EXPRESSION PROFILING OF microRNA GENES RELATED TO SALT STRESS IN BANANA (*MUSA ACUMINATA* CV. BERANGAN) ROOTS

LEE WAN SIN

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2015

EXPRESSION PROFILING OF microRNA GENES RELATED TO SALT STRESS IN BANANA (*MUSA ACUMINATA* CV. BERANGAN) ROOTS

LEE WAN SIN

THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2015

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: LEE WAN SIN (I.C./Passport No.:)

Registration/Matric No.:

Name of Degree: DOCTOR OF PHILOSOPHY

Title of Project/Research Report/Dissertation/Thesis ("this work"): EXPRESSION OF microRNA GENES RELATED TO SALT STRESS IN BANANA (*MUSA ACUMINATA* CV. BERANGAN) ROOTS

Field of Study: MOLECULAR BIOLOGY

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:
Designation:

ABSTRACT

As agricultural practices and climate change increase soil salinity, yields are reduced as crops succumb to salinity-induced abiotic stress. Physiological responses to stress are controlled by expression of genes, which in turn are regulated by microRNAs. Since most banana cultivars are salt-sensitive, improved understanding of genetic regulation of abiotic stress responses in banana can support future crop management and improvement.

Transcriptome libraries were constructed from salt-stressed and non-salt-stressed banana roots and then sequenced and annotated. Two banana root transcriptomes were assembled into 31,390 unigenes, of which 99.5% were mapped to the reference banana A-genome and 2,993 (9.5%) were found to be differentially expressed upon 300 mM NaCl treatment. Some of the salt stress responsive genes reported in plants were also observed in the salt-stressed banana roots in this study. The functions of these stress responsive genes include stress signaling, transcriptional regulation, DNA repair, transport, stress defence and cell wall modifications. Besides, genes that have not been reported to be responsive to salt stress or abiotic stress were also found in this study. The unannotated genes and genes coding for hypothetical proteins that were differentially expressed in the salt-stressed banana roots are putative novel salt stress responsive genes.

We identified 181 orthologous and 56 *Musa*-specific miRNAs, of which 59 miRNAs, 43 orthologous and 16 *Musa*-specific, were found to be differentially expressed in the salt-stressed banana roots. MiRNAs responded to salt stress in a dose-dependent manner. Banana roots expressed a unique set of miRNAs in response to salt stress and these miRNAs regulate diverse biological processes including stress signaling, transcriptional and translational regulations, stress defence, transport, cellular homeostasis, metabolisms and other stress-related functions. In addition, a number of

unannotated genes and genes coding for unknown proteins that are responsive to salt stress were predicted to be regulated by miRNAs. New miRNA-target modules related to salt stress response in banana roots were proposed in this study.

This present study may contribute to the understanding of gene regulation and abiotic stress response in banana roots and the high-throughput sequencing data sets generated in this present study may serve as important genetic resources for salt tolerance traits used for functional genomic studies and genetic improvement in banana.

ABSTRAK

Kegaraman (salinity) merupakan salah satu jenis tekanan abiotik utama yang mengancam penghasilan tanaman di seluruh dunia. Masalah kegaraman menjadi semakin serius akibat daripada perubahan iklim global dan amalan pertanian yang tidak betul. Mikro-RNA (microRNA) merupakan pengawal ekspresi gen yang penting. RNA ini mengawal pelbagai proses kehidupan termasuk pertumbuhan, perkembangan dan tindak balas tumbuhan terhadap tekanan persekitaran. Kebanyakan kultivar pokok pisang adalah peka terhadap kandungan garam yang tinggi dalam tanah. Peningkatan kandungan garam dalam tanah akan mengurangkan hasil buah daripada pokok pisang. Pemahaman ekspresi dan pengawalaturan gen terhadap tekanan abiotik pada pokok pisang adalah penting untuk pengurusan tanaman dan program peningkatan mutu genetik tanaman.

Dalam kajian ini, dua buah perpustakaan transriktom dan tiga buah perpustakaan untuk RNA jenis saiz kecil telah dibina daripada akar pokok pisang yang dirawat dengan garam dan juga akar yang tak terawat (kawalan). Penjujukan perpustakaan-perpustakaan tersebut telah dijalankan dengan menggunakan platform teknologi penjujukan Illumina. Sebanyak 31,390 unigen dapat dipasang daripada jujukan-jujukan DNA yang terhasil daripada dua buah perpustakaan transkriptom tersebut dan 99.5% daripada unigen-unigen ini dapat dipetakan pada genom rujukan pisang. Sebanyak 2,993 (9.5%) daripada unigen-unigen ini menunjukkan perbezaan tahap ekspresi yang ketara selepas rawatan garam. Unigen-unigen yang menunjukkan perbezaan ekspresi ketara tersebut adalah terlibat dalam isyarat tekanan, pengangkutan di dalam dan di antara sel, pengawalaturan transkripsi, pertahanan sel dan pengubahsuaian dinding sel.

MiRNA yang telah dikenalpasti di tumbuhan-tumbuhan lain dan miRNA yang hanya spesifik kepada *Musa* species dapat dijumpai dalam kajian ini. Sebanyak 181 miRNA ortolog dan 56 miRNA spesifik kepada *Musa* species telah dikenalpasti, di

mana 43 miRNA ortolog dan 16 miRNA spesifik kepada *Musa* species menunjukkan perubahan ekspresi yang ketara pada akar pokok pisang yang mengalami kegaraman dalam kajian ini. Sasaran-sasaran (mRNA) untuk miRNA dan laluan metabolik yang dikawalatur oleh miRNA dalam pokok pisang yang mengalami tekanan kegaraman dapat dijumpai di transkriptom akar pokok pisang. Dalam kajian ini, modul-modul baru untuk kawalan ekpresi gen oleh miRNA di akar pokok pisang telah dicadangkan.

Gabungan pendekatan penjujukan RNA bersaiz kecil dan penjujukan transkriptom dapat mendedahkan ekspresi miRNA dan pengawalaturan gen yang terlibat dalam tindak balas akar pokok pisang terhadap pengaraman.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my supervisors, Professor Dr. Jennifer Ann Harikrishna and Professor Dr. Norzulaani Khalid, for their guidance and patience with me throughout my doctorate study.

I would like to acknowledge the Ministry of Education Malaysia and University of Malaya for the sponsorship, research facilities and funding for my doctorate study.

I am thankful to Adjunct Professor Dr. Martti and my colleague, Mr. Ranga who have helped in the bioinformatic and statistical analyses. I would like to thank my colleagues in the BGM and CEBAR laboratories, Hui Li, Kai Swan, Su Ee, Fauziah, Wong, Jasdeep, Hana, Rezaul, and Tyson for their help in the lab. Many thanks to Mrs. Azlina Abdul Rahman and my colleagues in the Plant Biotechnology Research Laboratory, Dr. Tan Boon Chin, Chin Fong, Wendy Chin, Sher Ming, Tan HC, Fatin, Nabeel, Nazrin, Ain Wahid, Diyana, Nadiya, Akmal, Gayatri, Wani and Aiman, for their support and assistance. I am grateful to have been taught by Dr. Wong Wei Chee when I first joined the Plant Biotechnology research group. Special thanks to late Mr Abdul Razak Abu Bakar for his encouragement, motivation and kind words when I was feeling down during my doctorate study.

Many thanks to Lee LC, Teck Kai, Siew JJ, Tamil, Adriya, Hui Li, Chin Fong, Wendy Chin, Siaw YY, Jen Ni, Chia Yin, Su Ee, Ka Loo, Johnson and many others for their invaluable friendship and moral support. Last but not least, I would like to thank my beloved parents, sister, brothers, brother-in-law, sister-in-law, nephews and niece for their love and support.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

%	percentage
µl	microliter
A genome	*Musa acuminata* genome constituent
B genome	Musa *balbisiana* genome constituent
bp	base pairs
cDNA	complementary DNA
Cl^-	chloride ion
COG	Clusters of Orthologous Groups
contigs	contiguous sequences
CTAB	cetyltrimethylammonium bromide
CTR	control
DNA	deoxyribonucleic acid
dS m^{-1}	deciSiemens per meter
dsRNA	double-stranded RNA
EST	expressed sequence tags
FDR	false discovery rate
g	gram
GO	gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
KOG	Eukaryotic Orthologous Groups of proteins
m	meter
M	molar
M. acuminata	*Musa acuminata*
M. balbisiana	*Musa balbisiana*
mg	miligram

miRNA	microRNA
miRNA*	microRNA star stand
ml	mililitre
mM	milimolar
mRNA	messenger RNA
mRNA-Seq	mRNA sequencing
Na^+	sodium ion
NaCl	sodium chloride
Nr	GenBank non-redundant protein database
nt	nucleotides
PCR	polymerase chain reactions
PMRD	Plant MicroRNA Database
Pre-miRNA	precursor microRNA
Pri-miRNA	primary microRNA
PTGS	post transcriptional gene silencing
Q20	Phred score with 99% accuracy of the base call
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNAi	RNA interference
RNA-Seq	RNA sequencing
RPKM	reads per kilobase per million mapped reads
rRNA	ribosomal RNA
RT-PCR	reverse-transcriptase PCR
SAGE	serial analysis of gene expression
siRNA	small interfering RNA
snoRNA	small nucleolar RNA

sp.	species (singular)
spp.	species (plural)
sRNA-Seq	small RNA sequencing
ssp.	subspecies
SSR	short sequence repeat
ta-siRNA	trans-acting small interfering RNA
TPM	tags (or transcripts) per million
TR100	100 mM NaCl treatment
TR300	300 mM NaCl treatment
tRNA	transfer RNA
w/v	weight per volume

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

Bananas are perennial monocotyledonous plants taxonomically classified in the order Zingiberales and genus *Musa* (APG III, 2009). Being a giant herbaceous plant, bananas grow well in the humid tropical and sub-tropical regions and are planted mainly by small-scale farmers for home consumption and for sale in local and regional markets (Roux et al., 2008). Bananas are the fourth most important food crop after rice, wheat and maize in terms of gross production value in the developing world, and it has been one of the commodities with the highest world export volume and value in the fruit and vegetables category (Frison & Sharrock, 1998; FAO, 2003; Perrier et al., 2011). Besides being consumed in almost all regions of the world as sweet dessert fruit for their nutritional value, bananas are also an important starch-rich staple food providing a major daily source of calories for millions of people in the developing countries including the poorest regions in the world. Hence, bananas are an important component of the world's food security (Pillay & Tripathi, 2007; Roux et al., 2008).

Most cultivated bananas are sterile and polyploid, reproducing asexually with the fruits formed via parthenocarpy (Heslop-Harrison & Schwarzacher, 2007; Roux et al., 2008) and the majority of the global banana production is contributed by only a small number of genotypes. Therefore, cultivated bananas have a narrow or poor genetic base and are susceptible to damage by various pests and diseases such as fungi, bacteria, viruses, nematodes and insects, as well as physical stresses such as drought, salinity, cyclones and high winds, flood and poor soil fertility (Osborn, 2002; Pillay & Tripathi, 2007; Perrier et al., 2011). In addition, conventional breeding programmes are difficult due to the sterility of the cultivated bananas (Heslop-Harrison & Schwarzacher, 2007). Many of the devastative factors to banana plantations affect the roots of banana. For instance, soil-borne pathogens, nematodes and Fusarium wilt cause deterioration of the root system and enter the banana plant via roots. High water table and high salt and

sodium content cause the collapse of the banana root system (Gauggel et al., 2003). Therefore, growth and health of banana roots is one of the major factors in protecting the plants or reducing the risks of plant damage and loss in fruit production.

Soil salinity is defined as increased NaCl concentration and excess sodium in soil (Shapira et al., 2009). This phenomenon is caused by low precipitation, high evaporation, irrigation with saline water, poor cultural practice, weathering of native rocks and rising water tables, and is substantially found in the arid and semiarid regions with high evaporation rates (Dasgan et al. 2002; Ashraf & Foolad, 2007). Salinity affecting 7% of the current world's arable land and has been estimated to increase to 30% by 2025 and 50% by the year 2050 due to increased accumulation of salt delivered along with irrigation water and high evapotranspiration rate caused by climate change (Wang et al., 2003; Mahajan & Tuteja, 2005). Salt deposition in arable lands causes low soil porosity, reduced soil aeration and water conductance, and generates a low water potential zone that is difficult for plants to acquire water and nutrients (Mahajan & Tuteja, 2005). Besides, soil salinity also possesses secondary effects, such as oxidative stress and Na^+ toxicity in plants. In general, salt-stressed plants show reduced growth, early leaf senescence and appearance of chlorotic and necrotic spots on leaves (Tester & Davenport, 2003; Shapira et al., 2009). Therefore, soil salinity results in agricultural loss and yield reduction and poses a crisis to global food supply. Banana cultivars are considered salt sensitive with a yield reduction of about 50% (Israeli et al., 1986) and plant height reduction of about 80% (Yano-Melo et al., 2003) when the electrical conductivity of the irrigation water was raised from 1 to 7 dS m^{-1}.

In this study, high-throughput next generation sequencing technology was used to obtain high resolution gene expression profiles of salt-stressed banana roots. This included protein-coding transcripts and the regulatory small RNAs that shape the plant genome. A transcriptome is a set of transcripts from a cell or a population of cells,

which consists of protein-coding mRNAs and non-coding small RNAs, such as, rRNA, tRNA and miRNA, in defined growth conditions (Wang et al., 2010b). Transcriptomic approaches have been widely used to catalogue transcripts that are expressed in various cells, tissues or organs of plant species grown in different conditions. In bananas, EST libraries have been generated from leaves subjected to temperature stresses (hot and cold), and from fruits during ripening (Santos et al., 2005; Manrique-Trujillo et al., 2007). SuperSAGE has been used for characterization of banana leaf transcriptome, while microarray technology has been applied to characterize banana leaf transcriptome during drought stress (Coemans et al., 2005; Davey et al., 2009). High-throughput sequencing approaches have been used to profile gene expression in banana roots inoculated with *Fusarium oxysporum* f. sp. *Cubense* (Li et al., 2012a; Wang et al., 2012b; Li et al., 2013b) and banana leaf infected with *Mycosphaerella musicola* (Passos et al., 2013). However, functional genomic data (both transcriptome and regulatory microRNA) of salt-stressed banana roots have not been studied.

MicroRNAs (miRNA) are ~20-nt single-stranded functional RNA species derived from long stem-loop precursors transcribed from endogenous genomic DNA (Lee et al., 2004). The non-coding RNA transcribed from genomic DNA is then processed to form a long hairpin stem-loop precursor miRNA (pre-miRNA). This structure is further processed by a Dicer-like enzyme, to a shorter hairpin structure known as a primary miRNA (pri-miRNA) and then to a shorter double-stranded RNA (dsRNA) called miRNA:miRNA* duplex (Kurihara & Watanabe, 2004; Liu et al., 2005). In the cytoplasm, the active strand of the miRNA duplex incorporates into a RNA induced silencing complex (RISC) which guides the mature miRNA to the complementary target mRNA and either causes the degradation or the inhibition of translation of the target mRNA (reviewed in Zhang et al., 2006c). This mechanism of gene regulation is driven by an RNA interference (RNAi) machinery, and is commonly

known as post-transcriptional gene silencing (PTGS) in plants.

MicroRNAs have been reported to regulate diverse life processes in plants including the regulation of floral organ identity, flower timing, leaf shape, root development, organ boundary and polarity (Dugas & Bartel, 2004), embryogenesis (Luo et al., 2006) and auxin response pathways (Ru et al., 2006; Wu et al., 2006). MiRNAs also play important roles in regulating RNA interference pathways such as feedback regulation of miRNA biogenesis and directing biogenesis of a class of siRNA, known as trans-acting siRNA or ta-siRNA. A number of miRNA families have been shown to have involvement in various biotic and abiotic stress responses in plants. These include responses to oxidative stress, mechanical stress and nutrient stress (Jones-Rhoades & Bartel, 2004; Chiou, 2007; Zhao et al., 2007a). Therefore, study of miRNA is of importance in elucidation of gene regulatory networks of plants, and has important implications for the practical plant sciences.

Next generation sequencing technology applies sequencing of DNA nucleotides by synthesis. It is useful for massive parallel DNA sequencing producing millions of single molecule clusters in a short time. This sequencing technology enables us to obtain highly resolved structural information of RNA populations and to understand the functional elements within the genome, in order to discover novel developmental or environmental regulatory networks (Wang et al., 2010b). The next generation high-throughput sequencing technologies also offer an option to deeply sequence small RNA cDNA libraries with comparatively low cost and are less tedious compared to the older sequencing technologies.

Due to the increase of arable lands being salinised and the sensitivity of banana plants to salinity, it is of great importance to investigate the expressed genes, pathways and their regulating miRNAs in this plant, in response to salt stress. Being the first tissues exposed to salt stress injury and signal transduction, roots are an ideal target for

gene expression and regulation study in bananas. MicroRNAs were investigated in this study as they are crucial genome elements modulating gene expression for various life processes. Use of high-throughput sequencing technology to deeply sequence transcriptome and small RNA species (including microRNAs) in salt-stressed banana plants can potentially reveal gene regulation networks and important information of salt-stress sensitivity and tolerance in plants. Integration of the functional genomic data, together with other genomic, evolutionary and structural data will allow targeted breeding and transformation for banana genetic improvement in the future (Heslop-Harrison & Schwarzacher, 2007).

The objectives of this study were:

1. to determine the gene expression profile of salinity-stressed banana roots using a high-throughput mRNA sequencing platform
2. to determine the gene expression profile of salinity-stressed banana roots using a high-throughput sRNA sequencing platform
3. to determine candidate targets of miRNAs responding to salinity stress from banana root transcriptomes
4. to elucidate new or novel miRNA-mRNA interactions in salt-stressed banana roots
5. to reveal miRNA-mRNA interaction modules with potential application for engineering salt tolerant plants

CHAPTER 2: LITERATURE REVIEW

2.1 Bananas and plantains

2.1.1 Morphology, origin, taxonomy and diversity of bananas and plantains

Bananas and plantains (*Musa* spp.) are giant herbaceous monocotyledonous plants with a pseudostem that is made of tightly packed leaf sheaths. Leaves and bracts of the plants are spirally arranged to form the pseudostem. Each pseudostem begins as a sucker (Figure 2.1), a shoot bud that grows from an underground rhizome known as a corm (the true stem), and mature plants range from 2 to 15 meters high, with many fruit cultivars commonly 3 meters in height (reviewed in Heslop-Harrison & Schwarzacher, 2007; Roux et al., 2008). After a certain number of leaves are produced, the inflorescence (flowering head) also grows up through the centre of the pseudostem (Figure 2.1). The inflorescence is made up of several groups of tubular flowers enveloped in large reddish bracts. The first five to fifteen flower clusters are functionally female and form fruits, while subsequent flowers are transitional (functionally male) and male flowers with abortive ovary and well developed stamens (Simmonds, 1966; Roux, 2008). The female flowers (stigma) are orange to rich yellow in *Musa acuminata* and cream, pale white or pale pink in *Musa balbisiana*. Whereas the male flowers are creamy (*M. acuminata)* or variably flushed with pink (*M. balbisiana*) (Simmonds, 1962; Valmayor et al, 2000). The female flowers, which contain large ovaries, appear at the base of the inflorescence and will develop into fruits. At the tip of the inflorescence, a large purplish red bud containing male flowers continues to elongate (Figure 2.1). Each pseudostem produces a single bunch of bananas and will eventually die after fruiting (Simmonds, 1966; IPGRI–INIBAP/CIRAD, 1996).

Bananas and plantains grow in many parts of the humid tropics and subtropics. They are believed to have occurred from India to Polynesia with a centre of origin in South-East Asia where the wild parents of cultivated bananas are found (Simmonds,

1962) and the centre of diversity in Malaysia or Indonesia where the greatest diversity of edible bananas are observed (Daniells et al., 2001a). Bananas and plantains are classified in the Zingiberales order and Musaceae family. Zingiberales, which contains gingers and bananas, is grouped under the same clade (commelinids) as Poales that contains grasses such as rice, maize, wheat, sorghum and bamboo (APG III, 2009). Beside Musaceae, Zingiberales also contains seven other families including Zingiberaceae and Strelitziaceae, which are closely related to bananas (Figure 2.2). The Musaceae family consists of three genera, namely *Musa*, *Ensete* and *Musella* (Christelova et al., 2011; Perrier et al., 2011). Ensete has relatively minor importance as a crop and is used as starchy foodstufs and a boiled vegetable and for its fibre in East Africa. *Musa* is the most economically important genus in Musaceae. This genus is further devided into four (or five) sections: Australimusa, Callimusa, Eumusa and Rhodochlamys (Figure 2.2). Eumusa and Rhodochlamys have a basic chromosome number of $x = 11$, while Australimusa and Callimusa have chromosome numbers $x = 10$ and $x = 9$ or 10 respectively (reviewed in Heslop-Harrison & Schwarzacher, 2007).

Callimusa and Rhodoclamys do not produce edible fruits therefore they are only used as ornamental plants. Australimusa is used mainly for its fibre and less extensively for its fruit. Eumusa is the best known section as it is hardy and easy to grow. It is the biggest and most geographically widespread section (Stover & Simmonds, 1987). Eumusa section is known to have eleven wild species including *Musa acuminata* (A genome), *Musa balbisiana* (B genome) and *Musa schizocarpa* (S genome). Together with *Musa textilis* (T genome) from the Australimusa section, these wild species served as the progenitors of all cultivated bananas (Daniells et al., 2001a). The development of natural reproductive barriers due to chromosome structural changes or DNA recombination has resulted in nine subspecies within *Musa acuminata*, which are *banksii*, *burmannica*, *burmannicoides*, *malaccensis*, *microcarpa*, *truncata*, *siamea*,

zebrina and *errans* (Daniells et al., 2001a). With the evidence from DNA analyses carried out by various research groups, Häkkinen (2013) revised *Musa* to two sections: *Musa* sect. *Musa* and *Musa* sect. *Callimusa*. *Musa* sect. *Musa* has a basic chromosome number of $n = x = 11$ and comprises of species from sections Eumusa and Rhodochlamys. While *Musa* sect. *Callimusa* contains species from Callimusa, Australimusa and Ingentimusa, with basic chromosome numbers $n = x = 10/9/7$.

The majority of the cultivated bananas originate from the parent species, *Musa acuminata* (A genome) and *Musa balbisiana* (B genome) (Roux et al., 2008). The cultivated bananas, which are edible, lack the seeds that are found in the wild bananas. These edible cultivars are sterile diploid, triploid or tetraploid and propagate vegetatively. Fruits develop without pollination via a process called parthenocarpy (Heslop-Harrison & Schwarzacher, 2007). Most of the cultivated bananas appeared through natural mutations and/or hybridization involving one or both wild genomes from *M. acuminata* and *M. balbisiana* (reviewed in Heslop-Harrison & Schwarzacher, 2007). As agriculture developed, farmers further mixed and selected cultivars based on plant vigour, hardiness, yield, seed sterility (seedlessness) and fruit quality (Simmonds, 1962). Various ploidy and mixes of the A and B genomes are found in cultivated banana varieties: AA, AB, AAA, AAB, ABB, AAAA, AAAB, AABB and ABBB (INIBAP, 2006; Heslop-Harrison & Schwarzacher, 2007). Cavendish bananas (triploid AAA group) are currently the most well-known cultivated bananas and dominate international trade (Frison & Sharrock, 1998; INIBAP, 2006). In Malaysia, 'Pisang Berangan' (AAA genome) is the most commonly cultivated local banana cultivar. Berangan, with its bunch weight range 15 to 25 kg and sweet and aromatic pulp, make it an excellent dessert banana. It is also a banana variety for export to the neighbouring countries in Southeast Asia (Omar et al., 2012). Like Cavendish and Gros Michel, Berangan is a triploid A genome banana with haplotype lineages to wild bananas *Musa acuminata*

subspecies *malaccensis*, *zebrina* and *banksii*. Unlike Cavendish and Gros Michel, which probably originated from the same ancestral population, Berangan is clearly from a distinct population (personal communication with Dr. Hugo Alfried Volkaert).

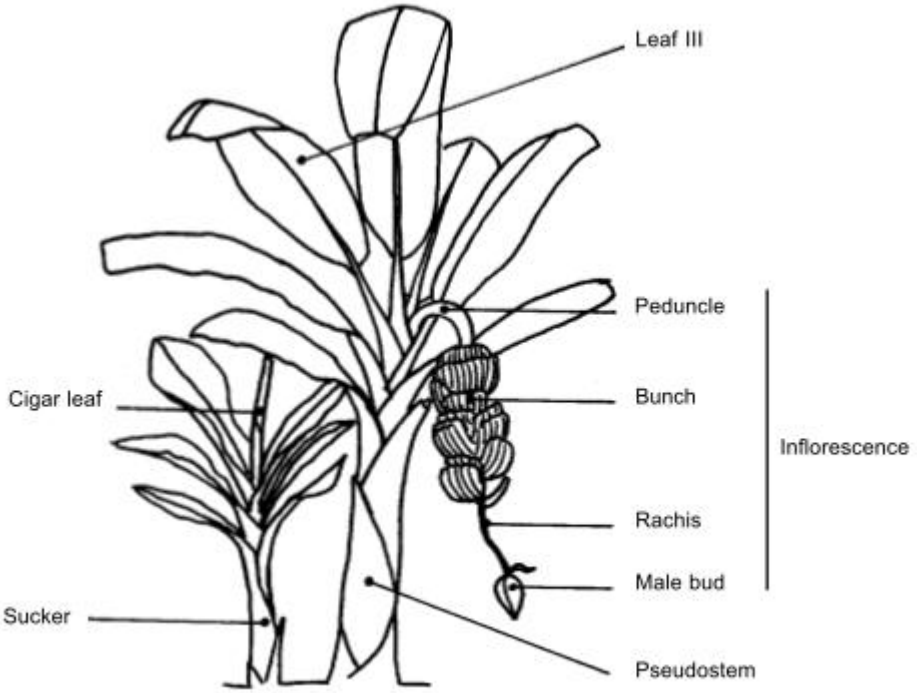


Figure 2.1: General morphology of bananas and plantains (IPGRI–INIBAP/CIRAD, 1996). Mature plant with pseudostem (right) and sucker (left) are shown here. Banana plant's inflorescence consists of the banana bunch (derived from female flowers) and male bud supported by peduncle and rachis respectively

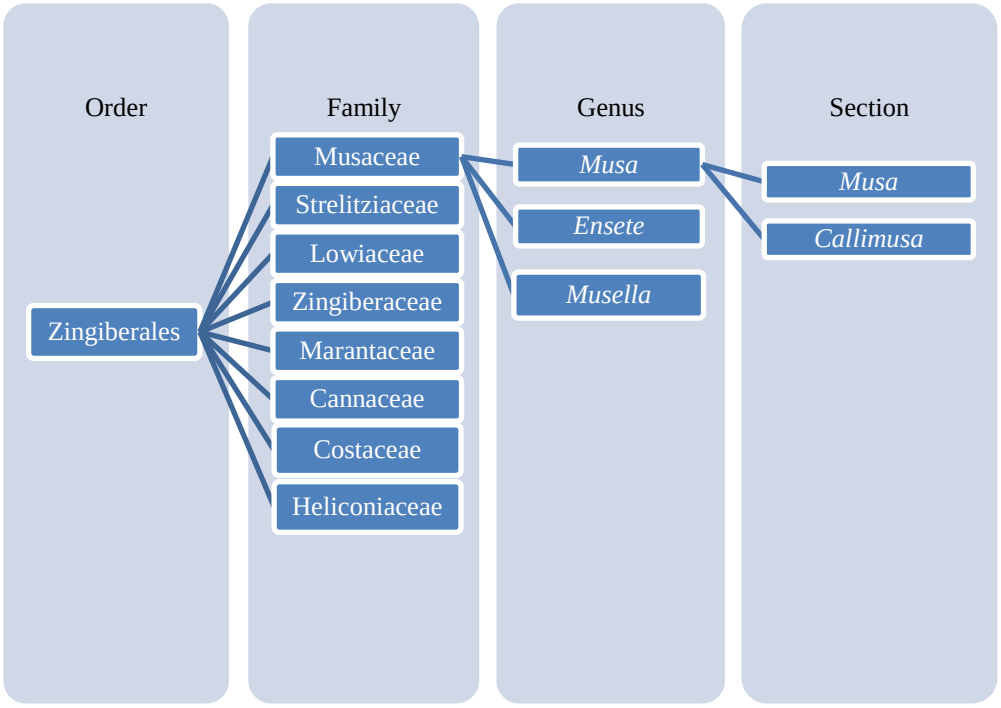


Figure 2.2: Taxonomic classification of bananas and plantains (APG III, 2009; Häkkinen, 2013). Bananas and plantains belong to Zingiberales order and Musaceae family. *Musa* genus is sub-divided into two sections, sect. *Musa* and sect. *Callimusa*. Sect. *Musa* comprises of species from Eumusa and Rhodochlamys. While sect. *Callimusa* contains species from Callimusa, Australimusa and Ingentimusa. *Musa acuminata* and *Musa balbisiana*, the progenitors of most cultivated bananas, belong to the *Musa* (Eumusa) section

2.1.2 Socio-economic importance of banana

Besides being a popular dessert fruit with high nutritional value, bananas and plantains are also an important staple food providing a cheap and easily produced source of energy to millions of people in the poorest regions of the world (Roux et al., 2008). Bananas and plantains are the world's most important food crop after rice, wheat and maize (INIBAP, 2006; Perrier et al., 2011). World gross production value of bananas and plantains in the year 2011 was 28,362 million USD compared to 188,652, 121,362 and 108,255 million USD in rice, maize and wheat respectively (FAOSTAT,

2011). World production of bananas and plantains increased steadily from 76.25 million tonnes in 1991 to 145.44 million tonnes in 2011, a 91% increase in production in the past 20 years (Figure 2.3).

Bananas are grown in all tropical regions and planted mainly by smallholders as a major income source (Roux et al., 2008). The smallholders grow over 85% of the bananas harvested in the world (FAO, 2003; INIBAP, 2006). As a staple food, bananas and plantains contribute to the food security of millions of people in the developing world. As a commodity, bananas and plantains are also a source of income and employment to rural populations and a key contributor to the economies of many low income food deficit countries (Frison & Sharrock, 1998; Pillay & Tripathi, 2007; Roux et al., 2008).

