MC (2006) Photosystem II gene sequences of psbB and psbC clarify the phylogenetic position of *Vanilla* (Vanilloideae, Orchidaceae). *Cladistics* **22(3)**: 239-248.

Cameron KM, Chase MW, Whitten WM, Kores PJ, Jarrell DC, Albert VA, Yukawa T, Hills HG and Goldman DH (1999) A phylogenetic analysis of the Orchidaceae: evidence from rbcL nucleotide. *American Journal of Botany* **86(2)**: 208-224.

Cammaerts S, Strazisar M, Rijk PD and Favero JD (2015) Genetic variants in microRNA genes: impact on microRNA expression, function and disease. *Front. Genet.* **6**: 186.

Castro-Bobadilla G and García-Franco JG (2007) Vanilla (*Vanilla planifolia* Andrews) crop systems used in the Totonacapan area of Veracruz, México: Biological and productivity evaluation. *Journal of Food Agriculture and Environment* **2**: 136-142.

Cavrak VV, Lettner N, Jamge S, Kosarewicz A, Bayer LM and Scheid OM (2014) How a Retrotransposon Exploits the Plant's Heat Stress Response for Its Activation, *PLoS Genetics* **10(1)**: e1004115.

Chai LC, Alderson PGA and Chin CF (2014) Genomic studies on somaclonal variations of *Vanilla planifolia*. In: *Plant Genomic Congress: Asia Conference*, Kuala Lumpur, Malaysia, Abstract, February 24-25.

Chandna R, Augustine R and Bisht NC (2012) Evaluation of candidate reference genes for gene expression normalization in *Brassica juncea* using real time quantitative RT-PCR. *PLOS One* **7(5)**: e36918.

Chang C, Chang WC (1998) Plant regeneration from callus culture of *Cymbidium ensifolium* var. misericors. *Plant Cell Report* **17**: 251–255.

Chang KH, Mestdagh P, Vandesompele J, Kerin MJ and Miller N (2010) microRNA expression profiling to identify and validate reference genes for relative quantification in colorectal cancer. *BMC Cancer* **10**: 173.

Chao CCT, Devanand PS and Chen J (2005) AFLP analysis of genetic relationships among *Calathea* species and cultivars. *Plant Sci* **168**: 1459–1469.

Chao YT, Su CL, Jean WH, Chen WC, Chang YC and Shih MC (2014) Identification and characterization of the microRNA transcriptome of a moth orchid *Phalaenopsis Aphrodite*. *Plant Mol Biol:* 1-20. doi: 10.1007/s11103-013-0150-0.

Charles DJ (2012) Vanilla. In: *Antioxidant Properties of Spieces, Herbs and Other Sources.* New York, USA: Springer.

Chen JT and Chang WC (2000) Efficient plant regeneration through somatic embryogenesis from callus cultures of Oncidium (Orchidaceae). *Plant Science* **160**: 87–93. doi: 10.1016/S0168-9452(00)00367-8.

Chen XM (2008) MicroRNA metabolism in plants. *Current Topics in Microbiology and Immunology* **320**: 117-136.

Chen ZH, Hsiao KC and Bornman CH (1994) Ploidy and division efficiency of petiolar protoplasts of *Brassica apus*. *Hereditas* **120**: 41-46.

Chen ZJ and Tian L (2007) Roles of dynamic and reversible histone acetylation in plant development and polyploidy. *Biochimica et Biophysica Acta* **1769**: 295–307.

Cheng XY, Gao MW, Liang ZQ, Liu GZ and Hu TC (1992) Somaclonal variation in winter wheat: frequency, occurrence and inheritance. *Euphytica* **64**: 1-10.

Chevreau E, Brisset MN, Paulin JP and James DJ (1998) Fire blight resistance and genetic trueness-to-type of four somaclonal variants from the apple cultivar Greensleeves. *Euphytica* **104**: 199–205.

Chomicki G, Bidel LPR, Ming F, Coiro M, Zhang X, Wang YF, Baissac Y, Jay-Allemand C, Renner SS (2014) The velamen protects photosynthetic orchid

roots against UV-B damage, and a large dated phylogeny implies multiple gains and losses of this function during the Cenozoic. *New Phytologist* **205**: 1330-1341.

Christenhusz MJM, Fay MF and Chase MW (2017) *Plants of the World: An Illustrated Encyclopedia of Vascular Plants*, pp. 154. Chicago, USA: University of Chicago Press.

Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M and Jacobsen SE (2008) Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* **452(7184)**: 215–219.

Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, Szcześniak MW, Gaffney DJ, Elo LL, Zhang XG and Mortazavi A (2016) A survey of best practices for RNA-seq data analysis. *Genome Biology* **17**: 13. doi: 10.1186/s13059-016-0881-8.

Cooper C, Crowther T, Smith BM, Isaac S and Collin HA (2006) Assessment of the response of carrot somaclones to *Pythium violae*, causal agent of cavity spot. *Plant Pathology* **55**: 427–432.

Costa C, Giménez-Capitán A, Karachaliou N and Rosell R (2013) Comprehensive molecular screening: from RT-PCR to RNA-seq. *Translational Lung Cancer Research* **2(2)**: 87-91.

Cote FX, Goue O, Domergue R, Panis B and Jenny C (2000) In-field behavior of banana plants (*Musa* spp.) obtained after regeneration of cryopreserved embryogenic cell suspensions. *Cryo-letters* **21**: 19-24.

Czechowski T, Stitt M, Altmann T, Udvardi MK and Scheible WR (2005) Genome wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Genome Analysis* **139**: 5–17.

D'Amato F (1975) The problem of genetic stability in plant tissue and cell cultures. In: OH Frankel OH and JG Hawkes (Eds.) *Crop genetic resources for today and tomorrow*, pp. 333–348. New York, USA: Cambridge University Press.

D'Amato F (1989) Polyploidy in cell differentiation. *Caryologia* **42**: 183-211.

da Fonseca RR, Albrechtsen A, Themudo GE, Ramos-Madrigal J, Sibbesen JA, Maretty L, Zepeda-Mendoza ML, Campos PF, Heller R and Pereira RJ (2016) Next-generation biology: Sequencing and data analysis approaches for non-model organisms. *Marine Genomics* **30:** 3-13.

Das A, Gosal SS, Sidhu JS and Dhaliwal HS (2000) Induction of mutations for heat tolerance in potato by using *in vitro* culture and radiation. *Euphytica* **114**: 205–209.

Davidonis G, Knorr D (1991) Callus formation and shoot regeneration in *Vanilla planifolia*. *Food Biotechnology* **5(1)**: 59–66.

Day ME, Greenwood MS and Diaz-Sala C (2002) Age- and size-related trends in woody plant shoot development: regulatory pathways and evidence for genetic control. *Tree Physiology* **22**: 507–513.

de la Cruz E, Rubluo IA, Palomino HG, Garcia AJM and Laguna CA (2007) Characterization of *Chenopodium* germplasm, selection of putative mutants and its cytogenetic study. In: Ochatt S and Jain MS (eds.). *Breeding of Neglected and Under-Utilized Crops, Spices and Herbs*. Science Publishers, Enfield, pp. 447. ISBN 978-1-57808-509-5.

De Laat AMM, Göhde W and Vogelzakg MJDC (1987) Determination of ploidy of single plants and plant populations by flow cytometry. *Plant Breeding* **99**: 303–307.

De Pauw MA, Remphrey WR, Palmer CE (1995) The cytokinin preference for *in vitro* germination and protocorm growth of *Cyperidium candidum. Annals of Botany* **75:** 267.

Debergh PC, Coster GD, Steurbaut W (1993) Carbendazim as an alternative plant growth regulator in tissue culture systems. *In vitro Cellular and Developmental Biology* **29(2)**: 89-91.

Deepak SA, Kottapalli KR, Rakwal R, Oros G, Rangappa KS, Iwahashi H, Masuo Y and Agrawal GK (2007) Real-time PCR: revolutionizing detection and expression analysis of genes. *Current Genomics* **8**: 234-251.

Demetriou K, Kapazoglou A, Tondelli A, Francia E, Stanca MA, Bladenopoulos K and Tsaftaris AS (2009) Epigenetic chromatin modifiers in barley: I. Cloning, mapping and expression analysis of the plant specific HD2 family of histone deacetylases from barley, during seed development and after hormonal treatment. *Physiol Plant* **136**: 358–368.

Deng ZN, Gentile A, Domina F, Nicolosi E and Tribulato E (1995) Selecting lemon protoplasts for insensitivity to *Phoma tracheiphila* toxin and regenerating tolerant plants. *J Am Soc Hortic Sci* **120**: 902–905.

Dequaire J (1976) L'amélioration du vanillier à Madagascar. *Journal D'Agriculture Tropicale Et De Botanique Appliquee* **23(7–12)**: 140–158.

Derveaux S, Vandesompele J and Hellemans J (2010) How to do successful gene expression analysis using real-time PCR. *Methods* **50**: 227-230.

Desvoyes B, Sanchez MP, Ramirez-Parra E and Gutierrez C (2010) Impact of nucleosome dynamics and histone modifications on cell proliferation during *Arabidopsis* development. *Heredity* **105**: 80–91.

Devanand PS, Chen J, Henny RJ and Chao CCT (2004) Assessment of genetic relationships among *Philodendron* cultivars using AFLP markers. *J Am Soc Hortic Sci* **129**: 690–697.

Devarumath RM, Nandy S, Rani V, Marimuthu S, Muraleedharan N and Raina SN (2002) RAPD, ISSR and RFLP fingerprints as useful markers to evaluate genetic integrity of micropropagated plants of three diploid and triploid elite tea clones representing *Camellia sinensis* (China type) and *C. assamica* ssp. assamica (Assam-India type). *Plant Cell Report* **21**: 166–173.

Diawara MM, Trumble JT, Lacy ML, White KK and Carson WG (1996) Potential of somaclonal celeries for use in integrated pest management. *J Econ Entomol* **89**: 218–223.

Diez MC, Osorio NW and Moreno F (2016) Effect of Dose and Type of Fertilizer on Flowering and Fruiting of Vanilla Plants. *Journal of Plant Nutrition* **39**: 1297-1310. doi: 10.1080/01904167.2015.1098673.
Ding C and Cantor CR (2004) Quantitative analysis of nucleic acids – the last few years of progress. *Journal Biochemistry and Molecular Biology* **37(1)**: 1-10.
Dolcet-Sanjuan R, Mok DWS and Mok MC (1992) Characterization and *in vitro* selection for iron efficiency in *Pyrus* and *Cydonia*. *In Vitro Cell Dev Biol Plant* **28**: 25–29.
Doležel J (1991) Flow cytometric analysis of nuclear DNA content in higher plants. *Phytochemical Analysis* **2**: 143–154.
Doležel J and Bartoš J (2005) Plant DNA flow cytometry and estimation of nuclear genome size. *Annals of Botany* **95**: 99–110.
Doležel J and Göhde W (1995) Sex determination in dioecious plants *Melandrium album* and *M. rubrum* using high-resolution flow cytometry. *Cytometry* **19**: 103–106.
Doležel J, Bartoš J, Voglmayr H, Greilhuber J (2003) Nuclear DNA content and genome size of trout and human. *Cytometry* **51A**: 127–128.
Doležel J, Binarová P and Lucretti S (1989) Analysis of nuclear DNA content in plant cells by flow cytometry. *Biologia Plantarum* **31:** 113–120.
Dolezel J, Sgorbati S and Lucretti S (1992) Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiologia Plantarum* **85**: 625 – 631.
Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, Dixon JE and Ecker JR (2012) Widespread dynamic DNA methylation in response to biotic stress. *PNAS* **109(32)**: 12858-12859.
Down TA, Rakyan VK, Turner DJ, Flicek P, Li H, Kulesha E, Graf S, Johnson N, Herrero J, Tomazou EM, Thorne NP, Bäckdahl L, Herberth M, Howe KL, Jackson DK, Miretti MM, Marioni JC, Birney E, Hubbard TJ, Durbin R, Tavaré S and Beck S (2008) A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nature Biotechnology* **26(7)**: 779–785.
Dreher K and Callis J (2007) Ubiquitin, hormones and biotic stress in plants. *Annals of Botany* **99(5)**: 787-822.
Dreier J, Störmer M and Kleesiek K (2007) Real-time polymerase chain reaction in transfusion medicine: application for detection of bacterial contamination in blood products. *Transfusion Medicine Reviews* **21(3)**: 237-254.
Dressler RL (2005) How many orchid species? *Selbyana* **26(1,2):** 155-158.
Dugdale LJ, Mortimer AM, Isaac S and Collin HA (2000) Disease response of carrot and carrot somaclones to *Alternaria dauci*. *Plant Pathol* **49**: 57–67.
Duke JA (2002) *Handbook of Medicinal Herbs*, 2nd Edition. Boca Raton, USA: CRC Press. ISBN: 0-8493-1284-1.
Duncan RR (1997) Tissue culture-induced variation and crop improvement. *Advances in Agronomy* **58**: 201–240.

Duran S and Karran P (2003) Vanillins: A novel family of DNA–PK inhibitors. *Nucleic Acids Research* **31**: 5501-5512.

Durand S, Bouche N, Perez Strand E, Loudet O and Camilleri C (2012) Rapid establishment of genetic incompatibility through natural epigenetic variation. *Current Biology* **22(4)**: 326–331.

Egesi CN, Pillay M, Asiedu R and Egunjobi JK (2002) Ploidy analysis in water yam, *Dioscoreaalata* L. germplasm. *Euphytica* **128:** 225–230.

Esmaiel NM, Al-Doss AA and Barakat MN (2012) *In vitro* selection for resistance to *Fusarium oxysporum f*. sp. dianthi and detection of genetic polymorphism via RAPD analysis in carnation. *J Med Plants Res* **6(23)**: 3997–4004.

Espy MJ, Uhl JR, Sloan LM, Buckwalter SP, Jones MF, Vetter EA, Yao JDC, Wengenack NL, Rosenblatt, Cockerill III FR and Smith TF (2006) Real-time PCR in clinical microbiology: applications for routine laboratory testing. *Clinical Microbiology Reviews* **Jan**: 165-256.

Etienne H and Bertrand B (2003) Somaclonal variation in *Coffea arabica*: effects of genotype and embryogenic cell suspension age on frequency and phenotype of variants. *Tree Physiology* **23**: 419–426.

Evans DA (1989) Somaclonal variation—genetic basis and breeding applications. *Trends Genet* **5**: 46–50.

Evrensel C, Yilmaz S, Temel A and Gozukirmizi N (2011) Variations in *BARE-1* insertion patterns in barley callus cultures. *Genet Mol Res* **10**: 980–987.

Excoffier L, Smouse PE and Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: applications to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.

Fan B, Borriss R, Bleiss W, Wu X (2012) Gram-positive rhizobacterium Bacillus amyloliquefaciens FZB42 colonizes three types of plants in different patterns. *The Journal of Microbiology* **50(1)**: 38-44.

Fang DQ and Roose ML (1997) Identification of closely related citrus cultivars with inter-simple sequence repeat markers. *Theoretical and Applied Genetics* **95**: 408–417.

Faria DC, Dias ACF, Melo IS, Carvalho-Costa FE (2013) Endophytic bacteria isolated from orchid and their potential to promote plant growth. *World Journal of Microbiology and Biotechnology* **29(2)**: 217-221.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783–791.

Feng S, Jacobsen SE and Reik W (2010b) Epigenetic reprogramming in plant and animal development. *Science* **330**: 622–627.

Ferdausi A, Nath UK, Das BL and Alam MS (2009) *In vitro* regeneration system in brinjal (*Solanum melongena* L.) for stress tolerant somaclone selection. *J Bangladesh Agric Univ* **7(2)**: 253–258.

Ferdous J, Li Y, Reid N, Langridge P, Shi B and Tricker PJ (2015) Identification of reference genes for quantitative expression analysis of microRNAs and mRNAs in barley under various stress conditions. *PLoS One* **10(3)**: e0118503.

Fiammengo R (2017) Can nanotechnology improve cancer diagnosis through miRNA detection? *Biomarkers in Medicine* **11(1)**: 69-86.

Fillion M (Ed.) (2012) *Quantitative real-time PCR in applied microbiology*. Poole, UK: Caister Academic Press. ISBN: 978-1-908230-01-0.

Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang HY, Dosztányi Z, El-Gebali S, Fraser M, Gough J, Haft D, Holliday GL, Huang H, Huang X, Letunic I, Lopez R, Lu S, Marchler-Bauer A, Mi H, Mistry J, Natale DA, Necci M, Nuka G, Orengo CA, Park Y, Pesseat S, Piovesan D, Potter SC, Rawlings ND, Redaschi N, Richardson L, Rivoire C, Sangrador-Vegas A, Sigrist C, Sillitoe I, Smithers B, Squizzato S, Sutton G, Thanki N, Thomas PD, Tosatto SC, Wu CH, Xenarios I, Yeh LS, Young SY and Mitchell AL (2016) InterPro in 2017 – beyond protein family and domain annotations. *Nucleic Acids Research* **45** (D1): D190–D199. doi:10.1093/nar/gkw1107.