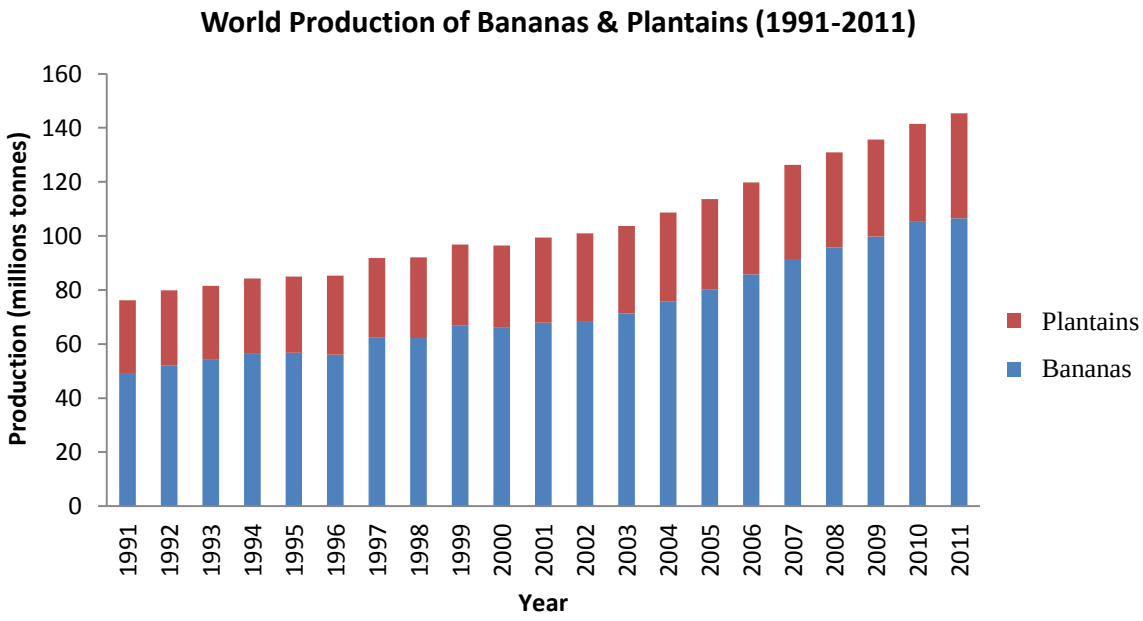


Figure 2.3: World production of bananas and plantains from year 1991 to 2011 (FAOSTAT, 2011). In FAOSTAT, 'bananas' is defined as *Musa sapientum*, *M. cavendishii* or *M. nana*, which are normally eaten raw. Whereas item named 'plantains' is defined as *Musa paradisiaca* and generally known as a cooking banana. FAOSTAT provides time-series and cross sectional statistics of food and agriculture for some 200 countries and is fully compliant with the principles of good practice governing international statistics. The data collectively indicates that the world production of bananas and plantains increases steadily over the past 20 years.

2.1.3 Nutritional value and usefulness of banana

Banana fruits are good sources of carbohydrates, minerals and vitamins. They are relatively high in starch content with 24% of the fresh banana fruit pulp being carbohydrate. Besides, banana fruits have a low to negligible level of fats, cholesterol and salt but high fibre content, high level of minerals especially potassium (400 mg/100 g pulp) and are rich in vitamins A, C and B_6 (Stover & Simmonds, 1987; Pillay &

Tripathi, 2007). These have made banana fruit an ideal staple food and source of energy, and an essential source of nutrition for several communities around the world.

Banana fruits can be fermented to produce beer, vinegar and other products. In East Africa, brewing using ripe fruits of highland banana cultivars to make beer is practiced (Stover & Simmonds, 1987). Banana fruits are also used to make processed food such as puree, chips or crisps, flour, powder and juice (Stover & Simmonds, 1987; Pillay & Tripathi, 2007).

Beside fruits, other parts of banana plant such as male buds, corms and shoots can be consumed as food. This is mainly practiced in Africa and Asia. Banana leaves can be used as plates and for wrapping food (Stover & Simmonds, 1987). The pseudostems of *Musa textilis* are harvested for fibre and used for textiles (Roux et al., 2008). Some *Musa* species possess variegated leaves or male floral bud with an attractive colour so are used as ornamental plants, for instance, *Musa acuminata* subspecies *zebrina*, *Musa ornata* and *Musa velutina* (Häkkinen & Sharrock, 2002).

2.1.4 Challenges to banana production

Most cultivated bananas are polyploid, seedless and sterile. They reproduce vegetatively and fruits are formed without fertilization, a process called parthenocarpy (Heslop-Harrison & Schwarzacher, 2007; Roux et al., 2008). Therefore, cultivated bananas have a narrow or poor genetic base and are susceptible to pests and diseases such as fungi, bacteria, viruses and nematodes (Table 2.1), as well as physical stresses such as wind, drought and flood (Osborn, 2002; Pillay & Tripathi, 2007; Perrier et al., 2011).

The most destructive disease to banana plantations worldwide is Fusarium wilt also known as Panama disease, caused by a soil-borne fungus (*Fusarium oxysporum* f. sp. *cubense* Race 1) that infects banana plants from the roots. Panama disease virtually

wiped out the banana variety that dominated the world's banana industry, Gros Michel during the 1950s (Heslop-Harrison & Schwarzacher, 2007). The successor cultivar, Cavendish which is resistant to Panama disease, however, is now being threatened by *Fusarium oxysporum* f. sp. *cubense* Tropical Race 4 (FOC4) (Stover & Simmonds, 1987; reviewed in Heslop-Harrison & Schwarzacher, 2007). Other destructive diseases are black Sigatoka, a foliar disease caused by soil borne fungus *Mycosphaerella fijiensis* (De Lapeyre de Bellaire, 2010), banana bunchy top virus (BBTV) which suppresses and retards the growth of banana plants, and burrowing nematodes which infect banana plants from the root and damage the root and vascular tissues of the plants (Heslop-Harrison & Schwarzacher, 2007).

A wide range of biotic stresses affect the yield and productivity of bananas, worldwide, as shown in Table 2.1. Nearly half of the world banana production relies on a single cultivar, Cavendish, and this practice poses high risk to the global banana industry as a pandemic disease outbreak could wipe out the current banana industry (Lescot et al., 2011; D'Hont et al., 2012). In addition, there is lack of simple screening methods for resistance across the available germplasm (Pillay & Tripathi, 2007). Continuous occurrence of new virulent diseases and rapid adaptation of diseases to the existing resistant banana varieties are among the challenges to the banana industry (Heslop-Harrison & Schwarzacher, 2007; De Lapeyre de Bellaire, 2010).

Banana plants are also prone to damage by environmental (abiotic) stresses. However, there is relatively little study on abiotic stress in banana compared to pests and diseases. Banana plants are vulnerable to strong wind due to their weak pseudostems, large leaves that trap wind and shallow root system (Pillay and Tripathi, 2007). A large number of plantations are devastated by wind causing a great loss in banana production every year (Pestana et al., 2011). Cultivated bananas are considered salt sensitive. An increase of electrical conductivity of the irrigation water from 1 to 7

dS m^{-1} was reported to cause about 50% yield reduction (Israeli et al., 1986) and about 80% reduction in height (Yano-Melo et al., 2003) in banana. In some parts of the world, banana cultivation is confined to the warmer coastal regions where soil and water salinity problems exist (Miri et al, 2009). Drought is one of the major abiotic factors limiting plant performance and banana production. As a giant herbaceous plant, banana is sensitive to cellular dehydration (Ravi et al., 2013; Sreedharan et al., 2013). Banana plants are also sensitive to cold stress by which injury is caused in episodes of chilling during winter and early spring (Kang et al., 2007). Abiotic stresses, including drought, salinity, strong wind, cold, heat and flooding, are increasingly significant in banana production due to global climate change (Heslop-Harrison & Schwarzacher, 2007). In addition, there is increased competition for land and clean water use and lack of arable land for agriculture. Therefore, abiotic stresses are likely to grow in importance for banana cultivation in the near future.

Despite being threatened by a wide range of biotic and abiotic stresses, the progress of banana genetic improvement to cope with the problems is still lagging behind. There is a need to enhance survival and production of banana plants under adverse environmental conditions. However, the low fertility nature of cultivated bananas is a challenge to producing improved banana cultivars via conventional breeding (Heslop-Harrison & Schwarzacher, 2007; Pillay & Tripathi, 2007).

Table 2.1: List of major pests and diseases in banana cultivation

	Agents	Disease & symptoms	References
Viruses	Banana Bunchy Top Virus (BBTV)	BBTV is an ssDNA virus that causes banana bunchy top disease, a serious widespread disease affecting bananas, and is transmitted by banana aphids. Infected plants either do not bear fruits or with stunted and unmarketable fruits.	Stover & Simmonds, 1987; Hooks et al, 2008
	Banana Streak Virus (BSV)	BSV is a Badnavirus with dsDNA genome that causes banana streak disease. Infected plants show chlorotic and necrotic streaks seen on leaves. It causes stunted plants that do not flower or with smaller bunches and abnormal fruit shape. BSV viral sequences can integrate into host genome.	Daniells et al., 2001b; Gayral et al., 2010
	Cucumber Mosaic Virus (CMV)	CMV is an ssRNA virus that causes severe chlorosis, mosaic and heart rot in banana plants. It is transmitted by aphid and infected plants may show either stunted, necrotic or no fruits.	Singh et al., 1995; Srivastava et al., 1995
	Banana Mild Mosaic Virus (BanMMV)	BanMMV is a filamentous virus. Infected plants with this virus will show mild chlorotic streak symptoms .	Reichel et al., 2003
	Banana Bract Mosaic Virus (BBrMV)	BBrMV is a Potyvirus with ssRNA genome and transmitted to host through aphids. Infected plants showed mosaic on the bracts of the inflorescence and streaks on pseudostem.	Hirimburegama, et al., 2004
Fungi	*Fusarium oxysporum* f. sp. *cubense*	This fungus causes Fusarium wilt or Panama disease in banana plants. It is a vascular disease causing disruption of translocation and systemic foliage symptoms. Infected plants show yellowing then wilting of leaves, splitting of the pseudostem base and vascular discoloration from pale brown to black, and eventually collapse of the crown and pseudostem.	Stover & Simmonds, 1987; Jeger et al., 1995
	Mycosphaerella musicola	*M. musicola* causes Yellow Sigatoka, a global disease epidemic of bananas. Infected bananas showed necrotic leaf lesions, reduced photosynthesis and crop yield.	Arzanlou et al., 2007
	Mycosphaerella fijiensis	*M. fijiensis* causes Black Sigatoka or Black Leaf Streak in banana, a foliar disease that affects leaf photosynthetic capacity, which, in turn, reduces the quality and quantity of fruits.	De Lapeyre de Bellaire, 2010; Churchill, 2011
	Mycosphaerella eumusae	*M. eumusae* causes Septoria Leaf Spot disease in banana plants. The symptoms are rather similar to Yellow Sigatoka and Black Sigatoka.	Arzanlou, et al., 2007

Table 2.1: Continued

	Agents	Disease & symptoms	References
Nematodes	*Radopholus similis*	Burrowing nematode is the most damaging and widespread nematode in bananas. It affects the root system causing wake anchorage and uprooting of banana plants.	Stover & Simmonds, 1987; Gowen, 1995
	Pratylenchus goodeyi & *Pratylenchus coffeae*	Lesion nematode that causes necrosis in banana root and rhizome.	Stover & Simmonds, 1987; Gowen, 1995
Insects	*Cosmopolites sordidus*	Borer weevil is found in majority of the banana planting area in the tropics. This nematode lays eggs at the pseudostem base and the larvae burrow into the rhizome leaving a network of tunnels which may kill young plants and increase toppling in the older plants.	Stover & Simmonds, 1987; Gowen, 1995
Bacteria	*Ralstonia solanacaerum*	Moko disease, an important bacterial disease of banana. This soil-borne plant vascular pathogen causes yellowing and necrotic leaves, arrested fruit bunch development and wilting of young suckers.	Jeger et al., 1995; Cellier et al., 2012
	Erwinia carotovora	Bacteria that causes head-rot (also known as rhizome rot) disease. Infected banana plants showed damage of rhizome.	Jeger et al., 1995

2.2 Abiotic stresses

In physical terms, stress is defined as mechanical force per unit area applied to an object. Whereas, in biological terms, stress is defined as an adverse force or a condition which inhibits the normal functioning and well-being of a biological system such as a plant (in, Mahajan & Tuteja, 2005). Various abiotic stresses have been described to pose adverse effects in plant growth and productivity: cold (chilling and frost), heat (high temperature), salinity (salt deposition), drought (water deficit), flooding (excess water), radiation (high intensity of ultra-violet and/or visible light), chemicals and pollutants (heavy metals, pesticides and aerosols), oxidative stress (reactive oxygen species, ozone and hypoxia), wind (sand and dust particles in the wind, and wind force) and nutrient stress (deprivation or excessive nutrients in soil) (Mahajan & Tuteja, 2005). Abiotic stresses cause agricultural and economical losses of over hundreds of million dollars yearly due to reduction in crop productivity, which in turn threatens sustainable agriculture and food security (Ciais et al., 2005; Mahajan & Tuteja,

2005). A series of stress-responsive mechanisms are activated upon stress in plants for stress adaptation, which is important for re-establishment of homeostasis and protection and repair of damaged proteins and membranes (Figure 2.4) (reviewed in Wang et al., 2003; Vinocur & Altman, 2005).

2.2.1 Salinity stress and its effects on plants

Salinity is a major environment stress that restricts plant growth and productivity. Salinity, which is defined as a soil condition with a high concentration of dissolved mineral salts (Munns & Tester, 2008), is caused by low rainfall, high evaporation, weathering of native rocks, rising water tables, use of saline irrigation water (marginal irrigation water) and poor water management (Ashraf & Foolad, 2007; Shapira et al., 2009). This phenomenon is profound in the arid and semiarid regions with high evaporation rates (Dasgan et al. 2002). Salinity affects 6-7% of the current world's arable land and has been estimated to increase to 30% by 2025 and 50% by the year 2050 due to increased accumulation of salt delivered along with irrigation water and high evapotranspiration rate caused by global climate change (Wang et al., 2003; Mahajan & Tuteja, 2005; Munns & Tester, 2008).

Salt deposition in arable lands causes decreased soil porosity, reduced soil aeration and water conductance, and generates a low water potential zone from which it is difficult for plants to acquire water and nutrients (Mahajan & Tuteja, 2005). The salts, once taken up and accumulated in the apoplast, may dehydrate the cells (Munns & Tester, 2008). Besides imposing water-deficit or hyperosmotic stress (a form of physiological drought), salinity also results in secondary effects, such as oxidative stress and Na^+ toxicity (ionic stress) on plants. Cytosolic enzymes are sensitive to salt in both glycophyes (salt-sensitive plants) and halophytes (salt-tolerant plants). High sodium ion levels in plant cells result in reduced photosynthesis and accumulation of reactive

oxygen species (ROS) as the key biochemical processes in the plant cells are inhibited by the competition by sodium for potassium-binding sites (Apse & Blumwald, 2002; Munns & Tester, 2008).

Salinity-stressed plants show reduced new shoot growth due to osmotic stress, and increased senescence of older leaves and appearance of chlorotic and necrotic spots on leaves due to Na^+ toxicity or ionic stress (Tester & Davenport, 2003; Munns & Tester, 2008; Shapira et al., 2009). Salt tolerance thresholds vary between plants. Most of the economically important crops, such as rice, maize, soybean and common beans, are considered sensitive to soil salinity (Mahajan & Tuteja, 2005). Soil salinity results in agricultural loss and yield reduction, and poses crisis to global food supply (Munns & Tester, 2008).

2.2.2 Salinity stress response and regulation

Recent advancement in plant molecular biology has greatly accelerated the progress in understanding abiotic stress responses, and many genes associated with stress adaptation have been discovered (Hirayama & Shinozaki, 2010; Mizoi & Yamaguchi-Shinozaki, 2013). These genes include regulatory genes, such as, transcription factors and protein kinases; structural genes, such as, enzymes for generating protective metabolites and proteins; and quantitative trait loci (QTLs) that are associated with higher stress tolerance (Mizoi & Yamaguchi-Shinozaki, 2013). The understanding of the molecular mechanisms in stress response is important for engineering stress tolerance in plants by enhancing or introducing stress tolerance mechanisms using molecular techniques (Mizoi & Yamaguchi-Shinozaki, 2013; Nakashima et al., 2014). With the advent and advancement of sequencing technologies, investigation of stress responses in plants other than Arabidopsis is now feasible, where

reference gene/genome sequences can be produced through genome sequencing, re-sequencing and whole transcriptome sequencing in shorter time and at lower cost.

Although the knowledge of some plant stress tolerance mechanisms in nature has been applied to improve abiotic stress tolerance in crops through gene transfer and marker-assisted breeding, the elucidation of genetic variation-associated gene functions and interactions between stress response mechanisms remains a challenge in understanding plant abiotic stress response and tolerance, in addition to effective application of this knowledge (Hirayama & Shinozaki, 2010; Nakashima, et al., 2014; Qi et al., 2011).

2.2.2.1 Signal perception and transduction

Stress is first perceived by the receptors present on the membrane of plant cells. The signal is then transduced downstream and results in the generation of second messengers (Figure 2.4) (Wang et al., 2003; Mahajan & Tuteja, 2005). Abscisic acid (ABA) is a well-characterized cellular signal important for root-to-shoot long distance signaling in plants. Besides, an ABA-independent signaling molecule, cytosolic Ca^{2+}, has also been characterized in plants. The accumulation of this cytosolic Ca^{2+} has been reported to be induced by increased Na^{+} level in *Arabidopsis* roots. This signaling molecule is perceived by a Ca^{2+} censor, calcineurin B-like protein (CBL4), which is also known as salt overly sensitive 3 (SOS3) and the signal is then transduced to the downstream components (reviewed in Mahajan & Tuteja, 2005; Munns & Tester, 2008). Other molecules responsible for stress-responsive signal perception and transduction in plants include G protein-coupled receptors (GPCRs), mitogen-activated protein kinases (MAPK), calcium dependent protein kinases (CDPKs), sucrose non-fermentation 1-related kinases (SnRKs) and lectin receptor-like kinases (LecRLKs) (reviewed in Hirayama & Shinozaki, 2010; Turan et al., 2012).

2.2.2.2 Transcriptional regulation

Transcription factors are the early induced genes, expressed within minutes of stress signal perception and responsible for activating an array of downstream stress-responsive proteins (Figure 2.4) (Mahajan & Tuteja, 2005). These transcription factors include dehydration-responsive transcription factors (DREB), C-repeat binding factors (CBF), heat shock factors (HSFs), ABF, MYB, MYC, AP2/ERF, NAC, WRKY and bZIP (reviewed in Wang et al., 2003; Turan et al., 2012).

Helicases, which are involved in gene regulation at various developmental stages as well as in stress conditions, also play important roles in salinity stress response. Among the reported stress-related helicases are pea DNA helicase 45 (PDH45), pea DNA helicase 47 (PDH47), pea DNA helicase MCM6, *Arabidopsis* DEAD box RNA helicase (LOS4) and *Medicago sativa* DEAD box helicase (MH1). Alternative splicing, which enables the production of different polypeptides (from spliced isoforms) from a gene, is also regulated by abiotic stress (Mahajan & Tuteja, 2005; Hirayama & Shinozaki, 2010; Turan et al., 2012). The roles of small RNA-dependent and epigenetic gene regulation in response to abiotic stress have also been described (reviewed in Hirayama & Shinozaki, 2010). The most established example is probably a nat-siRNA produced from overlapping mRNAs between Δ^1-pyrroline-5-carboxylate dehydrogenase (P5CDH) and SIMILAR TO RCD ONE 5 (SRO5) that regulates P5CDH and leads to accumulation of proline (an osmoprotectant) under salt stress (Borsani et al., 2005).

2.2.2.3 Ion homeostasis by ion channels and transporters

One of the mechanisms to maintain ion homeostasis in plant cells is through the activity of ion transporters (Figure 2.4). These ion transporters maintain a high K^+ to Na^+ ratio by reducing Na^+ entry into the cells, excluding Na^+ from the cells, and compartmentalizing Na^+ in the vacuoles (Apse & Blumwald, 2002).

Na^+/H^+ antiporters permit the exchange of Na^+ and H^+ across plasma membranes and regulate cytoplasmic pH, sodium levels and cell turgor. NHX1 and SOS1 are among genes encoding for Na^+/H^+ antiporters that have been isolated from plants (reviewed in Wang et al., 2003). Na^+ enters roots passively through the nonselective cation channels and other possible Na^+ transporters, Most of the Na^+ that enters root cells is pumped out from the cells by Na^+/H^+ antiporters (encoded by SOS1) in the plasma membrane. Whereas vacuolar Na^+/H^+ antiporter encoded by a Na^+/H^+ exchanger (NHX) family gene compartmentalizes excess Na^+ into vacuoles (reviewed in Mahajan & Tuteja, 2005; Munns & Tester, 2008). The functionality of the vascular Na^+/H^+ antiporters requires an electrochemical H^+ (proton) gradient to drive the active transfer of Na^+ into vacuole. This electrochemical gradient is generated by H^+-ATPase and H^+-pyrophosphatase (encoded by AVP1) (reviewed in Apse & Blumwald, 2002; Turan et al., 2012).

Beside SOS1 and NHX, HKT and CAX1 are also genes encoding important transporters for regulating salt stress responses (reviewed in Mahajan & Tuteja, 2005). High-affinity K^+ transporter (HKT) encodes for a low affinity Na^+ transporter that controls Na^+ entry into the root cells during salt stress, while CAX1 encodes for a vacuolar (tonoplast) Ca^{2+}/H^+ antiporter that regulates Ca^{2+} homeostasis (reviewed in Mahajan & Tuteja, 2005).

2.2.2.4 Effector proteins and metabolites

Beside the early induced genes as mentioned above, plants also express late induced genes, usually after a few hours of stress perception. These include the major stress responsive genes that encode and modulate proteins needed for synthesis of antioxidants, osmolytes, membrane stabilizing proteins and synthesis of LEA-like proteins (Figure 2.4) (reviewed in Mahajan & Tuteja, 2005). During stress, reactive oxygen species (ROS) are produced, including superoxide radicals (O_2-), hydrogen

peroxide (H_2O_2) and hydroxyl radicals (OH·), which can cause oxidative damage to cellular components such as membrane lipids, proteins and nucleic acids.

Plants deploy antioxidant protection strategies against ROS by producing a diverse array of enzymes such as superoxide dismutase (SOD), catalases (CAT), ascorbate peroxidases (APX), glutathione S-transferase (GST), aldehyde dehydrogenase (ALDH) and glutathione peroxidases (GPX), and non-enzyme molecules such as ascorbic acids, reduced glutathione, carotenoids, anthocyanins, peroxiredoxin and tocopherol (reviewed in Apse & Blumwald, 2002; Wang et al., 2003; Vinocur & Altman, 2005; Turan et al., 2012).

Plants also respond to salt stress by accumulating compatible solutes (osmolytes or osmo-protectants) to increase osmotic pressure inside the cells and to prevent intracellular water loss, which in turn, stabilizes proteins and cellular structures. The compatible solutes include metabolites with osmolyte function such as sugars (fructose and sucrose), sugar alcohols (mannitol and sorbitol) and complex sugars (trehalose and fructans), charged metabolites (proline, glycine betaine and ectoine), and polyamines (putrescine, spermidine and spermine) (reviewed in Apse & Blumwald, 2002; Mahajan & Tuteja, 2005; Vinocur & Altman, 2005; Turan et al., 2012). Among the well-studied biosynthesis pathways for compatible solutes are the synthesis of proline from glutamate by Δ^1-pyrroline-5-carboxylate synthase (P5CS) and Δ^1-pyrroline-5-carboxylate reductase (P5CR), and the synthesis of glycine betaine from choline by choline monooxygenase (COM) and betaine aldehyde dehydrogenase (BADH) (reviewed in Wang et al., 2003).

Molecular chaperones are a group of proteins involved in various cellular functions such as folding/unfolding, macromolecular assembly/disassembly, keeping proteins in their native state and preventing their aggregation under stress conditions, helping in protein synthesis/degradation and targeting to their cellular compartments

(Turan et al., 2012). Heat shock proteins (Hsps) and late embryogenesis abundant-type (LEA-type) proteins are two major types of stress-induced proteins that control the proper folding and conformation of proteins, thus protecting cell membranes and enzymes from denaturation by stress (reviewed in Vonocur & Altman, 2005).

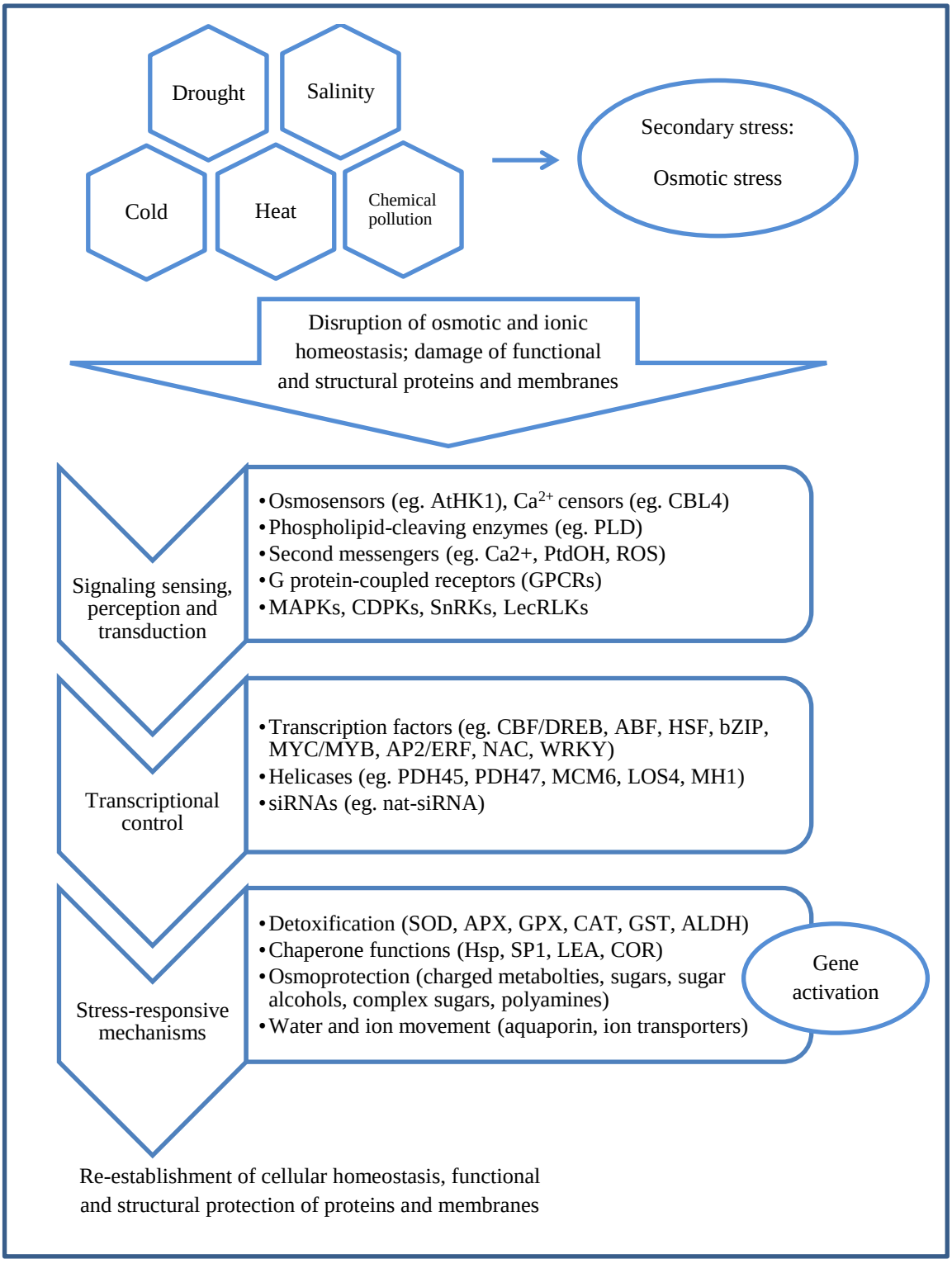


Figure 2.4: Abiotic stress responses in plants. Abiotic stresses trigger signaling processes and transcriptional controls which activate stress-responsive mechanisms to re-establish cellular homeostasis and confer protection to proteins and membranes (adapted from Vinocur & Altman, 2005)

2.3 Banana genetics and genomes

Various genetic and cytogenetic investigations have been carried out to understand the genome constituent, gene structure and gene function in bananas and plantains (Roux et al., 2008). Whole genome sequencing of a wild banana (*Musa acuminata*) has revealed genome structure and distribution of protein-coding genes, non-coding sequences, regulatory elements, transposable elements and repetitive DNA and so on in the banana genome (D'Hont, et al., 2012). Sets of genes or transcripts being expressed and levels of expression in particular growth and development stages of banana plants, and response to different environmental factors can be elucidated through transcriptomic studies. The orchestrated expression of genes is modulated by endogenous regulators, for instance, microRNAs and short interfering RNAs. Thus genetic and genomic information is important in formulating crop genetic improvement programmes to produce banana plants with desirable agronomic characteristics, for instance, resistance to biotic and abiotic stresses, and improved yield.

The advancement and progress made in *Musa* genomics, transcriptomics and genetics have been contributed by many *Musa* research groups, organizations and collaboration networks across the world, especially through international collaborations and sharing of resources.

Bioversity International is one of the most important players in research and development for *Musa*. This organization aims to apply agricultural biodiversity for sustainable agricultural development and global food and nutrition security through delivery of scientific evidence, management practices and policy options to use and safeguard agricultural biodiversity (http://www.bioversityinternational.org/). It currently coordinates three global networks for banana research, ProMusa, MusaNet and the Global Musa Genomics Consortium (GMGC).

The Global Musa Genomics Consortium is an international network that brings

together expertise from various institutions in different countries to apply genomics tools in banana research for improvements in breeding and management of banana. This consortium coordinates or promotes the access and use of a number of *Musa* genome resources such as, BAC, cDNA and gDNA libraries, and bioinformatics databases such as, genetic markers, genetic maps, physical map, and BAC, EST, GSS and small RNA sequences (http://www.musagenomics.org/) for genetic and genomic research in banana.

Whereas ProMusa is a knowledge-sharing platform on bananas that facilitates the exchange of information, knowledge and know-how on bananas by organizing symposia and managing integrated knowledge-sharing and communication tools (http://www.promusa.org/tiki-custom_home.php). ProMusa promotes dissemination of knowledge and exchange of information through InfoMus@ newsletter, Musapedia and three other databases, namely Musalit, Musarama and Musacontacts.

The first complete draft genome sequence of *Musa acuminata* is a result of collaboration between Genoscope (Centre National de Séquençage) and CIRAD (French Agricultural Research Centre for International Development), funded by ANR (The French National Research Agency). The sequence was analysed in collaboration with several teams in particular of the Global Musa Genomics Consortium (D'Hont et al., 2012). The *Musa* reference genome sequences, including re-sequencing project (Davey et al., 2013), are made publically accessible through the Banana Genome Hub (http://banana-genome.cirad.fr/). The Banana Genome Hub centralises databases of genetic and genomic data for the Musa species. This hub is developed by CIRAD and Bioversity International and supported by the South Green Bioinformatics platform. Data available are the complete genome sequence along with gene structure, gene product information, metabolism, gene families, transcriptomics (ESTs, RNA-Seq), genetic markers (SSR, DArT, SNPs) and genetic maps (http://banana-genome.cirad.fr/).

2.3.1 Banana genomes

Musa acuminata has a genome with 11 chromosomes in the haploid stage with a size of about 600Mb (reviewed in Heslop-Harrison & Schwarzacher, 2007). Recently, the sequenced genome of *Musa acuminata* ssp. *malaccensis* (accession DH-Pahang) showed that the double-haploid genotype of this wild subspecies has a genome value 1C = 523Mb estimated by flow cytometry (D'Hont, et al., 2012).

Before an annotated banana genome sequence was made available, various research had been carried out to reveal structures and components of the banana genomes. Faure et al. (1994) carried out an investigation on diploid crosses of *M. acuminata* subspecies and showed that wild bananas have a strong bias towards maternal transmission of chloroplast DNA and paternal transmission of mitochondrial DNA. Besides, the banana genome was estimated to contain 30% of repetitive DNA with Ty1/copia and Ty3/gypsy retroelements representing 16 and 7% of the genome respectively, while DNA transposons are rare (Hribova et al., 2010). Banana genomes were also found to contain viral sequences of an integrated pararetrovirus, the banana streak badnavirus virus (BSV). The integrated virus sequences are not only found in *Musa balbisiana* but also the cultivated bananas (*Musa acuminata*) containing B genome constituents (Harper et al., 1999; Safar et al., 2004).

A few bacterial artificial chromosome (BAC) libraries were also constructed and sequenced to survey the genome of *Musa* spp. (Vilarinhos et al., 2003; Aert et al., 2004; Safar, et al., 2004; Cheung & Town, 2007; Lescot et al., 2008). Cheung & Town (2007) reported the sequencing of 6,252 BAC end-sequences of banana and revealed the distribution of mitochondrial and chloroplastic sequences, transposons, repeat sequences, proteins and simple sequence repeats in banana genome. Through the survey of a sequenced BAC clone, Aert et al. (2004) found that banana genes are clustered in gene-rich regions separated by gene-poor DNA containing abundant transposons.

Meanwhile Safar et al. (2004) reported the first sequencing of BAC libraries for the banana B genome from *Musa balbisiana* and showed that *M. balbisiana* shares very high genome conservation with *M. acuminata*.

Recent whole genome sequencing of banana (*Musa acuminata* ssp. *malaccensis*) Double-haploid Pahang (DH-Pahang) has been reported by D'Hont et al. (2012). This work provides a higher resolution view of banana genome structure. In that study 90% of the genome was sequenced and 92% of the 36,542 predicted protein-coding genes could be positioned on the 11 chromosomes (Figure 2.5). Besides, 235 microRNA genes from 37 families were also identified in the sequenced banana genome. This DH-Pahang with AA genome (*M. acuminata*) was also reported to have integrated banana streak virus (BSV) which was previously only demonstrated in bananas with a B genome constituent. This BSV belonged to a different badnavirus phylogenetic group from endogenous BSV (eBSV) in *M. balbisiana*. Unlike *M. balbisiana*, the integrated BSV viral sequences in DH-Pahang are highly organized, fragmented and are thought unable to form free infectious particles. Besides, nearly half of the *M. acuminata* genome was found to be composed of transposable elements, with long terminal repeat retrotransposons (*Copia* and *Gypsy* elements) and long interspersed elements (LINEs) being the most abundant (Figure 2.5) (D'Hont, et al., 2012). Recently, draft genome sequences for *Musa balbisiana* and *Ensete ventricosum* have been produced by re-sequencing (Davey et al., 2013; Harrison et al., 2014). Davey et al. (2013) reported a *Musa balbisiana* ('Pisang Klutuk Wulug', diploid B-genome) draft genome that is 79% the size of the A-genome. This B-genome contains 36,638 predicted functional gene sequences and number of miRNA genes which are nearly identical to those in the A-genome, as well as additional novel miRNAs, detected in both the A- and B-genomes. The authors also demonstrated a 2:1 distribution of transcriptomic reads across the A- and B-genomes in AAB banana hybrid. The AAA banana hybrid contains regions of

significant homology to the B-genome, suggesting the historical interspecific recombination events between homologous A and B chromosomes in *Musa* hybrids. Harrison et al. (2014) reported a 547 megabases draft genome of ensete, which is a similar size to the A-genome (523 megabases). At least 1.8% of the annotated functional genes in A-genome are not conserved in ensete and ensete contains genes that are not present in banana. These reference genome sequences are important genetic resources for marker-assisted breeding and genetic improvement programmes (Davey et al., 2013; Harrison et al., 2014).

The Global *Musa* Genomics Consortium (2002) described bananas as a good monocotyledonous plant model for genomics beyond rice. Banana has a relatively small haploid genome size, which is only 25% larger than rice. It is different from many plant species with its bi-parental cytoplasmic inheritance, which is paternal inheritance of mitochondria but maternal inheritance of chloroplasts. Bananas are vegetatively propagated, thus genomically static in its centre of origin in Southeast Asia. This makes bananas a good model to study somaclonal variation and evolution in plants and plant-pathogen co-evolution. In addition, the variable ploidy levels in *Musa* spp. enable researchers to study the relationship of polyploid formation to phenotype changes and the causes and consequences of polyploidization to genome organization (The Global *Musa* Genomics Consortium, 2002; INIBAP, 2006).

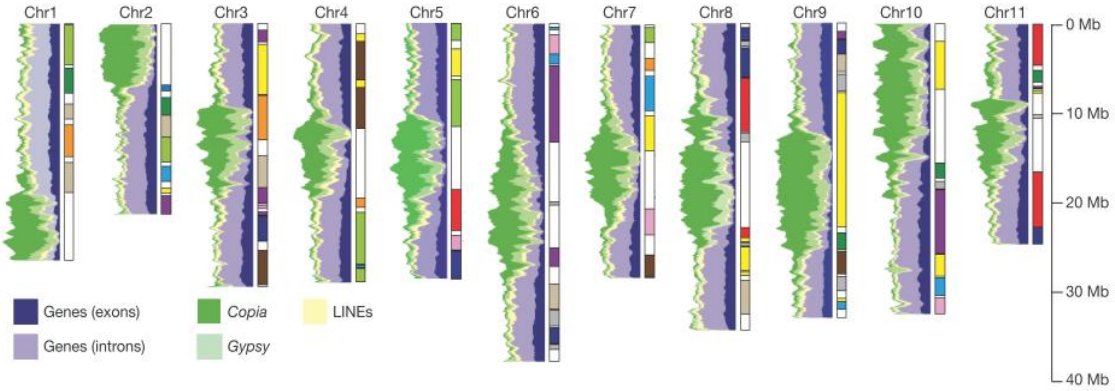


Figure 2.5: Banana genome and its components (D'Hont, et al., 2012). Banana has 11 chromosomes in haploid genome. Genes (exons and introns) and transposable elements (*Copia*, *Gypsy* & LINEs) are distributed differently in each chromosome (left). Twelve *Musa* α/β ancestral blocks represented by different colours (right)

2.4 Gene expression

A transcriptome is the complete set of transcripts in a cell or a population of cells, and their quantity, in a specific developmental stage and a defined physiological condition. The transcriptome consists of protein-coding mRNAs and non-coding small RNAs, such as, rRNA, tRNA, snoRNA, siRNA and miRNA (Wang et al., 2009b; Wang et al., 2010b). A mammalian transcriptome contains ~80% rRNA, ~15% tRNA, 2-4% mRNA and about 1% of intragenic and intergenic non-coding RNA sequences with regulatory functions (Lindberg & Lundeberg, 2010). Transcriptomic study allows the determination of gene expression profiles, transcriptional structure of genes (including transcriptional start sites and polyadenylation signals), splicing patterns and other post-transcriptional modifications (Wang et al., 2009b; Wang et al., 2010b). With the advent of high-throughput next generation sequencing technologies, it is now possible to obtain high resolution expression and structural profiles of RNA populations with lower cost and shorter time than using earlier methodology (Wang et al., 2010b).

2.4.1 Approaches for high-throughput transcriptomic studies

Transcriptomic approaches have been widely used to catalogue species of transcripts that are expressed in cells, tissues or organs of plant species grown at different stages or conditions and to quantify the expression level of the transcripts. Sanger-based sequencing (capillary sequencing) is one of the earliest automated-sequencing platforms used for transcripts discovery. This technology allowed sequencing of cDNA libraries (Expressed Sequence Tags, ESTs) and produced representative transcripts without the need of a sequenced genome. However, Sanger-based sequencing technology requires fragment-cloning steps and is relatively low-throughput, high cost and the sequencing data generated is non-quantitative (Wang et al., 2010b; Ozsolak & Milos, 2011).

Hybridization-based microarray technologies are among the initial high-throughput approaches to explore transcriptomes. Microarray involves the use of immobilized probes with sequence complementary to the fluorescently labelled DNA or RNA target molecules for high stringency probe-target hybridization to happen prior to detection of fluorescent signals. However, the application of microarray is limited to species with a sequenced genome or well-characterized transcriptome. The application of this approach is also restricted by the dynamic range of the probes/dyes and scanning instruments (resulting in a limited range for detection and quantification) and non-specific cross-hybridisation (resulting in high background signal) (Wang et al., 2009b; Costa et al., 2010; Wang et al., 2010b; Ozsolak & Milos, 2011). Besides various platforms for microarray, Serial Analysis of Gene Expression (SAGE), which is a tag-based sequencing approach, was also used for quantitative transcriptomic studies though it is less popular than microarray approach (Costa, et al., 2010). Unlike microarray platforms, a reference genome or prior knowledge of gene sequences are not required for SAGE (Harbers & Carninci, 2005). Short sequence tags (14-, 21- or 26-bp

depending on the restriction enzyme) unique to individual transcripts are generated and concatemerised into long DNA molecules and cloned prior to sequencing. Sequencing of these concatemers can reveal the sequence of the transcripts and allow the measurement of their absolute abundance (Powell, 2000; Velculescu et al., 2000; Vega-Sánchez et al., 2007).

One of the breakthroughs in DNA sequencing technology was recorded in 2004 when a massively parallel sequencing platform, Roche/454 Genome Sequencer (GC), was introduced. This 454 pyrosequencing platform can sequence several hundred thousand DNA fragments in parallel, with read length greater than 100 bp (Costa, et al., 2010). This was then followed by another sequencing-by-synthesis method, Illumina Genome Analyzer (GA), which was introduced in 2006, with the capacity of generating tens of millions of 32 bp reads. The next generation sequencing technologies continue to improve. Currently the GS FLX Titanium has the capacity to generate 1 million reads of 400bp; Illumina GAIIx generates 200 million 75-100 bp reads; Applied Biosystem Inc's platform, Sequencing by Oligo Ligation (SOLiD) could produce 400 million of 50 bp reads; Helico BioScience's Heliscope, a single molecule sequencer, produces 400 million of 25-35 bp reads (Costa, et al., 2010; Wang et al., 2010b). The use of these next generation sequencing (NGS) technologies for transcriptome sequencing, termed RNA-Seq (RNA sequencing), has the advantage of identifying, characterizing and cataloging, virtually, all of the transcripts expressed in specific cells/tissues and/or physiological conditions. The NGS technologies also enable the quantification of differential expression, capture of novel transcripts, discovery of novel 5' and 3' UTRs, determination of different splicing patterns and new splice isoforms and discovery of expressed Single Nucleotide Polymorphisms (SNPs) (Wang et al., 2009b; Costa, et al., 2010).

Next generation sequencing (NGS), also called non-Sanger-based high-

throughput sequencing, deep sequencing or ultra-high-throughput sequencing, relies on the generation of thousands and millions of 25-400bp short sequence reads in parallel (Wang et al., 2010b; Ozsolak & Milos, 2011). Intensive computational analyses involving various bioinformatics tools and algorithms are required for sequence alignment, assembly/clustering, mapping, normalization, expression measurement and functional annotation of the transcriptomic data produced by NGS (Egan et al., 2012). The enormous amount of short reads generated from RNA-Seq using the NGS platforms need to be assembled computationally to reconstruct the target RNA sequences (Figure 2.6) (Wang et al., 2009b; Martin & Wang, 2011). RNA-Seq reads will first be filtered for adaptor sequences and low quality reads before being used for downstream applications. FastQC is one of the tools used to perform filtering and produce high quality clean reads (Van Verk et al., 2013). Approaches for assembling the short reads to reconstruct transcripts are either *de novo* assembly, reference-based (also known as mapping-based or '*ab initio*') assembly or a combination of both approaches (Martin & Wang, 2011). Reference-based assembly of transcriptomes requires a good quality reference genome and reconstruction of short reads into transcripts is done by aligning the short reads to the reference genome and clustering of overlapping reads from each locus. In contrast, *de novo* assembly is performed for species that lack a sequenced or high-quality finished reference genome (Martin & Wang, 2011). Among the established and commonly used short read assemblers are SSAKE, SHARCGS and VCAKE (classified as Greedy Graph-based assemblers), Newbler and Celera Assembler (classified as Overlap/Layout/Consensus, OLC assemblers), Velvet, ABySS, ALLPATHS and SOAP*denovo* (classified as de Bruijn Graph approach) and Trinity (Greedy/deBruijn Graph-based assembler) while Scripture, Cufflinks, TopHat, Bowtie and BWA are among the publicly available software for mapping-based assembly of transcriptome using a reference genome (Miller et al., 2010; Wang et al., 2010b; Martin

& Wang, 2011; Ward et al., 2012; Van Verk, et al., 2013).

Due to its high sequencing depth, RNA-Seq is also a quantitative method with which quantifying and comparing the abundance of transcripts is possible. However, the RNA-Seq data needs to be normalized before any comparison can be done between different genes within the same sample and across samples (Costa, et al., 2010). Reads (or paired-end fragments) per kilobase (of exon model) per million mapped reads (RPKM/FPKM) is a widely used method for normalization of transcriptomes (Mortazavi et al., 2008). This method adjusts raw counts to the total gene length (per kilobase) and number of reads mapped (per million) within a sample to allow within- and cross-sample expression comparison of RNA-Seq data (Van Verk, et al., 2013). Other normalization methods, such as median count ratio (Med) and trimmed mean of M-values (TMM) have also been developed to perform RNA-Seq data normalization and each of these methods has its advantages (Van Verk, et al., 2013). Statistical tests are required to determine differentially expressed transcripts in RNA-Seq and among the commonly used statistical methods are DESeq, edgeR and baySeq. Significant ranking of genes is conducted with controlled false discovery rate (FDR) in these methods (Egan, et al., 2012; Van Verk, et al., 2013).

Assigning identity and function(s) to expressed RNA (transcripts) is an integral part of transcriptome study. Various reference gene and protein sequence databases are publically available for use in functional annotation of transcriptomes. For instance, NCBI non-redundant protein database (nr), UniProt, Gene Ontology (GO) database (Harris et al., 2004), Eukaryotic Orthologous Group (KOG) database (Tatusov et al., 2003), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database (Kanehisa, 2002). Several computational tools are available to carry out BLAST search of large number of transcriptomic data in batch against the sequence databases. Blast2GO is one of the popular tools for functional annotation of sequences and data

mining of the resulting annotations. This tool supports a variety of annotation, such as GenBank BLAST search, InterProScan and assigning KEGG pathways and gene ontology terms (Conesa & Gotz, 2008; Strickler et al., 2012). Functional annotation categorizes genes into functional groups with controlled vocabularies, which can facilitate the exchange of biological knowledge (Conesa & Gotz, 2008). The functional categorization of genes can assist with understanding the physiological meaning of large amount of genes and to assess the differences among the functional groups (Conesa et al., 2005).

The next generation sequencing (NGS) technologies are one of the most important tools used in advancing plant genomic research. The application of the NGS in transcriptomic studies can provide a deep insight into various aspects of life in plants, such as growth and development and stress responses and assist in deciphering the complex plant transcriptomes through the discovery of the functional elements and the developmental or environmental networks (Costa, et al., 2010; Wang et al., 2010b).

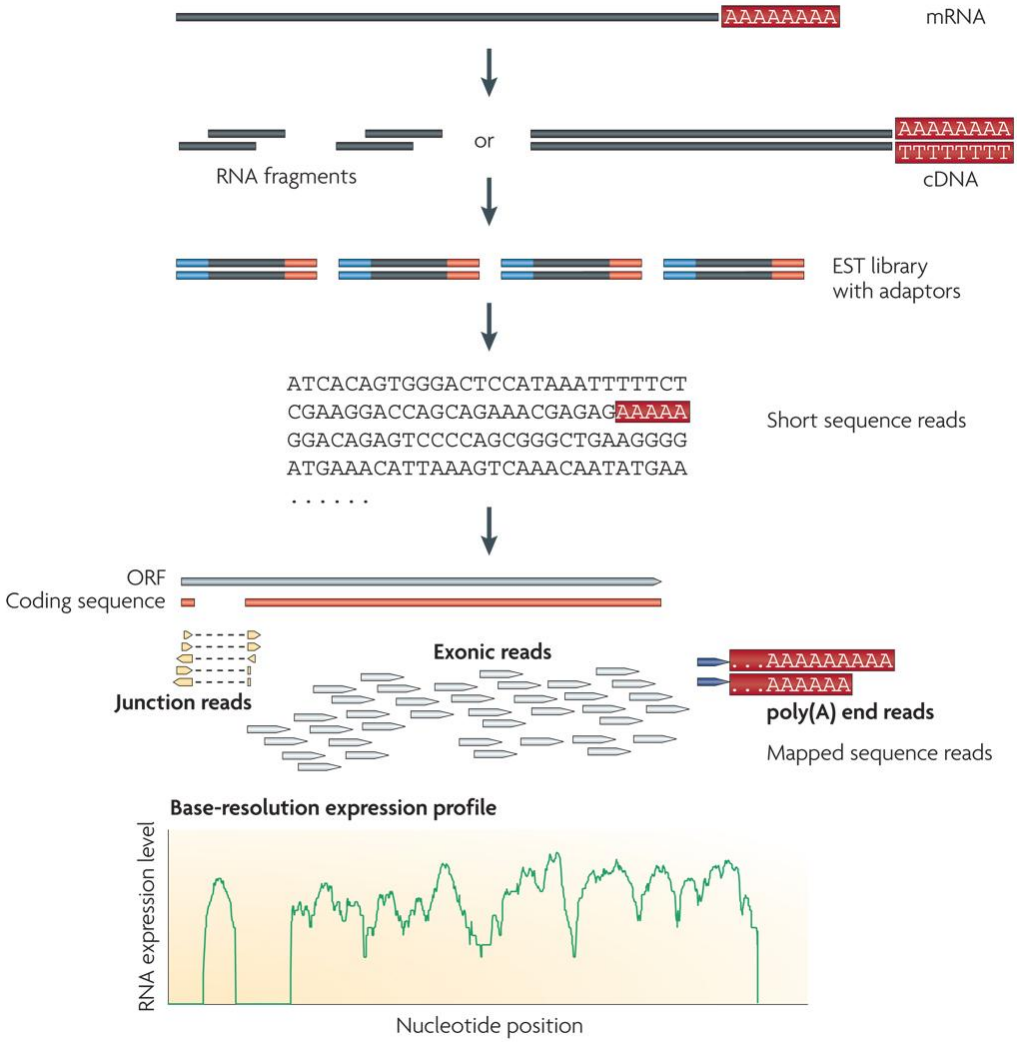


Figure 2.6: A typical RNA-Seq experiment (Wang et al., 2009b). A cDNA library is constructed from fragmented mRNA or cDNA. Sequencing of the library results in millions of short reads that can be assembled by *de novo* or map-based assembly to reconstruct the RNA transcripts from short reads. The short reads can be used for measuring RNA expression level

2.4.2 Transcriptomic studies in *Musa*

Despite being a social-economically important crop, there were only a few reports of global gene expression profiling in banana before the advent of NGS technologies and the availability of a reference *Musa* genome (Davey et al., 2009). These reports on banana transcriptomes were based on Sanger-based sequencing of

cDNA libraries generating ESTs and one report each on the use of SuperSAGE and microarray.

One of the earliest transciptome level investigations of banana was on fruit ripening (Clendennen & May, 1997). The authors reported the use of differential screening of cDNA libraries to identify 10^5 plaques containing labeled pulp cDNAs, in which 100 of the plaques had a strong signal intensity and only 38 cDNA clones were selected for sequencing after screening. Manrique-Trujillo et al. (2007) reported the use of suppression subtractive hybridization to identify differentially expressed genes in Cavendish banana fruit during the ripening process. The authors constructed and sequenced a subtraction cDNA (EST) library with 215 unigenes produced after sequence assembly. Santos et al. (2005) generated EST libraries from *Musa acuminata* ssp. *burmannicoides* Calcutta 4 leaves subjected to temperature stresses (hot and cold) and yielded total 2,880 ESTs that were assembled and clustered into 1,019 non-redundant transcripts, which included candidate abiotic stress responsive genes. Passos et al. (2012) produced 3,964 and 9,333 ESTs respectively from cDNA libraries constructed from leaf tissues of banana plants that are susceptible (*M. acuminata* subgroup Cavendish cv. Grande Nain) and resistant (*M. acuminata* ssp. *burmannicoides* Calcutta 4) to black leaf streak disease, after inoculation of *Mycosphaerella fijiensis*. These ESTs resulted in 3,995 non-redundant unigenes and 624 SSR-positive sequences after assembly. Meanwhile, SuperSAGE has been used for characterization of a young leaf transcripome from a wild diploid *Musa acuminata*, with 5,292 transcripts detected and their expression (abundance) quantified (Coemans et al., 2005). Davey et al. (2009) reported the use of microarray technology to characterize banana leaf transcriptomes during drought stress. An Affymetrix Rice Genome Array was used to profile expression of control and drought stressed leaves of *Musa* cultivar ‘Cachaco’ with 2,910 *Musa* gene homologues being drought responsive transcripts, after masking the rice sequence-

derived probes with weak or no hybridization signal with *Musa* gDNA.

High-throughput scale global transcriptomic investigation in banana began with the advent of NGS technologies. Transcriptome profiling using an Illumina sequencing platform was performed on resistant and susceptible Cavendish banana roots challenged with *Fusarium oxysporum* f. sp. cubense tropical race 4 (Foc TR4) and also Cavendish roots infected with Foc1 and Foc TR4 (Li et al., 2012a; Li et al., 2013b). Li et al. (2012a) reported the assembly of 88,161 from 103 million of 90-bp paired-end reads unigenes, with 5,008 genes assigned to plant-pathogen interactions. Li et al. (2013b) generated a transcriptome library from a mixture of RNA from various organs of banana cultivar 'Baxi' (Cavendish subgroup) with 47,411 non-redundant transcripts assembled from >26 million paired-end reads and 10,545 (~22.2%) of the transcripts can be mapped to the annotated genes in the *Musa* reference genome. The authors also used a tag-based approach (similar to SAGE) to generate libraries for Illumina sequencing, for digital gene expression analysis in Foc1- and Foc TR4-infected banana roots. A cDNA library was also constructed from root tissues of *Musa acuminata* L. AAA group, 'Brazilian', in which RNA samples of Foc TR4-infected and non-infected roots were pooled, to produce 25,158 distinct gene sequences (Wang et al., 2012b). In the same report, the authors also prepared tag libraries from banana roots 0, 2, 4 and 6 days after being inoculated with Foc TR4 to assess the transcriptional changes during infection. In another report, Passos et al. (2013) sequenced transcriptomes from Calcutta 4 and Grande Nain leaf materials uninfected and challenged with conidiospores of *Mycosphaerella fijiensis* producing 36,384 and 35,269 unigenes and 4,068 and 4,095 genic-SSR loci from Calcutta and Grande Nain respectively. Calcutta 4 and Grande Nain bananas are known for their contrasting resistance to *M. fijiensis* with the former being resistant and the latter being susceptible. These ESTs/unigenes and SSR markers derived from RNA-Seq serve as valuable resources for functional genomics and genetic

improvement in banana (Passos et al., 2013).

To date, high-throughput functional genomic data (both transcriptome and sRNAome) for salt stress or abiotic stress in banana have not been reported.

2.5 microRNA

In 1991 and 1993, microRNA (miRNA) was first observed as small RNA produced from the lin-4 transcript that had an impact on lin-14 mRNA, modulating a temporal developmental switch in *C. elegens*. These small RNAs were named as 'microRNA' (miRNA) in the year 2001 when its origin, biogenesis and mode of action were uncovered. This class of small RNA (miRNA) was described in plants in the following year (Table 2.2).

2.5.1 Biogenesis of microRNA and RNAi machinery

The miRNAs are ~22-nt short single stranded functional RNA species derived from long stem-loop precursors transcribed from endogenous genomic DNA (predominantly in intergenic regions for plants) by RNA polymerase II (Bartel, 2004; Lee et al., 2004). This non-coding RNA transcript is then capped, spliced, polyadenylated and folded into a long hairpin stem-loop precursor miRNA (pre-miRNA) (Figure 2.7). In plants, this structure is further processed by an RNase III, Dicer-like enzyme, to a shorter hairpin structure known as a primary microRNA (pri-miRNA), as illustrated in Figure 2.7 (Kurihara & Watanabe, 2004; Chen, 2005b; Liu et al., 2005; Zhang et al., 2006b). The pri-miRNA is further trimmed into a shorter double-stranded RNA (dsRNA) by the Dicer-like enzyme, forming a miRNA:miRNA* duplex which is then transported from the nucleus to the cytoplasm (Figure 2.7) (Papp et al., 2003). In the cytoplasm, the active strand of the miRNA duplex incorporates into a RNA induced silencing complex (RISC or miRISC) which guides the mature miRNA to the

complementary target mRNA and either causes degradation of target mRNA (the major mode of action in plants), translational inhibition of the target mRNA or transcriptional inhibition through deposition of repressive chromatin marks (Figure 2.7 and 2.8) (reviewed in Zhang et al., 2006c; Cuperus et al., 2011; Meng et al., 2011; Khraiwesh et al., 2012; Axtell, 2013). In plants, most well-documented miRNAs have extensive complementarity, with less than 4 imperfect paired bases, to their target mRNAs (Jones-Rhoades, 2012). The mechanism of gene regulation by miRNA requires an RNA interference (RNAi) pathway, which was commonly known as post-transcriptional gene silencing (PTGS) in plants or 'quelling' in fungi (Table 2.2).

RNA interference or RNAi is an RNA silencing process that involves a group of effector proteins incorporated with the small regulatory RNA that is derived from hairpin or longer double-stranded RNA. Classification of small RNA in plants is complex, where different small RNA biogenesis and silencing effector modules are involved. Axtell (2013) suggested the plant small RNAs to be classified into two groups: hairpin RNAs (hpRNAs) and small interfering RNAs (siRNAs). hpRNAs are grouped into 'miRNAs' and 'other hpRNAs' while siRNAs can further be classified into heterochromatic siRNAs, secondary siRNA and natural antisense siRNAs (NAT-siRNAs), depending on their biogenesis and/or function (Figure 2.8). MiRNA is distinct from other small RNAs for its endogenous origin and formation of a hairpin stem-loop RNA structure for further processing into miRNA:miRNA* duplexes (Figure 2.7). miRNAs are typically 21 nt in size though some being 20 or 22 nt, whereas long miRNAs, another class of miRNA, are 23-24 nt in size but function similar to heterochromatic siRNAs (Figure 2.8) (Axtell, 2013). In contrast, siRNA has both endogenous (intergenic regions and/or repetitive regions) and exogenous (transgenes, RNA viruses or introduced dsRNA) origins, with a dsRNA precursor rather than single-stranded RNA-derived hpRNA like miRNA precursor. For biogenesis, miRNA (20-22 nt)

and long miRNA (23-24 nt) are transcribed from genomic DNA by RNA polymerase II and they both fold into a RNA stem-loop hairpin structure. miRNA require DCL1-clade DCL for maturation and AGO1-clade AGO for function through slicing (AGO-catalyzed cleavage of targeted RNAs). Whereas long miRNA requires DCL3-clade DCL for maturation then enters into heterochromatic siRNA effector pathway and direct chromatin modifications at their target genes (Axtell, 2013), where the siRNA-bound AGO 4 direct histone modification and DNA methylation processes. Other important proteins that are involved in miRNA biogenesis and RNA silencing in plants include: a double-stranded RNA binding protein named Hyponastic Laves 1 (HYL1) and an ethylene zinc finger protein Serrate (SE) that assist DCL1 in cleaving pri-miRNA in the nucleus; sRNA-specific methyltransferase Hua Enhancer 1 (HEN1) that methylates miRNA:miRNA* duplexes, which is thought to protect sRNAs from uridylation and adenylation and subsequent degradation; an Exportin-5 homolog named Hasty (HST) that exports miRNA:miRNA* duplex from the nucleus to the cytoplasm (Figure 2.7) (reviewed in Eamens et al., 2008; Meng et al., 2011; Khraiwesh et al., 2012).

miRNA-mediated gene silencing is an ancient evolutionary mechanism for global gene expression regulation. The last common ancestor of plant miRNAs can be traced back to bryophytes (the most ancient land plants) and seed plants and there is no evidence yet that plant and animal miRNAs have a common ancestor (reviewed in Zhang et al., 2006b; Axtell, 2008; Cuperus et al., 2011). One of the proposed mechanisms of how miRNA emerged is the duplication of protein-coding sequences, in which miRNA genes originated from their target genes through the formation of inverted duplications (reviewed in Zhang et al., 2006b; Cuperus et al., 2011; Jones-Rhoades, 2012).