Fitzgerald DJ, Stratford M and Narbad A (2003) Analysis of the inhibition of food spoilage yeasts by vanillin. *International Journal of Food Microbiology* **86**: 113-122.

Fleige S and Pfaffl MW (2006) RNA integrity and the effect on the real-time qRT-PCR performance. *Molecular Aspects of Medicine* **27**: 126-139.

Fluminhan A, de Aguiar-Perecin Jr MLR. and Dos Santos JA (1996) Evidence for heterochromatin involvement in chromosome breakage in maize callus culture. *Annals of Botany* **78**: 73-81.

Fock-Bastide I, Palama TL, Bory S, Lecolier A, Noirot M and Joet T (2014) Expression profiles of key phenylpropanoid genes during *Vanilla planifolia* pod development reveal a positive correlation between PAL gene expression and vanillin biosynthesis. *Plant Physiology and Biochemistry* **74**: 304–314.

Forestan C, Cigliano RA, Farinati S, Lunardon A, Sanseverino W and Varotto S (2016) Stress-induced and epigenetic-mediate maize transcriptome regulation study by means of transcriptome reannotation and differential expression analysis. *Scientific Reports* **6**: 30446. doi: 10.1038/srep30446.

Fouché JG and Jouve L (1999) *Vanilla planifolia*: history, botany and culture in Reunion island. *Agronomie* **19(8)**: 689-703.

Fraga CND, Couto DR, Pansarin ER (2017) Two new species of *Vanilla* (*Orchidaceae*) in the Brazilian Atlantic Forest. *Phytotaxa* **296(1)**: 63-72.

Fraga M, Caňal M and Rodríguez R (2002) Phase-change related epigenetic and physiological changes in *Pinus radiata* D. Don. *Planta* **215**: 672–678.

Francischini JHMB, Kemper EL, Costa JB, Manechini JRV and Pinto LR (2017) DNA methylation in sugarcane somaclonal variants assessed through methylation-sensitive amplified polymorphism. *Genetics and Molecular Research* **16(2)**: 1-12. doi: 10.4238/gmr16029585.

Franklin CI, Mott RL and Vuke TM (1989) Stable ploidy levels in long-term callus cultures of loblolly pine. *Plant Cell Reports* **8(2)**: 101-104. doi: 10.1007/BF00716849. PMID: 24232995.

Fras A and Maluszynska J (2003) Regeneration of diploid and tetraploid plants of *Arabidopsis thaliana* via callus. *Acta Biologica Cracoviensia Series Botanica* **45(2)**:145-152.
Fras A and Maluszynska J (2004) The correlation between the chromosome variation in callus and genotype of explants of *Arabidopsis thaliana. Genetica* **121**: 145-154.
Fras A, Juchimiuk J, Siwinska D and Malunszynska J (2007) Cytological events in explants of *Arabidopsis thaliana* during early callogenesis. *Plant Cell Reports* **26**: 1933-1939.
Fu W, Wu K and Duan J (2007) Sequence and expression analysis of histone deacetylases in rice. *Biochem Biophys Res Commun* **356**: 843–850.
Fuchs J, Demidov D, Houben A and Schubert I (2006) Chromosomal histone modification patterns — from conservation to diversity. *Trends in Plant Science* **11**: 199–208.
Fukai E, Umehara Y, Sato S, Endo M, Kouchi H, Hayashi M, Stougaard J and Hirochika H (2010) Derepression of the plant Chromovirus *LORE1* induces germline transposition in regenerated plants. *PLoS Genet* **6**: e1000868.
Fulneček J and Kovařík A (2014) How to interpret Methylation Sensitive Amplified Polymorphism (MSAP) profiles? *BMC Genet.* **15**: 2. doi: 10.1186/1471-2156-15-2.
Fulnecek J, Matyasek R, Kovarik A and Bezdek M (1998) Mapping of 5-methylcytosine residues in *Nicotiana tabacum* 5S rRNA genes by genomic sequencing. *Molecular Genetics and Genomics* **259(2)**: 133–141.
Funk C and Brodelius P (1994) *Vanilla planifolia* Andrews: *In vitro* biosynthesis of vanillin and other phenylpropanoid derivatives. In: Bajaj YPS (Ed.) *Biotechnology in Agriculture and Forestry. Medicinal and Aromatic Plants VI*, Volume 26, pp. 377–402. Berlin, Germany: Springer.
Funk C and Brodelius P (1990) Influence of growth regulators and an elicitor on phenylpropanoid metabolism in suspension cultures of *Vanilla planifolia*. *Phytochemistry* **29**: 845–848.
Galbraith DW (1984) Flow cytometric analysis of the cell cycle in higher plants. In: Vasil IK (ed.), *Cell culture and somatic cell genetics of plant*. New York; Academic Press, 765–777.
Galbraith DW (2009) Simultaneous flow cytometric quantification of plant nuclear DNA contents over the full range of described angiosperm 2C values. *Cytometry* **75A**: 692-698.
Galbraith DW, Harkins KR and Knapp S (1991) Systemic endopolyploidy in *Arabidopsis thaliana*. *Plant Physiology* **96**: 985-989.
Galbraith DW, Harkins KR Maddox JM, Aryes NM, Sharma DP and Firoozabady E. (1983) Rapid flow cytometric analysis of the cell cycle in intact plant-tissues. *Science* **220**: 1049–1051.
Galbraith DW, Lambert GM, Macas J and Doležel J (1998) Analysis of nuclear DNA content and ploidy in higher plants. In: Robinson, JP, Darzynkiewicz, Z, Dean, PN, Dressler, LG, Orfao, A, Rabinovitch, PS, Stewart, CC, Tanke, HJ and

Wheeless, LL (*eds.*), pp. 7.6.1.-7.6.22. *Current protocols in cytometry*. John Wiley & Sons, New York.
Gallage NH and Møller BL (2015) Vanillin - bioconversion and bioengineering of the most popular plant flavor and its de novo biosynthesis in the vanilla orchid. *Molecular Plant* **8(1)**: 40-57.
Gamborg OL, Miller RA and Ojima K (1968) Nutrient requirements of suspension cultures of soybean root cells. *Exp. Cell Res.* **50**: 151-158.
Gao X, Yang D, Cao D, Ao M, Sui X, Wang Q, Kimatu J and Wang L (2010) *In vitro* micropropagation of *Freesia hybrida* and the assessment of genetic and epigenetic stability in regenerated plantlets. *Journal of Plant Growth Regulation* **29**: 257–267.
García M, Esteve-Zarzoso B, Crespo J, Cabellos JM and Arroyo T (2017) Yeast monitoring of wine mixed or sequential fermentations made by native strains from D.O. "Vinos de Madrid" using real-time quantitative PCR. *Frontiers in Microbiology* **8**: Article 2520.
García PC, Ruiz JM, Rivero RM, Lopez-Lefebre LR, Sanchez E, Romero L (2002) Is the Application of Carbendazim Harmful to Healthy Plants? Evidence of Weak Phytotoxicity in Tobacco. *Journal of Agricultural and Food Chemistry* **50**: 279-283.
Garg R, Sahoo A, Tyagi AK and Jain M (2010) Validation of internal control genes for quantitative gene expression studies in chickpea. *Biochemical and Biophysical Research Communications* **396**: 283–288. doi: 10.3389/fmicb.2017.02520.
Gaspar T, Kevers C, Penel C, Greppin H, Reid DM and Thorpe TA (1996) Plant hormones and plant growth regulators in plant tissue culture. *In Vitro Cellular & Developmental Biology - Plant* **32(4)**: 272-289.
Gautheret RJ (1942) *Manuel Technique de Culture des Tissue Vegetaux*. Paris: Masson Cie.
Gee HE, Buffa FM, Camps C, Ramachandran E, Leek R, Taylor M, Patil M, Sheldon H, Betts G, Homer J, West C, Ragoussis J and Harris AL (2011) The small-nucleolar RNAs commonly used for microRNA normalisation correlate with tumour pathology and prognosis. *British Journal of Cancer* **104**: 1168-1177.
Geer L *et al.* (2002) CDART: protein homology by domain architecture. *Genome Res.* **12(10)**: 1619-1623.
Geetha S, Shetty SA (2000) *In vitro* propagation of *Vanilla planifolia*, a tropical orchid. *Current Science* **79**: 886-889.
George PS, Ravishankar GA (1997) *In vitro* multiplication of *Vanilla planifolia* using axillary bud explants. *Plant Cell Report* **16**: 490-494.
Gesteira AS, Otoni WC, Barros EG and Moreira MA (2002) RAPD based detection of genomic instability in soybean plants derived from somatic embryogenesis. *Plant Breeding* **121**: 269–271.
Ghaemi MA, Sarrafi A and Alibert G (1993) Influence of genotype and culture conditions on the production of embryos from anthers of tetraploid wheat (*Triticum turgidum*). *Euphytica* **65**: 81-85.

Ghag SB, Shekhawat UK and Ganapathi TR (2014) Characterization of Fusarium wilt resistant somaclonal variants of banana cv. Rasthali by cDNA-RAPD. *Mol Biol Rep* **41(12)**: 7929–7935.

Giacomelli JI, Weigel D, Chan RL and Manavella PA (2012) Role of recently evolved miRNA regulation of sunflower HaWRKY6 in response to temperature damage. *New Phytologist* **195**: 766–773.

Gilissen LJW, Staveren-Van MJ, Hakkert JC and Smulders MJM (1996) Competence for regeneration during tobacco intermodal development. *Plant Physiology* **111**: 1243-1250.

Giménez C, de García E, de Enrech NX and Blanca I (2001) Somaclonal variation in banana: cytogenetic and molecular characterization of the somaclonal variant CIEN BTA-03. *In Vitro Cellular & Developmental Biology - Plant* **37**: 217–222.

Gimenez MD, Yañez AM, Paz RC, Quiroga MP, Marfil CF, Conci VC and García-Lampasona SC (2016) Assessment of genetic and epigenetic changes in virus-free garlic (*Allium sativum* L.) plants obtained by meristem culture followed by *in vitro* propagation. *Plant Cell Report* **35(1)**: 129-141.

Giridhar P and Ravishankar GA (2004) Efficient micropropagation of *Vanilla planifolia* Andr. under the influence of thidiazuron, zeatin and coconut milk. *Indian Journal of Biotechnology (IJBT)* **3**: 113-118.

Giridhar P, Obun Reddy B, Ravishankar GA (2001) Silver nitrate influences *in vitro* shoot multiplication and root formation in *Vanilla planifolia* Andr. *Current Science* **81(9)**: 1166-1170.

Godwin ID, Aitken EAB and Smith LW (1997) Application of inter-simple sequence repeat (ISSR) markers to plant genetics. *Electrophoresis* **18**: 1524–1528.

Gopi C, Vatsala TM, Ponmurugan P (2006) *In vitro* multiple shoot proliferation and plant regeneration of *Vanilla planifolia* Andr. - A commercial spicy orchid. *Journal of Plant Biotechnology* **8(1)**: 37-41.

Gopinath C (1994) In: *Secret of vanilla – Farmer's notebook on vanilla* (pp. 31) Indian Spice Associates, Puttur, Karnataka.

Grafi G, Ben-Meir H, Avivi Y, Moshe M, Dahan Y and Zemach A (2007) Histone methylation controls telomerase-independent telomere lengthening in cells undergoing dedifferentiation. *Developmental Biology* **306**: 838–846.

Grandbastien MA (2015) LTR retrotransposons, handy hitchhikers of plant regulation and stress response. *Biochimica et. Biophysica Acta (BBA)* **1849**: 403-416.

Green MR and Sambrook J (2012) Extraction, purification, and analysis of RNA from eukaryotic cells. In: *Molecular cloning: a laboratory manual* (Fourth Edition), pp. 345-454. New York, USA: Cold Spring Harbor Laboratory Press. ISBN: 978-1-936113-41-5.

Griesbach RJ (1989) Selection of a dwarf *Hemerocallis* through tissue culture. *HortScience* **24**: 1027–1028.

Grosser JW, Gmitter FG Jr and Chandler JL (1996) Somaclonal variation in sweet orange — a tool for cultivar improvement. In: *Proceedings of International Society of Citriculture VIII Congress*, Sun City, South Africa, pp 203–206.

Gu C, Chen S, Liu Z, Shan H, Luo H, Guan Z and Chen F (2011) Reference gene selection for quantitative real-time PCR in *Chrysanthemum* subjected to biotic and abiotic stress. *Molecular Biotechnology* **49**: 192–197.

Gu Z, Arditti J, Nyman LP (1987) *Vanilla planifolia*: Callus induction and plantlet production *in vitro*. *Lindleyana* **2**: 48–52.

Guenin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C and Gutierrez L (2009) Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *Journal of Experimental Botany* **60**: 487–493.

Guo J, Ling H, Wu Q, Xu L and Que Y (2014) The choice of reference genes for assessing gene expression in sugarcane under salinity and drought stresses. *Scientific Reports* **4**: 7042.

Guo W, Wu R, Zhang Y, Liu X, Wang H, Gong L, Zhang Z and Liu B (2007) Tissue culture-induced locus-specific alteration in DNA methylation and its correlation with genetic variation in *Codonopsis lanceolata* Benth. et Hook. f. *Plant Cell Reports* **26**: 1297–1307.

Guo WL, Chang YCA, Kao CY (2010) Protocorm-like bodies initiation from root tips of *Cyrtopodium paranaense* (Orchidaceae). *HortScience* **45(9)**: 1365-1368.

Gupta M, Chyi YS, Romero-Severson J and Owen JL (1994) Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. *Theoretical and Applied Genetics* **89**: 998–1006.

Gupta PK (1998) Chromosomal basis of somaclonal variation in plants. In: Jain SM, Brar DS, Ahloowalia BS (eds) *Somaclonal variation and induced mutations in crop improvement*. Kluwer Academic Publishers, Dordrecht, pp 149–168.

Gupta PK, Rustgi S and Mir RR (2008) Array-based high-throughput DNA markers for crop improvement. *Heredity* **101**: 5–18.

Halim NAAH, Tan BC, Midin MR, Madon M, Khalid N and Yaacob JS (2017) Abscisic acid and salinity stress induced somaclonal variation and increased histone deacetylase (HDAC) activity in Ananas comosus var. MD2. Plant Cell, Tissue and Organ Culture (PCTOC): 1-13. doi: 10.1007/s11240-017-1367-4.

Hall HK, Skirvin RM and Braam WF (1986) Germplasm release of 'Lincoln Logan', a tissue culture-derived genetic thornless 'Loganberry'. *Fruit Var J* **40**: 134–135.

Hammerschlag FA (2000) Resistant responses of peach somaclone 122–1 to *Xanthomonas campestris* pv. pruni and to *Pseudomonas syringae* pv. syringae. *HortScience* **35**: 141–143.

Hammerschlag FA and Ognjanov V (1990) Somaclonal variation in peach: screening for resistance to *Xanthomonas camperstris* pv. pruni and *Pseudomonas syringae* pv. syringae. *Acta Hortic* **280**: 403–408.

Hao YJ and Deng XX (2002) Occurrence of chromosomal variations and plant regeneration from long-term-cultured citrus callus. *In Vitro Cellular & Developmental Biology - Plant* **38:** 472–476.
Hao YJ, Liu QL and Deng XX (2001) Effect of cryopreservation on apple genetic resources at morphological, chromosomal and molecular levels. *Cryobiology* **43(1)**: 46-53. doi: 10.1006/cryo.2001.2339.
Hasbún R, Valledor L, Berdasco M, Santamaría E, Caňal MJ and Rodríguez R (2007) Dynamics of DNA methylation in chestnut development. *Acta Horticulturae* **760**: 563–567.
Hashmi G, Hammerschlag FA, Huettel RN and Krusberg LR (1995) Growth, development and response of peach somaclones to the root-knot nematode, *Meloidogyne incognita*. *J Am Soc Hortic Sci* **120**: 932–937.
Havkin-Frenkel D and Belanger FC (2011) *Handbook of Vanilla Science and Technology*. Wiley-Blackwell, Oxford.
Heath-Pagliuso S and Rappaport L (1990) Somaclonal variant UC-T3: the expression of Fusarium wilt resistance in progeny arrays of celery, *Apium graveolens* L. *Theor Appl Genet* **80**: 390–394.
Hebbes TR, Thorne AW and Cranerobinson C (1988) A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO Journal* **7**: 1395–1402.
Heid CA, Stevens J, Livak KJ and Williams PM (1996) Real time quantitative PCR. *Genome Research* **6**: 986-994.
Heller FO (1973) DNS-Bestimmung an Keimwurzeln von Vicia faba L. mit Hilfe der Impulscytophotometrie. *Bericht der Deutschen Botanischen Gesellschaft* **86**: 437–441.
Henny RJ and Chen J (2011) *Tropical foliage plant development: origin of new cultivars. ENH1092.* Environmental Horticulture Department, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida. http://edis.ifas.ufl.edu.
Henny RJ, Chen J and Norman DJ (2003) 'Diamond Bay' and 'Emerald Bay'Aglaonema. *HortScience* **38**: 1446–1447.
Henny RJ, Poole RT and Conover CA (1992) 'Silver Bay' Aglaonema. *HortScience* **27**: 1238.
Herman E (1989) Non-axenic plant tissue culture: possibilities and opportunities. *Acta Horticulturae* **280**: 233–238.
Hernández-Hernández J (2011) "Vanilla Diseases". in Handbook of Vanilla Science and Technology (eds D. Havkin-Frenkel and F. C. Belanger), Wiley-Blackwell, Oxford, UK. doi: 10.1002/9781444329353.ch9.
Hewelt A, Prinsen E, Schell J, Van OH, Schmulling T (1994) Promoter tagging with a promoterless *ipt* gene leads to cytokinin-induced phenotypic variability in transgenic tobacco plants: implications of gene dosage effects. *The Plant Journal* **6**: 879-891.