Table 2.2: The history of microRNA discovery

Year	Event	References
1991	Description of possible negative regulatory sequences on 3'UTR of lin-14 mRNA modulating temporal developmental switch in *C. elegans.*	Arasu et al., 1991; Lee et al., 1993
1993	First descriptions of small RNA-mediated posttranscriptional gene regulation. Small RNAs from lin-4 transcripts in *C. elegans* complementary to 3'UTR of lin-14 mRNA down-regulate its translation via an antisense RNA-RNA interaction.	Lee, et al., 1993; Wightman et al., 1993
1996	First observation of RNA silencing phenomenon in plants (as 'posttranscriptional gene silencing') and fungi (as 'quelling').	Baulcombe, 1996; Cogoni et al., 1996; English et al., 1996
1998	First description of mechanism of dsRNA triggering RNAi gene silencing in *C. elegans.*	Fire et al., 1998
2000	Description of 21-nucleotide RNA from let-7 mRNA regulating transition timing in development in *C.elegans* and this small RNA is conserved in animals.	Pasquinelli et al., 2000; Reinhart et al., 2000
2001	Description of the processing machinery that guides RNAs and mediates RNAi in *C.elegans*.	Grishok et al., 2001; Tabara et al., 2002
	Introduction to the concept of microRNA. Recognition of the lin-4 and let-7 (also known as small temporal RNA, stRNA), and other small expressed transcripts as microRNA in *C.elegans*. Some of these microRNAs showed evolutionary conservation in animals.	Lagos-Quintana et al., 2001; Lau, et al., 2001; Lee & Ambros, 2001
2002	First description of endogenous and silencing-associated small RNAs resembled animal microRNA in plants.	Llave et al., 2002a
	First descriptions of microRNA in plants.	Llave et al., 2002b; Reinhart et al., 2002

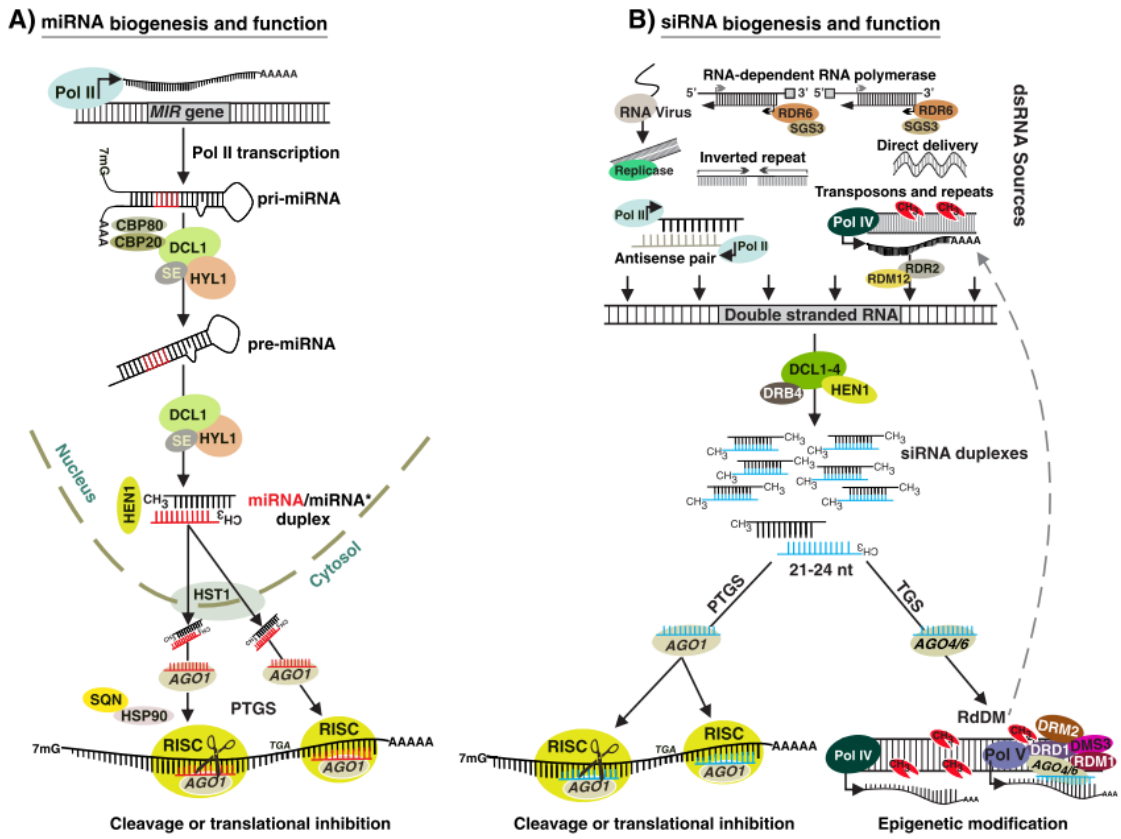


Figure 2.7: Biogenesis and gene silencing machinery of miRNA and siRNA (Khraiwesh et al., 2012). The difference and similarity in biogenesis and function between miRNA and siRNA demonstrated in plants

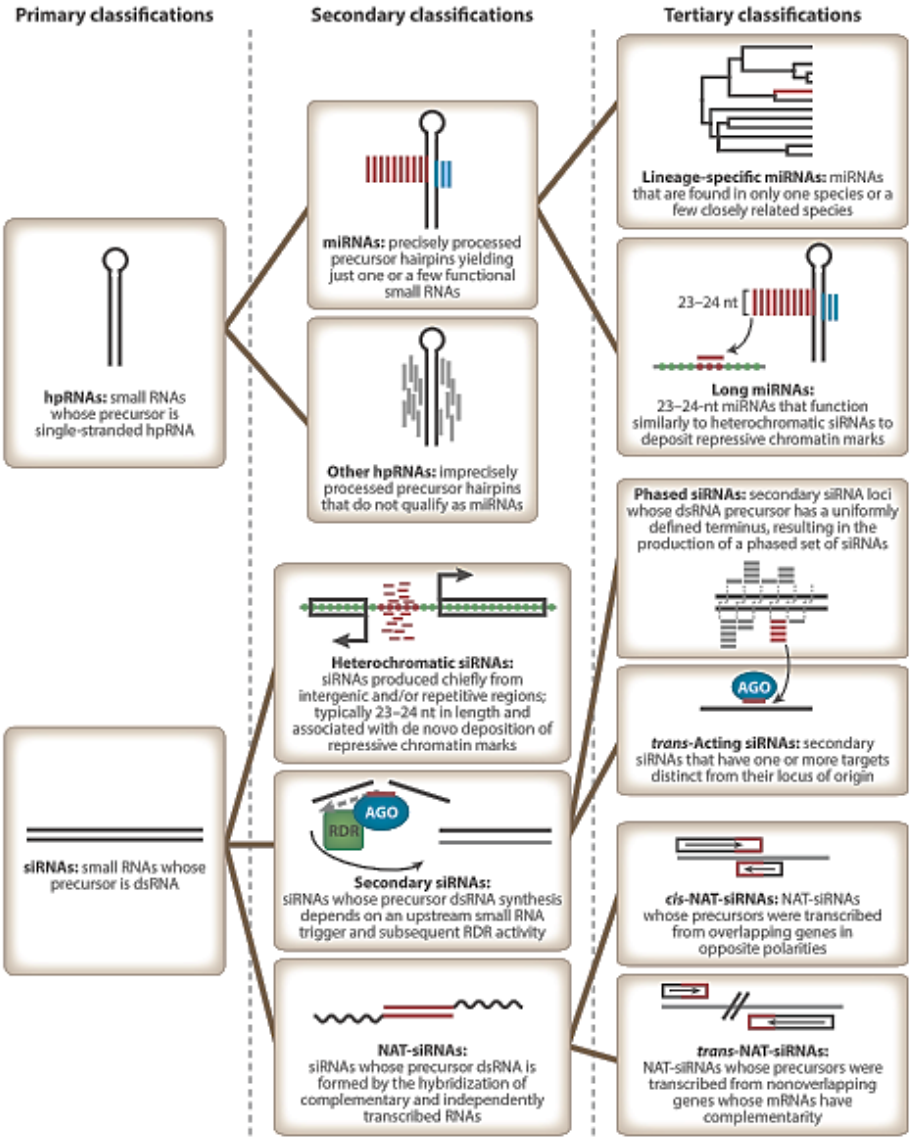


Figure 2.8: Types of endogenous small RNAs in plants (Axtell, 2013). MiRNA is classified as a hairpin RNA (hpRNA), which lineage-specific miRNA and long miRNA are also included in the hpRNA group. As opposed to hpRNA that forms from single-stranded RNA, small RNAs derive from dsRNA precursors are classified as siRNAs

2.5.2 Functions of microRNA

MicroRNAs have been established to regulate diverse life processes in plants, such as development, biotic and abiotic stress responses, feedback regulation of small RNA pathways, biogenesis of some siRNAs, signal transduction and metabolism (reviewed in Mallory & Vaucheret, 2006; Rubio-Somoza & Weigel, 2011; Sun, 2012; Sunkar et al., 2012; Eldem et al., 2013).

2.5.2.1 Plant development

The functions of miRNA in regulating plant development are summarized in Table 2.3. These functions include regulation of development in various organs such as flower, leaf, meristem, root, nodule, vascular and seed and phenomena such as phase transition, sex determination and plant fertility (Table 2.3). Some miRNAs and their targets are conserved across plant species (reviewed in Zhang et al., 2006b; Cuperus et al., 2011; Jones-Rhoades, 2012). Examples of conserved miRNAs and targets found in multiple plant species include: miR156/miR157 (SQUAMOSA PROMOTER BINDING PROTEIN-LIKE [SPL]), miR159/319 (Gibberellin- and abscisic acid-regulated MYB [GAMYB] and TEOSINTE BRANCHED1, CYCLOIDEA and PCF [TCP]), miR160 and miR167 (AUXIN RESPONSE FACTOR [ARF]), miR164 (CUP-SHAPED COTYLEDON [CUC] of the NAM/ATAF/CUC[NAC]-domain transcription factor family), miR165/miR166 (class III homeodomain leucine zipper [HD-ZIP]), miR171 (SCARECROW-LIKE [SCL], a GRAS domain family gene), miR172 (APETALA2 [AP2] and AP2-like genes) and miR396 (GROWTH-REGULATING FACTOR [GRF]). Each of these miRNAs was also shown to regulate multiple developmental processes, for examples, miR172 alone was shown to regulate flowering time, floral development, floral meristem, sex determination and vegetative phase change, and miR165/166 was shown to control leaf development, leaf polarity, apical

meristem, vascular development and root and nodule development (Table 2.3).

2.5.2.2 Biotic stress

To date, only two reports demonstrate a direct correlation of miRNA to biotic stress. Navarro et al. (2006) reported that flagellin-derived peptide induces a plant miRNA (miR393) that negatively regulates F-box auxin receptors TIR1, AFB2 and AFB3TIR1, which represses auxin signalling and further restricts the growth of *Pseudomonas syringae* in *Arabidopsis*. Feng et al. (2013) reported that miR164-NAC21/22 module is involved in regulating resistance of wheat plants to stripe rust. TaNAC21/22, which is located in nucleus, functions as transcriptional activator and negatively regulates resistance to stripe rust (Feng et al., 2013).

Jagadeeswaran et al. (2009) reported the down-regulation of miR398 in response to biotic stress caused by *Pseudomonas syringae* although no direct correlation of this miR398 to the infection. Using next generation high-throughput sequencing technologies, Xin et al. (2010) reported 24 miRNAs that are responsive to powdery mildew infection in *Triticum aestivum*. Zhang et al. (2011a) and Chen et al. (2011b) found a number of differentially expressed miRNAs in response to *Pseudomonas syringae* infection in *Arabidopsis* and canker disease pathogen inoculation in *Populus* respectively. Twelve miRNAs (miR156, miR159, miR160, miR164, miR166, miR168, miR172, miR319, miR398, miR408, miR1448, and miR1450) were upregulated in the stem bark of *Populus trichocarpa* induced with the poplar stem canker pathogen, *Botryosphaeria dothidea*, but no downregulated miRNAs were found (Zhao et al., 2012).

Cucumber mosaic virus (CMV) and tomato aspermy virus (TAV) were found to interfere with the miRNA pathway in tomato by causing spatial and temporal changes in miRNA and target mRNA expression levels (Feng et al., 2011). Sattar et al. (2012)

reported the altered expression profiles of conserved and cucurbit-specific miRNAs found in small RNA libraries constructed from bulked leaf tissues of a resistant virus aphid transmission (Vat) melon line following early and late aphid infestations. Deeply sequenced small RNA libraries of cassava tissue infected and non-infected with *Xanthomonas axonopodis* pv. manihotis (Xam) revealed that some miRNAs known to mediate defense by targeting auxin-responding factors were upregulated and some miRNAs involved in copper regulation and targeting disease resistance genes were repressed, suggesting the roles of miRNAs in defense (Perez-Quintero et al., 2012). Yin et al. (2013b) described three miRNAs (miR160, miR393 and miR1510) that are involved in plant resistance from deeply sequenced small libraries of soybean infected by soybean mosaic virus (SMV).

Table 2.3: Functions of microRNA in various plant developmental processes

Function	miRNA	Target gene	Species	References
Flowering time regulation	miR156/157	SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) transcription factors	*Arabidopsis* & *Oryza sativa*	Xie et al., 2006; Gandikota et al., 2007; Wang et al., 2009a; Kim et al., 2012b
	miR159	Gibberellin- and abscisic acid-regulated MYB (GAMYB)	*Arabidopsis* & gloxiania	Achard et al., 2004; Li et al., 2013e
	miR169	CCAAT-binding transcription factor (Cbf-b/Nf-ya)	Opium poppy	Unver et al., 2010
	miR172	APETALA2 (AP2) and AP2-like genes	*Arabidopsis*	Aukerman & Sakai, 2003; Schmid et al., 2003; Mathieu et al., 2009
	miR399	PHOSPHATE 2 (PHO2)	*Arabidopsis*	Kim et al., 2011
Floral development	miR172	AP2-like genes: SUPERNUMERARY BRACT (SNB) and Os03g60430	*Oryza sativa*	Zhu et al., 2009
	miR171	SCARECROW-LIKE (SCL6)	Opium poppy	Unver et al., 2010
Floral meristem	miR165/166	Type III homeodomain-leucine zipper (HD-ZIP III) transcription factors	*Arabidopsis*	Ji et al., 2011
	miR172	APETALA2 (AP2) transcription factor	*Arabidopsis*	Zhao et al., 2007b; Ji, et al., 2011
Petal number	miR164	CUC1 and CUC2, a NAC [no apical meristem (NAM), *Arabidopsis* transcription activation factor (ATAF) and cup-shaped cotyledon (CUC)] family transcription factor	*Arabidopsis*	Baker et al., 2005
	miR319	TEOSINTE BRANCHED1, CYCLOIDEA and PCF 4 (TCP4) transcription factor	*Arabidopsis*	Nag et al., 2009
Anther development	miR156	SQUAMOSA PROMOTER BINDING PROTEIN-LIKE8 (SPL8)	*Arabidopsis*	Xing et al., 2010
	miR159	GAMYB transcription factor	*Arabidopsis*	Achard, et al., 2004
Sex determination	miR172	Indeterminate spikelet1 (ids1), an APETALA2 gene	*Zea mays*	Chuck et al., 2007
Plant fertility	miR167	AUXIN RESPONSE FACTOR 6 and 8 (ARF6 and ARF8) transcription factors	*Arabidopsis*	Ru et al., 2006; Wu et al., 2006
Leaf development	miR159	TEOSINTE BRANCHED1, CYCLOIDEA and PCF (TCP) transcription factors	*Arabidopsis*	Palatnik et al., 2003
	miR319	LANCEOLATE (La), a transcription factor from the TCP family	*Solanum lycopersicum*	Ori et al., 2007
Leaf polarity	miR165/166	HD-ZIP III transcription factors (PHABULOSA, PHAVOLUTA and REVOLUTA)	*Arabidopsis*	Mallory et al., 2004; Yao et al., 2009
	miR166	rolled leaf1 (rld1), an HD-ZIP III family member	*Zea mays*	Juarez et al., 2004
	miR396	GROWTH-REGULATING FACTORS (GRFs)	*Arabidopsis*	Wang et al., 2011b
Shoot apical meristem	miR164	NAC transcription factor family, CUC1 and CUC2	*Arabidopsis*	Laufs et al., 2004
	miR165/166	HD-ZIP III genes including ATHB-9/PHV, ATHB-14/PHB and ATHB-15	*Arabidopsis*	Williams et al., 2005; Zhou et al., 2007a

Table 2.3: Continued

Function	miRNA	Target gene	Species	References
Axillary meristem formation	miR164	CUC genes and LATERAL SUPPRESSOR	*Arabidopsis*	Raman et al., 2008
Cell proliferation	miR396	GROWTH-REGULATING FACTOR (GRF)	*Arabidopsis*	Rodriguez et al., 2010
Shoot brunching	miR171	SCARECROW-LIKE6-II (SCL6-II), SCL6-III, and SCL6-IV	*Arabidopsis*	Wang et al., 2010c
Stomatal development	miR824	Agamous-like16 (AGL16), a member of the MADS box protein family	*Arabidopsis*	Kutter et al., 2007
Vascular development	miR166	ATHB15 mRNA, an HD-ZIP III transcription factor	*Arabidopsis*	Kim et al., 2005
Vegetative phase transition	miR156	SPL transcription factors	*Arabidopsis*, *Zea mays*, *Nicotiana benthamiana* & trees	Wu & Poethig, 2006; Smith et al., 2009; Jung et al., 2011; Wang et al., 2011; Yang et al., 2011
	miR172	miR156-SPL module	*Arabidopsis*	Jung et al., 2011
	miR172	glossy15 (gl15), an AP2-like gene	*Zea mays*	Lauter et al., 2005
	miR534	Transcripts encoding BLADE-ON-PETIOLE (BOP) transcriptional coactivators	*Physcomitrella patens*	Saleh et al., 2011
Seed development and germination	miR159	GAMYB-like genes, MYB33 and MYB65; MYB33 and MYB101	*Arabidopsis*	Reyes & Chua, 2007; Alonso-Peral et al., 2010
	miR160	ARF10, 16 and 17	*Arabidopsis*	Liu et al., 2007; Liu et al., 2010
	miR166	HB15L mRNA	*Pinus taeda*	Oh et al., 2008
	miR167	ARF8L mRNA	*Pinus taeda*	Oh et al., 2008
	miR395c/e	ATP sulfurylase APS1, APS3, APS4 and sulfate transporter SULTR2;1 mRNA	*Arabidopsis*	Kim et al., 2010b
	miR402	DEMETER-LIKE protein3 mRNA	*Arabidopsis*	Kim et al., 2010a
Lateral root development	miR164	NAC1 transcription factor	*Arabidopsis*	Guo et al., 2005
	miR390	ARF3/4, mediated by tasiRNA from TAS3 precursor	*Arabidopsis*	Marin et al., 2010; Yoon et al., 2010
Adventitious root development	miR160	ARF17	*Arabidopsis*	Gutierrez et al., 2009
	miR167	ARF6 and ARF8	*Arabidopsis*	Gutierrez et al., 2009
Root cap development	miR160	ARF10 and ARF16	*Arabidopsis*	Wang et al., 2005
Root cell fate	miR165/166	HD-ZIP III transcription factor	*Arabidopsis*	Boualem et al., 2008; Carlsbecker et al., 2010; Miyashima et al., 2011
Root and nodule development	miR169	MtHAP2-1, a CCAAT-binding family transcription factor	*Medicago truncatula*	Combier et al., 2006
	miR172	AP2 transcription factor	Soybean	Yan et al., 2013
Nodule and mycorrhizal signalling	miR171	Nsp2 (encoding for GRAS transcription factor)	*Medicago truncatula*	Branscheid et al., 2011
Root system architecture	miR393	AFB3	*Arabidopsis*	Vidal et al., 2010
Tuberization	miR156	StSPL3, StSPL6, StSPL9, StSPL13 and StLIGULELESS1	Potato	Bhogale et al., 2013

This review will focus on the roles of miRNA in abiotic stress.

2.5.2.3 Abiotic stress

Khraiwesh, et al. (2012) proposed a network in which biotic and abiotic stress factors interacting with miRNA-mRNA modules triggers physiological, metabolic and morphological adaptation, such as, stomatal closure, ROS inactivation and lateral root formation in *Arabidopsis* (Figure 2.9). Among the stress-responsive miRNAs, miR398 is proposed to be directly linked to the plant stress regulatory network and widely adopted in regulation of stresses such as oxidative stress, water deficit, salt stress, abscisic acid stress, ultraviolet stress, copper and phosphate deficiency, high sucrose and bacterial infection (reviewed in Zhu et al., 2011). Down-regulation of miR398 by stresses results in up-regulation of its targets, Cu/Zn superoxide dismutases (CSD1 and CSD2), which detoxify reactive oxygen species (Figure 2.9). Accumulation of NFYA5 due to down-regulation of miR169 causes stomatal closure during drought and ABA stress (Figure 2.9). Repression of TIR1 and ARF10 (which control auxin perception and signalling) by up-regulation of miR393 and miR160 respectively causes growth attenuation that may enhance tolerance to biotic and abiotic stress (reviewed in Sunkar et al., 2012) (Figure 2.9).

A number of miRNA families have been shown to be involved in various biotic and abiotic stress responses and nutrient homeostasis in plants (Figure 2.9 and Figure 2.10).

(a) Nutrient stress

The roles of miR395 in sulfate homeostasis through the regulation of sulfate intake, accumulation, allocation and assimilation in *Arabidopsis* have been reported. Regulation of ATP sulfurylase by miR395 was found to be induced during upon sulfate

starvation in *Arabidopsis* suggesting its role in regulating nutrient stress in plant (Jones-Rhoades & Bartel, 2004). MiR395 targets ATP sulfurylases (encoded by APS gene) and sulfate transporter 2;1 (SULTR2;1), both of which are involved in the sulfate metabolism pathway (Liang et al., 2010; Liang & Yu, 2010). Besides, miR395 was also demonstrated to regulate ATP sulfurylase and responsive to sulphate deficiency in *Brassica napus* (Huang et al., 2010).

MiR398 regulates copper homeostasis by down-regulating the expression of copper/zinc-superoxide dismutase (CSD), a scavenger of superoxide radicals (reviewed in Ding & Zhu, 2009). Yamasaki et al., (2007) reported that miR398 is a key factor in copper homeostasis in *Arabidopsis* by directing the degradation of Cu/Zn superoxide dismutase (CSD) mRNA when copper is limited. miR397, miR398, miR408 and miR957 were shown to regulate copper proteins during limited copper supply in *Arabidopsis* (Abdel-Ghany & Pilon, 2008). In *Populus trichocarpa*, miR397, miR398, miR408 and miR1444 targeting laccases, CSD, plastocyanin-like proteins and polyphenol oxidases respectively, were shown to be responsive to copper changes (Lu et al., 2011; Ravet et al., 2011).

MiR399 is highly induced during phosphate (Pi) starvation and targeting 5'UTR of ubiquitin-conjugating E2 enzyme (encoded by UBC24) mRNA in *Arabidopsis* (Fujii et al., 2005; Bari et al., 2006; Chiou et al., 2006). miR399 was found to present in the phloem sap of rapeseed and pumpkin and its abundance was strongly and specifically increased in phloem sap during phosphate deprivation and long distance shoot-to-root transport of mature miR399 was demonstrated (Pant et al., 2008). MiR399 was also demonstrated as a negative regulator of ubiquitin E2 conjugase (PHO2) in *Phaseolus vulgaris* (Valdes-Lopez et al., 2008). Pi homeostasis in plants is also regulated by miR827, which targets NITROGEN LIMITATION ADAPTATION (NLA), in a nitrate-dependent manner (Kant et al., 2011). Expression of *Arabidopsis* miR169 was strongly

reduced under nitrogen starvation and this miRNA targets Nuclear Factor Y, subunit A (NFYA) family members (Zhao et al., 2011).

MiR393 and miR167, which regulate AFB3 and ARF8 respectively, are also known as nitrogen responsive miRNAs. These nitrogen responsive miRNAs regulate auxin signaling and subsequently regulate root system architecture (Gifford et al., 2008; Vidal et al., 2010). Liang et al., (2012) used a next generation sequencing platform to investigate expression of small RNAs under N-sufficient and -deficient conditions. They reported that miR160, miR167, and miR171 could be responsible for the development of *Arabidopsis* root systems under N-starvation conditions.

Kong & Yang (2010) investigated iron-deficiency in *Arabidopsis* and identified 8 conserved miRNA genes belong to 5 families that were up-regulated during Fe deficiency. In *Arabidopsis*, miR397a, miR398a, and miR398b/c) are among the miRNAs reported as iron deficiency-responsive miRNAs, which regulate transcripts of Cu-containing proteins (Waters et al., 2012)

(b) Drought stress

Liu et al. (2008) and Liu et al. (2010) identified and reported four drought-regulated miRNAs, miR167, miR168, miR171 and miR396 in *Arabidopsis* using microarray analysis, in which stress-related elements can be found in their promoter regions, whereas 30 miRNAs were differentially expressed under drought stress in rice using the same assay (Zhou et al., 2010). miR156, miR166, miR171 and miR408 were detected as dehydration stress-responsive miRNAs and shown to have a positive correlation between miRNA expression and target mRNA suppression in dehydration-stressed barley (Kantar et al., 2010).

Using microarray assay, miR169g was found as the only member induced by drought among the miR169 family in rice and two dehydration-responsive elements

(DREs) were found in the upstream of the miR169g promoter (Zhao et al, 2007a). miR169, which targets nuclear factor Y subunit A (NYFA5) in *Arabidopsis*, was found to be down-regulated by drought stress through an ABA-dependent pathway. Overexpression of an miR169-resistant NYFA mutant enhances drought tolerance by promoting stomatal closure under drought stress (Li et al., 2008). Zhang et al. (2011b) reported that miR169 is highly-regulated in drought-stressed tomato plants and it targets three nuclear factor Y subunit genes (SINF-YA1/2/3) and one multidrug resistance-associated protein gene (SIMRP1). This miR169 responds to drought stress by negatively regulating stomatal movement (Zhang et al., 2011b). Similar results were also shown in soybean, in which overexpression of GmNFYA3, a target of miR169, resulted in *Arabidopsis* with reduced leaf water loss and enhanced drought tolerance (Ni et al., 2013).

Overexpression of soybean miR394a reduced the transcript level of an F-box gene (At1g27340) and conferred tolerance to drought in transgenic *Arabidopsis* (Ni et al., 2012). Frazier et al. (2011) reported miR395 as a highly sensitive miRNA to drought stress in tobacco, in which an upregulation by 616-fold was induced by 1% PEG. miR398 and miR408 are up-regulated in response to water deficit in *Medicago truncatula*. miR398 targets copper proteins COX5b (subunit 5b of mitochondrial cytochrome c oxidase) and miR408 targets plantacyanin (Trindade et al., 2010). In *Medicago truncatula*, miR162 (which targets DCL1 mRNA) and miR168 (which targets AGO1 mRNA) showed expression reduction in the roots of plants subjected to water deficit and their targets showed increase in transcript level (Capitao et al., 2011). The application of high-throughput next generation small RNA sequencing has revealed greater number of drought-responsive miRNAs in plants, such as, cowpea (Barrera-Figueroa et al., 2011), *Populus euphratica* (Li et al., 2011a) and *Medicago truncatula* (Wang et al., 2011e).

(c) Salinity stress

Sunkar & Zhu (2004) reported that miR393 was strongly up-regulated in NaCl-stressed *Arabidopsis*, whereas miR397 and miR402 were slightly up-regulated. Using a microarray approach, Liu et al. (2008) reported 12 miRNAs that were regulated by 300 mM NaCl in *Arabidopsis*. They are miR396, miR168, miR167, miR165, miR319, miR159, miR394, miR156, miR393, miR171, miR158 and miR169 (fold change in decreasing order). Two members of miR169 family in rice, miR169g and miR169n are highly salinity-responsive miRNAs, in which the miR169 regulates NF-YA genes, a CCAAT-box binding transcription factor (Zhao et al., 2009).

Macovei & Tuteja (2012) demonstrated that osa-MIR414, osa-MIR164e and osa-MIR408 were experimentally validated to target OsABP (ATP-Binding Protein), OsDBH (DOB1/SK12/helY-like DEAD-box Helicase) and OsDSHCT (DEAD-Box Helicase) respectively. The negative correlation between the miRNAs and their targets in relation to salt stress was shown. Wang et al. (2013b) reported 5 miRNA:mRNA pairs, miR156-SPL2, miR162-DCL1, miR159-TCP3, miR395-APS1 and miR396-GRF1, that showed significant regulation relationship under salinity stress in cotton.

Creeping bentgrass expressing a rice miR319 (which targets TCP transcription factors) constitutively showed morphological changes that enhanced drought and salinity tolerance (Zhou et al., 2013a). miR393 was predicted to target some abiotic stress-related genes in rice and *Arabidopsis* and overexpression of miR393 caused higher sensitivity to salt and alkaline conditions (Gao et al., 2011). In tobacco, miR395 was shown to be highly sensitive (upregulation at 2810-folds) to salinity at 171 mM NaCl (Frazier et al., 2011). miR396a in rice was found to target several transcription factors related to growth, development and stress tolerance and miR396c-overexpressing rice plants showed reduced salt and alkaline stress tolerance (Gao et al., 2010). miR398 was induced in poplars upon ABA or salt stress and it targets Cu/Zn

superoxide dismutase (CSD1) in *Arabidopsis* (Jia et al., 2009b). miR398, targeting CSD1, was also reported by Jagadeeswaran et al. (2009) to be responsive to salinity in *Arabidopsis*. In *Hevea brasiliensis*, a negative co-regulation between miR398b with its target mRNA, chloroplastic Cu/Zn superoxide dismutase was observed, in response to salinity (Gebelin et al., 2013). miR417 was found as a negative regulator of seed germination in *Arabidopsis* under salt stress (Jung et al., 2007).

Using microarray analysis, Ding et al. (2009) reported 98 miRNAs (from 27 plant miRNA families) that were differentially expressed in maize roots after salt treatment, of which 18 of the miRNAs were only expressed in the salt-tolerant maize line and 25 miRNAs showed a delayed regulation pattern in the salt-sensitive line. The application of next generation sequencing technologies has enabled the discovery of more miRNAs from salt-stressed plants. Xie et al. (2013) reported 11 miRNA clusters containing 29 miRNAs that are differentially or uniquely expressed during salinity or drought stress in switchgrass. Sequencing of small RNA transcriptome from sugarcane revealed a number of miRNAs that regulate salt stress through the regulation of salinity tolerance-related miRNA targets that encode transcription factors, metabolic enzymes and hormone signalling proteins (Carnavale Bottino et al., 2013). Li et al. (2013a) investigated small RNAome, degradome and transcriptome in salt stress treated *Populus euphratica* by deep sequencing and reported reverse expression changing pattern in 15 miRNA-target pairs under salt stress.

(d) Alkaline stress

Two miRNAs have been reported to modulate both salt and alkaline stress in rice. MiR396c in rice showed dramatic change in expression under salt and alkali stress conditions. This miRNA targets several transcription factors related to growth, development and stress tolerance (Gao et al., 2010). *Oryza sativa* miR393 expression

was altered under salinity and alkaline stress and this miRNA was predicted to target abiotic-related genes. Gao et al. (2011) reported that rice and *Arabidopsis* plants that over-expressed osa-miR393 were more sensitive to salt and alkali treatment.

Analysis of *Glycine soja* root transcriptome under alkaline stress roots (subjected to $NaHCO_3$ treatment) revealed 11 miRNAs regulated by $NaHCO_3$ stress, including miR156, miR162, miR166, miR167, miR168 and miR398 (Ge et al., 2010).

(e) Submergence stress

MiR166, miR167, miR171 and miR396, which target HD-ZIP, ARF, SCL and the WRKY domain protein respectively, were up-regulated in maize root cells upon submergence, whereas miR159, miR395, miR474 and miR528, which target starch synthase, invertase, malic enzyme and ATPase respectively, were down-regulated (Zhang et al., 2008b).

(f) Metal stress

MiR395, which has a low-affinity sulphate transporter and a family of ATP sulphurylases as targets, is responsive to cadmium-stress in rapeseed (Huang et al., 2010). miR395 was also reported to involve in detoxification of cadmium in *Brassica napus* (Zhang et al., 2013a). Besides, miR393 and miR171 were also reported as gene expression regulators of plants under cadmium stress (reviewed in Ding & Zhu, 2009). Ding et al. (2011) reported 19 miRNAs identified in cadmium-stressed rice using miRNA microarray assay and the targets of these miRNAs are transcription factors and proteins associated with metabolic processes or stress responses. Metal stress-responsive cis-elements can be found in the promoter regions of the Cd-responsive miRNAs (Ding et al., 2011). In other reports, a number of cadmium stress-responsive

miRNAs were found in rice, *Brassica napus* and radish using the next generation sequencing platform (Huang et al., 2009; Zhou et al., 2012c; Xu et al., 2013b).

Lima et al. (2011) investigated aluminium-stress responsive miRNAs in rice roots using qPCR. Zeng et al. (2012) sequenced aluminium-treated and aluminium-free small RNA libraries and also degradomes, and reported 30 miRNAs that are responsive to aluminium stress in wild soybean. Burklew et al. (2012) reported that miR395, miR397, miR398, and miR399 increased strongly in expression in tobacco during exposure to 1% Al_2O_3 nanoparticles.