Higuchi R, Fockler C, Dollinger G and Watson R (1993) Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Nature Biotechnology* **11**: 1026-1030.

Hintze J (2004) *NCSS and PASS*. Kaysville, Utah: Number Cruncher Statistical Systems.

Hirochika H (1993) Activation of tobacco retrotransposons during tissue culture. *EMBO J* **12**: 2521–2528.

Hirochika H, Sugimoto K, Otsuki Y, Tsugawa H and Kanda M (1996) Retrotransposons of rice involved in mutations induced by tissue culture. *Proc Natl Acad Sci USA* **93**: 7783–7788.

Hoang NH, Kane ME, Radcliffe EN, Zettler LW and Richardson LW (2017) Comparative seed germination and seedling development of the ghost orchid, *Dendrophylax lindenii* (Orchidaceae), and molecular identification of its mycorrhizal fungus from South Florida. *Annals of Botany* **119**: 379-393. doi:10.1093/aob/mcw220.

Hodel RGJ, Gitzendanner MA, Germain-Aubrey CC, Liu XX, Crowl AA, Sun M, Landis JB, Segovia-Salcedo MC, Douglas NA, Chen S, Soltis DE and Soltis PS (2016) A new resource for the development of SSR markers: Millions of loci from a thousand plant transcriptome. *Applications in Plant Sciences* **4(6)**: 1600024.

Holst-Jensen A, Røning SB, Løvseth A and Berdal KG (2003) PCR technology for screening and quantification of genetically modified organisms (GMOs). *Analytical and Bioanalytical Chemistry* **375**: 985-993. doi: 10.1007/x00216-003-1767-7.

Horáček J, Švábová L, Šarhanová P and Lebeda A (2013) Variability for resistance to *Fusarium solani* culture filtrate and fusaric acid among somaclones in pea. *Biol Plant* **57(1)**: 133–138.

Hossain MA, Konisho K, Minami M and Nemoto K (2003) Somaclonal variation of regenerated plants in chili pepper (*Capsicum annuum* L.). *Euphytica* **130**: 233–239.

Hřibová E, Holušova K, Trávníček P, Petrovská B, Ponert J, Šimková H, Kubátová B, Jersáková J, Čurn V, Suda J, Doležel J and Vrána J (2016) The enigma of progressively partial endoreplication: new insights provided by flow cytometry and next-generation sequencing. *Genome Biology and Evolution* **8(6)**: 1996-2005.

Hu R, Fan C, Li H, Zhang Q and Fu YF (2009) Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. *BMC Molecular Biology* **10**: 93–104.

Hu Y, Qin F, Huang L, Sun Q, Li C, Zhao Y and Zhou DX (2009) Rice histone deacetylase genes display specific expression patterns and developmental functions. *Biochem Biophys Res Commun* **388**: 266–271.

Huan, LVT, Takamura T, Tanaka M (2004) Callus formation and plant regeneration from callus through somatic embryo structures in *Cymbidium* orchid. *Plant Science* **166**: 1443-1449.

Huang J, Zhang K, Shen Y, Huang Z, Li M, Tang D, Gu M and Cheng Z (2009) Identification of a high frequency transposon induced by tissue culture, *nDaiZ*, a member of the hAT family in rice. *Genomics* **93**: 274–281.

Hülgenhof E, Weidhase RA, Schlegel R and Tewes A (1988) Flow cytometric determination of DNA content in isolated nuclei of cereals. *Genome* **30**: 565–569.

Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S, De Castro E, Coggill P, Corbett M, Das U, Daugherty L, Duquenne L, Finn RD, Fraser M, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C and McDowall J (2011) InterPro in 2011: New developments in the family and domain prediction database. *Nucleic Acids Research* **40** (Database issue): D306–D312. doi:10.1093/nar/gkr948.

Hussain A, Qarshi IA, Nazir H and Ullah I (2012) Plant Tissue Culture: Current Status and Opportunities. In: Dr. Annarita Leva (Ed.) *Recent Advances in Plant in vitro Culture*. InTech. DOI: 10.5772/50568. Available from: https://www.intechopen.com/books/recent-advances-in-plant-in-vitro-culture/plant-tissue-culture-current-status-and-opportunities

Hwang S (2002) Somaclonal variational approach to breeding Cavendish banana for resistance to Fusarium wilt race 4. In: *Global conference on banana and plantain*, Bangalore, India, Abstract 57, October 28–31.

Hwang SC and Ko WH (1992) Breeding for resistance to Fusarium wilt of Cavendish banana by using tissue culture method. *Spec Publ Taichung Dist Agric Improv Stn **29***: 229–237.

Iacona C, Cirilli M, Zega A, Frioni A, Silvestri C and Muleo R (2013) A somaclonal myrobalan rootstock increases waterlogging tolerance to peach cultivar in controlled conditions. *Sci Hortic* **156**: 1–8.

Ichiyanagi T, Ichiyanagi K, Miyake M and Sasaki H (2013) Accumulation and loss of asymmetric non-CpG methylation during male germ-cell development. *Nucleic Acids Research* **41(2)**: 738–745.

Ikeuchi M, Sugimoto K and Iwase A (2013) Plant Callus: Mechanisms of Induction and Repression. *The Plant Cell* **25**: 3159-3173.

IPNI (2017) *International Plant Names Index*. Published on the Internet: http://www.ipni.org, The Royal Botanic Gardens, Kew, Harvard University Herbaria & Libraries and Australian National Botanic Gardens. [Retrieved 20 December 2017].

Israeli Y, Reuveni O and Lahav E (1991) Qualitative aspects of somaclonal variations in banana propagated by *in vitro* techniques. *Scientia Horticulturae* **48**: 71–88.

Jaccard P (1908) Nouvelle recherches sur La distribution florale. *Bull Soc Vaud Sci Nat* **44**: 223–270.

Jaccoud D, Peng K, Feinstein D and Kilian A (2001) Diversity Arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Research* **29(4)**: e25.

Jagadeeswaran G, Li YF and Sunkar R (2014) Redox signaling mediates the expression of a sulfate-deprivation inducible microRNA395 in *Arabidopsis. Plant Journal* **77**: 85–96.

Jain M, Nijhawan A, Tyagi AK and Khurana JP (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *Biochemical and Biophysical Research Communications* **345**: 646–651.

Jain SM (1997) Micropropagation of selected somaclones of *Begonia* and *Saintpaulia. J Biosci* **22**: 582–592.

Jaligor E, Alder S, Debladis E, Beule T, Richaud F, Ilbert P, Finnegan EJ and Rival A (2011) Epigenetic imbalance and the floral developmental abnormality of the *in vitro*-regenerated oil palm *Elaeis guineensis*. *Annals of Botany* **108**: 1453-1462.

Jaligot E, Beulé T, Baurens FC, Billotte N and Rival A (2004) Search for methylation-sensitive amplification polymorphisms associated with the 'mantled' variant phenotype in oil palm (*Elaeis guineensis* Jacq.). *Genome* **47**: 224–228.

Jaligot E, Rival A, Beulé T, Dussert S and Verdeil JL (2000) Somaclonal variation in oil palm (*Elaeis guineensis* Jacq.): the DNA methylation hypothesis. *Plant Cell Reports* **19**: 684–690.

James AC, Peraza-Echeverria S, Herrera-Valencia VA and Martinez O (2004) Application of the amplified fragment length polymorphism (AFLP) and the methylation-sensitive amplification polymorphism (MSAP) techniques for the detection of DNA polymorphism and changes in DNA methylation in micropropagated bananas. In: SM Jain and R Swennen (Eds.) *Banana improvement: cellular, molecular biology, and induced mutations*, pp. 287-306. Enfield, US: Science Publishers Inc.

Janarthanam B, Seshadri S (2008) Plantlet regeneration from leaf derived callus of *Vanilla planifolia* Andr. *In Vitro Cellular & Developmental Biology - Plant* **44**: 84-89.

Jaroszeski MJ and Radcliff G (1999) Fundamentals of flow cytometry. *Molecular Biotechnology* **11**: 37-54.

Jaya RS, Rao PS and Mhatre M (2002) Suitability of RAPD for analysing spined and spineless variant regenerants of pineapple (*Ananas comosus* L., Merr.). *Plant Mol Biol Rep* **20**: 307.

Jevremović S, Subotic´ A, Miljkovic´ D, Trifunovic´ M, Petric´ M and Cingel A (2012) Clonal fidelity of Chrysanthemum cultivars after long term micropropagation by stem segment culture. *Acta Hortic* **961**: 211–216.

Jha S, Sen J and Sen S (1989) Stable regenerants from long-term callus cultures of *Ruscus hypophyllum* L. *Cytologia* **54**: 687-691.

Jhang T, Kaur M, Kalia P and Sharma TR (2010) Efficiency of different marker systems for molecular characterization of subtropical carrot germplasm. *Journal of Agricultural Science* **148(2)**: 171-181.

Jia X, Ren L, Chen QJ, Li R and Tang G (2009) UV-B-responsive microRNAs in *Populus tremula*. *Journal of Plant Physiology* **166**: 2046–2057.
Jiang N, Bao Z, Zhang X, Hirochika H, Eddy SR, McCouch SR and Wessler SR (2003) An active DNA transposon family in rice. *Nature* **421**: 163–167.
Jin S, Mushke R, Zhu H, Tu L, Lin Z, Zhang Y and Zhang X (2008) Detection of somaclonal variation of cotton (*Gossypium hirsutum*) using cytogenetics, flow cytometry and molecular markers. *Plant Cell Report* **27**: 1303–1316.
Jones N, Ougham H, Thomas H and Pašakinskienė I (2009) Markers and mapping revisited: finding your gene. *New Phytologist* **183**: 935–966.
Jones W, Kuehnle A and Arumuganathan K (1998) Nuclear DNA content of 26 orchid (Orchidaceae) genera with emphasis on *Dendrobium*. *Annals of Botany* **82**: 189 – 194.
Jones WE and Kuehnle AR (1998) Ploidy identification using flow cytometry in tissues of *Dendrobium* species and cultivars. *Lindleyana.* **13**: 11–18. doi: 10.1017/S0268416098003051.
Jones-Rhoades MW and Bartel, DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Molecular Cell* **14**: 787–799.
Jover-Gil S, Candela H and Ponce MR (2005) Plant microRNAs and development. *The International Journal of Developmental Biology* **49**: 733-744.
Kalendar R, Grob T, Regina M, Suoniemi A and Schulman A (1999) IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. *Theoretical and Applied* Genetics **98(5)**: 704-711.
Kalimuthu K, Senthilkumar R, Murugalatha N (2006) Regeneration and mass multiplication of *Vanilla planifolia* Andr. – a tropical orchid. *Current Science* **91**: 1401-1403.
Kanehisa M (2000) In: *Post-genome Informatics*. England, UK: Oxford University Press.
Kanehisa M and Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* **28**: 27-30.
Kanehisa M, Furumichi M, Tanabe M, Sato Y and Morishima K (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research* **45**: D353-D361.
Kanehisa M, Sato Y, Kawashima M, Furumichi M and Tanabe M (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research* **44**: D457-D462.
Kar B, Kuanar A, Singh S, Mohanty S, Joshi RK, Subudhi E and Nayak S (2014) In vitro induction, screening and detection of high essential oil yielding somaclones in turmeric (*Curcuma longa* L.). *Plant Growth Regul* **72(1)**: 59–66.
Karp A (1995) Somaclonal variation as a tool for crop improvement. *Euphytica* **85**: 295-302.
Kaundun SS and Matsumoto S (2003) Development of CAPS markers based on three key genes of the phenylpropanoid pathway in tea, *Camellia sinensis* (L.)

O. Kuntze, and differentiation between assamica and sinensis varieties. *Theoretical and Applied Genetics* **106(3)**: 375-383.

Kausch AP and Horner HT (1984) Increased nuclear DNA content in raphid crystal idioblasts during development in *Vanilla planifolia* L. (*Orchidaceae*). *European Journal of Cell Biology* **33**: 7-12.

Keb-Llanes M, González G, Chi-Manzanero B and Infante D (2002) A rapid and simple method for small-scale DNA extraction in *Agavaceae* and other tropical plants. *Plant Molecular Biology Reporter* **20**: 299a-299e.

Keshava C, Keshava N, Ong TM and Nath J (1998) Protective effect of vanillin on radiation-induced micronuclei and chromosomal aberrations in V79 cells. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **397(2)**: 149-159. doi: 10.1016/S0027-5107(97)00203-0.

Khan SA, Rahman LU, Shanker K and Singh M (2014) *Agrobacterium tumefaciens*-mediated transgenic plant and somaclone production through direct and indirect regeneration from leaves in *Stevia rebaudiana* with their glycoside profile. *Protoplasma* **251(3)**: 661–670.

Kidner CA and Martienssen RA (2005) The developmental role of microRNA in plants. *Curr Opin Plant Biol* **8**: 38–44.

Kidwell KK and Osborn TC (1993) Variation among alfalfa somaclones in copy number of repeated DNA sequences. *Genome* **36**: 906-912.

Kikuchi K, Terauchi K, Wada M and Hirano HY (2003) The plant MITE *mPing* is mobilized in anther culture. *Nature* **421**: 167–170.

Kirti PB, Hadi S, Kumar PA and Chopra VL (1991) Production of sodium-chloride-tolerant *Brassica juncea* plants by *in vitro* selection at the somatic embryo level. *Theoretical and Applied Genetics* **83**: 233-237.

Kitchen RR, Kubista M and Tichopad A (2010) Statistical aspects of quantitative real-time PCR experiment design. *Methods* **50**: 231-236.

Kitimu SR, Taylor J, March TJ, Tairo F, Wilkinson MJ and Lopez CMR (2015) Meristem micropropagation of cassave (*Manihot esculenta*) evokes genome-wide changes in DNA methylation. *Frontier in Plant Sciences* **6 (590)**: 1-12. doi: 10.3389/fpls.2015.00590.

Kneifel W and Leonhardt W (1992) Testing of different antibiotics against gram-positive and gram-negative bacteria isolated from plant tissue culture. *Plant Cell Tissue Organ Cul* **29**: 139-144.

Knudson L (1950) Germination of seeds of *Vanilla. American Orchid Society Bulletin* **41**: 443-445.

Kolano B, Siwinska D and Maluszynska J (2008) Endopolyploidy patterns during development of chenopodium quinoa. *Acta Biologica Cracoviensia Series Botanica* **51/2**: 85-92.

Komatsu M, Shimamoto K and Kyozuka J (2003) Two-step regulation and continuous retrotransposition of the rice LINE-type retrotransposon *Karma*. *Plant Cell* **15**: 1934–1944.