Valdes-Lopez et al. (2010) reported miRNAs responsive to manganese toxicity in common bean plants using macroarrays with probes complimentary to 68 known miRNAs. Using an Illumina deep sequencing approach, Zhou et al. (2012d) identified 12 miRNAs induced specifically by mercury (Hg) exposure in *Medicago truncatula* seedlings. While Liu & Zhang (2012) identified 67 arsenite-responsive miRNAs and their potential targets, some transcription factors, protein kinases, and DNA- or metal ion-binding proteins that are associated with cellular and metabolic processes, pigmentation and stress responses in rice seedling roots.

(g) Oxidative stress

MiR398 regulates two closely related Cu/Zn superoxide dismutases (cytosolic CSD1 and chloroplastic CSD2) that function in superoxide radicals detoxification in *Arabidopsis* (Sunkar et al., 2006; Jagadeeswaran et al., 2009). Zhang et al. (2013a) demonstrated that miR395-over-expressing *Brassica napus* showed a lower degree of Cd-induced oxidative stress than wild type, indicating the possible role of miR395 in oxidative stress in plants. A set of 7 conserved and 32 novel miRNAs were found to express in rice seedlings in response to hydrogen peroxide (H_2O_2) treatment (Li et al., 2011c). Iyer et al. (2012) used a plant miRNA array to identify 22 miRNA families that

are differentially expressed within one hour of ozone fumigation. A majority of the target genes of ozone responsive miRNAs were associated with developmental processes, indicating the regulation of developmental processes to cope with the oxidative stress (Iyer et al., 2012).

(h) Cold stress

A set of miRNAs (rather than a single miRNA) are responsive to cold stress in *Arabidopsis* and individual members of a miRNA family have different responses to cold stress. Stress-related cis-regulatory elements could be found in the promoter regions of these miRNAs (Sunkar & Zhu, 2004; Liu et al., 2008; Zhou et al., 2008). Microarray assay was used to discover 11 cold (4°C) stress-responsive miRNAs in *Arabidopsis* (Liu et al., 2008) and 19 cold-stress responsive miRNA from 14 miRNA families in *Populus* (Lu et al., 2008). In *Brachypodium,* 3 conserved miRNAs (miR397, miR169 and miR172) and 25 predicted novel miRNAs were found to be strongly responsive to cold stress (Zhang et al., 2009a). Cold stress-responsive novel miRNAs were also described in rice with 18 cold-responsive miRNAs determined using microarrays. Several novel miRNAs response to more than one stress treatment inclusive cold stress as observed in rice (Jian et al., 2010; Lv et al., 2010). Using high-throughput sequencing, Chen et al. (2012b) reported 21 miRNAs were down-regulated and nine miRNAs were up-regulated in response to cold stress in *Populus tomentosa*. Sunkar & Zhu (2004) reported miR393 was strongly up-regulated by cold stress in *Arabidopsis*. miR319 was reported to be highly regulated in leaves and roots of cold-stressed sugarcane (Thiebaut et al., 2011), which this miRNA has also been reported as cold responsive in *Arabidopsis* (Sunkar & Zhu, 2004).

(i) Heat stress

Guan et al. (2013) and Lu et al. (2013) reported the role of miR398 in regulating heat stress. Heat-inducible miR398 targets and down-regulates CSD genes and their copper chaperone CCS to achieve thermotolerance to protect the reproductive tissues (Guan et al., 2013). Wang et al. (2012a) reported that TamiR159 overexpression rice lines were more sensitive to heat stress relative to the wild type, suggesting that TamiR159 (which targets TaGAMYB1 and TaGAMYB2 by cleavage) contributed to the heat stress tolerance in plants. Xin et al. (2010) reported that 12 miRNAs were found to be responsive to heat stress in *Triticum aestivum*, and Yu et al. (2012a) reported five miRNA families that were responsive to heat stress in *Brassica rapa*.

(j) UV-B stress

Zhou et al. (2007b) reported 21 miRNA genes from 11 miRNA families responsive to UV-B stress in *Arabidopsis*. Jia et al. (2009a) identified 24 miRNAs responsive to UV-B stress in *Populus tremula*. Casati (2013) investigated UV-B regulated miRNAs and targets in maize leaves using microarrays and reported 17 differentially expressed miRNAs in response to UV-B, which the putative target mRNAs showed strong negative correlation to the miRNA in expression. Wang et al. (2013a) reported six miRNAs highly responsive to UV-B stress in wheat, including miR159, miR167a and miR171, which were significantly upregulated.

(k) Mechanical stress

A set of miRNAs responding to mechanical stress was reported in *Populus trichocarpa*, including miR408. miR408 targets laccase suggesting its role in structural and machanical fitness of this woody plant (Lu et al., 2005b). Lin et al. (2012) reported that wounding induced expression of miR828, which targets IbMYB and IbTLD by

cleavage in in sweet potato.

2.5.2.4 Other functions

Besides development and stress responses, miRNAs are also involved in regulating RNAi pathways such as feedback regulation of miRNA biogenesis, and directing the biogenesis of other small RNAs. DCL1, the key enzyme for biogenesis of miRNA, is a target of miR162 (Xie et al., 2003). Production of miR162 is controlled by DCL1 and level of DCL1 is modulated by miR162, forming a feedback circuit to maintain DCL1/miR162 homeostasis (Xie et al., 2003). ARGONAUTE1 (AGO1), the important functional component of RISC, is regulated by miR168 (Vaucheret et al., 2004). miR173 and miR390 were found to target trans-acting siRNA (ta-siRNA)-generating transcripts and guide the formation of pre-ta-siRNA (Allen et al., 2005; Bartel, 2005).

The miRNA-mediated control of signal transduction was observed in plants, especially in the hormone signalling pathways. Genes important for hormone signalling and response, gibberellin- and abscisic acid-regulated MYB (GAMYB), AUXIN RESPONSE FACTORS (ARF), NAC1 and transport inhibitor response (TIR), were regulated by miR159, miR160/167, miR164 and miR393 respectively (reviewed in Zhang et al., 2006c).

MiRNA also regulates secondary metabolism in plants. MiR156 negatively regulates SPL genes controlling anthocyanin accumulation in *Arabidopsis* via destabalisation of MYB-bHLH-WD40 transcriptional activation complex (Gou et al., 2011).

2.5.2.5 Single microRNA with multiple functions

Most plant miRNAs do not work independently. Instead, the miRNA-mediated gene expression regulation acts in overlapping regulatory networks (reviewed in Mallory and Vaucheret, 2006; Rubio-Somoza & Weigel, 2011). A single miRNA family may have multiple targets with different regulatory roles and multiple miRNAs may be involved in regulating a single development process or stress response (Table 2.3, Figure 2.9 and Figure 2.10). For example, miR160 regulates auxin response, leaf development and floral identity, while floral identity is determined by multiple miRNAs comprise of miR160, miR164, miR172 and miR319 (reviewed in Mallory and Vaucheret, 2006). Some miRNAs have opposite roles in regulating a single function, for instance, miR160 and miR167 regulate targets that have opposite roles in auxin regulation. ARF6 induces miR160 that represses ARF17 and this ARF17 may activate miR167 to modulate ARF6 and ARF8 expression (Gutierrez et al., 2009).

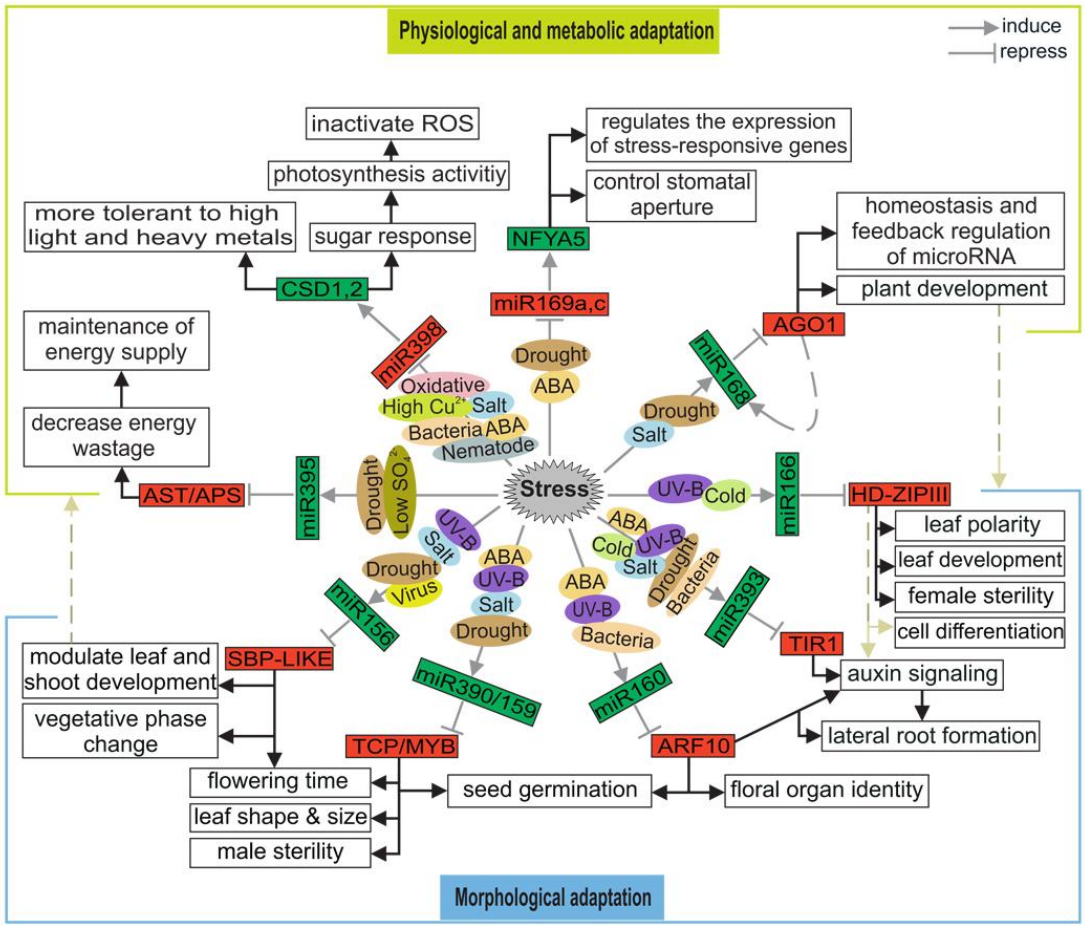


Figure 2.9: Regulation networks of stress-responsive miRNAs in *Arabidopsis*. Green boxes: up-regulated RNAs; red boxes: down-regulated RNAs. Source: Khraiwesh, et al. (2012)

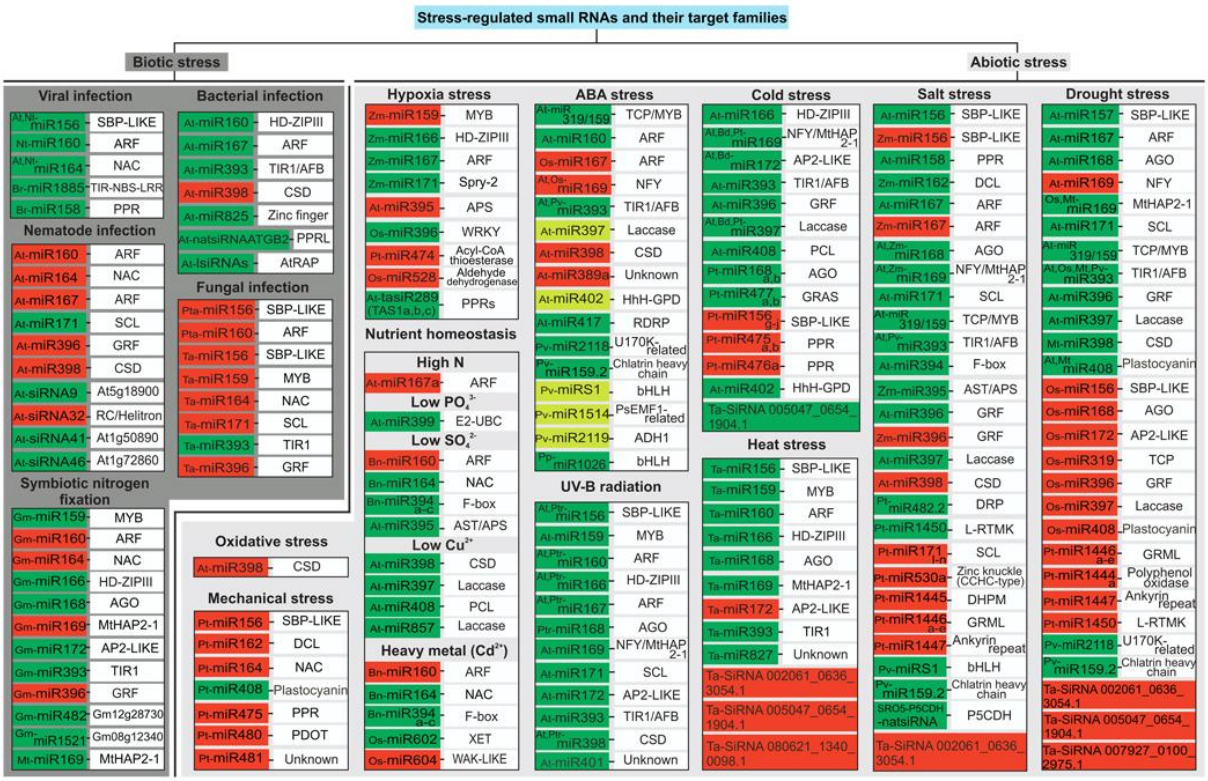


Figure 2.10: Stress-regulated miRNAs and their target gene families summarised in Khraiwesh et al. (2012). Grey box: abiotic stress; dark grey: biotic stress; green boxes: up-regulated miRNA/siRNA; light green boxes: slightly up-regulated miRNA/siRNA; red boxes: down-regulated miRNA/siRNA; At: *Arabidopsis thaliana*; Bd: *Brachybodium distachyon*; Bn: *Brassica napus*; Br: *Brassica rapa*; Gm: *Glycine max*; Mt: *Medicago truncatula*; Nt: *Nicotiana tabacum*; Os: *Oryza sativa*; Pt: *Populus trichocarpa*; Pta: *Pinus taeda*; Ptr: *Pinus tremula*; Pp: *Physcomitrella patens*; Pv: *Phaseolus vulgaris*; Ta: *Triticum aestivum*; Zm: *Zea mays*. Source: Khraiwesh, et al. (2012)

2.5.3 Approaches for microRNA studies

2.5.3.1 Discovery of microRNAs

(a) Experimental approaches for miRNA discovery

Several approaches have been used towards miRNA discovery. Genetic screening, also known as a forward genetics approach, involves the production of random mutations in an organism to inspect genes that cause a phenotypic change. It was used in the discovery of the founding members of miRNA, *lin-4* and *let-7* from *Caenorhabditis elegans* (Lee, et al., 1993; Wightman, et al., 1993; Pasquinelli, et al., 2000; Reinhart, et al., 2000). However, this approach is relatively expensive in terms of labour and resources whilst the number of miRNA that can be discovered is very limited.

Construction of cDNA libraries for small RNA species followed by conventional Sanger sequencing (direct cloning and sequencing approach) is a feasible way to discover small RNA sequences including miRNAs, expressed in different tissues, developmental stages and environmental conditions in an organism (Reinhart e al., 2002; Llave et al., 2002; Sunkar & Zhu, 2004; Sunkar et al., 2005). This approach requires cloning techniques, in which oligo-nucleotide adapters are ligated to the 5' and 3' ends of isolated small RNA species by T4 RNA ligase, followed by RT-PCR for amplification. The amplification products can either be directly ligated into a cloning vector or concatemerized before being ligated into a cloning vector prior to DNA sequencing (Wang et al., 2004a; Arazi et al., 2005; Sunkar et al., 2005). The former approach requires high cost to sequence the short insert of each clone in the cDNA library, whereas the latter requires multiple steps to produce concatemers for use in cloning before Sanger sequencing.

Microarray is hybridization-based method for high-throughput and parallel quantitative detection of an array of DNA or RNA sequences. The probes with known sequences are immobilized on a solid surface in a chip where samples of study with

complementary sequence will bind and signal released from the labeled targets can be detected quantitatively. In miRNA profiling, oligonucleotides with sequence complementary to the target miRNAs serve as the probes for miRNA detection. A microarray approach to study miRNA profiles in plant was first reported in Liang et al. (2005). Its use has been continously reported in various plant miRNA studies such as drought stress in rice (Zhao et al., 2007a), validation of miRNAs predicted from tomato (Zhang et al., 2008a), submergence stress in maize root (Zhang et al., 2008b), salt stress in maize roots (Ding et al., 2009) and cotton (Yin et al., 2012), phosphorous deficiency (Zeng et al., 2010), disease resistance (Guo et al., 2011) and shoot apical meristem (Wong et al., 2011) in soybean, cadmium and low-N stress in rice (Ding et al., 2011; Nischal et al., 2012) and arsenic stress in *Brassica juncea* (Srivastava et al., 2013). MiRNA microarray analysis has also been incorporated with high-throughput sequencing and quantitative PCR analysis for expression comparison and validation (Lu et al., 2008; Li et al., 2011b; Zhou et al., 2012a; Li et al., 2012b). Although being a high-throughput approach for the detection and expression profiling of miRNAs, miRNA microarray is limited for profiling of miRNAs with known (homologous) sequences only.

High-throughput sequencing or the next-generation sequencing technologies such as MPSS, 454 Life Sciences and Illumina (formerly Solexa) to sequence small RNA cDNA libraries can produce thousands to millions of sequence tags or reads in parallel (Lu et al., 2005a; Barakat et al., 2007; Fahlgren et al., 2007; Sunkar et al., 2008). Next generation high-throughput sequencing offers an option to deeply sequence small RNA cDNA libraries with comparatively lower cost, in less time and is less tedious than Sanger sequencing (Sun, 2012). The NGS technologies, with the capacity of deeply sequencing sRNAomes generating 10^5 to more than 10^7 sequences in a single experiment, have a high possibility to recover rare transcripts (Jones-Rhoades, 2012).

Lu et al. (2005a) was probably the first report of the application of NGS for plant sRNAome sequencing. They applied massively parallel signature sequencing (MPSS) to produce more than 2 million small RNA reads from four libraries of seedlings and inflorescence of *Arabidopsis thaliana*. The NGS technologies are now widely used to sequence sRNAomes of various plant species, tissues, stages and/or growth conditions (Table 2.4).

Being able to produce a huge amount of data, the use of next generation sequencing technologies for miRNA discovery requires extensive computational analyses and relies on available genomic and/or transcriptomic resources (Figure 2.11). Small RNA sequences generated from next generation sequencing platforms are annotated using publicly available annotated small RNA sequence databases, especially miRBase (Griffiths-Jones et al., 2006), a miRNA registry containing all sequences regardless of species, PMRD, a plant microRNA database (Zhang et al., 2010c) and Rfam, a database containing annotated non-coding RNAs and structured RNA elements (Gardner et al., 2011). sRNA-Seq reads are then mapped to a reference genome (or transcriptome) to find homologous miRNAs as well as novel miRNAs, in which the reads mapped to the genome regions can form characteristic miRNA precursors or hairpin stem-loop structures. Metpally et al. (2013) compared the most cited software packages for mapping reads to reference genome, miRDeep2, miRNAKey and miRExpress, and found that miRDeep2 detected and aligned more miRNAs than the other two miRNA mapping tools. The core algorithm of miRDeep2 is the use of RNAfold tool to predict RNA secondary structures, and the structures and signature of each potential miRNA precursor is evaluated by assigning a score that reflects the likelihood of a precursor being a genuine miRNA (Metpally et al., 2013).

Table 2.4: Application of sRNA-Seq in plants

Category	Subject	References
Species	*Arabidopsis thaliana*	Lu et al., 2006; Fahlgren et al., 2007
	Basal eudicots	Barakat et al., 2007
	Triticum aestivum	Yao et al., 2007
	Oryza sativa	Sunkar et al., 2008; Jeong et al., 2011
	Solanum lycopersicum	Moxon et al., 2008
	Pinus contorta	Morin et al., 2008
	Zea mays	Zhang et al., 2009b
	Vitis vinifera	Pantaleo et al., 2010
	Arachis hypogaea	Zhao et al., 2010
	Hevea brasiliensis	Gebelin et al., 2012
	Brassica napa	Kim et al., 2012a
	Raphanus sativus	Xu et al., 2013a
	Punus persica	Zhu et al., 2012
	Setaria italica	Yi et al., 2013
Tissues/Organs/ Developmental stages	Grains of *Oryza sativa*	Zhu et al., 2008; Xue et al., 2009
	Grains of *Triticum aestivum*	Meng et al., 2013
	Roots and nodules of *Medicago truncatula*	Lelandais-Briere et al., 2009
	Ovule and fibre development in *Gossypium hirsutum*	Pang et al., 2009
	Red flesh of sweet orange	Xu et al., 2010
	Developing pollens in *Oryza sativa*	Wei et al., 2011
	Heterosis in the seeds of *Zea mays*	Ding et al., 2012a
	Developing ears of *Zea mays*	Ding et al., 2013
	Male sterile *Gossypium hirsutum*	Wei et al., 2013
	Male sterile *Brassica juncea*	Yang et al., 2013
Stress response	Biotic stress in *Nicotiana sp.*	Pandey et al., 2008
	Cold stress in *Brachybodium*	Zhang et al., 2009a
	Powdery mildew infection and heat stress in *Triticum aestivum*	Xin et al., 2010
	Drought stress in *Vigna* sp.	Barrera-Figueroa et al., 2011
	Pathogen stress in *Populus* spp.	Chen et al., 2011b
	Drought stress in *Populus* spp.	Li et al., 2011a
	Cold stress in *Populus* spp.	Chen et al., 2012b
	Biotic and abiotic stresses in *Glycine max*	Kulcheski et al., 2011; Li et al., 2011b
	Ethylene response in *Medicago truncatula* roots	Chen et al., 2012a
	Heavy metal stress in *Medicago truncatula*	Zhou et al., 2012d
	CO_2 and temperature stress in *Arabidopsis*	May et al., 2013
	Cadmium stress in *Brassica napus*	Zhou et al., 2012c
	Cadmium stress in *Raphanus sativus*	Xu et al., 2013b
	salt stress in *Thellungiella salsuginea*	Zhang et al., 2013b
	Nutrient stress in *Zea mays*	Zhao et al., 2013c

Normalization is an important step in analysing next generation sequencing data to make relative change in miRNA expression between different samples (libraries) comparable. At present, scaling to library size is the standard procedure to normalize sRNA-Seq experiments (Meyer et al., 2010). No single best normalization method for miRNA-Seq has been described to date (Zhou et al., 2013b), and the difference between sequencing depths is the primary consideration in normalization (McCormick et al., 2011). Among the methods commonly tested for miRNA-Seq normalization, namely global normalization, Lowess normalization, Trimmed Mean of Method (TMM), quantile normalization, scaling normalization, variance stabilization (VSN) and invariant method (INV), Garmire & Subramaniam (2012) suggested the Lowess and quantile normalization methods perform the best, although Zhou et al. (2013b) found it arguable. The most commonly used packages for analyzing differential expression in sRNA-Seq data are DESeq and EdgeR (Metpally et al., 2013). DESeq uses a negative binomial model for normalization, while EdgeR uses trimmed mean of M-values (TMM) method, prior to computing significant differentially expressed miRNAs and estimating *p*-values (reviewed in Garmire & Subramaniam, 2012; Metpally et al., 2013).

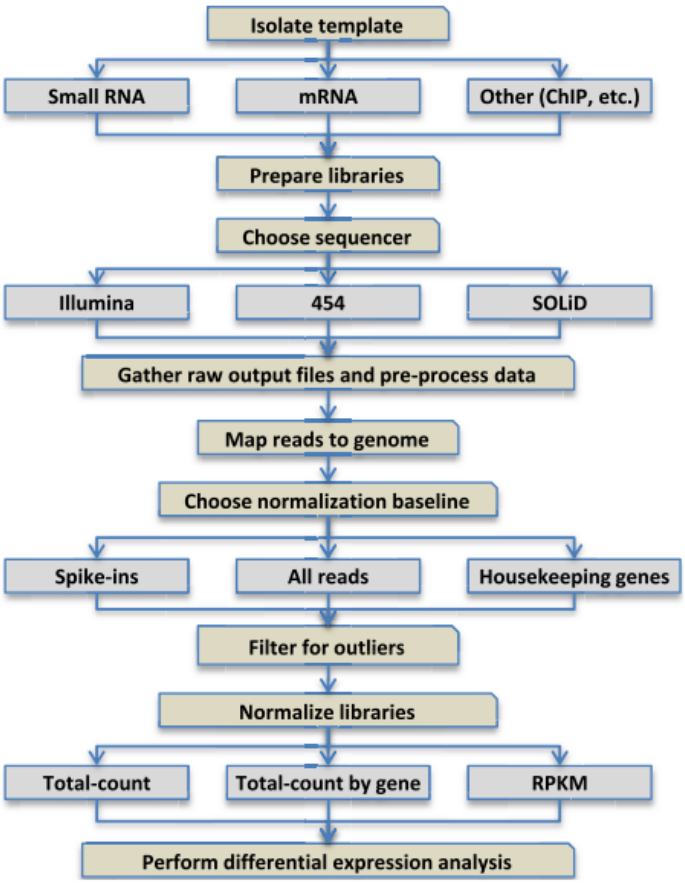


Figure 2.11: A typical sRNA-Seq experiment and analysis workflow from library preparation to library normalization and differential expression analysis (McCormick et al., 2011)

(b) Computational approaches for miRNA discovery

Computational approachs can be used to identify characteristic miRNA genes from annotated complete genome sequences or from sets of gene sequences, such as, EST and GSS, for species without an annotated genome (Bonnet et al., 2004a; Jones-Rhoades & Bartel, 2004; Wang et al., 2004b; Adai et al., 2005; Zhang et al., 2005a). Computational prediction methods have been developed to identify homologous miRNAs across plant species based on conservation of the known miRNAs (Zhang et al., 2006a). In addition, extra filtering criteria, such as properties of the secondary structure (hairpin), are also applied in the miRNA prediction algorithm (reviewed in Unver et al., 2009). This approach is also relatively fast and affordable and the

application of this approach to miRNA prediction from EST and GSS databases has been conducted for various plant species (Bonnet et al., 2004a; Zhang et al., 2005a; Guo et al., 2007; Nasaruddin et al., 2007; Qiu et al., 2007; Xie et al., 2007; Gleave et al., 2008; Jin et al., 2008; Han et al., 2010; Lee et al., 2011; Sablok et al., 2011; Yu et al., 2011). Whilst analysis of EST data can give some indication of gene expression, miRNA is generally represented at a rather low levels of around 1-10 per 10,000 EST (Zhang et al., 2005a, Nasaruddin et al., 2007; Qiu et al., 2007; Xie et al., 2007; Jin et al., 2008; Lee et al., 2011; Sablok et al., 2011).

2.5.3.2 Detection and expression measurement of microRNAs

Northern blot analysis is an approach to determine the presence and relative quantity of a target RNA transcript in transcriptome. In miRNA studies, northern blotting is commonly used to detect the presence of miRNAs, including the mature miRNA (leading strand), miRNA* and precursor (pre-miRNA) sequences, to confirm computationally predicted miRNAs (Elbashir et al., 2001; Molnar et al., 2007). The northern blot analysis has the advantages of indicating both the size of miRNA (mature miRNAs and/or precursors) and its expression level (qualitative or semi-quantitative). While a radioactive probe labeling approach is commonly practiced due to its sensitivity, a non-radioactive approach with high detection sensitivity for small RNA (miRNA) has been developed (Ramkissoon et al., 2006; Kim et al., 2010c). The detection of small RNA is improved using a novel cross-linking step using 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) (Pall et al., 2007). While the specificity of hybridization has also been improved by using LNA (locked nucleic acids) oligonucleotides as probes (Valoczi et al., 2004). Besides, in situ hybridisation (ISH) is also a useful hybridisation-based method for miRNA study (Eldem et al., 2013). Synthetic probes with sequence complementary to the target miRNAs are used to detect

spatial and/or temporal miRNA expression (localisation) and determine relative expression levels of miRNA within the fixed cells, tissues, organs and even whole organisms (Eldem et al., 2013)

Quantitative RT-PCR (RT-qPCR) is the most commonly used method to measure or compare miRNA expression levels quantitatively and accurately. It is also used to validate miRNA expression from computational prediction, microarray and high-throuput sequencing data. miRNA expression profiling using RT-qPCR can be done at two levels: mature miRNAs (Raymond et al., 2005) and miRNA precursors (inclusive of pri-miRNA and pre-miRNA) (Schimittgen et al., 2004). Among the RT-qPCR methods for quantifying mature miRNA are poly(A)-tailing RT-qPCR, miQPCR and stem-loop RT-qPCR (Benes & Castoldi, 2010; Mou et al., 2013). After poly(A)-tailing of mature miRNA (~20-nt), upon binding of an oligo-d(T) primer in an RT reaction, allows reverse transcription of this small-sized RNA into cDNA prior to quantitative PCR. miQPCR involves the use of T4 RNA ligase to attach an universal linker to the ~20-nt miRNA, followed by reverse transcription and quantitative PCR (Benes & Castoldi, 2010). The more popular approach nowadays, stem-loop RT-qPCR, is the use of a stem-loop primer containing complementary sequence of the target miRNA seed sequence for reverse transcription, then quantitative PCR (Chen et al., 2005a). RT-qPCR for miRNAs can be done using SYBR® Green or TaqMan® chemistries (Chen et al., 2005a; Raymond et al., 2005). However, the TaqMan stem-loop RT-qPCR miRNA assay has advantages over SYBR Green as it has been shown to discriminate mature miRNAs that differ by a single nucleotide and the expression quantification is not affected by genomic DNA contamination (Chen et al., 2005a; Shen et al., 2010; Chen et al., 2011a).

2.5.3.3 Study of microRNA targets

(a) Computational approaches for miRNA target prediction

Complementary between miRNA and target sites (miRNA:target duplex) is the key feature for computational analysis of miRNA target genes in plants (reviewed in Ding et al., 2012b). The computational miRNA target prediction tools are developed based on a scoring schema for the canonical seed region of miRNA:target pair (reviewed in Ding et al., 2012b). Several bioinformatics tools have been developed for miRNA target prediction from plants using available genomic and/or transcriptomic resources. These tools include PatScan, miRNAassist, miRU, WMD3, TAPIR, UEAEA sRNA, Target-align, Target Finder, pTAREF, psRNATarget and imiRTP (reviewed in Chen et al., 2010b; Ding et al., 2012b).

(b) Experimental approaches for determination and validation of miRNA targets

Before 2008, reported target findings for miRNAs are mainly based on genetic screening, computational prediction, RT-qPCR expression profiling of miRNA:target and a modified 5'RLM-RACE approach (Chen et al., 2010b; Ding et al., 2012b). The modified 5' RNA ligase mediated-rapid amplification of cDNA ends (5' RLM-RACE) involves the ligation of an RNA adapter to the free 5'phosphate of an uncapped mRNA, a product from AGO-mediated cleavage. The ligation product is then reverse transcribed and amplified using primers designed based on the linker (forward primer) and gene of interest (reverse primer) (Llave et al., 2002b). Thus, 5' RLM-RACE can be used to validate the cleavage sites of a specific miRNA target. Genetic screening and 5'RACE are not efficient ways for target determination as only limited targets can be found, while computational approaches may possess the risk of over- or under- prediction. The advancement in next-generation sequencing has allowed high-throughput sequencing of

degradomes, also known as parallel analysis of RNA ends (PARE). Deep sequencing of degradomes containing only mRNA transcripts without an intact 5'cap, enables high-throughput study and discovery of miRNA target mRNAs and their cleavage sites (Addo-quaye et al., 2008; German et al., 2008; German et al., 2009). Northern blot, RT-qPCR and western blot analysis are also the methods suitable for profile expression patterns of the miRNA targets (at mRNA or protein levels), to determine the correlation between the miRNA and its specific targets (reviewed in Chen et al., 2010b; Ding et al., 2012b). The use of the modified 5' RLM-RACE and PARE are limited to the finding of cleaved mRNA products. MiRNA targets that are regulated transciptionally (DNA methylation) and translationally (translational inhibition) will be missed from the analysis. This problem can be solved by using an expression profiling method such as northern and RT-qPCR (for the transcriptionally-regulated target genes), and western blot (for the translationally-regulated target genes) (reviewed in Chen et al., 2010b; Ding et al., 2012b).

MiRNA functions can also be investigated in plants by overexpression or lowered expression of miRNA genes. If transcriptional inhibition and mRNA cleavage are the mode of action, overexpression of a miRNA gene (precursor sequence) will see a decrease in its target mRNA, and vice versa. Besides, mutations in miRNA genes or miRNA target genes (engineering miRNA-resistant targets) are also very useful and commonly practiced in *Arabidopsis* to confirm the interaction of specific miRNA and target mRNA (reviewed in Mallory and Vaucheret, 2006; Zhang et al., 2006c; Unver et al., 2009; Chen et al., 2010b).

CHAPTER 3: MATERIALS & METHODS

3.1 Plant materials and sample treatments

Clonal, tissue culture-derived plantlets of banana, *Musa acuminata* cultivar Berangan (AAA genome) with healthy roots were used in this study. Plantlets of 6-8 cm in height were selected for stress treatments. *In vitro* banana plantlets were treated on Murashige and Skoog Medium (Murashige & Skoog, 1962) (Table S1) supplemented with 0 mM (CTR), 100 mM (TR100) and 300 mM (TR300) NaCl respectively for 48 hours prior to RNA isolation. Root tissues were pooled randomly from 3-4 plantlets in each treatment group before RNA isolation.

3.2 RNA isolation

3.2.1 Total RNA extraction from banana roots

A CTAB-based RNA isolation method described by Kiefer et al. (2000) was used for total RNA isolation from banana roots. A slight modification was made, in which an additional extraction step using phenol-chloroform-isoamylalcohol (25:24:1) was performed prior to extraction using chloroform-isoamylalcohol (24:1).

In brief, fresh root tissue of banana was ground into powder in liquid nitrogen before transferring to homogenization buffer containing 2% (w/v) CTAB, 100 mM Tris-HCl (pH8.0), 2% (w/v) PVP-40, 2 M NaCl, 25 mM EDTA (pH8.0), 2% (v/v) β-mercaptoethanol (Sigma-Aldrich®, St. Louis, MO, USA) and mixing well by inversion repeatedly. About 0.2 g of plant tissue was ground and put into 1 mL of homogenization buffer contained in a 2 mL microcentrifuge tube. An equal volume of phenol-chloroform-isoamylalcohol (25:24:1) (from Sigma-Aldrich®, St. Louis, MO, USA; Fisher Scientific International Inc., Hampton, NH, USA; and Amresco LLC, Solon, OH, USA respectively) was added into the mixture of tissue and homogenization buffer, and mixed well by continuous inversion for 5 minutes. The mixtures were spun at 10,000 *g*

for 10 minutes. This step was repeated once, followed by an extraction step using chloroform-isoamylalcohol (24:1) (from Fisher Scientific International Inc., Hampton, NH, USA; and Amresco LLC, Solon, OH, USA, respectively). The nucleic acids were then precipitated overnight in an equal volume of isopropanol (Sigma-Aldrich®, St. Louis, MO, USA) at -80°C. Overnight-precipitated RNA was spun at 10,000 *g* for 30 minutes at 4°C. The resulting RNA pellet was air-dried and then dissolved in nuclease-free water.

3.2.2 Determination of RNA quantity, purity and integrity

The presence of RNA was determined on a 1% (w/v) agarose gel. Absorbance at 260 nm and 280 nm was measured spectrophometrically using a BioPhotometer (Eppendorf AG, Hamburg, Germany) and RNA integrity was determined uisng an Agilent 2100 Bioanalyzer (Agilent Technologies Inc., Santa Clara, CA, USA). Only RNA samples with absrobance ratios A_{260nm}/A_{280nm} ranging from 1.8 to 2.2, A_{230nm}/A_{260nm} ratio higher than 1 and an RNA integrity number (RIN) higher than 8 were used for library construction and sequencing using Illumina sequencing platforms (Illumina Inc., San Diego, CA, USA).

3.3 Transcriptome sequencing

3.3.1 Library construction and sequencing

Illumina DNA sequencing technology was used for high-throughput transcriptional profiling. The Illumina RNA sequencing (RNA-Seq) service was provided by the Beijing Genomics Institute (BGI), Shenzhen, China (http://www.genomics.cn/index.php). Two purified mRNA samples from banana roots, namely CTR (control) and TR300 (300 mM NaCl treatment) were sequenced using Illumina HiSeq^{TM} 2000 platform (Illumina Inc., San Diego, CA) according to the

manufacturer's instructions. Briefly, Sera-mag Magnetic Oligo (dT) Beads (Illumina Inc.) were used to isolate poly(A) mRNA from 20 µg of total RNA. mRNA was interrupted into short fragments (100 to 400 nt) at 94°C for 5 min in 1 x Fragmentation Buffer (Ambion, Inc. Austin, TX, USA) that contains divalent cations. Random hexamer primers (Illumina Inc.) and SuperScript II reverse transcriptase (Invitrogen, Camarillo, CA, USA) were used to synthesize the first-strand cDNA from these short fragments. The second-strand cDNA was synthesized in reaction buffer containing RNaseH, DNA polymerase I and dNTPs. The cDNA was purified using QIAQuick PCR extraction kit (Qiagen, Hilden, Germany) and then end-repaired, phosphorylated and 3'-adenylated using Illumina's kit according to the manufacturer's instruction, prior to ligation with Illumina sequencing adapters. After gel electrophoresis, cDNA fragments 200 ± 25 bp were purified from agarose gel and enriched by 15 cycles of PCR amplification. The insert size for the libraries was approximately 200 bp and both ends of the libraries were sequenced using an Illumina HiSeq™ 2000 platform.

3.3.2 Sequence generation and processing

Raw paired-end reads sized 90-bp were generated from Illumina HiSeq^{TM} 2000 sequencing platform and processed to remove adapter sequences, low quality reads with ambigous sequences ('N') and sequence reads containing more than 10% bases with Phred score, Q<20. The resulting clean reads, provided as FASTQ files by BGI, were used for further analysis. The subsequent bioinformatics analyses were carried out using workstations and technical supports in the University of Malaya. The paired-end Illumina sequencing reads of banana root transcriptomes described in this study are available under NCBI's BioProject accession PRJNA246442.

3.3.3 Sequence assembly and mapping to reference genome

After removal of adapters and low quality sequences, the clean reads were assembled into contigs, scaffolds and unigenes (illustrated in Figure 3.1) using a short reads assembling program, SOAP*denovo* assembler version 1.05 (http://soap.genomics.org.cn/soapdenovo.html) (Li et al., 2010b). The clean reads were assembled using the 63-mer version of SOAP*denovo* and a K-mer size of 51. The SOAP*denovo* assembler was used to combine sequence reads to form longer fragments without an ambigous nucleotides ('N'). The assembled sequences containing no 'N's are called contigs. Then SOAP*denovo* was used to connect and link the contigs into scaffolds. The unknown bases in the scaffolds were filled with 'N's. Paired-end reads were used again for gap filling of scaffolds by using GapCloser v.12 for SOAPdenovo (http://soap.genomics.org.cn/soapdenovo.html) to obtain sequences with the least 'N's that could not be extended on either end. Such sequences are defined as unigenes. The assembled unigenes from both the transcriptomes CTR and TR300 were clustered together using TGI clustering tool 2.1 (Pertea et al., 2003) to form a set of non-redundant unigenes. After clustering, the *de novo* assembled unigenes were mapped to the reference *Musa* genome (D'Hont et al., 2012) using BWA version 0.6.1 (Li & Durbin, 2009), with mismatch score less than 3.

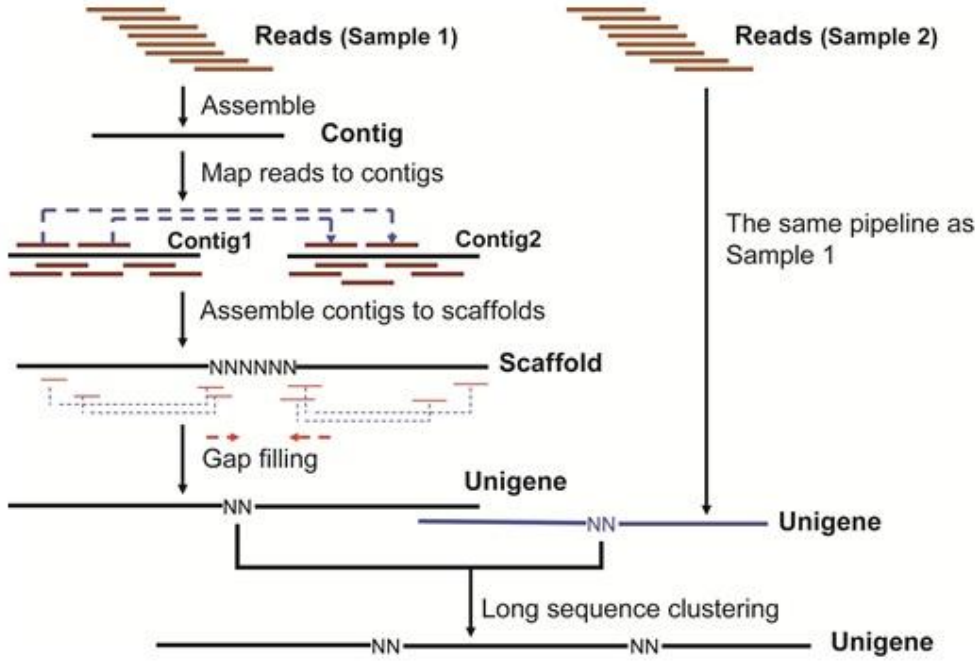


Figure 3.1: *De novo assembly* of high-throughput sequencing short reads into contigs, scaffolds and unigene (Huang et al., 2012)

3.3.4 Functional annotation of transcripts

The *de novo* assembled unigenes were assigned with putative identities and/or functional annotations by similarity searches against the publicly available reference protein databases. Standalone BLASTX was carried out using NCBI's BLASTALL tool for similarity searches against the NCBI non-redundant (nr) protein database (Dec 29, 2011, 4:42 PM, Version 4, with 16,785,757 sequences) with a stringent cut-off E-value of $1e^{-10}$. The unigenes were also searched against the Uni-Prot database (Feb 28, 2012, 7:17 PM, Version 4, with 534,242 sequences) using standalone BLASTX with the same cut-off E-value. Blast2GO program (http://www.blast2go.com/b2ghome) (Conesa et al., 2005) version 2.7 was used to assign Gene Ontology (GO, http://www.geneontology.org/) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG, http://www.genome.jp/kegg/) pathways to the unigenes. Eukaryotic Orthologous Groups of Proteins (KOG) assignment was carried out using the NCBI's BLASTALL tool.

3.3.5 Quantification and determination of differentially expressed transcripts

Library normalization was conducted prior to expression comparison of transcripts between the two libraries (CTR and TR300). Transcripts were normalized to transcript per million (TPM) (Audic & Claverie, 1997). An R package, DEGSeq (version 1.15, http://bioinfo.au.tsinghua.edu.cn/software/degseq) (Wang et al., 2010a) was used to identify differentially expressed unigenes based on read count. The transcripts normalization and differential gene expression were calculated using an R script (Appendix B). Only unigenes with at least 10 read counts present in the library were used for transcript quantification as sequences with less than 10 counts might be sequencing artifacts. FDR < 0.05 and $\log_2$ ratio ≥1 were set as threshold values to select for differentially expressed transcripts.

3.3.6 Identification of simple sequence repeats (SSR)

Identification of simple sequence repeats from unigenes (cDNA-based SSR, cSSR) was carried out using method and parameters described in Wang et al. (2010e) and Franchini et al. (2011). The assembled unigenes were screened for SSRs using a Perl script known as MIcroSAtellite identification tool (MISA, http://pgrc.ipk-gatersleben.de/misa/misa.html). The minimum lengths for SSRs were at least 8 for dinucleotide, 12 for trinucleotide, 16 for tetranucleotide, 20 for pentanucleotide and 24 for hexanucleotide repeats.

3.4 Small RNA profiling

3.4.1 Direct cloning and sequencing approach

3.4.1.1 Isolation of small RNA from total RNA

Total RNAs were separated by high resolution 15% denaturing polyacrylamide gel electrophoresis containing 15% (w/v) 19:1 acrylamide: bis-acrylamide (Ambion, Inc. Austin, TX, USA), 42% (w/v) urea (Ambion, Inc. Austin, TX, USA), 0.07% (w/v) ammonium persulphate (Promega Corporation, Madison, WI, USA), 0.035% (v/v) TEMED (Bio-Rad Laboratories, Inc., Hercules, CA, USA) in 0.5X TBE (Lau et al., 2001; Fu et al., 2005; Ho et al.,2006; Zhao et al., 2007). Alternatively 5% (w/v) low melting agarose gel (Promega Corporation, Madison, WI, USA) was used for high resolution RNA separation. The gels were stained with 0.5 µg/mL ethidium bromide (Promega Corporation, Madison, WI, USA) to visualize the RNA bands. Small RNAs sized 15 to 35 nt according to a 10 bp DNA Ladder (Fermentas, Thermo Fisher Scientific Inc., Waltham, MA) were excised and eluted from the polyacrylamide or agarose gel overnight in 500 µL of 0.3M NaCl followed by precipitation using isopropanol.

3.4.1.2 Direct cloning and sequencing of small RNA species

Adapter ligation-based methods described by Fu et al. (2005) and Ho et al. (2006) were used to clone and sequence small RNAs (Figure 3.2). In the method described by Fu et al. (2005), small RNAs were polyadenylated using poly(A) polymerase (Takara Bio Inc., Otsu, Shiga, Japan) at 37°C for 30 min. Reaction buffer for the poly(A) polymerase contains 50mM Tris-HCl (pH7.9), 10mM $MgCl_2$, 2.5mM $MnCl_2$, 250mM NaCl, 1mM DTT, 0.05% (w/v) bovine serum albumin (BSA), 400 µg/mL tRNA and 0.1mM ATP. The polyadenylated small RNAs were purified and recovered by phenol-chloroform-isoamylalcohol (25:24:1) (from Sigma-Aldrich®, St.

Louis, MO, USA; Fisher Scientific International Inc., Hampton, NH, USA; and Amresco LLC, Solon, OH, USA, respectively) extraction and precipitated in an equal volume of isopropanol (Sigma-Aldrich). The polyadenylated small RNAs were then ligated with a 5' adapter (5'-N6accctcttggcacccactAAA-3') using T4 RNA ligase (Promega Corporation, Madison, WI, USA). First strand cDNA synthesis was carried using Reverse Transcriptase (Promega Corporation, Madison, WI, USA) with a poly-d(T) adapter as RT primer (5'-ATT CTA GAG GCC GAG GCG GCCGAC ATG-d(T)30 (A, G or C) (A, G, C or T). PCR was carried out using primers targeting 5' adapter and seed sequence on the poly-d(T) adapter. PCR amplification was carried out at annealing temperature of 50°C and total PCR cycles are 25. PCR product with size about 100 bp was purified from electrophoresis gel. The PCR products were then ligated into a commercial cloning vector, pGEM-T Easy vector (Promega Corporation, Madison, WI, USA), followed by *E. coli* transformation, colonies screening and plasmid purification prior to DNA sequencing. The DNA sequences were then processed to remove the vector sequence as well as the adapter sequences and were searched against databases of NCBI GenBank and miRBase (http://www.mirbase.org/) for homologous sequences.

In the method described by Ho et al. (2006), hybrid adapters that were used had the sequences as follows: 5' adapter (5'-N6accctcttggcacccactAAA-3') and 3' adapters (5'-UUUaccaggcacccagcaatgN3-3') in which, upper case: RNA; lower case: DNA; N6: amino 6-carbon spacer for blocking the 5' end; and N3: amino modifier for blocking the 3' end. The hybrid adapters were ligated to the isolated small RNAs using T4 RNA ligase (Promega Corporation, Madison, WI, USA) and the ligation product was amplified by RT-PCR. PCR amplification products with sizes of about 60 to 70 bp were purified from agarose gels after electrophoresis. The cDNA was then cloned into the pGEM-T Easy vector (Promega Corporation, Madison, WI, USA). Positive clones (white colonies) were then sequenced and the small RNA sequences were used to search

for homologous sequences in the GenBank (NCBI) and miRBase (http://www.mirbase.org/) databases.

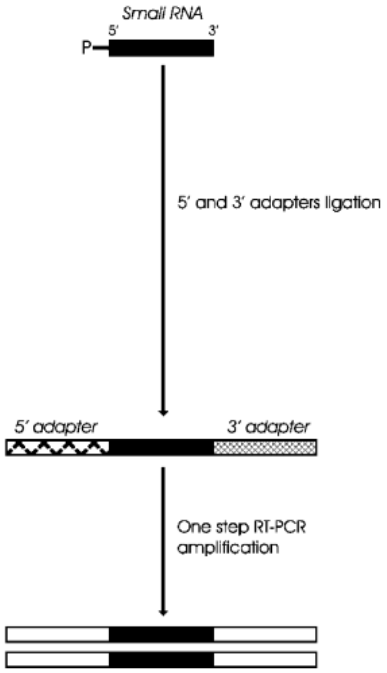


Figure 3.2: Direct cloning of small RNA by one-step adapters-ligation and RT-PCR amplification (Ho et al., 2006)

3.4.2 Illumina high-throughput sequencing approach

Illumina sequencing technology was used for high-throughput in-parallel sequencing of small RNA cDNA libraries. This service was provided by the Beijing Genomics Institute (BGI), Shenzhen, China (http://www.genomics.cn/index.php).

3.4.2.1 Library construction and sequencing

Small RNA library construction for Illumina sequencing was carried out using Illumina's kit according to the manufacturer's recommendations. The procedures of small RNA cDNA library construction are shown in Figure 3.3. In brief, small RNAs of 16 to 28 nt were recovered from high resolution gel (15% (w/v) PAGE) and then ligated

with 5’ and 3’ Illumina adapters using T4 RNA ligase. The small RNA-adapters ligation products were amplified by RT-PCR using Illumina’s small RNA primer set and the cDNA was sequenced using Illumina Genome Analyzer, GA IIx platform following the manufacturer’s instructions.

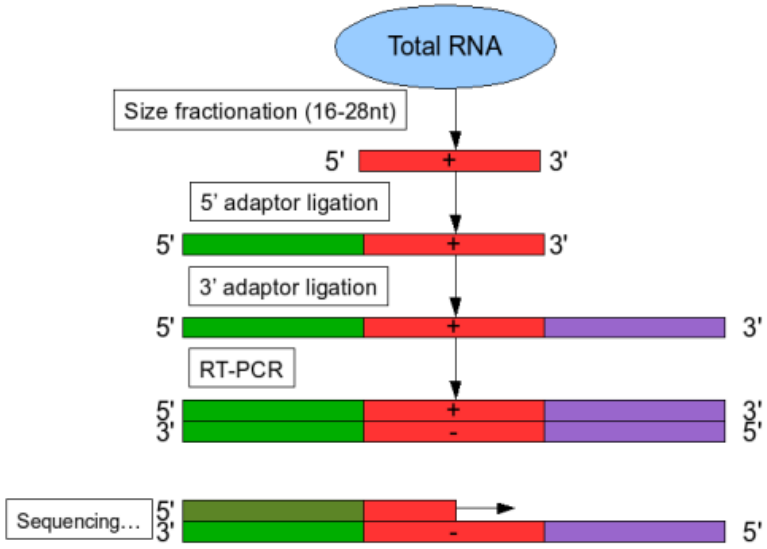


Figure 3.3: Procedures of small RNA cDNA libraries construction for sequencing using Illumina high-throughput sequencing platform (Source: Beijing Genomics Institute, BGI, http://www.genomics.hk/SamllRna.htm)

3.4.2.2 Sequence generation and processing

Sequence reads generated from the Illumina GA IIx platform and provided as FASTQ format files by BGI, were filtered for adapter, low quality reads and contaminant sequences. The subsequent bioinformatics analyses were carried out using workstations and technical supports in the University of Malaya. The Illumina small RNA sequence reads of banana root described in this study are available under NCBI’s BioProject accession PRJNA246442. Short Read Mapping Package (SHRiMP version

2.2.3, http://compbio.cs.toronto.edu/shrimp/) (Rumble et al., 2009) was used to map the filtered reads to Rfam 11.0 (http://rfam.xfam.org/) (Griffiths-Jones et al., 2005) with no mismatch allowed. The small RNA sequences that matched with the non-protein-coding RNA sequences including rRNA, tRNA, small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA) in Rfam were removed from further analysis. Sequences below 19 and above 24 nucleotides were also removed from the data set and not used in further analysis.

3.4.2.3 Functional annotation of small RNAs

3.4.2.3.1 Plant orthologous microRNA

The clean reads from small RNA sequencing (section 3.4.2.2) were used to search for orthologous miRNAs. A non-redundant dataset was produced from the Plant microRNA Database (PMRD June 11, 2012 update, http://bioinformatics.cau.edu.cn/PMRD/) (Zhang et al., 2010c) and used as an orthologous plant miRNA reference set. A Python script (Appendix C) was used to align the small RNAs with the reference set and only one mismatch was allowed for the alignment. The orthologus miRNA matches were named according to the orignal gene name except the species name, which was replaced with *Musa acuminata* (mac-).

3.4.2.3.2 *Musa*-specific microRNA

After removal of sequences that matched to entries in the Rfam 11.0 and PMRD databases (June 11, 2012 update), the remaining small RNA sequences were used for the prediction of *Musa*-specific miRNAs, which were sequences not reported as miRNA in species other than *Musa* spp., from the *Musa* reference genome (D'Hont, et al., 2012). MiRDeep2 tool (Friedlander et al., 2012) was used and criteria for miRNA prediction from the plant genome were set according to Meyers et al, (2008). The small RNA

sequences were mapped to the reference A-genome (D'Hont, et al., 2012) by using Bowtie (Langmead et al., 2009) in the miRDeep2 pipeline. Using default parameters, three hundred nucleotides spanning the matched small RNA sequences in the reference genome were excised for stem-loop structure prediction using RNAfold (Hofacker, 2003). Then, the p-values were calculated for the miRNA precursors predicted by miRDeep2 using Randfold (Bonnet et al., 2004b). The criteria used to annotate *Musa*-specific candidate miRNA were: (i) transcripts (unigenes) could form a stem-loop structure of 75-nt with a bulge-loop size less than 6 nt; (ii) small RNA reads fell within the stem region of the precursor; (iii) a maximum of 3 mismatches was allowed between the miRNA:miRNA* duplex; miRNA and miRNA* formed a duplex with 3' overhangs; (iv) predicted minimum folding energy (MFE) was between -15 kcal/mol to -47.2 kcal/mol. *Musa*-specific miRNAs were arbitrarily named starting at '1' and using the miRBase gene nomenclature (Griffiths-Jones et al., 2006).

3.4.2.4 MicroRNA quantification and differential expression

MiRNA read counts were normalized to tags per million (TPM) in order to compare miRNA expression across libraries (CTR, TR100 and TR300). The TPM was calculated as follows: normalized expression, TPM = (actual miRNA count/total clean read) x 1,000,000. Differential miRNA expression was determined using DEGSeq (Wang, et al., 2010a) with $\log_2$Fold change ≥ 1 and cut-off false discovery rate (FDR) value less than 0.05. Expression change was plotted in heatmaps using gplots package (http://cran.r-project.org/web/packages/gplots/index.html) from bioconductor in R.

3.4.2.5 MicroRNA target prediction

The salt-stressed banana root transcriptomes (mRNA-Seq) in this study were used as reference data to predict miRNA targets, according to the method reported by

(Davey et al., 2013). MiRNA targets were predicted using psRNAtarget online server (http://plantgrn.noble.org/psRNATarget/) (Dai & Zhao, 2011) with default options.

CHAPTER 4: RESULTS

4.1 RNA quality and yield

This study is registered with the NCBI's BioProject database as accession PRJNA 246442 and the samples are registered in the NCBI's BioSample database with accession names SAMN0267607, SAMN0267608 and SAMN0267609 for 'Berangan' plantlets samples named CTR (control), TR100 (100 mM NaCl treatment) and TR300 (300 mM NaCl treatment), respectively.

Using a CTAB-based RNA isolation method (modified from Kiefer et al. (2000), was able to produce intact total nuleic acids from *in vitro* banana roots (Figure 4.1) RNA showed 28S rRNA band intensity about twice that of the 18S rRNA as shown in Figure 4.2.

Figure 4.1: *In vitro* banana cultivar 'Berangan' (AAA genome) plantlets

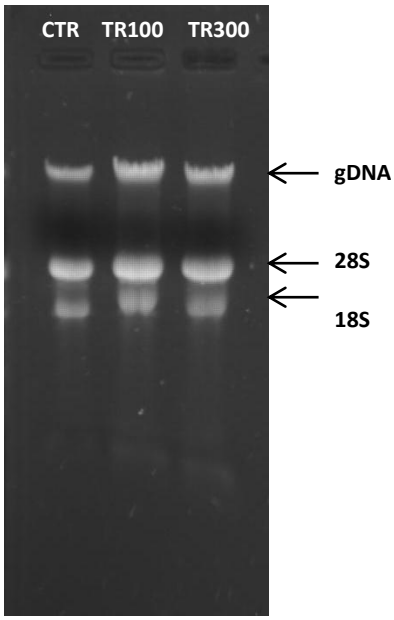


Figure 4.2: Gel electrophoresis image for banana root RNA samples. CTR: untreated control (0 mM NaCl); TR100: 100 mM NaCl-treated banana roots; TR300: 300 mM NaCl-treated banana roots

4.2 Banana salt stress transcriptomes

4.2.1 Purity and integrity of RNA used for RNA-Seq

Both samples (CTR and TR300) used for mRNA-Seq showed A_{260nm}/A_{280nm} of about 2.0 and A_{260nm}/A_{230nm} greater than 1 (Table 4.1). The 28S:18S rRNA ratio of the isolated RNA was about 2 and both the samples showed RIN values higher than 8 (Table 4.2 and Figure 4.3).

Table 4.1: Optical density readings for samples used for mRNA-Seq

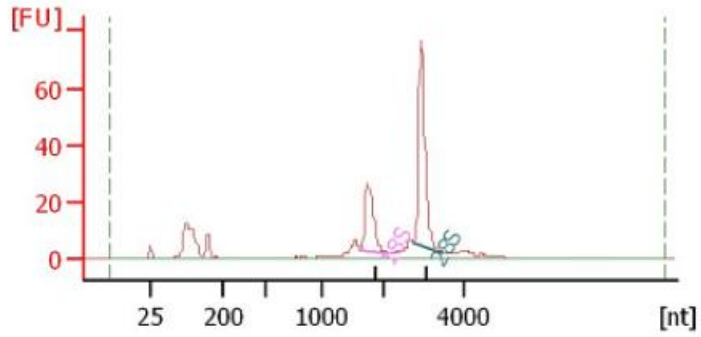
Sample	A_{260nm}/A_{280nm}	A_{260nm}/A_{230nm}
CTR	2.01	1.33
TR300	1.95	1.39

Table 4.2: Output of Agilent's Bioanalyzer assay for RNA samples used for mRNA-Seq

Sample	Concentration (ng/µl)	Volume (µl)	Amount (µg)	28S:18S	RIN
CTR	438	81	35.478	2.1	8.7
TR300	780	81	63.180	2	8.9

28S: 28S ribosomal RNA; 18S: 18S ribosomal RNA; RIN: RNA integrity number

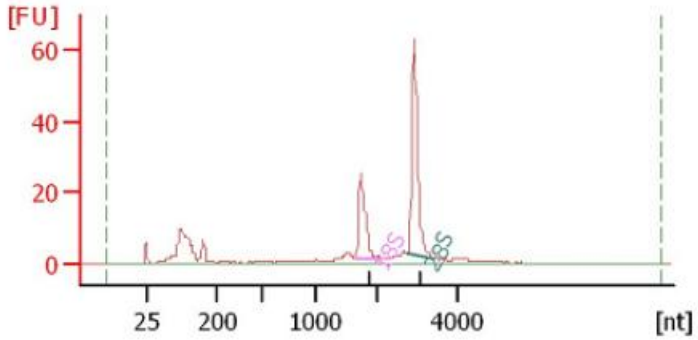
A – CTR



Overall Results for sample 3 :

RNA Area:	336.2
RNA Concentration:	219 ng/µl
rRNA Ratio [28s / 18s]:	2.1
RNA Integrity Number (RIN):	8.7 (B.02.07)
Result Flagging Color:	
Result Flagging Label:	RIN: 8.70

B – TR300



Overall Results for sample 4 :

RNA Area:	240.7
RNA Concentration:	156 ng/µl
rRNA Ratio [28s / 18s]:	2.0
RNA Integrity Number (RIN):	8.9 (B.02.07)
Result Flagging Color:	
Result Flagging Label:	RIN: 8.90

Figure 4.3: Bioanalyzer (Agilent 2100) RNA analysis. A: CTR RNA sample; B: TR300 RNA sample. CTR: untreated control (0 mM NaCl); TR300: 300 mM NaCl-treated banana roots

4.2.2 Sequence quality information

Over 12.2 million and 11.2 million, 90-bp paired-end reads were generated from non-stressed control and 300 mM NaCl treated roots, respectively (Table 4.3). All together these reads accounted for 2.1 Gb of transcriptomic sequence data. Both transcriptomes contained more than 92% of high quality reads and more than 93% of high quality bases (Phred score Q>20) (Table 4.3).

Table 4.3: Paired-end transcriptome sequencing (mRNA-Seq) output

	CTR	TR300
Total number of reads	12,279,060	11,274,596
Average read length	90	90
Total number of HQ reads*	11,352,903	10,504,837
Percentage of HQ reads	92.46%	93.17%
Total number of bases	1,105,115,400	1,014,713,640
Total number of HQ bases**	1,037,037,224	956,651,654
Percentage of HQ Bases	93.84%	94.28%

CTR: untreated control (0 mM NaCl); TR300: 300 mM NaCl-treated banana roots; HQ: high quality

* >70% of bases in a read with Phred Quality Score > 20

** Bases with Phred Quality Score > 20

4.2.3 Sequence assembly and clustering

The 11,352,903 and 10,504,837 high quality reads (Table 4.4) from each data set were assembled into 69,441 and 74,525 contigs representing 49,576 and 56,572 unigenes from the control and NaCl-treated roots, respectively and used to form a set of non-redundant representative transcripts containing 31,390 unigenes (‘reference assemblies’). Mean coverage per base (CPB) per unigene of the assembled transcriptomes is 45.33X in CTR and 38.57X in TR300, as shown in Table 4.5.