Koně M, Patat-Ochatt EM, Conreux C, Sangwan RS and Ochatt SJ (2007) *In vitro* morphogenesis from cotyledon and epicotyl explants and flow cytometry

distinction between landraces of Bambara groundnut [*Vigna subterranea* (L.) Verdc] an Underutilised grain legume. *Plant Cell Tissue Organ Cult* **88**: 61–75.
Kononowicz H and Janick J (1984) *In vitro* propagation of *Vanilla planifolia*. *HortScience* **19**: 58–59.
Koressaar T and Remm M (2007) Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23(10)**: 1289-1291.
Kornet N and Scheres B (2009) Members of the GCN5 histone acetyltransferase complex regulate PLETHORA-mediated root stem cell niche maintenance and transit amplifying cell proliferation in Arabidopsis. *The Plant Cell* **21**: 1070–1079.
Koroban NV, Kudryavtseva AV, Krasnov GS, Sadritdinova AF, Fedorova MS, Snezhkina AV, Bolsheva NL, Muravenko OV, Dmitriev AA and Meinikova NV (2016) The role of microRNA in abiotic stress response in plants. *Molecular Biology* **50(3)**: 337-343.
Kou SJ *et al*. (2012) Selection and validation of suitable reference genes for miRNA expression normalization by quantitative RT-PCR in citrus somatic embryogenic and adult tissues. *Plant Cell Reports* **31**: 2151–2163.
Kouzarides T (2007) Chromatin modifications and their function. *Cell* **128**: 693–705.
Krishna H, Alizadeh M, Singh D, Singh U, Chauhan N, Eftekhari M and Sadh RK (2016) Somaclonal variations and their applications in horticultural crops improvement. *3 Biotech* **6**: 54. doi: 10.1007/s13205-016-0389-7.
Kuanar A, Nayak PK, Subudhi E and Nayak S (2014) *In vitro* selection of turmeric somaclone resistant to *Fusarium oxysporum* f. sp. Zingiberi. *Proc Natl Acad Sci India Sect B Biol Sci* **84**:1077–1082.
Kubista M, Andrade JM, Bengtsson M, Forootan A, Jonák J, Lind K, Sindelka R, Sjöback R, Sjögreen B, Strömbom L, Ståhlberg A and Zoric N (2006) The real-time polymerase chain reaction. *Molecular Aspects of Medicine* **27**: 95-125.
Kuderova A, Urbankova I, Valkova M, Malbeck J, Nemethova D, Hejatko J (2008) Effects of conditional IPT-dependent cytokinin overproduction on root architecture of *Arabidopsis* seedlings. *Plant and Cell Physiology* **49**: 570-582.
Kudo N and Kimura Y (2001a) Flow cytometric evidence for endopolyploidy in seedlings of some Brassica species. *Theoretical and Applied Genetics* **102**: 104–110. doi: 10.1007/s001220051624.
Kudo N and Kimura Y (2001b) Flow cytometric evidence for endopolyploidization in cabbage (*Brassica oleracea* L.) flowers. *Sexual Plant Reproduction* **13:** 279–283. doi: 10.1007/s004970100066
Kudo N and Kimura Y (2002) Nuclear DNA endoreduplication during petal development in cabbage: relationship between ploidy levels and cell size. *Journal of Experimental Botany* **53**: 1017–1023. doi: 10.1093/jexbot/53.371.1017.
Kudo N and Kitamura Y (2001) Patterns of endopolyploidy during seedling development in cabbage (*Brassica oleracea* L.). *Annals of Botany* **87**: 275-281.

Kukreja AK, Dhawan OP, Ahuja PS, Sharma S and Kumar S (2000) Yield potential and stability behaviour of *in vitro* derived somaclones of Japanese mint (*Mentha arvensis* L.) under different environments. *J Genet Breed* **54(2)**: 109–115.

Kukreja AK, Dhawan OP, Mathur AK, Ahuja PS and Mandal S (1991) Screening and evaluation of agronomically useful somaclonal variations in Japanese mint (*Mentha arvensis* L.). *Euphytica* **53**: 183–191.

Kuksova VB, Piven NM and Gleba YY (1997) Plant cell variation and *in vitro* induced mutagenesis in grapevine. *Plant Cell Tissue Organ Cult* **49**: 17–27.

Kukurba KR and Montgomery SB (2015) RNA sequencing and analysis. *Cold Spring Harbour Protocols* **2015(11)**: 951-969. doi: 10.1101/pdb.top084970**.**

Kunitake H, Nakashima T, Mori K and Tanaka M (1998) Somaclonal and chromosomal effects of genotype, ploidy and culture duration in *Asparagus officinalis* L. *Euphytica* **102**: 309-316.

Kuznetsova OI, Ash OA, Hartina GA and Gostimskij SA (2005) RAPD and ISSR analyses of regenerated pea *Pisum sativum* L. plants. *Russian Journal of Genetics* **41**: 60–65.

Lakshmanan V (2007) Molecular analysis of genetic stability in long-term micropropagated shoots of banana using RAPD and ISSR markers. *Electronic Journal of Biotechnology* **10(1)**: 107-113. doi: 10.2225/vol10-issue5-fulltext-12.

Landey RB, Cenci A, Guyot R, Bertrand B, Georget F, Dechamp E, Herrera JC, Aribi J, Lashermes P and Etienne H (2015) Assessment of genetic and epigenetic changes during cell culture ageing and relations with somaclonal variation in *Coffea arabica*. *Plant Cell, Tissue and Organ Culture* **122(3)**: 517-531.

Larkin P and Scowcroft W (1981) Somaclonal variation a novel source of variability from cell cultures for plant improvement. *Theoretical and Applied Genetics* **60**: 197-214.

Larkins BA, Dikes BP, Dante RA, Coelho CM, Woo YM and Liu Y (2001) Investigating the hows and whys of DNA endoreduplication. *Journal of Experimental Botany* **52**: 183-192.

Law RD and Suttle JC (2005) Chromatin remodeling in plant cell culture: patterns of DNA methylation and histone H3 and H4 acetylation vary during growth of asynchronous potato cell suspensions. *Plant Physiology and Biochemistry* **43**: 527–534.

Lee M and Phillips RL (1987) Genomic rearrangements in maize induced by tissue culture. *Genome* **29**: 122-128.

Lee M and Phillips RL (1988) The chromosomal basis of somaclonal variation. *Annual Review of Plant Physiology and Plant Molecular Biology* **39**: 413–437.

Lee SY, Su YU, Chou CS, Liu CC, Chen CC and Chao CP (2011) Selection of a new somaclone cultivar 'Tai-Chiao No. 5' (AAA, Cavendish) with resistance to Fusarium wilt of banana in Chinese Taipei. *Acta Hortic* **897**: 391–397.

Lee-Espinosa HE, Murguía–González J, García-Rosas B, Córdova-Contreras AL, Laguna-Cerda A, Mijangos-Cortés JO, Barahona-Pérez LF, Iglesias-

Andreu LG, Santana-Buzzy N (2008) *In vitro* clonal propagation of vanilla (*Vanilla planifolia* 'Andrews'). *HortScience* **43(2)**: 454–458.

Lefever S, Hellemans J, Pattyn F, Przybylski DR, Taylor C, Geurts R, Untergasser A and Vandesompele J (2009) RDML: structured language and reporting guidelines for real-time quantitative PCR data. *Nucleic Acids Research* **37(7)**: 2065-2069.

Leitch I, Kahandawala I, Suda J, Hanson L, Ingrouille M, Chase M, and Fay M (2009) Genome size diversity in orchids: Consequences and evolution. *Annals of Botany* **104**: 469 – 481.

Leitch IJ and Dodsworth S (2017) Endopolyploidy in Plants. In: eLS. John Wiley & Sons, Ltd: Chichester. doi: 10.1002/9780470015902.a0020097.pub2.

Lepers-Andrzejewski S, Brunschwig C, Collard FX and Dron M (2010) Morphological, chemical, sensory and genetic specificities of Tahitian vanilla. In E. Odoux and M. Grisoni *(eds.), Vanilla*, pp. 205 – 228. CRC Press, Boca Raton, Florida, USA.

Lepers-Andrzejewski S, Siljak-Yakovlev S., Brown SC, Wong M and Dron M (2011) Diversity and dynamics of plant genome size: An example of polysomaty from a cytogenetic study of Tahitian vanilla (*Vanilla x tahitensis*, *Orchidaceae*). *American Journal of Botany* **98(6)**: 986-997.

Leus L, Laere K, Dewitte A and Huylenbroeck J (2009). Flow cytometry for plant breeding. *Acta Horticulturae* **836**: 221-226. doi: 10.17660/ActaHortic.2009.836.31.

Leva AR, Petruccelli R and Rinaldi LMR (2012) Somaclonal variation in tissue culture: a case study with olive. In: Leva AR, Rinaldi LMR (eds) *Recent advances in plant in vitro culture.* INTECH Open Access Publisher, Croatia, pp 123–150.

Li B, Carey M and Workman JL (2007) The role of chromatin during transcription. *Cell* **128**: 707–719.

Li G and Quiros CF (2001) Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in *Brassica*. *Theoretical and Applied Genetics* **103(2-3)**: 455-461.

Li R and Stelly DM (1989) Cytogenetic abnormalities in Cotton (*Gossipium hirsutum* L.) cell cultures. *Genome* **32**: 1128-1134.

Li R, Qu R, Bruneau AH and Livingston DP (2010) Selection for freezing tolerance in St. Augustine grass through somaclonal variation and germplasm evaluation. *Plant Breed* **129**: 417–421.

Li X, Mo X, Shou H and Wu P (2006) Cytokinin-mediated cell cycling arrest of pericycle founder cells in lateral root initiation of *Arabidopsis*. *Plant and Cell Physiology* **47**: 1112-1123.

Li X, Yu X, Wang N, Feng Q, Dong Z, Liu L, Shen J and Liu B (2007) Genetic and epigenetic instabilities induced by tissue culture in wild barley (*Hordeum brevisubulatum* (Trin.) Link). *Plant Cell, Tissue and Organ Culture (PCTOC)* **90**: 153–168.

Li Y, Zhao SL, Li JL, Hu XH,Wang H, Cao XL, Xu YJ, Zhao ZX, Xiao ZY, Yang N and Fan J (2017b) Osa-miR169 negatively regulates rice immunity against the blast fungus *Magnaportheoryzae*. *Front. Plant Sci.* **8**: 2.
Lim U and Song MA (2012) Dietary and lifestyle factors of DNA methylation. *Methods in Molecular Biology* **863**: 359–376.
Lim WL and Loh CS (2003) Endopolyploidy in *Vanda* Miss Joaquim (*Orchidaceae*). *New Phytology* **159:** 279-287.
Linacero R and Vazquez AM (1992a) Genetic analysis of chlorophyll deficient somaclonal variants in rye. *Genome* **35**: 981-984.
Linacero R and Vazquez AM (1992b) Cytogenetic variation in rye regenerated plants and their progeny. *Genome* **35**: 428-430.
Ling H, Wu Q, Guo J, Xu L and Que Y (2014) Comprehensive selection of reference genes for gene expression normalization in sugarcane by real time quantitative RT-PCR. *PLoS ONE* **9:** e97469.
Lister R and Ecker JR (2009) Finding the fifth base: Genome-wide sequencing of cytosine methylation. *Genome Res*earch **19(6)**: 959–966.
Lister R, Gregory BD and Ecker JR (2009) Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. *Current Opinion in Plant Biology* **12**: 107–118.
Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH and Ecker JR (2008) Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* **133**: 523–536.
Litz RE, Mathews WH, Hendrix RC and Yurgalevitch C (1991) Mango somatic cell genetics. *Acta Hortic* **291**: 133–140.
Liu HH, Tian X, Li YJ, Wu CA and Zheng CC (2008) Microarray-based analysis of stress regulated microRNAs in *Arabidopsis thaliana*. *RNA* **14**: 836–843.
Liu W, Yu W, Hou L, Wang X, Zheng F, Wang WX, Liang D, Yang HL, Jin Y and Xie XM (2014) Analysis of miRNAs and Their Targets during Adventitious Shoot Organogenesis of *Acacia crassicarpa*. *PLoS ONE* **9(4)**: e93438. doi: 10.1371/journal.pone.0093438.
Long B, Niemiera AX, Cheng ZY, Long CL (2010) *In vitro* propagation of four threatened *Paphiopedilum* species (*Orchidaceae*). *Plant Cell, Tissue and Organ Culture (PCTOC)* **101**: 151–162.
Loureiro J, Rodriguez E, Doležel J and Santos C (2006a) Flow cytometric and microscopic analysis of the effect of tannic acid on plant nuclei and estimation of DNA content. *Annals of Botany* **86**: 309-316.
Loureiro J, Rodriguez E, Doležel J and Santos C (2006b) Comparison of Four Nuclear Isolation Buffers for Plant DNA Flow Cytometry. *Annals of Botany* **98**: 679-689.
Lu S, Sun YH and Chiang VL (2008) Stress-responsive microRNAs in Populus. *Plant Journal* **55**: 131-151.
Lu Y, Shen YJ, Warren W and Walter R (2016) Next Generation Sequencing in Aquatic Models. In: Jerzy K (Ed.) *Next Generation Sequencing - Advances,*

Applications and Challenges. InTech. doi: 10.5772/61657. Available from: https://mts.intechopen.com/books/next-generation-sequencing-advances-applications-and-challenges/next-generation-sequencing-in-aquatic-models.

Lubinsky P (2003) Conservation of wild vanilla. In: *First International Congress on the Future of the Vanilla Business*. Princeton, NJ, USA, 11-12 Nov 2003.

Lucia G, Castiglione MR, Turrini A, Ronchi VN and Geri C (2011) Cytogenetic and histological approach for early detection of "mantled" somaclonal variants of oil palm regenerated by somatic embryogenesis: first results on the characterization of regeneration system. *Caryologia* **64(2)**: 223-234.

Luo M, Gao Z, Li H, Li Q, Zhang C, Xu W, Song S, Ma C and Wang S (2018) Selection of reference genes for miRNA qRT-PCR under abiotic stress in grapevine. *Scientific Reports* **8**: 4444.

Ma XJ, Lv SB, Zhang C, Yang CP (2013) Histone deacetylases and their functions in plants. *Plant Cell Reports* **32**: 465-478.

Mackay IM (2004) Real-time PCR in the microbiology laboratory. *Clinical Microbiology and Infection* **10**: 190-212.

Magyar Z, De Veylder L, Atanassova A, Bako L, Inze D and Bogre L (2005) The role of the Arabidopsis E2FB transcription factor in regulating auxin-dependent cell division. *The Plant Cell* **17**: 2527-2541.

Mahdieh N and Rabbani B (2013) An overview of mutation detection methods in genetic disorders. *Iranian Journal of Pediatrics* **23(4)**: 375-388.

Malabadi RB, Nataraja K (2007) Genetic transformation of *Vanilla planifolia* by *Agrobacterium-tumefaciens* using shoot tip sections. *Research Journal of Botany* **2(2)**: 86-94.

Mallory A and Bouché N (2008) MicroRNA-directed regulation: to cleave or not to cleave. *Trends in Plant Science* **13(7)**: 359-367.

Mallory AC, Reinhart BJ, Jones-Rhoades MW, Tang G, Zamore PD, Barton MK and Bartel DP (2004) MicroRNA control of PHABULOSA in leaf development: important of pairing to the microRNA 5' region. *EMBO Journal* **23**: 3356-3364.

Mani BK, Murthy V, Boland M, Yee K (2017) Analysis of constituents in different fractions collected during distillation of Cardamom oil for flavour and fragrance applications. *Journal of Applied Pharmaceutical Science* **7(01)**: 177-183. doi: 10.7324/JAPS.2017.70125.

Marchler-Bauer A and Bryant SH (2004) CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* **32**: W327-331.

Marchler-Bauer A *et al.* (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* **39**: D225-229.

Marchler-Bauer A *et al.* (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res.***4**: D222-226.

Marchler-Bauer A *et al.* (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* **45(D1)**: D200-D203.

Marino G, Beghelli S, Rombola` AD and Cabrini L (2000) *In vitro* performance at high culture pH and *in vitro* responses to Fe deficiency of leaf-derived quince BA29 (*Cydonia oblonga*) somaclones regenerated at variable pH. *J Hortic Sci Biotechnol* **75**: 433–440.

Maroufi A, Bockstaele EV and Loose MD (2010) Validation of reference genes for gene expression analysis in chicory (*Cichorium intybus*) using quantitative real-time PCR. *BMC Molecular Biology* **11:** 15–27.

Martin K, Pachathundikandi S, Zhang C, Slater A and Madassery J (2006) RAPD analysis of a variant of banana (*Musa* sp.) cv. Grande Naine and its propagation via shoot tip culture. *In Vitro Cellular & Developmental Biology - Plant* **42**: 188–192.

Mason MG, Mathews DE, Argyros DA, Maxwell BB, Kieber JJ, Alonso JM, Joseph RE, Schaller GE (2005) Multiple type-B response regulators mediate cytokinin signal transduction in *Arabidopsis*. *Plant Cell* **17**: 3007-3018.

Massey N (2017) Fine Vanillas & Flavors Crop Report: October 2017. Published on the Internet: https://nielsenmassey.wpengine.com/wp-content/uploads/2017/11/nmv-crop-report-oct-2017.pdf [Retrieved 20 December 2017].

Mathur AK (2010) Biotechnological studies in Cymbopogons. In: Akhila A (ed) *Essential oil-bearing grasses: the genus Cymbopogon*. CRC Press, Taylor & Francis Group, LLC, Boca Raton, pp 107–133.

Mathur AK, Ahuja PS, Pandey B, Kukreja AK and Mathur A (1988) Development of superior strains of an aromatic grass—*Cymbopogon winterianus* Jowitt—through somaclonal variation. In*: International conference on research in plant science and its relevance to future*, Abstract, University Press, Delhi, p 160.

Matsunaga W, Ohama N, Tanabe N, Masuta Y, Masuda S, Mitani N, Yamaguchi-Shinozaki K, Ma JF, Kato A and Ito H (2015) A small RNA mediated regulation of a stress-activated retrotransposon and the tissue specific transposition during the reproductive period in *Arabidopsis*. *Front. Plant Sci.* **6**: 48. https://doi.org/10.3389/fpls.2015.00048.