More than 80% of the contigs assembled in CTR and TR300 were sized between 100 to 400 nt (Table S2). These banana transcriptomic sequences showed increased mean length and N50 after further assembly and clustering into unigenes (Figure 4.4 Table S3 and Table S4). More than 80% of the assembled scaffolds and unigenes are of high quality with zero ambiguous nucleotides ('gap') (Figure 4.5 and Table S5). The reference assemblies have a mean length of 517-nt and N50 of 669-nt as shown in Table 4.4.

Table 4.4: *De novo* assembly of banana root transcriptomes

	CTR			**TR300**			**All**		
Category	Total count	Mean length (nt)	N50 (nt)	Total count	Mean length (nt)	N50 (nt)	Total count	Mean length (nt)	N50 (nt)
Contigs	69,441	263	315	74,525	265	313	-	-	-
Scaffolds	49,576	349	441	56,572	334	411	-	-	-
Unigenes	49,576	350	443	56,572	334	411	31,390	517	669

CTR: untreated control (0 mM NaCl); TR300: 300 mM NaCl-treated banana roots; All: clustering of both CTR and TR300 into a reference assembly

Table 4.5: Coverage of the assembled transcriptomes

Sample	**Number of Reads***	**Number of Bases****	**Coverage*****
CTR	8,738,948	17,350,046	45.33X
TR300	8,099,430	18,899,141	38.57X

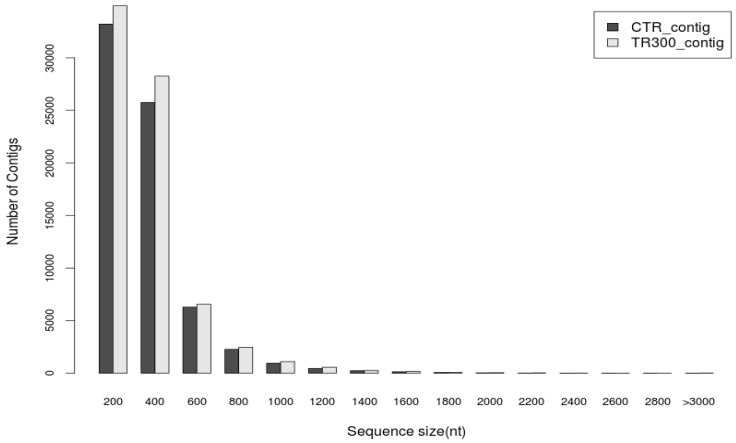
CTR: untreated control (0 mM NaCl); TR300: high salinity-stressed (300 mM) banana roots

*Number of reads that can be assembled into unigenes

**Number of bases in the assembled transcriptome

***Coverage = Read count x Read length / Number of bases in the assembled transcriptome

A – Contigs



B – Scaffolds

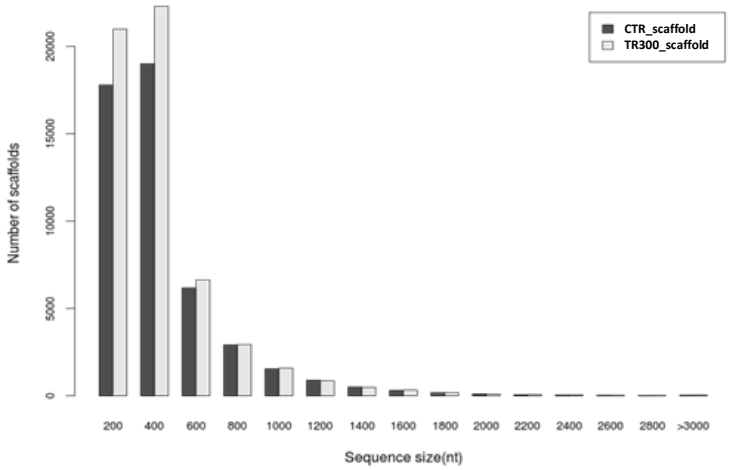
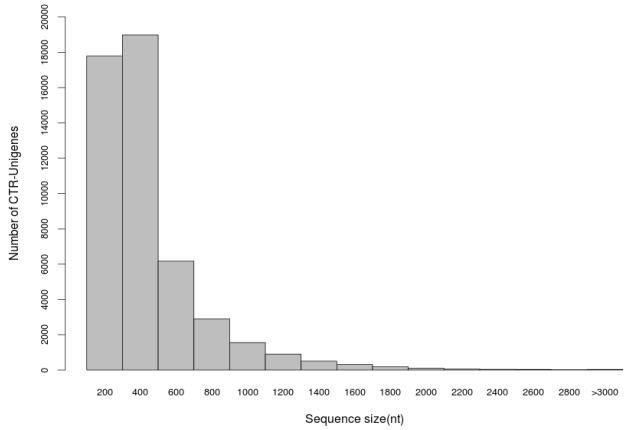
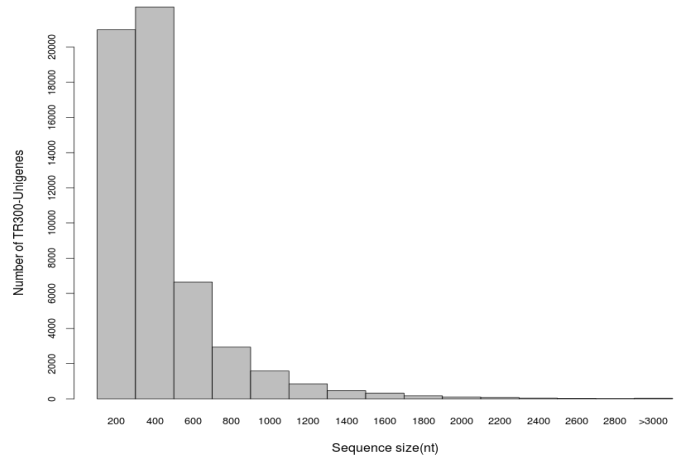


Figure 4.4: Overview of the length distribution of the assembled contigs, scaffolds and unigenes. A: Contigs in CTR and TR300; B: Scaffolds in CTR and TR300; C: Unigenes in CTR; D: Unigenes in TR300; E: The assembled unigenes from both CTR and TR300 were clustered to form a set of reference assemblies. CTR: untreated control (0 mM NaCl); TR300: 300 mM NaCl-treated banana roots

C – CTR



D – TR300



E – All.unigenes

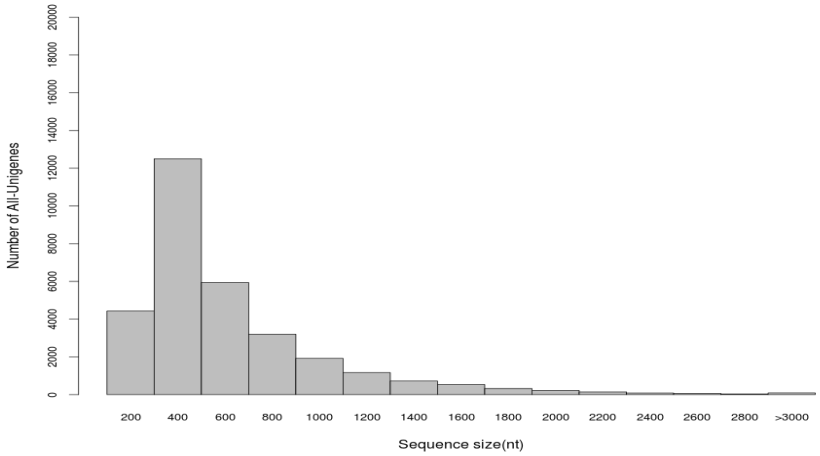


Figure 4.4: Continued

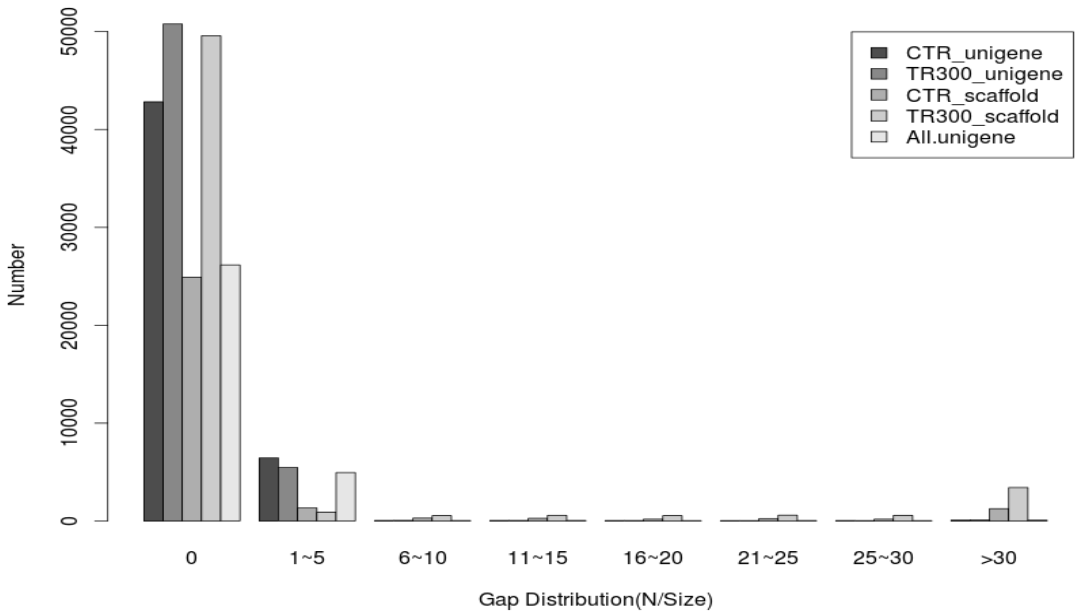


Figure 4.5: Quality of unigenes indicated by gap percentage (N/size of scaffolds and unigenes in the *de novo* assembled banana root transcriptomes. N: ambiguous nucleotide; CTR_scaffold: assembled scaffolds in CTR sample; TR300_scaffold: assembled scaffolds in TR300 sample; CTR_unigene: assembled unigenes in CTR sample; TR300_unigene: assembled unigenes in TR300 sample; All.unigenes: assembled unigenes after clustering of both set CTR and TR300 unigenes. CTR: untreated control (0 mM NaCl); TR300: 300 mM NaCl-treated banana roots

4.2.4 Mapping against the reference banana genome

Over 99.5% of the *de novo* assembled unigenes could be mapped to the *Musa* genome (D'Hont et al., 2012) (Figure 4.6 and Table 4.6). These unigenes were distributed along the 11 chromosomes and unrandom sequences of the *Musa* reference genome. The number of unigenes that could be mapped to each chromosome ranged from 2,000 to 3,200, or 6 to 10% of the unigenes (Table 4.6).

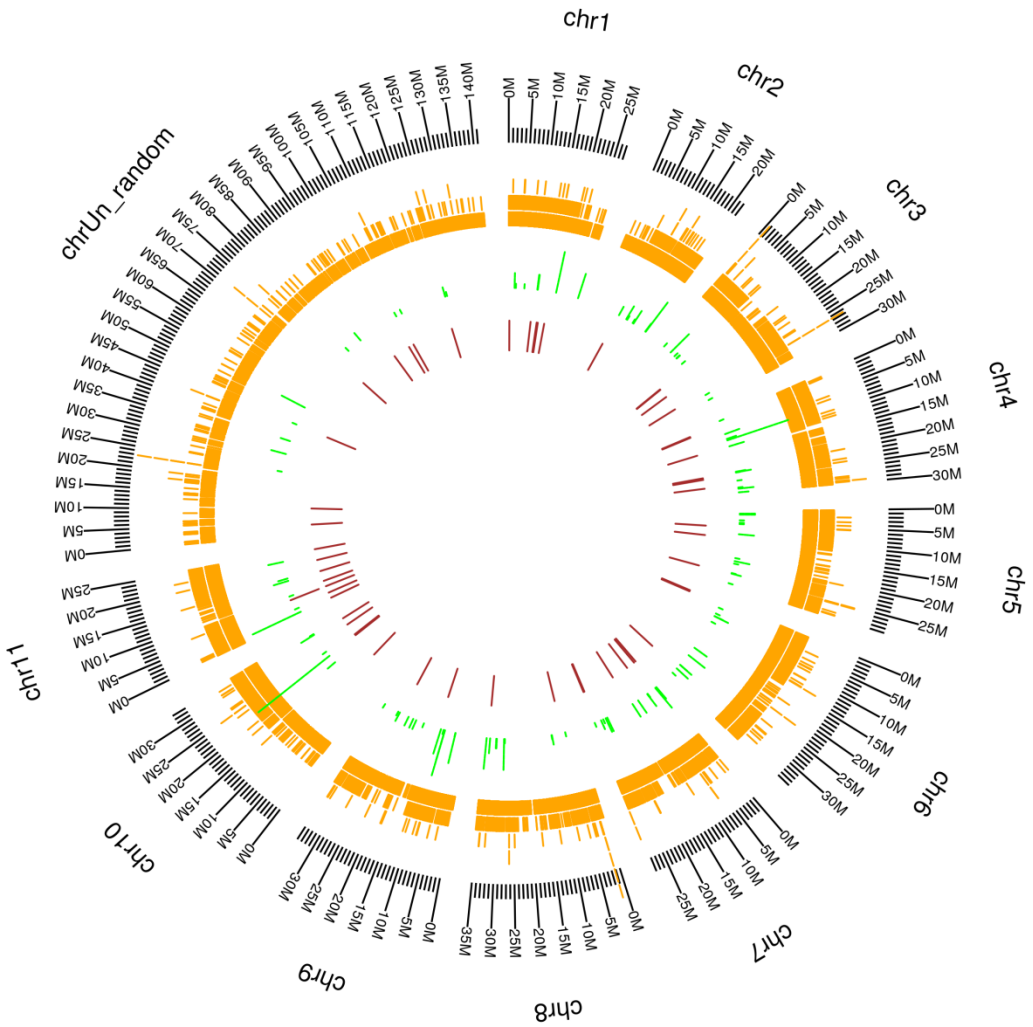


Figure 4.6: Distribution of expressed transcripts on a reference *Musa* Genome (D'Hont et al., 2012). Black (outer ring): chromosomes of the reference *Musa* A-genome, scale in Mb; Orange: *de novo* assembled unigenes; Green: orthologous miRNAs; Maroon: *Musa*-specific miRNAs. Note: Mapping of miRNAs to *Musa* reference genome is discussed later in section 4.3.2.3

Table 4.6: Mapping of *de novo* assembled unigenes to banana genome sequences

Chromosome	**Number**	**Percentage (%)**
1	2,552	8.13
2	2,021	6.44
3	2,886	9.2
4	2,848	9.08
5	2,594	8.27
6	3,204	10.21
7	2,545	8.11
8	3,026	9.65
9	2,608	8.31
10	2,732	8.71
11	2,269	7.23
Unrandom	1,965	6.26
Mapped	31,250	99.56
Unmapped	140	0.54

4.2.5 Functional annotation of the unigenes

Of the 31,390 clustered unigenes, 70.8% (22,231) have a BLAST hit in the GenBank non-redundant protein database (Nr) and 47.5% (14,913) in the UniProt database (Table 4.7). A majority of the BLAST hits are *Oryza sativa* protein sequences in the Nr database, followed by *Vitis vinifera*, *Glycine max*, *Populus trichocarpa*, *Arabidopsis thaliana* and *Zea mays* (Figure 4.7). About 56% or 17,617 of the assembled unigenes could be assigned to at least one Gene Ontology (GO) term, 30% or 9,449 to 25 categories in Eukaryotic Orthologous Groups of Proteins (KOG), and 22.2% or 6,997 to 136 pathways in Kyoto Encyclopaedia of Genes and Genomes (KEGG) (Table 4.7).

A majority of the unigenes in the reference assembly derived from banana root transcriptomes were assigned to cellular process and metabolite process terms in the biological process category, binding and catalytic activity terms in the molecular function category, and cell, organelle and membrane terms in the cellular component category (Figure 4.8). Annotation with KOG showed that the *de novo*-assembled banana root transcriptomes in this study were classified most abundantly in the general function prediction group, followed by posttranslational modification, protein turnover, chaperones group, signal transducer meachanisms group and the intracellular trafficking, secreation, and vesicular transport group (Figure 4.9). Meanwhile, gene annotation using KEGG pathways showed that the banana root reference assembly in this study was assigned most abundantly in the starch and sucrose metabolism and purine metabolism pathways (Figure 4.10).

Table 4.7: Functional annotation of banana root transcriptome

	Number of unigene	Percentage (%)
Total	31,390	100
Nr	22,231	70.8 %
UniProt	14,913	47.5%
GO	17,617	56.1%
KEGG	6,997	22.2%
KOG	9,449	30%

Nr: GenBank non-redundant protein database; UniProt: Universal Protein Resource; GO: Gene Ontology; KEGG: Kyoto Encyclopaedia of Genes and Genomes; KOG: Eukaryotic Orthologous Groups of Proteins

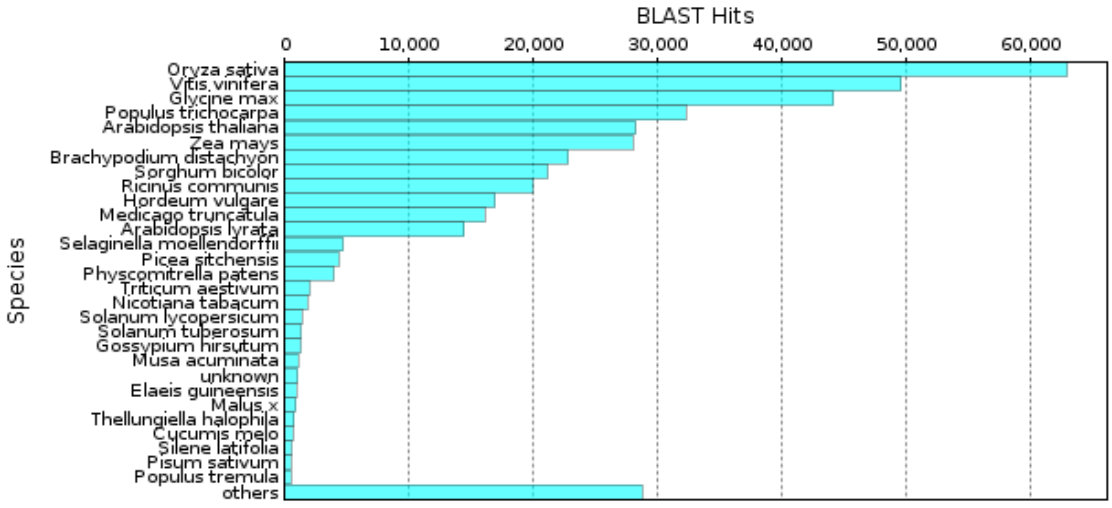


Figure 4.7: BLAST hits of the *de novo* assembled unigenes. Assembled unigenes from banana root transcriptomes searched against sequences from various plant species deposited in the GenBank non-redundant protein database

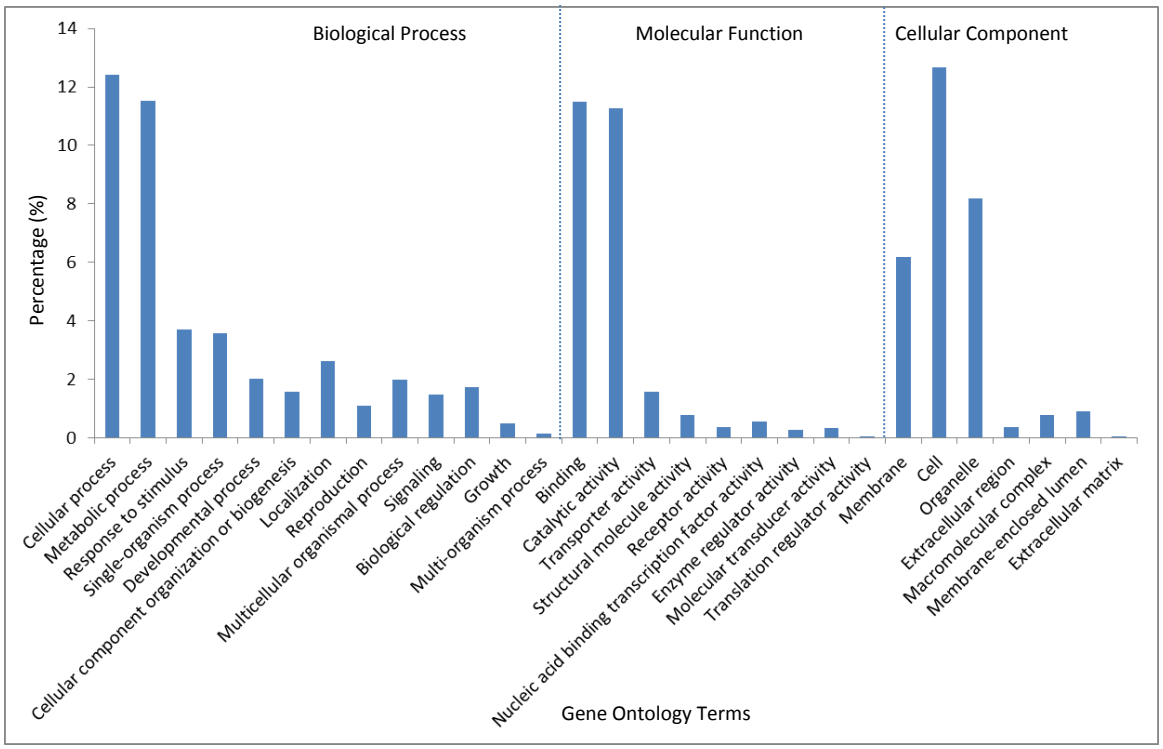


Figure 4.8: Gene Ontology (GO) classification of the *de novo* assembled unigenes of banana root transcriptomes

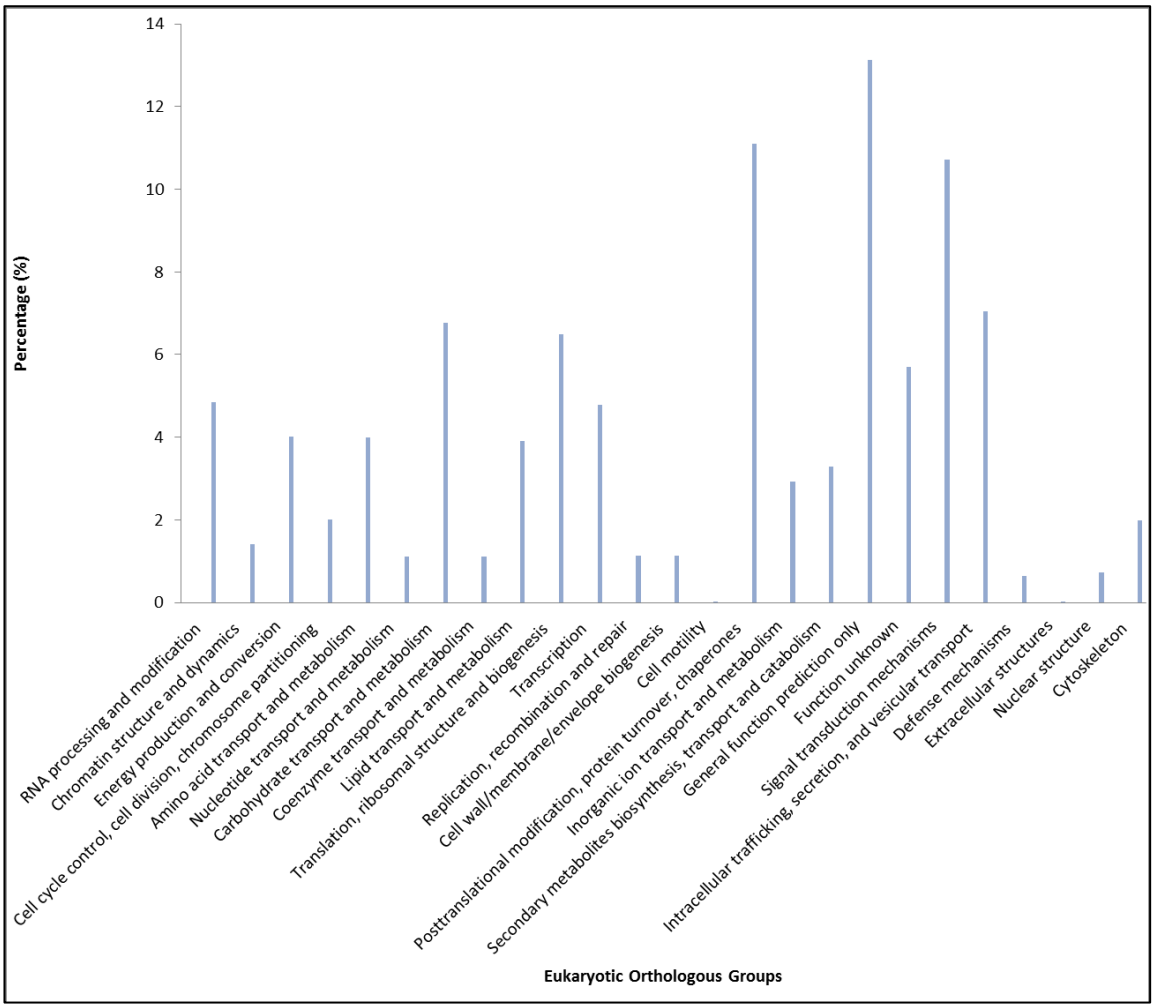


Figure 4.9: Eukaryotic orthologous groups (KOG) classification of the *de novo* assembled unigenes in the banana root transcriptomes

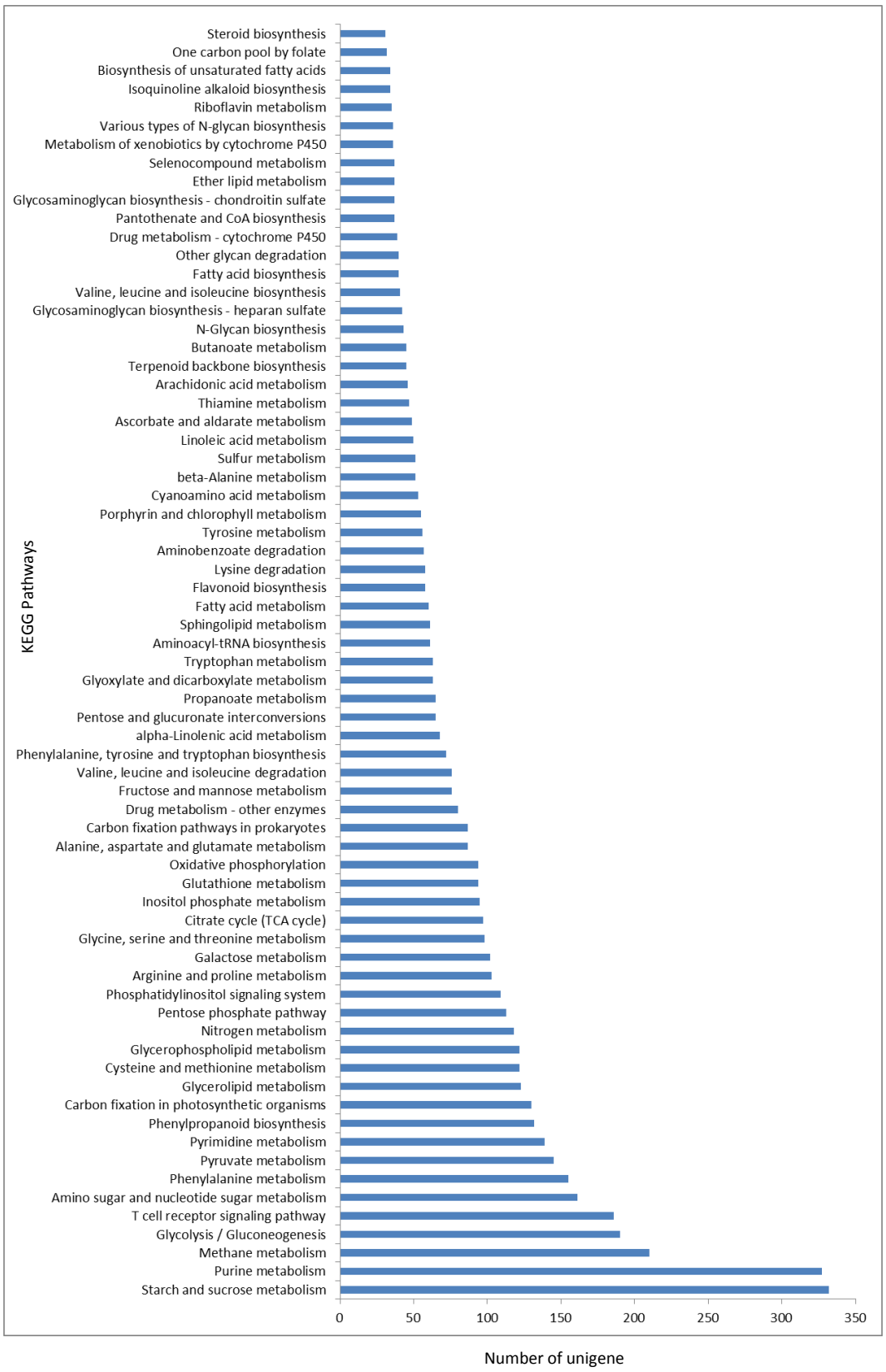


Figure 4.10: KEGG pathway assignment. The *de novo* assembled unigenes of banana root transcriptomes were assigned with KEGG pathways. Only KEGG pathways with more than 30 unigenes assigned are presented

4.2.6 Transcripts quantification and differential expression

After being normalized to transcript per million (TPM), 9.5% or 2,993 of the *de novo* assembled unigenes were observed to be differently expressed in salt-stressed banana root with $\log_2(CTR/TR300) \geq 1$ and False Discovery Rate < 0.05 (Table 4.8 and Figure 4.11). More transcripts (unigenes) were up-regulated (57.5% or 1,720) than down-regulated (42.4% or 1,273) in the banana roots upon 300 mM NaCl treatment (Table 4.8).

Among the differentially expressed unigenes, the largest proportion was assigned to the catalytic activity term (molecular function) followed by metabolic process (biological process), cellular process (biological process), cell (cellular component) and binding (molecular function) terms of the Gene Ontology (Figure 4.12). Meanwhile, the non-differentially-expressed unigenes have best matches that are mainly not annotated. The annotated matches were assigned chiefly to cell, membrane and organelle terms (cellular component category) of the Gene Ontology (Figure 4.12).

A majority of the differentially expressed unigenes in salt-stressed banana roots in this study were assigned to carbohydrate transport and metabolism, general function prediction, and post-translational modifications, protein turnover, chaperones of the Eukaryotic Orthologous Groups of Proteins (KOG) (Figure 4.13). In KEGG pathways assignment, the differentially expressed unigenes were predominantly classified in the starch and sucrose metabolism pathways, followed by phenylalanine metabolism, cysteine and methionine metabolism, and phenylpropanoid biosynthesis pathways (Figure 4.14).