Matzk F, Meister A and Schubert I (2000) An efficient screen for reproductive pathways using mature seeds of monocot and dicots. *The Plant Journal* **21**: 97–108.

May P, Liao W, Wu Y, Shuai B, Richard McCombie W, Zhang MQ and Liu QA (2013) The effects of carbon dioxide and temperature on microRNA expression in *Arabidopsis* development. *Nat. Commun.* **4**: 2145. https://doi.org/10.1038/ncomms3145.

Mc Coy TJ and Phillips RL (1982) Chromosome stability in maize (*Zea mays* L.) tissue culture and sectoring in some regenerated plants. *Canadian Journal of Genetics and Cytology* **24**: 559-565.

Medford JI, Horgon R, El-Sawi Z, Klee HJ (1989) Alterations of endogenous cytokinins in transgenic plants using chimeric isopentenyl transferase gene. *Plant Cell* **1:** 403-413.

Megha S, Basu U and Kav NNV (2018) Regulation of low temperature stress in plants by microRNAs. *Plant Cell Environ.* **41**: 1–15. https://doi.org/10.1111/pce.12956.

Mehta YR and Angra DC (2000) Somaclonal variation for disease resistance in wheat and production of dihaploids through wheat maize hybrids. *Genetics and Molecular Biology* **23**: 617–622.

Mestdagh P, Vlierberghe PV, Weer AD, Muth D, Westermann F, Speleman F and Vandesompele J (2009) A novel and universal method for microRNA RT-qPCR data normalization. *Genome Biology* **10**: R64. doi: 10.1186/gb-2009-10-6-r64.

Meyer SU, Pfaffl MW and Ulbrich SE (2010) Normalization strategies for microRNA profiling experiments: a 'normal' way to a hidden layer of complexity? *Biotechnology Letters* **32**: 1777-1788. doi: 10.1007/s10529-010-0390-z.

Meyer W, Mitchell TG, Freedman EZ and Vilgays R (1993) Hybridization probes for conventional DNA fingerprinting used as single primers in the polymerase chain reaction to distinguish strains of *Cryptococcus neoformans*. *Journal of Clinical Microbiology* **31**: 2274–2280.

Michelmore RW, Paran I and Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *PNAS* **88(21)**: 9828-9832.

Miguel C and Marum L (2011) An epigenetic view of plant cells cultured *in vitro*: somaclonal variation and beyond. *Journal of Experimental Botany* **62(11)**: 3713-3725. doi: 10.1093/jxb/err155.

Miler N and Zalewska M (2014) Somaclonal variation of chrysanthemum propagated *in vitro* from different explants types. *Acta Sci Pol Hortorum Cultus* **13(2)**: 69–82.

Millar AA and Waterhouse PM (2005) Plant and animal microRNAs: similarities and differences. *Functional & Integrative Genomics* **5**: 129-135.

Minerva G and Kumar S (2013) Micropropagation of gerbera (*Gerbera jamesonii* Bolus). In: Lambardi M, Ozudogru EA, Jain SM (eds) Protocols for micropropagation of selected economically-important horticultural plants. *Methods Mol Biol* **994**: 305–316.

Minoo D (2002) Seedling and somaclonal variation and their characterization in Vanilla. *PhD thesis*. Calicut University, Kerala, India.

Minoo D, Babu KN, Ravindran PN and Peter KV (2006) Interspecific hybridization in vanilla and molecular characterization of hybrids and selfed progenies using RAPD and AFLP markers. *Scientia Horticulturae* **108**: 414 – 422.

Minoo D, Sajina A, Nirmal Babu K and Ravindran PN (1997) Ovule culture of vanilla and its potential in crop improvement. In: Edison S, Ramana KV, Sasikumar B, Nirmal Babu K and Santhosh JE (Eds.). *Biotechnology of Spices,*

Medicinal and Aromatic Plants (pp. 112-118). Indian Society for Spices, Calicut, India.

Mocellin S, Rossi CR, Pilati P, Nitti D and Marincola FM (2003) Quantitative real-time PCR: a powerful ally in cancer research**.** *TRENDS in Molecular Medicine* **9(5)**: 189-195.

Modgil M, Guleria N, Ghani M and Sharma JN (2012) Identifying somaclonal variants of the apple rootstock Malling 7 resistant to white root rot. *Sci Hortic* **137**: 148–155.

Mohanty S, Panda M, Subudhi E and Nayak S (2008) Plant regeneration from callus culture of *Curcuma aromatica* and *in vitro* detection of somaclonal variation through cytophotometric analysis. *Biologia Plantarum* **52**: 783–786.

Molina MdC and Garcia MD (1992) Phenotypic and chromosomic aberrations in maize plants regenerated from callus. *Maize Genetics Cooperation Newsletter* **66**: 64-65.

Molina MdC and Garcia MD (1998) Analysis of genetic variability in long-term callus cultures and regenerated plants of maize. *Cytologia* **63**: 183-190.

Momose M, Abe Y and Ozeki Y (2010) Miniature inverted-repeat transposable elements of *Stowaway* are active in potato. *Genetics* **186**: 59–66.

Mondal TK, Bantawa P, Sarkar B, Ghosh P and Chand PK (2013) Cellular differentiation, regeneration, and secondary metabolite production in medicinal *Picrorhiza* spp. *Plant Cell Tissue Organ Cult* **112(2)**: 143–158.

Monteuuis O, Doulbeau S and Verdeil JL (2008) DNA methylation in different origin clonal offspring from a mature *Sequoiadendron giganteum* genotype. *Trees-Structure and Function* **22**: 779–784.

Morel GM (1965) Clonal propagation of orchids by meristem culture. *Cymb. Soc. News* **20**: 3-11.

Moreno S, Martin JP and Ortiz JM (1998) Inter-simple sequence repeats PCR for characterization of closely related grapevine germplasm. *Euphytica* **101**: 117–125.

Morrison RA, Loh WHT, Green SK, Griggs TD and McLean BT (1989) Tissue culture of tomato and pepper: new tools for plant breeding. Tomato and pepper production in the tropics. In: *Proceedings of international symposium on integrated management practices,* Tainen, Taiwan. Shanhua, Taiwan, AVRDC, pp 44–50.

Mull L, Ebbs ML and Bender (2006) A histone methylation-dependent DNA methylation pathway is uniquely impaired by deficiency in *Arabidopsis* S-adenosylhomocysteine hydrolase. *Genetics* **174**: 1161-1171.

Müller B, Sheen J (2008) Cytokinin and auxin interplay in root stem-cell specification during early embryogenesis. *Nature* **453(7198)**: 1094-1097.

Mullis K, Faloona F, Scharf S, Saiki R, Horn G and Erlich H (1986) Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology* **LI**: 263-273.

Munshi A, Ahuja YR and Bahadur B (2015) Epigenetic Mechanisms in Plants: An Overview. In: Bahadur B *et al.* (eds.) *Plant Biology and Biotechnology:*

Volume II: Plant Genomics and Biotechnology, pp. 265-275. doi: 10.1007/978-81-322-2283-5_12.

Murashige T and Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiologia Plantarum* **15**: 473

Nacheva LR, Gercheva PS, Andonova MY, Panayotova DV and Dzhuvinov VT (2014) Somaclonal variation: a useful tool to improve disease resistance of pear rootstock 'Old Home 9 Farmingdale' (OHF 333) (*Pyrus communis* L.). *Acta Hortic* **1056**: 253–258.

Nair RR and Ravindran PN (1994) Somatic association of chromosomes and other mitotic abnormalities in *Vanilla planifolia* Andrews. *Caryologia* **47**: 65 – 73.

Naqvi AR, Sarwat M, Hasan S and Roychodhury N (2012) Biogenesis, functions and fate of plant microRAs. *Journal of Cellular Physiology* **227**: 3163-3168.

Nassar AM, Kubow S, Leclerc YN and Donnelly DJ (2014) Somatic mining for phytonutrient improvement of 'Russet Burbank' potato. *Am J Potato Res* **91**: 89–100.

Nayak S, Debata BK, Srivastava VK and Sangwan NS (2003) Evaluation of agronomically useful somaclonal variants in *Jamrosa* (a hybrid Cymbopogon) and detection of genetic changes through RAPD. *Plant Sci* **164**: 1029–1035.

Neelakandan AK and Wang K (2012) Recent progress in the understanding of tissue culture-induced genome level changes in plants and potential applications. *Plant Cell Rep* **31(4)**: 597-620. doi: 10,1007/s00299-011-1202-z.

Neelannavar VS (2006) In vitro propagation studies on vanilla (Vanilla planifolia Andr.). *Master of Science Thesis*. University of Agricultural Sciences Dharwad.

Negishi O, Sugiura K, Negishi Y (2009) Biosynthesis of Vanillin via Ferulic Acid in *Vanilla planifolia*. *Journal of Agricultural and Food Chemistry* **57**: 8856-9961.

Nehra NS, Kartha KK, Stushnott C and Giles KL (1992) The influence of plant growth regulator concentrations and callus age on somaclonal variation in callus culture regenerants of strawberry. *Plant Cell, Tissue and Organ Culture (PCTOC)* **29**: 257–268.

Nei M and Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. U.S.A.* **76**: 5269- 5273.

Ngezahayo F, Dong YS and Liu B (2007) Somaclonal variation at the nucleotide sequence level in rice (*Oryza sativa* L.) as revealed by RAPD and ISSR markers and by pairwise sequence analysis. *Journal of Applied Genetics* **48(4)**: 329-336 doi: 10.1007/BF03195229.

Nhut DT, Hai NT, Thu PTM, Thi NN, Hien TTD, Tuan TT, Nam NB, Huy NP, Chien HX and Jain SM (2013) Protocol for inducing flower color somaclonal variation in torenia (*Torenia fournieri* Lind.). *Methods Mol Biol* **11013**: 455–462.

Nicot N, Hausman JF, Hoffmann L and Evers D (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of Experimental Botany* **56:** 2907–2914.

Nitsch JP and Nitsch C (1969) Haploid plants from pollen grains. *Science* **163**: 85-87.

Nogoy FM, Niño MC, Song JY, Jung YJ, Kang KK, Nou I and Cho YG (2018) Plant microRNAs in molecular breeding. *Plant Biotechnol. Rep.* **12**: 15–25. https://doi.org/10.1007/s11816-018-0468-9.

Noirot M, Barre P, Duperray C, Louarn J and Hamon S (2003) Effects of caffeine and chlorogenic acid on propidium iodide accessibility to DNA: consequences on genome size evaluation in coffee tree. *Annals of Botany* **92**: 259–264.

Noirot M, Barre P, Louarn J, Duperray C and Hamon S. (2000) Nucleus–cytosol interactions—A source of stoichiometric error in flow cytometric estimation of nuclear DNA content in plants. *Annals of Botany* **86**: 309–316.

Novak FJ (1981) Chromosomal characteristics of long-term callus cultures of *Allium sativum* L. *Cytologia* **46**: 371-379.

Nussbaumer T, Martis MM, Roessner SK, Pfeifer M, Bader KC, Sharma S, Gundlach H and Spannagl M (2013) MIPS PlantsDB: a database framework for comparative plant genome research. *Nucleic Acids Research* **41**(Database issue): D1144-51. doi: 10.1093/nar/gks1153.

Ochatt SJ, Mousset-DeÂclas C and Rancillac M (2000) Fertile pea plants regenerate from protoplasts when calluses have not undergone endoreduplication. *Plant Science* **156**: 177-118. PMID: 10936524.

Odoux E and Brillouet JM (2009) Anatomy, histochemistry and biochemistry of glucovanillin, oleoresin and mucilage accumulation sites in green mature vanilla pod (*Vanilla planifolia*; Orchidaceae): a comprehensive and critical reexamination. *FRUITS* **64(4)**: 221-241.

Odoux E and Grisoni M (2011) Vanilla. In: Odoux M (Ed.) *Medicinal and aromatic plants*, pp. 387. Boca Raton, USA: CRC Press.

Ojemann LM, Nelson WL, Shin DS, Rowe AO and Buchanan RA (2006) Tian ma, an ancient Chinese herb, offers new options for the treatment of epilepsy and other conditions. *Epilepsy & Behavior* **8**: 376–383.

Ok CY, Singh RR and Salim AA (2016) Single-cell quantitative PCR: advances and potential in cancer diagnostics. In: R Luthra, RR Singh and KP Patel (Eds.) *Clinical Applications of PCR* (Third Edition), pp. 153-160. New York, USA: Humana Press. doi: 10.1007/978-1-4939-3360-0_14.

Okada Y, Yamagata K. Hong K, Wakayama T and Zhang Y (2010) A role of elongator in zygotic paternal genome demethylation. *Nature* **463(7280):** 554-558.

Ong-Abdullah M, Ordway JM, Jiang N *et al.* (2015) Loss of *Karma* transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature* **525(7570)**: 533-537. doi: 10.1038/nature15365.

Oono K (1985) Putative homozygous mutations in regenerated plants of rice. *Molecular Genetics and Genomics* **198**: 377–384.
Opabode JT and Adebooye OC (2005) Application of biotechnology for the improvement of Nigerian indigenous leaf vegetables. *Afr J Biotechnol* **4(3)**: 138–142.
Otto F (1992) Preparation and staining of cells for high-resolution DNA analysis. In: Radbruch A, ed. *Flow cytometry and cell sorting*. Berlin: Springer-Verlag, 101-104.
Ozeki Y, Davies E and Takeda J (1997) Somatic variation during long term subculturing of plant cells caused by insertion of a transposable element in a phenylalanine ammonia-lyase (PAL) gene. *Mol Gen Genet* **254**: 407–416.
Palama TL, Menard P, Fock I, Choi YH, Bourdon E, Govinden-Soulange J, Bahut M, Payet B, Verpoorte R, Kodja H (2010) Shoot differentiation from protocorm callus cultures of *Vanilla planifolia* (Orchidaceae): proteomic and metabolic responses at early stage. *BMC Plant Biology* **10**: 82.
Palomino G, Doležel J, Mendez I and Rublou A (2003) Nuclear genome size analysis of *Agave tequilana* Weber. *Caryologia* **56**: 37–46.
Pansarin ER and Miranda MR (2016) A new species of *Vanilla* (Orchidaceae: Vanillaloideae) from Brazil. *Pytotaxa* **267(1)**: 84-88.
Pantaleo V *et al*. (2016) Novel functional microRNAs from virus-free and infected *Vitis vinifera* plants under water stress. *Scientific Reports* **6**, 20167.
Paran I and Michelmore RW (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theoretical and Applied Genetics* **85(8)**: 985-993.
Park SY Murthy HN, Chakrabarthy D and Paek KY (2008) Detection of epigenetic variation in tissue-culture-derived plants of *Doritaenopsis* by methylation-sensitive amplification polymorphism (MSAP) analysis. *In Vitro Cellular Developmental Biology - Plant* **45**: 104–108. doi: 10.1007/s11627-008-9166-6.
Parra-Quezada RA, Gonzalez-Rosas H, Engleman EM (1990). *In vitro* culture of Vanilla (*Vanilla planifolia* Andrews) ovules [Spanish]. *Agrociencia***.**
Paszwlaoski J and Grossniklaus U (2011) Selected aspects of transgenerational epigenetic inheritance and resetting in plants. *Current Opinion in Plant Biology* **14**: 1–9.
Patnaik J, Sahoo S and Debata BK (1999) Somaclonal variation in cell suspension culture-derived regenerants of *Cymbopogon martini* (Roxb.) Wats var. Motia. *Plant Breed* **118**: 351–354.
Paul S, Datta SK and Datta K (2015) miRNA regulation of nutrient homeostasis in plants. *Frontiers in Plant Science* **6**: 232.
Paun O, Bateman RM, Fay MF, Hedren M, Civeyrel L and Chase MW (2010) Stable epigenetic effects impact adaptation in allopolyploid orchids (*Dactylorhiza*: *Orchidaceae*). *Molecular Biology and Evolution **27*****(11):*** 2465–2473.

Peakall R and Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* **6**: 288–295.

Pepke S, Wold B and Mortazavi A (2009) Computation for ChIP-seq and RNA-seq studies. *Nature Methods* **6**: S22-32.

Peraza-Echeverria S, Herrera-Valencia VA and James-Kay A (2001) Detection of DNA methylation changes in micropropagated banana plants using methylation-sensitive amplifcation polymorphism (MSAP). *Plant Science* **161**: 359-367.

Pereira MA, Malta FSV, Freire MCM and Couto PGP (2017) Application of next-generation sequencing in the era of precision medicine. In: *Applications of RNA-seq and omics strategies - from microorganims to human health*.

Peres LEP and Kerbauy G (1999) High cytokinin accumulation following root tip excision changes the endogenous auxin-to-cytokinin ratio during root-to-shoot conversion in *Catasetum fimbriatum* Lindl (*Orchidaceae*). *Plant Cell Reports* **18:** 1002-1006.

Peretti AL, Antunes JS, Lovison K, Kunz RI, Castor LRG, Brancalhao RMC, Bertolini GRF and Ribeiro LFC (2017) Action of vanillin (*Vanilla planifolia*) on the morphology of tibialis anterior and soleus muscles after nerve injury. *Einstein* **15(2)**: 186-191.

Pérez G, Yanez E, Isidrón M and Lorenzo JC (2009) Phenotypic and AFLP characterization of two new pineapple somaclones derived from *in vitro* culture. *Plant Cell Tissue Organ Cult* **96**: 113–116.

Pérez G, Yanez E, Mbogholi A, Valle B, Sagarra F, Yabor L, Aragón C, González J, Isidrón M and Lorenzo JC (2012) New pineapple somaclonal variants: P3R5 and dwarf. *Am J Plant Sci* **3**: 1–11.

Peschke VM and Phillips RL (1991) Activation of the maize transposable element *Suppressor-mutator* (*Spm)* in tissue culture. *Theor Appl Genet* **81**: 90–97.

Peschke VM and Phillips RL (1992) Genetic implications of somaclonal variation in plants. *Advances in Genetics* **30**: 41–75.

Peschke VM, Phillips RL and Gengenbach BG (1987) Discovery of transposable element activity among progeny of tissue culture ederived maize plants. *Science* **238**: 804–807.

Pfaffl MW (2001) A new mathematical model for relative quantification of real time RT-PCR. *Nucleic Acids Research* **29(9)**: 2003-2007.

Pfaffl MW (2010) The ongoing evolution of qPCR. *Methods* **50**: 215-216.

Pfaffl MW and Hageleit M (2001) Validities of mRNA quantification using recombinant RNA and recombinant DNA external calibration curves in real-time RT-PCR. *Biotechnology Letters* **23**: 275-282.

Pfaffl MW, Tichopad A, Prgomet C and Neuvians TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. *Biotechnology Letters* **26**: 509-515.

Pfaffl MW, Vandesompele J and Kubista M (2009) Data Analysis Software. In: J Logan, K Edwards and N Saunders (Eds.) *Real-Time PCR: Current Technology and Applications*, pp. 65-83. Poole, UK: Caister Academic Press. ISBN: 978-1-904455-39-4.

Pfosser M, Amon A, Lelley T and Herberle-Bors E (1995) Evaluation of sensitivity of flow cytometry in detecting aneuploidy in wheat using disomic and ditelosomic wheat-rye addition lines. *Cytometry* **21**: 387-393.

Philip VJ and Nainar AZ (1986) Clonal propagation of *Vanilla planifolia* (Salisb.) Ames using tissue culture. *Journal of Plant Physiology* **122**: 211–215.

Phua QY, Chin CK, Asri ZRM, Lam DYA, Subramaniam S, Chew BL (2016) The callugenic effects of 2,4-dichlorophenoxyl acetic acid (2,4-D) on leaf explants of sabah snake grass (Clinacanthus nutans). *Pakistan Journal of Botany* **48(2)**: 561-566.

Picault N, Chaparro C, Piegu B, Stenger W, Formey D, Llauro C, Descombin J, Sabot F, Lasserre E, Meynard D, Guiderdoni E and Panaud O (2009) Identification of an active LTR retrotransposon in rice. *Plant J* **58**: 754–765.

Pillay M, Ogundiwin E, Tenkouano A and Doležel J (2006) Ploidy and genome composition of *Musa* germoplasm at the International Institute of Tropical Agriculture (IITA). *African Journal of Biotechnology* **5**: 1224–1232.

Pinto G, Loureiro J, Lopes T and Santos C. 2004. Analysis of the genetic stability of *Eucalyptus globulus* Labill. somatic embryos by flow cytometry. *Theoretical and Applied Genetics* **109**: 580–587.

Pirttilä A, Joensuu P, Pospiech H, Jalonen J, Hohtola A (2004) Bud endophytes of Scots pine produce adenine derivatives and other compounds that affect morphology and mitigate browning of callus cultures. *Physiologia Plantarum* **121**: 304-312.

Planckaert F and Walbot V (1989) Molecular and genetic characterization of *Mu* transposable elements in *Zea mays*: behavior in callus culture and regenerated plants. *Genetics* **123**: 567–578.

Podwyszynska M, Kuras A and Korbin M (2010) Somaclonal variation in micropropagated tulips as a source of novel genotypes—field and molecular characteristic. *Acta Hortic* **855**: 225–232.

Pombo MA, Zheng Y, Fei Z Martin G and Rosli HG (2017) Use of RNA-seq data to identify and validate RT-Qpcr reference genes for studying the tomato-*Pseudomonas* pathosystem. *Scientific Reports* **7**: 44905.

Pontaroli AC and Camadro EL (2005) Somaclonal variation in *Asparagus officinalis* plants regenerated by organogenesis from long-term callus cultures. *Genetics and Molecular Biology* **28(3)**: 423-430.

Popescu AN, Isac VS and Coman MSR (1997) Somaclonal variation in plants regenerated by organogenesis from callus cultures of strawberry (*Fragaria x ananassa*). *Acta Horticulturae* **439**: 89–96.

Porras-Alfaro A and Bayman P (2007) Mycorrhizal fungi of *Vanilla*: diversity, specificity and effects on seed germination and plant growth. *Mycologia* **99(4)**: 510-525.

Pouteau S, Huttner E, Grandbastien MA and Caboche M (1991) Specific expression of the tobacco *Tnt1* retrotransposon in protoplasts. *EMBO J* **10**: 1911–1918.

Prado MJ, Rodriguez E, Rey L, GonzaÂlez MV, Santos C and 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**:49-59.

Price HJ and Johnston JS (1996) Influence of light on DNA content of *Helianthus annuus* Linnaeus. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* **93**: 11264-11267.

Pring DR, Conde MF and Gengenbach BG (1981) Cytoplasmic genome variability in tissue culture-derived plants. *Environmental and Experimental Botany* **21**: 369-377.

Puite KJ and Ten Broeke WRR (1983) DNA staining of fixed and non-fixed plant protoplasts for flow cytometry with Hoechst 33342. *Plant Science Letters* **32**: 79–88.

Puolimatka M and Karp A (1993) Effect of genotype on chromosome variation in tissue culture of inbred and outbred rye. *Heredity* **71**: 138-144.

Purseglove JW (1972) In: *Tropical Crops: Monocotyledons*, Longman, London.

Purseglove JW, Brown EG, Green CL and Robbins SRJ (1988) In: *Spices*, Vol. 2, Scientific and Technical, Longman.

QIAGEN (2011) *miScript® PCR system handbook* **10/2011**. Available at: https://www.qiagen.com/at/resources/download.aspx?id=7954ef25-3a39-4b0a-a27e-42689dbb4f5f&lang=en (Assessed: 21st June 2017).

QIAGEN (2014) *miScript® Plant qPCR system handbook* **06/2014**. Available at: https://www.qiagen.com/no/resources/resourcedetail?id=9e04163e-a9f3-45b3-a765-380606eee504&lang=en (Assessed: 21st June 2017).

Quambusch M, Pirttilä AM, Tejesvi MV, Winkelmann T, Bartsch M (2014) Endophytic bacteria in plant tissue culture: differences between easy- and difficult-to-propagate *Prunus avium* genotypes. *Tree Physiology* **34(1)**: 524-533.

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R and Lopez R (2005) InterProScan: protein domains identifier. *Nucleic Acids Research* **33** (Web Server issue): W116–W120. doi:10.1093/nar/gki442.

Qureshi JA, Huel P and Kartha KK (1992) Is somaclonal variation a reliable tool for spring wheat improvement? *Euphytica* **60**: 221-228.

Rabiei K, Khodambashi N and Omidi M (2011) Use of somaclonal variation on improvement of drought resistant lines of carrot (*Daucus carota* L.). *J Hortic Sci (Agric Sci Technol)* **25**:156–169.

Radonic A, Thulke S, Mackay IM, Landt O, Siegert W and Nitsche A (2004) Guideline to reference gene selection for quantitative real-time PCR. *Biochemical and Biophysical Research Communications* **313:** 856–862.

Ragupathy R, Ravichandran S, Mahdi MSR, Huang D, Reimer E, Domaratzki M and Cloutier S (2016) Deep sequencing of wheat sRNA transcriptome reveals

distinct temporal expression pattern of miRNAs in response to heat, light and UV. *Sci. Rep.* **6**: 39373. https://doi.org/10.1038/srep39373.

Rai K, Jafri IF, Chidester S, James SR, Karpf AR, Cairns BR and Jones DA (2010) Dnmt3 and G9a cooperate for tissue-specific development in zebrafish. *The Journal of Biological Chemistry* **285(6)**: 4110–4121.

Raimondi JP, Camadro EL and Masuelli RW (2001) Assessment of somaclonal variation in asparagus by RAPD fingerprinting and cytogenetic analyses. *Sci Hortic* **90**: 19-29.

Rain P (2004) *Vanilla: The Cultural History of the World's Favorite Flavor and Fragrance*. New York, Penguin Group.

Rakocevic A, Mondy S, Tirichine L, Cosson V, Brocard L, Iantcheva A, Cayrel A, Devier B, Abu El-Heba GA and Ratet P (2009) MERE1, a low-copy-number copia-type retroelement in *Medicago truncatula* active during tissue culture. *Plant Physiol* **151**: 1250–1263.

Ramakers C, Ruijter JM, Deprez RHL and Moorman AFM (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters* **339**: 62-66.

Ramakrishnan M, Caesar SA, Duraipandiyan V and Ignacimuthu S (2014) Efficient plant regeneration from shoot apex explants of maize (*Zea mays*) and analysis of genetic fidelity of regenerated plants by ISSR markers. *Plant Cell, Tissue and Organ Culture* **119(1)**: 183-196. doi: 10.1007/s11240-014-0525-1.

Rambaut A (2008) FigTree v1.1.1: Tree figure drawing tool. http://tree.bio.ed.ac.uk/software/figtree/

Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP and Jaenisch R (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* **97(10)**: 5237–5242.

Ramulu SK and Dijkhuis P (1986) Flow cytometric analysis of polysomaty and in vitro genetic instability in potato. *Plant Cell Reports* **5**: 234–237.

Rana S, Dhar N, Bhat WW, Razdan S, Khan S, Dhar RS, Dutt P and Lattoo SK (2012) A 12-deoxywithastramonolide-rich somaclonal variant in *Withania somnifera* (L.) Dunal—molecular cytogenetic analysis and significance as a chemotypic resource. *In Vitro Cell Dev Biol Plant* **48(5)**: 546–554.

Ranadive AS (1994) Vanilla-cultivation, curing, chemistry, technology and commercial products. In: Charalambous G (Ed.) *Spices, Herbs and Edible Fungi*, pp. 517–577. Amsterdam, Netherland: Elsevier Science Publisher B.V..

Rani V and Raina S (2000) Genetic fidelity of organized meristem derived micropropagated plants: a critical reappraisal. *In Vitro Cellular & Developmental Biology - Plant* **36**: 319–330.

Rao SR and Ravinshankar GA (2000) Vanilla flavour: production by conventional and biotechnological routes. *Journal of the Science of Food and Agriculture* **80(3)**: 289-304.

Rao XL, Krom N, Tang YH, Widiez T, Havkin-Frenkel D, Belanger FC, Dixon RA and Chen R (2014) A deep transcriptomic analysis of pod development in

the vanilla orchid (*Vanilla planifolia*). *BMC Genomics* **15**: 964. doi:10.1186/1471-2164-15-964.

Rasmussen HN, Dixon KW, Jersáková, J and Těšitelová T (2015) Germination and seedling establishment in orchids: a complex of requirements. *Annals of Botany* **116(3)**: 391-402.

Ratnaparkhe MB, Tekeoglu M and Muehlbauer FJ (1998) Intersimple-sequence-repeat (ISSR) polymorphisms are useful for finding markers associated with disease resistance gene clusters. *Theoretical and Applied Genetics* **97**: 515–519

Ravindra NS, Ramesh SI, Gupta MK, Jhang T, Shukla AK, Darokar MP and Kulkarni RN (2012) Evaluation of somaclonal variation for genetic improvement of patchouli (*Pogostemon patchouli*), an exclusively vegetatively propagated aromatic plant. *J Crop Sci Biotechnol* **15**: 33–39.

Ray T, Dutta I, Saha P, Das S and Roy SC (2006) Genetic stability of three economically important micropropagated banana (*Musa* spp.) cultivars of lower Indo-Gangetic plains, as assessed by RAPD and ISSR markers. *Plant Cell, Tissue and Organ Culture (PCTOC)* **85**: 11–21.

Reddy MP, Sarla N and Siddiq EA (2002) Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. *Euphytica* **128(1)**: 9-17.

Renfree MB, Suzuki S and Kaneko-Ishino T (2013) The origin and evolution of genomic imprinting and viviparity in mammals. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **368(1609)**: 20120151.

Resende GA, Mangolin CA and Machado MDF (2010) Somaclonal variation in *Cereus peruvianus* Mill. (*Cactaceae*): its potential to generate new varieties and broaden the species' genetic basis. *J Basic Appl Genet* **21**: 33–42.

Retheesh, ST, Bhat AI (2011) Genetic transformation and regeneration of transgenic plants from protocorm-like bodies of vanilla (*Vanilla planifolia* Andrews) using *Agrobacterium tumefaciens*. *Journal of Plant Biochemistry and Biotechnology* **20(2)**: 262-269.

Reuveni O and Israeli Y (1990) Measures to reduce somaclonal variation in *in vitro* propagated bananas. *Acta Horticulturae* **275**: 307–313.

Reyes JL and Chua NH (2007) ABA induction of miR159 controls transcript levels of two MYB factors during *Arabidopsis* seed germination. *Plant Journal* **49**: 592–606.

Reyna-Lopez GE, Simpson J and Ruiz-Herrera J (1997) Differences in DNA methylation patterns are detectable during the dimorphic transition of fungi by amplification of restriction polymorphisms. *Molecular Genetics and Genomics* **253(6)**: 703–710. doi: 10.1007/s004380050374.

Rhee Y, Lin H, Buell R, Childs K and Kaeppler S (2009) *Ac2* allele of maize identified in regenerant-derived progeny from tissue culture results from insertion of a novel transposon. *Maydica* **54**: 429–437.

Rhodes CA, Phillips RL and Green CE (1986) Cytogenetic stability of aneuploid maize tissue culture. *Canadian Journal of Genetics and Cytology* **28**: 374-384.

Ricci U, Marchi C, Previderè C and Fattorini P (2006) Quantification of human DNA by real-time PCR in forensic casework. *International Congress Series* **1288**: 750-752.

Rodolphe G, Séverine B, Michel G and Pascale B (2011) Biodiversity and evolution in the *Vanilla* Genus. In: Grillo O (Ed.) *The dynamical processes of biodiversity - Case studies of evolution and spatial distribution*. London, UK: IntechOpen.

Rodrigues PHV, Tulmann Neto A, Cassieri Neto P and Mendes BMJ (1998) Influence of the number of subcultures on somaclonal variation in micropropagated Nanico (*Musa* spp., AAA group). *Acta Horticulturae* **490**: 469–473.

Rogers K and Chen XM (2013) Biogenesis, turnover, and mode of action of plant microRNAs. *The Plant Cell* **25**: 2383-2399.

Roldan-Ruiz I, Dendauw J, Vanbockstaele E, Depicker A, De Loose M (2000) AFLP markers reveal high polymorphic rates in ryegrasses (Lolium spp.). Mol Breed 6:125–134.

Rosati P, Mezzetti B, Ancherajni M, Foscolo S, Predieri S, and Fasolo F (1990) *In vitro* selection of apple rootstock somaclones with *Phytophthora cactorum* culture filtrate. *Acta Hortic* **280**: 409–416.

Roux N, Doležel J, Swennen R and Zapata-Arias FJ (2001) Effectiveness of three micropropagation techniques to dissociate cytochimeras in *Musa* spp. *Plant Cell, Tissue and Organ Culture (PCTOC)* **66**: 189–197.

Roux N, Toloza A, Radecki Z, Zapata-Arias FJ and Doležel J (2003) Rapid detection of aneuploidy in *Musa* using flow cytometry. *Plant Cell Reports* **21**: 483–490.

Roy AR, Patel RS, Patel VV, Sajeev S and Deka BC (2011) Asymbiotic seed germination, mass propagation and seedling development of *Vanda coerulea* Griff ex. Lindl. (Blue Vanda): An *in vitro* protocol for an endangered orchid. *Scientia Horticulturae* **128**: 325–331.

Roy J and Banerjee N (2003) Induction of callus and plant regeneration from shoot-tip explants of *Dendrobium fimbriatum* Lindl. Var. oculatum Hk. F. *Scientia Horticulturae* **97**: 333-340.

Růžička K, Šimáškvá M, Duclercq J, Petrášek J, Zažímalová E, Simon S, Friml J, Montagu MCEV and Benková E (2009) Cytokinin regulates root meristem activity via modulation of the polar auxin transport. *Proceedings of the National Academy of Sciences* (*PNAS)* **106 (11)**: 4284-4289.

Ságі L, Remy S and Swennen R (1998) Genetic transformation for the improvement of banana - a critical assessment. In: *INIBAP annual report*. INIBAP Montpellier, France, pp. 33-36.

Sahijram L (2015) Somaclonal variation in micropropagated plants. In: Bahadur B *et al.* (eds.) *Plant Biology and Biotechnology: Volume II: Plant Genomics and Biotechnology*, pp.407-416. doi: 10.1007/978-81-322-2283-5_12.

Sahijram L, Soneji J and Bollamma K (2003) Analyzing somaclonal variation in micropropagated bananas (*Musa* spp.). *In Vitro Cellular & Developmental Biology - Plant* **39**: 551–556.

Sahni S, Prasad BD and Kumar P (2017*) Plant Biotechnology, Volume 2: Transgenic, Stress Management, and Biosafety Issues*. CRC Press. ISBN: 9781351819473.

Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA and Arnheim N (1985) Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**: 1350-1354.

Sakhanokho HF, Witcher AL, Pounders CT and Spiers JM (2012) 'Ramata': a new dwarf variegated *Hedychium* (ornamental ginger) cultivar. *HortSci* **47**: 803–805.

Sambrook JF and Russell DW (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sanchez MdlP, Caro E, Desvoyes B, Ramirez-Parra E and Gutierrez C (2008) Chromatin dynamics during the plant cell cycle. *Seminars in Cell and Developmental Biology* **19**: 537–546.

Sankar AA and Moore GA (2001) Evaluation of inter-simple sequence repeat analysis for mapping in *Citrus* and extension of genetic linkage map. *Theoretical and Applied Genetics* **102**: 206–214.

Santner A and Estelle M (2011) The ubiquitin-proteasome system regulates plant hormone signaling. *Plant Journal* **61(6)**: 1029-1040.

Sarwat M, Nabi G, Das S and Srivastava PS (2011) Molecular Markers in Medicinal Plant Biotechnology: Past and Present. In: *Critical Rev Biotech*. doi:10.3109/07388551.2011.551872.

Sathyanarayana BN and Varghese DB (2007). *Plant tissue culture practices and new experimental protocols,* pp.38-39. I.K. International Publishing House Pvt. Ltd, New Delhi, India.

Sato M, Kawabe T, Hosokawa M, Tatsuzawa F and Doi M (2011) Tissue culture-induced flower-color changes in *Saintpaulia* caused by excision of the transposon inserted in the *flavonoid 30,50 hydroxylase* (*F3050H*) promoter. *Plant Cell Reports* **30**: 929–939.

Saxena G, Verma PC, Rahman LU, Banerjee S and Kumar S (2008) Field performance of somaclones of rose scented geranium (*Pelargonium graveolens* L' Her Ex Ait.) for evolution of their essential oil yield and composition. *Ind Crops Prod* **27**: 86–90.

Schaller GE, Bishopp A, Kieber JJ (2015) The Yin-Yang of Hormones: Cytokinin and Auxin Interactions in Plant Development. *The Plant Cell* **27**: 44-63.

Schenk RV and Hilderbrandt AC (1972) Medium and techniques for induction and growth of monocotyledonous and dicotyledonous plant cell cultures. *Can. J. Bot.* **50**: 199-204.

Schmittgen TD, Lee EJ, Jiang J, Sarkar A, Yang L, Elton TS and Chen C (2008) Real-time PCR quantification of precursor and mature microRNA. *Methods* **44(1)**: 31-38.

Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M and Ragg T (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology* **7(3)**: 1-14. doi: 10.1186/1471-2199-7-3.

Schwab W, Lange BM and Wüst M (2018) In: *Biotechnology of Natural Products*. Cham, Switzerland: Springer International Publishing. ISBN: 978-3-319-67903-7.

Schweizer L, Yerk-Davis DL, Phillips RL, Srienc F and Jones RJ (1995) Dynamics of maize endosperm development and DNA endoreduplication. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* **92**: 7070-7074.

Sedov KA, Fomenkov AA Solov'yova AI, Nosov AV and Dolgikh YI (2014) The level of genetic variability of cells in prolonged suspension culture of *Arabidopsis thaliana*. *Biology Bulletin* **41(6)**: 493-499.

Sengupta J, Jha S and Sen S (1988) Karyotype stability in long-term callus derived plants of *Crepis tectorum* L. *Biologia Plantarum* **30**: 247-251.

Setter TL and Flannigan BA (2001) Water deficit inhibits cell division and expression of transcripts involved in cell proliferatio and endoreduplication in maize endosperm. *Journal of Experimental Botany* **52**: 1401-1408.

Shakeel M, Rodriguez A, Tahir UB and Jin F (2018) Gene expression studies of reference genes for quantitative real-time PCR: an overview in insects. *Biotechnology Letters* **40(2)**: 227-236. doi: 10.1007/s10529-017-2465-4.

Sharp AJ, Stathaki E, Migliavacca E, Brahmachary M, Montgomery SB and Dupre Y, Antonarakis SE (2011) DNA methylation profiles of human active and inactive X chromosomes. *Genome Research* **21(10)**: 1592–1600.

Shchukin A, Ben-Bassat D and Israeli Y (1997) Plant regeneration *via* somatic embryogenesis in Grand Naine banana and its effect on somaclonal variation. *Acta Horticulturae* **447**: 317-318.

Shchukin A, Ben-Bassat D and Israeli Y (1998) Somaclonal variation and horticultural performance of 'Grand Naine' bananas multiplied *via* somatic embryogenesis or shoot-tip culture. In: *Plant Biotechnology and In Vitro Biology in the 21st Century*, pp. 14-19. Jerusalem, Middle East: International Association for Plant Tissue Culture.

Shearman JR, Jantasuriyarat C, Sangsrakru D, Yoocha T, Vannavichit A, Tragoonrung S and Tangphatsornruang S (2013) Transcriptome analysis of normal and mantled developing oil palm flower and fruit. *Genomics* **101(5)**: 306–312.

Shen X, Chen J, Kane M and Henny R (2007) Assessment of somaclonal variation in *Dieffenbachia* plants regenerated through indirect shoot organogenesis. *Plant Cell Tissue Organ Cult* **91**: 21–27.

Shields R, Robinson SJ, Anslow PA (1984) Use of fungicides in plant tissue culture. *Plant Cell Report* **3**: 33-36.

Shilatifard A (2006) Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annual Review of Biochemistry* **75**: 243–269.

Shu ZQ, Row S and Deng WM (2018) Endoreplication: The Good, the Bad, and the Ugly. *Trends in Cell Biology* **28(6)**:465-474. doi: 10.1016/j.tcb.2018.02.006.

Singh C and Roy-Chowdhuri S (2016) Quantitative real-time PCR: recent advances. In: R Luthra, RR Singh and KP Patel (Eds.) *Clinical Applications of PCR* (Third Edition), pp. 161-176. New York, USA: Humana Press. doi: 10.1007/978-1-4939-3360-0_14.

Skirvin RM and Janick J (1976) Tissue culture-induced variation in scented *Pelargonium* spp. *J Am Soc Hortic Sci* **101**: 281–290.

Skoog F and Miller CO (1957) Chemical regulation of growth and organ formation in plant tissues culture *in vitro*. *Symposia of the Society for Experimental Biology* **11**: 118-131.

Slater A, Scott NW, Fowler MR (2008) "Plant tissue culture*." Plant Biotechnology – the genetic manipulation of plants.* Ed. Second. New York: Oxford University Press, pp. 37-53. ISBN: 9780199282616

Sliwinska E and Lukaszewska E (2005) Polysomaty in growing *in vitro* sugar-beet (*Beta vulgaris* L.) seedlings of different ploidy level. *Plant Science* **168**: 1067-1074.

Smigocki AC (1991) Cytokinin content and tissue distribution in plants transformed by a reconstructed isopentenyl transferase gene. *Plant Molecular Biology* **16**: 105-115.

Smith RH (2000) *Plant tissue culture: techniques and experiments*. Houston, Texas, USA: Gulf Professional Publishing.

Solis MT, Rodriguez-Serrano M, Meijon M, Canal MJ, Cifuentes A, Risueno MC and Testillano PS (2012) DNA methylation dynamics and MET1a-like gene expression changes during stress-induced pollen reprogramming to embryogenesis. *Journal of Experimental Botany* **63(18):** 6431–6444.

Soto-Arenas M and Cribb P (2010) A new infrageneric classification and synopsis of the Genus *Vanilla* Plum. ex. Mill. (Orchidaceae: Vanillinae). *Lankesteriana* **9(3)**: 355-398.

Spannagi M, Noubibou O, Haase D, Yang L, Gundlach H, Hindemitt T, Klee K, Haberer G, Schoof H and Mayer KF (2007) MIPSPlantsDB - plant database resource for integrative and comparative plant genome research. Nucleic Acids Res. 35 (Database issue): D834-40. doi: 10.1093/nar/gkl945.

Spannagi M, Nussbaumer T, Bader KC, Martis MM, Seidel M, Kugler KG, Gundlach H and Mayer KF (2016) PGSB PlantsDB: updates to the database framework for comparative plant genome research. *Nucleic Acids Res.* **44(D1)**: D1141-7. doi: 10.1093/nar/gkv1130.

Sreedhar RV, Venkatachalam L and Bhagyalakshmi N (2007) Genetic fidelity of long-term micropropagated shoot cultures of vanilla (*Vanilla planifolia*

Andrews) as assessed by molecular markers. *Biotechnology Journal* **2**: 1007-1013. doi: 10.1002/biot.200600229.

Ståhlberg A and Bengtsson M (2010) Single-cell gene expression profiling using reverse transcription quantitative real-time PCR. *Methods* **50**: 282-288.

Stelly DM, Altman DW, Kohel RJ, Rangan TS and Commiskey E (1989) Cytogenetic abnormalities of cotton somaclones from callus cultures. *Genome* **32**: 762-770.

Stover RH (1987) Somaclonal variation in Grande Naine and Saba bananas in the nursery and in the field. In: GJ Persley, E De Langhe (Eds.) *ACIAR proceeding* **21**. Canberra, Australia.

Su CL, Chao YT, Chang YCA, Chen WC, Chen CY, Lee AY, Hwa KT and Shih MC (2011a) De Novo Assembly of Expressed Transcripts and Global Analysis of the *Phalaenopsis aphrodite* Transcriptome. Plant & Cell Physiology 52(9): 1501-1514.

Su CL, Chao YT, Yen SH, Chen CY and Shih MC (2017) Orchidstra 2.0 – A Transcriptomics Resource for the Orchid Family. *Plant Cell Physiol.* **pii**: pcw220. doi: 10.1093/pcp/pcw220.

Su CL, Chao YT, Yen SH, Chen CY, Chen CW, Chang YC and Shih MC (2013) Orchidstra: an integrated orchid functional genomics database. *Plant & cell physiology* **54(2)**: e11. doi: 10.1093/pcp/pct004.

Su Z, Han L and Zhao Z (2011b) Conservation and divergence of DNA methylation in eukaryotes: new insights from single base-resolution DNA methylomes. *Epigenetics* **6(2**): 134–140.

Suda J (2004) An employment of flow cytometry into plant biosystematics. PhD thesis, Charles University, Czech Republic. Available at http://www.ibot.cas.cz/fcm/suda/presentation/dissertation.pdf.

Suman S (2017) Plant tissue culture: a promising tool of quality material production with special reference to micropropagation of banana. *Biochemical and Cellular Archives* **17(1)**: 1-26.

Sun Y, Hou R, Fu XT, Sun CS, Wang S, Wang C, Li N, Zhang LL and Bao ZM (2014) Genome-wide analysis of DNA methylation in five tissues of Zhikong scallop, *Chlamys farrer*. *PLOS One* **9(1)**: e86232. doi: 10.1371/journal.pone.0086232.

Sunkar R (2010) MicroRNAs with macro-effects on plant stress responses. *Seminars in Cell & Developmental Biology* **21**: 805-811.

Sunkar R and Zhu JK (2004) Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell* **16**: 2001–2019.

Svejstrup JQ (2007) Elongator complex: how many roles does it play? *Curr Opin Cell Biol.* **19**: 331– 336.

Takahashi H, Matsumoto T and Takai T (1993) Somaclonal variants from strawberry cv. Morioka-16, M16-AR 1, 2 and 3 resistant to several isolates of *Alternaria alternata* strawberry pathotype occurring in Tohoku and Hokkaido [Japan]. *J Jpn Soc Hortic Sci* **61**: 821–826.

Tan BC, Chin CF, Alderson P (2011a) Optimisation of plantlet regeneration from leaf and nodal derived callus of *Vanilla planifolia* Andrews. *Plant Cell, Tissue and Organ Culture (PCTOC)* **105**: 457-463.
Tan BC, Chin CF, Alderson P (2011b) An improved plant regeneration of *Vanilla planifolia* Andrews. *Plant Tissue Culture & Biotechnology* **21(1)**: 27-33.
Tang CY, Liu CC and Hwang SC (2000) Improvement of the horticultural traits of Cavendish banana (*Musa* spp., AAA group I). Selection and evaluation of a semi-dwarf clone resistant to Fusarium wilt. *J Chin Soc Hortic Sci* **46**: 173–182.
Tanurdzic M, Vaughn MW, Jiang H, Lee T-J, Slotkin RK, Sosinski B, Thompson WF and Doerge RW, Martienssen RA (2008) Epigenomic consequences of immortalized plant cell suspension culture. *PLoS Biology* **6**: e302.
Tautz D and Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research* **12**: 4127–4138.
Thellin O, Zorzi W, Lakaye B, Borman BD, Coumans B, Hennen G, Grisar T, Igout A and Heinen E (1999) Housekeeping genes as internal standards: use and limits. *Journal of Biotechnology* **75**: 291-295.
Thieme R and Griess H (2005) Somaclonal variation in tuber traits of potato. *Potato Res* **48**: 153–165.
Thorn EC (1992) The influence of genotype and environment on seed and embryo development in barley (*Hordeum vulgare* L.) after crossing with *Hordeum bulbosum* L. *Euphytica* **59**: 109-118.
Tichopad A, Bar T, Pecen L, Kitchen RR, Kubista M and Pfaffl MW (2010) Quality control for quantitative PCR based on amplification compatibility test. *Methods* **50**: 308-312.
Tomlinson JA, Barker I and Boonham N (2007) Faster, simpler, more-specific methods for improved molecular detection of *Phytophthora ramorum* in the field. *Applied and Environmental Microbiology* **73(12)**: 4040-4047.
Torres KC (1989) Stages of Micropropagation. In: *Tissue Culture Techniques for Horticultural Crops*. Springer, Boston, MA, pp. 52-65.
Toyoda H, Horikoshi K, Yamano Y and Ouchi S (1991) Selection of Fusarium wilt disease resistance from regenerant derived from callus of strawberry. *Plant Cell Rep* **10**: 167–170.
Traas J, Hulskamp M, Gendreau E and Hofte H (1998) Endoreduplication and development: rule without dividing? *Current Opinion in Plant Biology* **1**: 498-503.
Trávníček P, Ponert J, Urfus T, Jersáková J, Vrána J, Hřibová E, Doležel J and Suda J (2015) Challenges of flow-cytometric estimation of nuclear genome size in orchids, a plant group with both whole-genome and progressively partial endoreplication. *Cytometry Part A* **87A**: 958-966.
Tripathi RK, Ram S (1982) Induction of growth and differentiation of carrot callus cultures by carbendazim and denximidazole. *Indian Journal of Experimental Biology* **20**: 674-677.

Trotter (2000) *WinMDI version 2.8. Available at http://facs.scripps.edu/software.html*
Tsumura Y, Ohba K and Strauss SH (1996) Diversity and inheritance of inter-simple sequence repeat polymorphisms in Douglasfir (*Pseudotsuga menziesii*) and sugi (*Cryptomeria japonica*). *Theoretical and Applied Genetics* **92**: 40–45.
Tuomi JM, Voorbraak F, Jones DL and Ruijter JM (2010) Bias in the C_q value observed with hydrolysis probe based quantitative PCR can be corrected with the estimated PCR efficiency value. *Methods* **50**: 313-322.
Ulrich I and Ulrich W (1991) High-resolution flow cytometry of nuclear DNA in higher plants. *Protoplasma* **165**: 212–215.
Ungaro A, Pech N, Martin JF, McCairns RJS, Me/vy JP, Chappaz R and Gilles A (2017) Challenges and advances for transcriptome assembly in non-model species. *PLOS One* **12(9)**: e0185020.
Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M and Rozen SG (2012) Primer3 - new capabilities and interfaces. *Nucleic Acids Research* **40 (15)**: e115.
Valledor L, Meijón M, Hasbún R, Jesús Cañal M and Rodríguez R (2010) Variations in DNA methylation, acetylated histone H4, and methylated histone H3 during *Pinus radiata* needle maturation in relation to the loss of *in vitro* organogenic capability. *Journal of Plant Physiology* **167**: 351–357.
van der Wurff AWG, Chan YL, van Straalen NM and Schouten J (2000) TE-AFLP: combining rapidity and robustness in DNA fingerprinting. *Nucleic Acids Research* **28(24)**: e105.
van Eijk MJ, Broekhof JL, van der Poel HJ, Hogers RC, Schneiders H, Kamerbeek J, Verstege E, van Aart JW, Geerlings H, Buntjer JB, van Oeveren AJ and Vos P (2004) SNPWave™: a flexible multiplexed SNP genotyping technology. *Nucleic Acids Research* **32(4)**: e47.
Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A and Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* **3(7)**: research0034.1-research0034.11.
Vandesompele J, Kubista M and Pfaffl MW (2009) Reference gene validation software for improved normalization. In: J Logan, K Edwards and N Saunders (Eds.) Real-Time PCR: Current Technology and Applications, pp. 47-64. Poole, UK: Caister Academic Press. ISBN: 978-1-904455-39-4.
Varshney RK, Chabane K, Hendre PS, Aggarwal RK and Graner A (2007) Comparative assessment of EST-SSR, EST-SNP and AFLP markers for evaluation of genetic diversity and conservation of genetic resources using wild, cultivated and elite barleys. *Plant Sci* **173**: 638–649.
Vázquez AM and Linacero R (2010) Stress and Somaclonal Variation. In: *Plant Developmental Biology - Biotechnological Perspectives* Vol. 2 (eds. Pua EC and Davey MR). Springer-Verlag, Berlin Heidelberg. doi: 10.1007/978-3-642-04670-4_3.

Veilleux RE and Johnson AAT (1998) Somaclonal variation: Molecular analysis, transformation interaction and utilization. In: J Janick (Ed.) *Plant Breeding Reviews* (Vol. **16**), pp. 229-256. New Jersey, USA: John Wiley and Sons Inc.

Veitia-Rodriguez N, Francisco-Cardoso J, Perez JN, Garcia-Rodriguez L, Bermudez-Caraballosos I, Garcia-Rodriguez L, Padron-Montesinos Y, Orellana-Perez P, Romero-Quintana C and Hernandez N (2002) Evaluations in field of somaclones of Irish potatoes (*Solanum tuberosum* Lin.) of the variety Desiree obtained by somaclonal variation and *in vitro* mutagenesis. *Biotechnol Veg* **2**: 21–26.

Velankar MH and Heble MR (2004) Biotransformation of Externally Added Vanillin Related Compounds by Multiple Shoot Cultures of *Vanilla planifolia* L. *Journal of Plant Biochemistry & Biotechnology* **13**: 153-156.

Venkatachalam L, Sreedhar RV and Bhagyalakshmi N (2007) Genetic analyses of micropropagated and regenerated plantlets of banana as assessed by RAPD and ISSR markers. *In Vitro Cellular & Developmental Biology - Plant* **43**: 267–274.

Vidal BC, Mello MLS, Illg RD, de Campos Vidal B, Campos De and Vidal B (1993) Chromosome number and DNA content in cells of a biotechnologically selected somaclone of garlic (*Allium sativum* L). *Rev Bras Genet* **16(2)**: 347–356.

Vidal MDC and De García E (2000) Analysis of a *Musa* spp. somaclonal variant resistant to yellow Sigatoka. *Plant Molecular Biology Reporter* **18**: 23–31.

Viseur J (1990) Evaluation of fire blight resistance of somaclonal variants obtained from the pear cultivar 'Durondeau'. *Acta Hortic* **273**: 275–284.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23(21)**: 4407–4414.

Walker D, Moňino I and Correal E. 2006. Genome size in *Bituminaria bituminosa* (L.) C.H. Stirton (Fabaceae) populations: separation of 'true' differences from environmental effects on DNA determination. *Environmental and Experimental Botany* **55**: 258–265.

Wang H, Chai Y, Chu XC, Zhao YY, Wu Y, Zhao JH, Ngezahayo F, Xu CM and Liu B (2009a) Molecular characterization of a rice mutator-phenotype derived from an incompatible cross-pollination reveals transgenerational mobilization of multiple transposable elements and extensive epigenetic instability. *BMC Plant Biology* **9**: 1-12. doi: 10.1186/1471-2229-9-63.

Wang K, Zhang S, Weber J, Baxter D, Galas DJ (2010) Export of microRNAs and microRNA-protective protein by mammalian cells. *Nucleic Acids Research* **38(20)**: 7248–7259. doi:10.1093/nar/ gkq601.

Wang SC and Frey PA (2007) S-adenosylmethionine as an oxidant: the radical SAM superfamily. *Trends Biochem Sci.* **32**: 101–110.

Wang X, Elling AA, Li X, Li N, Peng Z, He G, Sun H, Qi Y, Liu XS and Deng XW (2009b) Genome-wide and organ-specific landscapes of epigenetic

modifications and their relationships to mRNA and small RNA transcriptomes in maize. *The Plant Cell* **21**: 1053–1069.

Wang X, Song M, Gao C, Dong B, Zhang Q, Fang H, Yu Y (2009c) Carbenazim induces a temporary change in soil bacterial community structure. *Journal of Environmental Sciences (China)* **21(12)**: 1679-1683.

Wang X, Song M, Wang Y, Gao C, Zhang Q, Chu X, Fang H, Yu Y (2011) Response of soil bacterial community to repeated applications of carbendazim. *Ecotoxicology and Environmental Safety* **75(1)**: 33-39.

Wang X, Wu R, Lin XY, Bai Y, Song CD, Yu XM, Xu CM, Zhao N, Dong YZ and Liu B (2013) Tissue culture-induced genetic and epigenetic alterations in rice pure-lines, F1 hybrids and polyploids. *BMC Plant Biology* **13**: 1-12. doi: 10.1186/1471-2229-13-77.

Wang YS, Huang YJ, Chen WC, Yen JH (2009d) Effect of carbendazim and pencycuron on soil bacterial community. *Journal of Hazardous Materials* **172(1)**: 84-91.

Wani K and Aldape KD (2016) PCR techniques in characterizing DNA methylation. In: R Luthra, RR Singh and KP Patel (Eds.) *Clinical Applications of PCR* (Third Edition), pp. 177-186. New York, USA: Humana Press. doi: 10.1007/978-1-4939-3360-0_14.

Ward DC, Reich E and Goldberg IH (1965) Base specificity in the interaction of polynucleotides with antibiotic drugs. *Science* **149**: 1259-1263.

Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BT, Powell W (1997) Genetic distribution of BARE-1 retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). *Molecular Genetics and Genomics* **253**: 687-694.

Welsh J and McClelland M (1990) Fingerprinting genomes using PCR with arbitary primers. *Nucleic Acids Research* **18(24)**: 7213-7218.

Welsh J and McClelland M (1991) Genomic fingerprinting using arbitarily primed PCR and a matrix of pairwise combinations of primers. *Nucleic Acids Research* **19(19)**: 5275-5279.

Werner T, Motyka V, Laucou V, Smets R, Van OH, Schmulling T (2003) Cytokinin-deficient transgenic Arabidopsis plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. *Plant Cell* **15**: 2532-2550.

Wheeler DA, Srinivasan M, Egholm M, Shen YF, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski JR, Chinault C, Song XZ, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock GM, Gibbs RA and Rothberg JM (2008) The complete genome of an individual by massively parallel DNA sequencing. *Nature* **452**: 872-876.

White JF, Torres MS, Sullivan RF, Jabbour RE, Chen Q, Tadych M, Irizarry I, Bergen MS, Havkin-Frenkel D, Belanger FC (2014) Occurrence of *Bacillus amyloliquefaciens* as a systemic endophyte of vanilla orchids. *Microscopy Research and Technique* **77**: 874-885.

White PR (ed.) (1963) *The Cultivation of Animal and Plant Cells*. 2nd ed. New York: Ronald Press.

Whitehouse AB, Johnson AW, Passey AJ, McLeary KJ and Simpson DW (2014) Serenity: a paler skin-coloured somaclonal variant of the short-day cultivar 'Florence. *Acta Hortic* **1049**: 819–821.

Wildermuth MC, Steinwand MA, McRae AG, Jaenisch J and Chandran D (2017) Adapted Biotroph Manipulation of Plant Cell Ploidy. *Annu. Rev. Phytopathol*. **55**: 23.1-23.28.

Wilkinson TC and Thompson SA (1987) Genotype, medium and genotype medium effects on the stablishment of regenerable maize callus. *Maydica* **32**: 89-105.

Williams L, Zhao J, Morozova N, Li Y, Avivi Y and Grafi G (2003) Chromatin reorganization accompanying cellular dedifferentiation is associated with modifications of histone H3, redistribution of HP1, and activation of E2F-target genes. *Developmental Dynamics* **228**: 113–120.

Wittschieben B, Fellows J, Du W, Stillman DJ and Svejstrup JQ (2000) Overlapping roles for the histone acetyltransferase activities of SAGA and Elongator *in vivo. The EMBO Journal* **19(12)**: 3060-3068.

Wittschieben B, Otero G, Bizemont T, Fellows J, Erdjument-Bromage H, Ohba R, Li Y, Allis D, Tempst P and Svejstrup JQ (1999) A novel histone acetyltransferase is an integral subunit of elongating RNA polymerase II Holoenzyme. *Molecular Cell* **4**: 123-128.

Wu K, Jones R, Dannaeberger L and Scolnik PA (1994) Detection of microsatellite polymorphisms without cloning. *Nucleic Acids Research* **22**: 3257–3258.

Wyman J, Brassard N, Flipo D, Laliberte'e S (1992) Ploidy level stability of callus tissue, axillary and adventitious shoots of *Larix* × *eurolepis* Henry regenerated *in vitro*. *Plant Science* **85(2)**: 189-196.

Xu J, Hou QM, Khare T, Verma SK and Kumar V (2019) Exploring miRNAs for developing climate-resilient crops: A perspective review. *Science of the Total Environment* **653**: 91-104.

Xu Y, Zhao Q, Mei S and Wang J (2012) Genomic and transcriptomic alterations following hybridisation and genome doubling in trigenomic allohexaploid *Brassica carinata* x *Brassica rapa*. *Plant biology (Stuttg, Germany)* **14(5)**: 734–744.

Yaacob JS, Loh HS and Taha RM (2013a) Protein profiling and histone deacetylation activities in somaclonal variants of oil palm (*Elaeis guineensis* Jacq.). *The Scientific World Journal* **2013**: Article ID 613635. doi: 10.1155/2013/613635.

Yaacob JS, Taha RM, Esmaeili AK (2013b) Comparative studies on cellular behaviour of carnation (*Dianthus caryophyllus* Linn. cv. Grenadin) grown *in vivo* and *in vitro* for early detection of somaclonal variation. *The Scientific World Journal* **2013**: 686752. doi: 10.1155/2013/686752.

Yamashita H and Tahara M (2006) A LINE-type retrotransposon active in meristem stem cells causes heritable transpositions in the sweet potato genome. *Plant Molecular Biology* **61**: 79–94.
Yang C, Zhang M, Niu W, Yang R, Zhang Y, Qiu Z, Sun B and Zhao Z (2011) Analysis of DNA methylation in various swine tissues. *PLoS One* **6(1)**: e16229.
Yang MC and Loh CS (2004) Systemic endopolyploidy in *Spathoglottis plicata* (*Orchidaceae*) development. *BMC Cell Biology* **5**: 33.
Yang S, Yu H, Xu Y, Goh CJ (2003) Investigation of cytokinin-deficient phenotypes in *Arabidopsis* by ectopic expression of orchid DSCKX1. *FEBS Letters* **555**: 291-296.
Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S and Madden T (2012) Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13**: 134.
Yeh FC, Yang RC, Boyle TBJ, Ye ZH and Mao JX (1999) POPGENE 3.2, User-friendly shareware for population genetic analysis. http://ualberta.ca/wfeyeh
Yoder JA, Walsh CP and Bestor TH (1997) Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics* **13(8)**: 335–340.
Yu B and Chen XM (2010) Analysis of miRNA modifications. *Methods in Molecular Biology* **592**: 137-148. doi: 10.1007/978-1-60327-005-2_10.
Yu B, Yang ZY, Li JJ, Minakhina S, Yang MC, Padgett W, Steward R and Chen XM (2007) Methylation as crucial step in plant microRNA biogenesis. *Science* **307**: 932-934.
Yu C, Chen Y, Cao Y, Chen H, Wang J, Bi YM, Tian F, Yang F, Rothstein SJ, Zhou X and He C (2018) Overexpression of miR169o, an overlapping microRNA in response to both nitrogen limitation and bacterial infection, promotes nitrogen use efficiency and susceptibility to bacterial blight in rice. *Plant Cell Physiol.* **59**: 1234–1247. https://doi.org/10.1093/pcp/pcy060.
Zand J, Spreen A and LaValle J (1999) *Smart Medicine for Healthier Living*, pp. 44. New York, USA: Avery Publishing Group.
Zebrowska JI (2010) *In vitro* selection in resistance breeding of strawberry (*Fragaria x ananassa* Duch.). *Commun Agric Appl Biol Sci* **75**: 699–704.
Zhang J, Zhang H, Srivastava AK, Pan Y, Bai J, Fang J, Shi H and Zhu JK (2018) Knockdown of rice microRNA166 confers drought resistance by causing leaf rolling and altering stem xylem development. *Plant Physiol.* **176**: 2082–2094. https://doi.org/10.1104/pp.17.01432.
Zhang LQ, Cheng ZH, Khan MA and Zhou YL (2012) *In vitro* selection of resistant mutant garlic lines by using crude pathogen culture filtrate of *Sclerotium cepivorum* Australas. *Plant Pathol* **41**: 211–217.
Zhang N, Yang J, Wang Z, Wen Y, Wang J, He W, Liu B, Si H and Wang D (2014) Identification of novel and conserved microRNAs related to drought stress in potato by deep sequencing. *PLoS ONE* **9**: e95489.
Zhang NG, Yong JWH, Hew CS and Zhou X (1995) The production of cytokinin, abscisic acid and auxin by CAM orchid aerial roots. *Journal of Plant Physiology* **146:** 371-377.

Zhang Y, Rohde C, Tierling S, *et al.* (2009) DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. *PLoS Genet* **5**: e1000438.

Zhang Z, Wei L, Zou X, Tao Y, Liu Z and Zheng Y (2008) Submergence-responsive microRNAs are potentially involved in the regulation of morphological and metabolic adaptations in maize root cells. *Annals of Botany* **102**: 509–519.

Zhao S and Fernald RD (2005) Comprehensive algorithm for quantitative real-time polymerase chain reaction. *Journal of Computational Biology* **12(8)**: 1047-1064. doi: 10.1089/cmb.2005.12.1047.

Zhao YY, Mo BX and Chen XM (2012) Mechanisms that impact microRNA stability in plants. *RNA biology* **9(10)**: 1218-1223.

Zhou JH, Fu FY and Lei HP (1965) Recent advances in the Pharmacological research of medicinal plants in the people's Republic of China. In: *Proceedings of the 2nd International Pharmacology Meeting*, Prague.

Zhou TS and Sakaguchi S (2003) Somaclonal variation and the ploidy level in *Phalaenopsis alliance. Plant Morphology* **15(1)**: 2-7.

Zhou X, Wang G, Sutoh K, Zhu JK and Zhang W (2008) Identification of cold-inducible microRNAs in plants by transcriptome analysis. *Biochim Biophys Acta* **1779**: 780–788.

Zietkiewicz E, Rafalski A and Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* **20**:176–183.

Zonneveld B, Leitch I and Bennett M (2005) First nuclear DNA amounts in more than 300 angiosperms. *Annals of Botany* **96**: 229 – 244.

Zonneveld BJM (2001). Nuclear DNA contents of all species of *Helleborus* (Ranunculaceae) discriminate between species and sectional divisions. *Plant Systematics and Evolution* **229**: 125–130.

Zotz G, Winkler U (2013) Aerial roots of epiphytic orchids: the velamen radicum and its role in water and nutrient uptake. *Oecologia* **171(3)**: 733-741.

LIST OF PUBLICATIONS

Proteomic analysis on root callus and protocorm-like body of *Vanilla planifolia* Andrews [International Proteomics Conference 2017]

Genomic studies on somaclonal variations of *Vanilla planifolia* [Plant Genomic Congress: Asia (February 2014)]

An investigation on genetic and epigenetic instability in tissue culture of *Vanilla planifolia* [Cold Spring Harbor Asia Conference (June 2013]]

Detection of somaclonal variation in plant tissue culture using Random Amplified Polymorphic DNA (RAPD) [19th Malaysian Society for Biochemistry and Molecular Biology (MSMBB) Scientific Meeting (October 2012]]

Exogenous cytokinin induces callus and protocorm-like-bodies (PLBs) formation in *in vitro* root tips of *Vanilla planifolia* Andrews [Manuscript to be submitted]

Systemic variations of cellular endopolyploidy in different tissues of *V. planifolia* and its implications in micropropagation [Manuscript in preparation]

Integrated mRNA and microRNA analysis of *in vitro* vanilla callus culture [Manuscript in preparation]