Table 4.9 and Table 4.10 show partial lists of differentially-expressed unigenes with the highest expression fold change in the salt-stressed banana roots. Among the highly up-regulated genes are nodulin-like protein, histone proteins, peroxidases, expansins, lipid transfer protein, inorganic phosphate, phosphate transporter, cell wall-

associated hydrolase, BTB-POZ domain-containing protein, xylanase inhibitor, glutamate synthase, late embryogenesis abundant (LEA) protein and a number of hypothetical proteins (Table 4.9). While among the genes that are highly down-regulated are pectin esterase-1 precursor, AP2 ERF domain-containing transcription factor, WD-40 repeat family protein, permease, HSP70, pyruvate kinase, alcohol dehydrogenase, gibberillin 2-oxidase, ethylene response factor, pathogenesis-related protein, hexose transporter, dirigent-like protein, phytocyanin and many unannotated transcripts and hypothetical proteins (Table 4.10).

Table 4.8: Differentially-expressed unigenes in salt-stressed banana roots

	Number of unigenes	**Percentage (%)**
Total	31,390	100
Differentially expressed unigenes*	2,993	9.5
Up-regulated	1,720	57.5
Down-regulated	1,273	42.5

*Differentially-expressed genes with $\log_2$ fold change ≥ 1 and false discovery rate (FDR) < 0.05

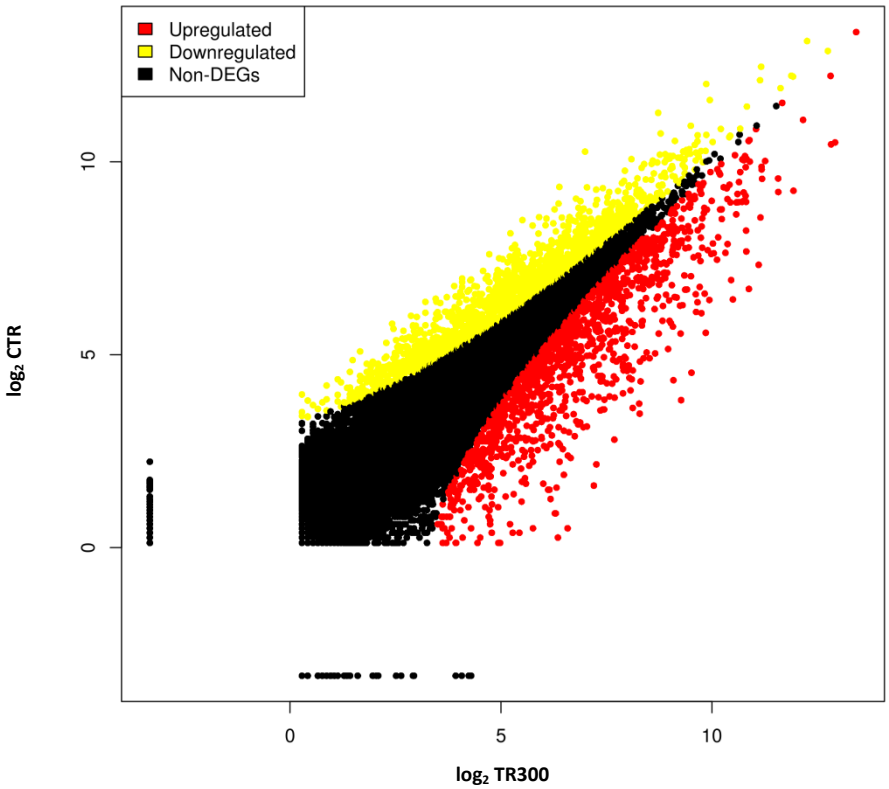


Figure 4.11: Expression changes of unigenes in the salt-stressed banana roots in comparison to the control. Red: up-regulated unigenes with $\log_2$ fold change ≥ 1; yellow: down-regulated unigenes with $\log_2$ fold change ≥ 1; black: non-differentially-expressed unigenes. CTR: control; TR300: 300 mM NaCl treatment

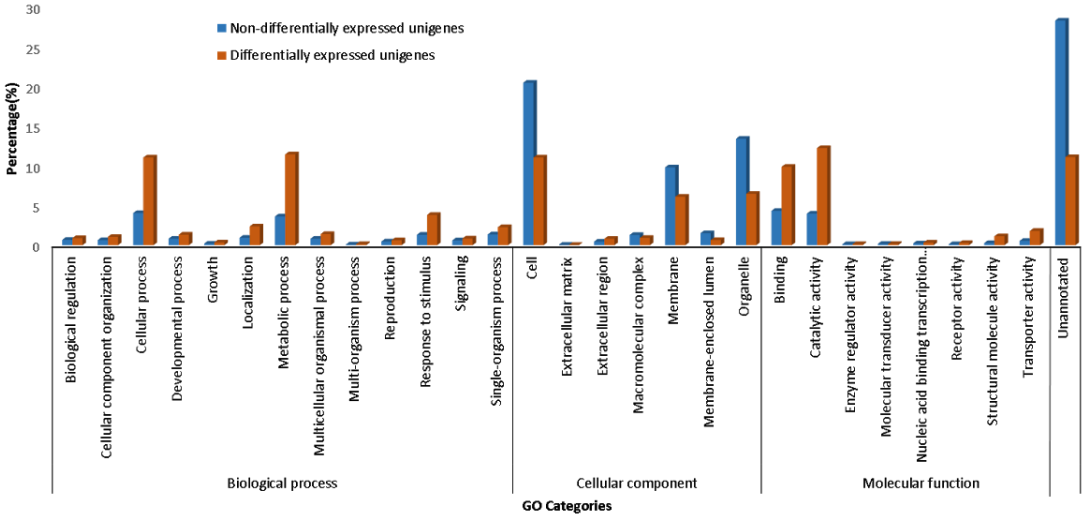


Figure 4.12: Gene Ontology (GO) assignments for unigenes differentially-expressed and non-differentially expressed in the salt stressed banana roots

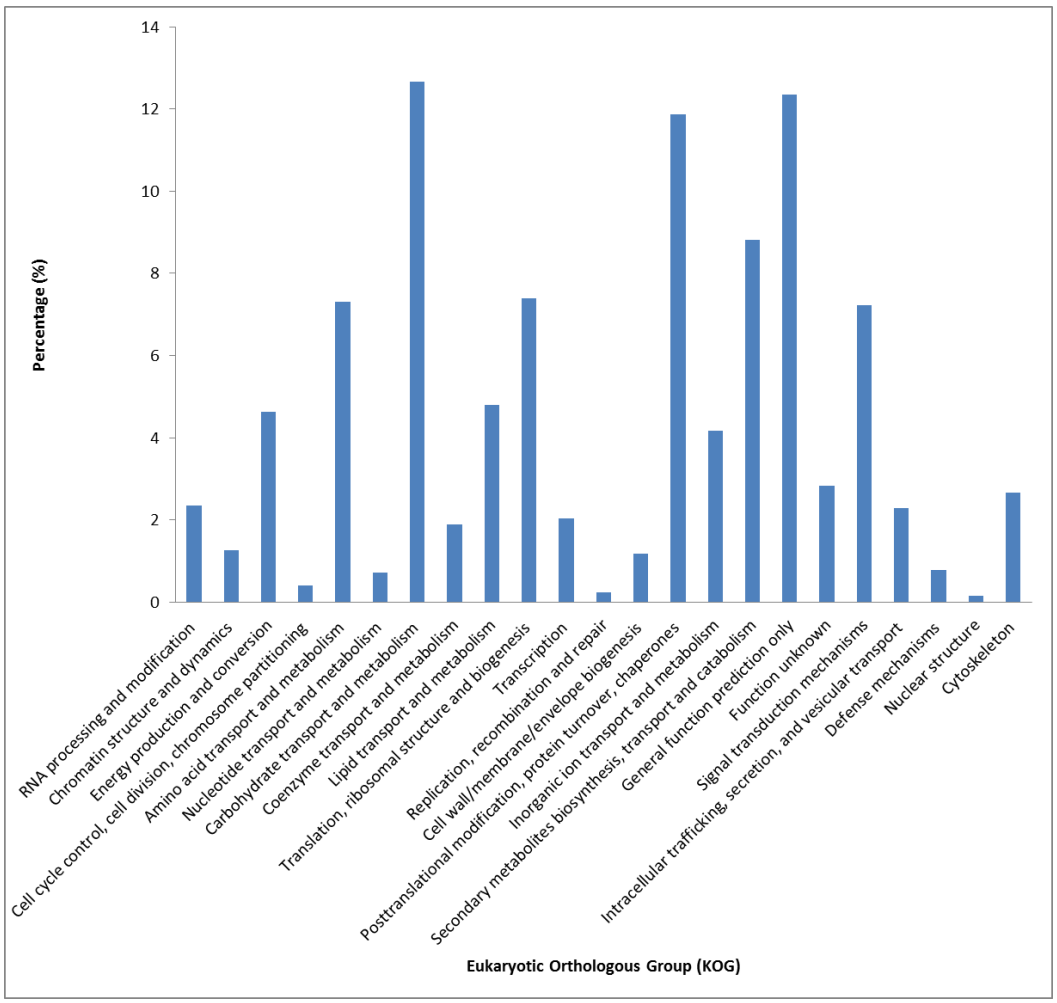


Figure 4.13: Eukaryotic Orthologous Group (KOG) annotation of the differentially expressed unigenes in salt-stressed banana roots

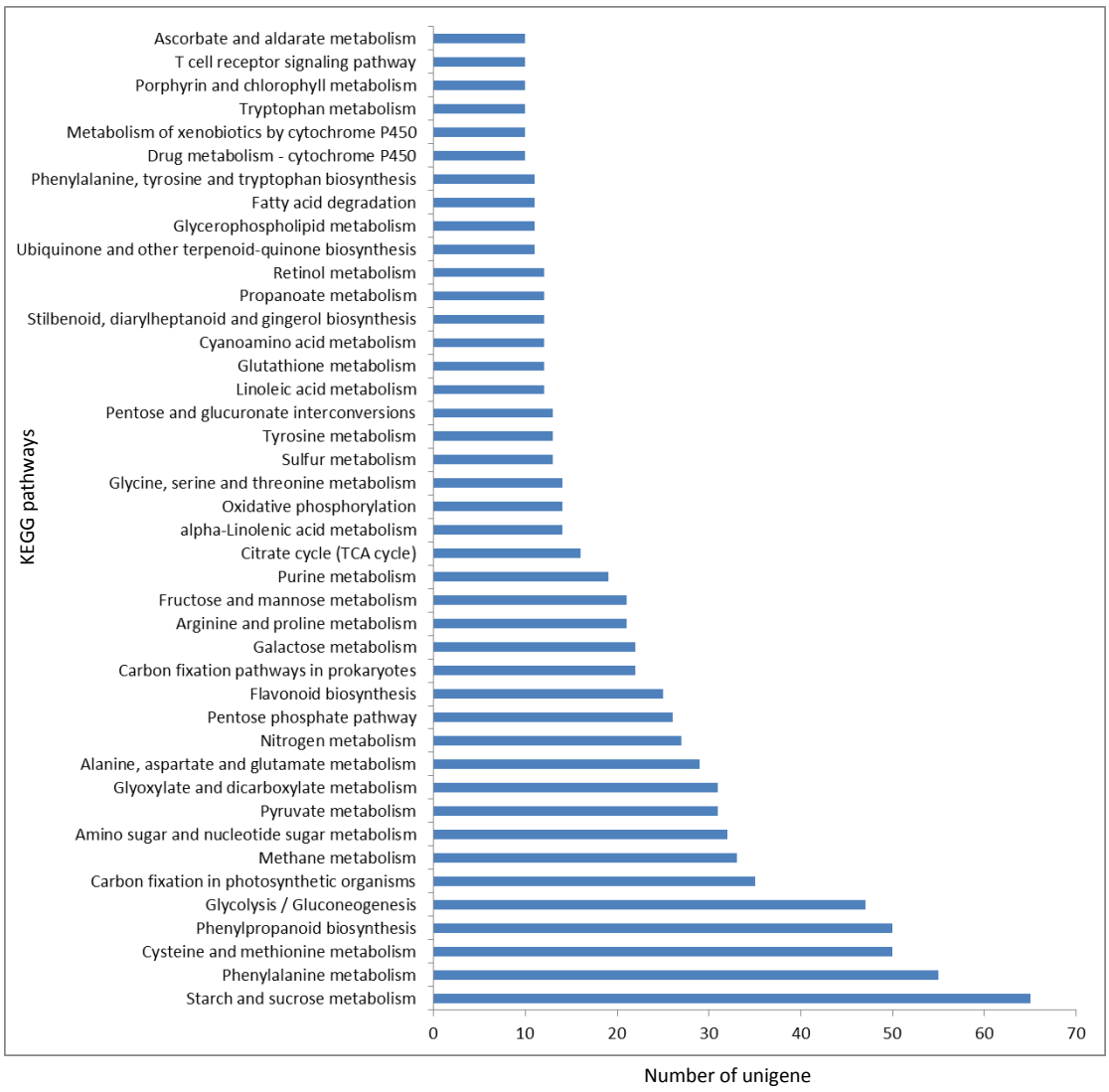


Figure 4.14: KEGG pathway assignment of Differentially expressed unigenes in salt-stressed banana roots. Only KEGG pathways with at least 10 unigenes assigned are presented

Table 4.9: Partial list of differentially-expressed unigenes with the highest expression fold change, up-regulated in salt-stressed banana roots

Unigene	log_2 Fold change	Putative identity
C70598	3.686298747	Hypothetical protein
CL966Contig1	3.423264341	Nodulin-like protein
CL1Contig7761	3.390097477	N/A
CL324Contig2	3.381685238	PNGase
CL1Contig6044	3.375435816	Inorganic phosphate
C132686	3.332290681	Galactosyltransferase family protein
CL1Contig2862	3.271439858	Conserved domain protein
C135398	3.230619263	Fasciclin-like arabinogalactan protein
C79164	3.230619263	Histone H3
C109102	3.216813464	Phosphate transporter
CL9046Contig1	3.160229935	Histone H2A
CL1419Contig1	3.148157103	Receptor-like protein kinase
CL7275Contig1	3.140864611	Auxin-induced protein 5NG4-like
C121298	3.133757724	DNA-3-methyladenine glycosylase I
C91380	3.13108359	N/A
Contig723	3.074190104	Hypothetical protein
CL1Contig4954	3.024168386	Peroxidase 47-like
CL8180Contig1	2.997500435	Alpha-expansin 7
C114160	2.996731203	Hypothetical protein
C134662	2.98227839	Hypothetical protein
C117792	2.978783064	Cell wall-associated hydrolase
scaffold6818	2.975805364	BTB/POZ domain-containing protein
C96944	2.973816811	ORF16-lacZ fusion protein
CL1Contig1515	2.965299733	Ycf68 protein
C101212	2.963832723	Beta-tubulin
CL8448Contig1	2.94610513	Beta-expansin 3
scaffold5711	2.932277989	Protein phosphatase 2c 4
scaffold1398	2.930560348	N/A
Contig746	2.923274652	Cell wall-associated hydrolase
CL942Contig1	2.904154388	3-ketoacyl-synthase 4-like
CL7314Contig1	2.895533917	Pollen ole e 1 allergen and extensin family protein
CL5244Contig1	2.887704081	Lipid transfer protein
CL347Contig1	2.88168828	Xylanase inhibitor
CL1Contig4915	2.878086707	Late embryogenesis abundant protein

Table 4.9: Continued

Unigene	log_2 Fold change	Putative identity
CL1Contig501	2.861141817	N/A
CL1Contig6328	2.855110128	Aspartic proteinase nepenthesin-1 precursor
CL9171Contig1	2.844107462	N/A
scaffold5384	2.828248176	Homeobox-leucine zipper protein
scaffold3317	2.826620035	Cadmium zinc-transporting ATPase 3-like
C85860	2.820743469	N/A
CL1Contig7270	2.800178295	Peroxidase 3-like
scaffold1645	2.788913814	Hypothetical protein
CL3948Contig1	2.785834421	N/A
Contig1335	2.779485614	VAMP protein SEC22
C131698	2.775157858	Glutamine synthetase
CL1601Contig1	2.767218743	Hypothetical protein
CL1Contig5300	2.765451275	Hypothetical protein
CL1Contig6810	2.756247625	Aspartic proteinase nepenthesin-1
scaffold3601	2.75447025	Beta-glucanase precursor
scaffold5915	2.735208348	Galacturonosyltransferase-like 3-like

N/A: identity not available

PNGase: FN4-(N-acetyl-beta-glucosaminyl)asparagine amidase-like

The 50 most differentially-expressed unigens are shown

Table 4.10: Partial list of differentially-expressed unigenes with the highest expression fold change down-regulated in salt-stressed banana roots

Unigene	log_2 Fold change	Putative identity
C124675	-6.090283826	N/A
C125837	-6.079643944	Putative pectinesterase-1 precursor
CL1Contig2893	-5.593941767	Hypothetical protein
CL1Contig2841	-5.447100378	Hexose transporter
C145430	-5.41185722	AP2/ERF domain-containing transcription factor
C124777	-5.391647932	WD-40 repeat family protein
CL1Contig1647	-5.258919967	N/A
CL1Contig1234	-5.103354173	Amino acid permease
CL1Contig3633	-5.05246909	N/A
Contig1064	-4.981097692	N/A
CL1437Contig1	-4.916585662	70 kDa peptidyl-prolyl isomerase-like
CL797Contig1	-4.894148273	F-box and WD-40 domain
CL9167Contig1	-4.887064417	Lignin-forming anionic peroxidase
CL1Contig3838	-4.856578352	N/A
CL1908Contig2	-4.817049988	N/A
CL1Contig6581	-4.815702305	Hypothetical protein
CL1000Contig1	-4.762161754	Heat shock protein 70
Contig1187	-4.748598465	Hypothetical protein
CL1Contig3371	-4.648610515	Hypothetical protein
C146862	-4.618605377	Heptahelical protein 4
C93870	-4.60700867	N/A
CL1Contig214	-4.603665872	N/A
CL5Contig2	-4.545544066	HIN1-like protein
C144766	-4.50265108	Amino acid transporter AAP4
CL1Contig4116	-4.396914621	Acyl acyl-carrier-protein thioesterase type B
C78114	-4.345376316	N/A
C143622	-4.328134175	Hypothetical protein
C126491	-4.323585842	Pyruvate kinase
CL1Contig2550	-4.296857743	Protein kinase Xa21
CL1Contig7831	-4.284348084	Alcohol dehydrogenase 1
CL1Contig1207	-4.249161001	N/A
CL1Contig503	-4.200617149	CBS domain-containing protein CBSX5-like
CL31Contig2	-4.170135545	Dirigent-like protein

Table 4.10: Continued

Unigene	log_2 Fold change	Putative identity
C50810	-4.163793578	N/A
CL289Contig1	-4.135702951	Calmodulin-binding protein
CL1Contig4602	-4.063613731	Phytocyanin
C36240	-4.036167277	N/A
scaffold7670	-4.001524888	Hypothetical protein
CL3071Contig1	-3.994751458	N/A
CL1Contig5347	-3.933073785	N/A
CL1Contig390	-3.916585662	Plant viral-response family
CL1049Contig1	-3.912622211	Pectinesterase/pectinesterase inhibitor 41-like
CL1Contig3634	-3.909891009	N/A
CL7278Contig1	-3.897327306	Gibberellin 2-oxidase
CL1663Contig1	-3.892023877	Stem-specific protein TSJT1-like
CL4762Contig1	-3.8853142	Ethylene response factor 2
CL2555Contig1	-3.864154223	Hypothetical protein
CL5975Contig1	-3.860504946	GEM-like protein 5
CL2483Contig1	-3.835665666	Pathogenesis-related protein 4
C137896	-3.831549558	Phenylalanine ammonia-lyase

N/A: identity not available

The 50 most diffentially-expressed unigenes are shown

4.2.7 Identification of simple sequence repeats (SSR)

Out of the 31,390 assembled unigenes, 6,469 contained short sequence repeats (termed cSSR) when a minimum of four contiguous repeat units for di-, tri-, tetra-, penta- and hexa-nucleotide were allowed. These 6,469 unigenes contained 8,428 SSRs, of which a majority were dinucleotide and trinucleotide repeats (Table 4.11).

Repeats of 4 to 25 units were observed in the dinucleotide SSRs, whereas trinucleotide SSRs were observed to have repeats of 4-15 units only (Table 4.12). The maximum number of repeat unit in tetra-, penta- and hexa-nucleotide SSRs was 12, 6 and 7 respectively (Table 4.12).

Table 4.11: Statistics of the SSRs derived from banana root transcriptomes

Searching items	Numbers
Total unigenes examined	31,390
Total size of sequences examined (nt)	16,246,671
Total number of SSRs identified	8,428
Number of SSR-containing unigenes	6,469
Number of unigene containing more than 1 SSR	1,491
Number of SSR present in compound formation	1,028
Frequency of SSRs	
Di-nucleotide	4,582 (54.4%)
Tri-nucleotide	3,623 (43.0%)
Tetra-nucleotide	107 (1.3%)
Penta-nucleotide	27 (0.3%)
Hexa-nucleotide	89 (1.1%)

Table 4.12: Number of repeat unit in different SSR repeat motifs

Number of repeat unit	Repeat motifs				
	Di-	**Tri-**	**Tetra-**	**Penta-**	**Hexa-**
4	3,147	2,206	69	19	69
5	488	744	22	5	14
6	176	327	10	3	5
7	129	193	3	0	1
8	130	75	1	0	0
9	103	33	1	0	0
10	74	18	0	0	0
11	60	18	0	0	0
12	67	2	1	0	0
13	48	3	0	0	0
14	35	2	0	0	0
≥15	125	2	0	0	0

4.3 Small RNA profiling

4.3.1 Direct cloning and sequencing approach

Separation of small RNAs from the high molecular weight RNA species after being resolved in 5% (w/v) low melting preparative agarose is shown in Figure 4.15. The use of method described by Ho et al. (2006) resulted in PCR products of ~45 bp as shown in Figure 4.16 and the resulting sequences are shown in Table S7. Using the same small RNA preparation steps (Figure 4.15), PCR amplification carried out using the method described by Fu et al. (2005) resulted in PCR products sized ~100bp (Figure 4.17) and the resulting sequences are shown in Table S8.

Clones selected randomly and sequenced from two libraries constructed using approaches described by Ho et al. (2006) and Fu et al. (2005) yielded sequences as shown in Table S7 and Table S8, respectively. The application of the method described by Ho et al. (2006) in this study produced sequences with an average length of 24-nt Classes of small RNA found in the cDNA libraries were mainly ribosomal RNA sequences (32.5%), adapter sequences (‘no insert’, 25%), and one to a few chloroplast sequences (10%), mRNA (10%), mitochondrial sequences (2.5%), retrotransposon (2.5%), *Musa* BAC sequence (2.5%) and unknown RNA sequences (12.5%). Only one miRNA sequence (2.5%) i.e. miR408 was identified (Table 4.13 and Table S7).

Application of the method described by Fu et al. (2005) in this study yielded sequences with a longer average size of about 40-nt. Transfer RNA (tRNA) appeared to be the most abundant small RNA cloned among all the sequences, representing 81.1% of clones surveyed. This small RNA cDNA library also contained degraded ribosomal RNAs, messenger RNA (mRNA), mitochondrial sequences and sequences with unknown identity, but not miRNA (Table 4.13 and Table S8).

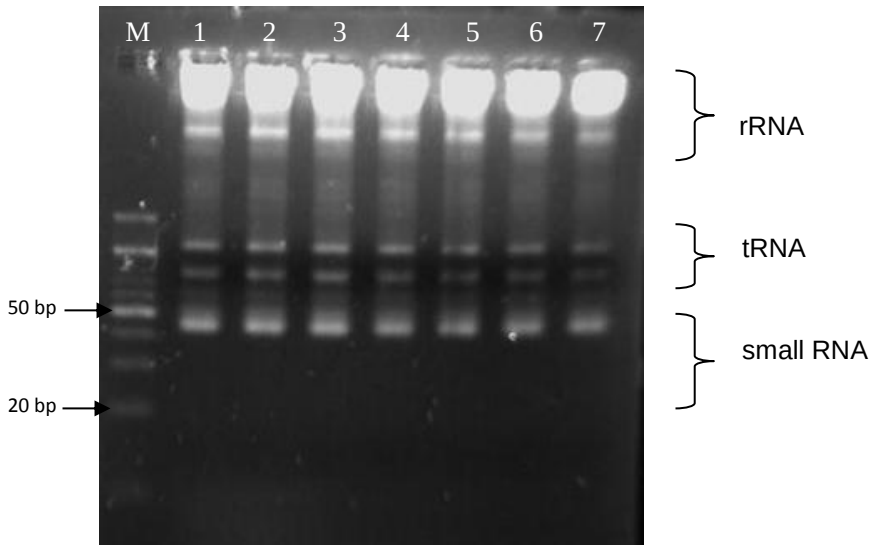


Figure 4.15: Separation of total RNA on 5% (w/v) low melting preparative agarose (Promega). M: 10 bp DNA marker (Fermentas); 1-7: RNA samples from mixed banana tissues

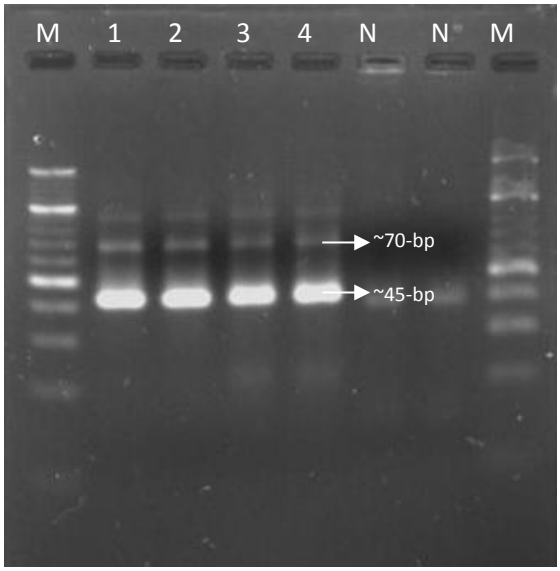


Figure 4.16: RT-PCR after 5' and 3' RNA-DNA hybrid adapters were ligated to the purified small RNA species. M: O'RangeRulerTM 10 bp DNA Ladder (Fermentas, Thermo Scientific); 1-4: PCR products; N: negative control for PCR reaction (faint product of ~45-bp is probable primer-dimers)

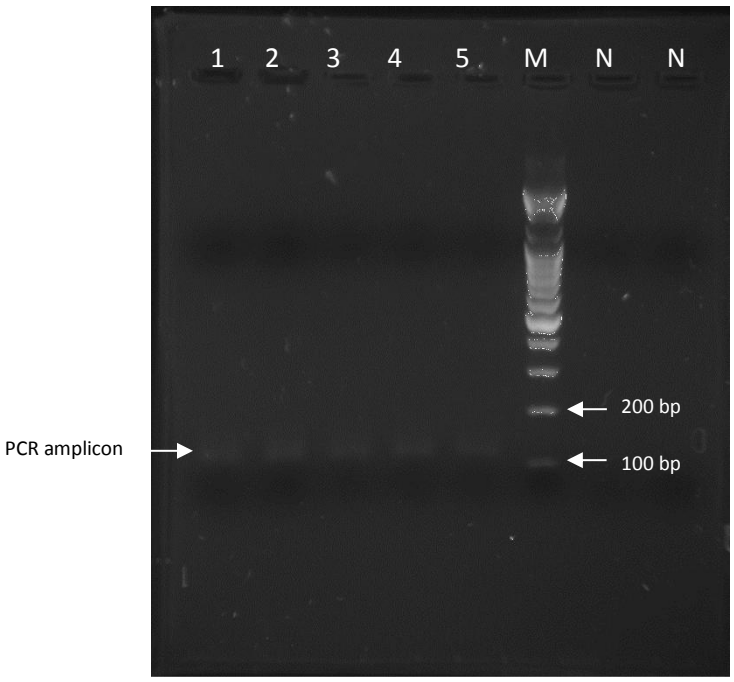


Figure 4.17: RT-PCR after polyadenylation and 3' RNA-DNA hybrid adapter was ligated to the purified small RNA species. 1-5: PCR products; M: VC 100 bp Plus DNA Ladder (Vivantis Technologies Sdn Bhd); N: negative control

Table 4.13: Sequence identity of the cloned small RNAs

Sequence identity	**Percentage (%)**	
	Method Ho et al. (2006)	**Method Fu et al. (2005)**
miRNA sequence	2.5	0
rRNA sequence	32.5	5.4
tRNA sequence	0	81.1
Mitochondrial sequence	2.5	8.1
Chloroplast sequence	10	0
Banana BAC sequence	2.5	0
mRNA	10	2.7
Unknown	12.5	2.7
Adapter sequence	25	0
TOTAL	100	100

4.3.2 Illumina high-throughput sequencing approach

4.3.2.1 Purity and integrity of RNA used for RNA-Seq

All three samples (CTR, TR100 and TR300) used for small RNA sequencing (sRNA-Seq) showed A260nm/A280nm of around 1.8 (CTR was slightly lower than 1.8) and A260nm/A230nm greater than 1 (Table 4.14). The 28S:18S rRNA ratio of the isolated RNA ranged 1.7 to 1.8 and all the three samples showed RIN values close to 9 (Table 4.15 and Figure 4.18). Small RNA species (< 200 nt) were present in very low abundance in the total RNA as shown in Figure 4.18.

Table 4.14: Optical density readings for RNA samples

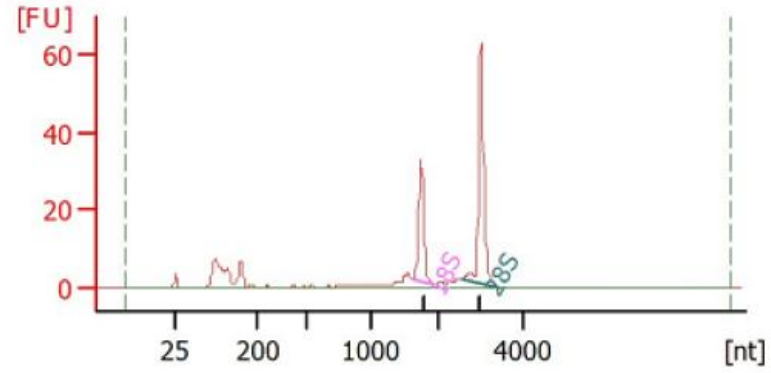
Sample	A_{260nm}/A_{280nm}	A_{260nm}/A_{230nm}
CTR	1.74	1.43
TR100	1.81	1.67
TR300	1.85	1.81

Table 4.15: Output of Agilent's Bioanalyzer assay for RNA samples

Sample	Concentration (ng/µl)	Volume (µl)	Amount (µg)	28S:18S	RIN
CTR	1,400	30	42	1.8	9.2
TR100	2,010	30	60.3	1.6	8.9
TR300	1,455	30	43.65	1.7	8.9

28S: 28S ribosomal RNA; 18S: 18S ribosomal RNA; RIN: RNA integrity number

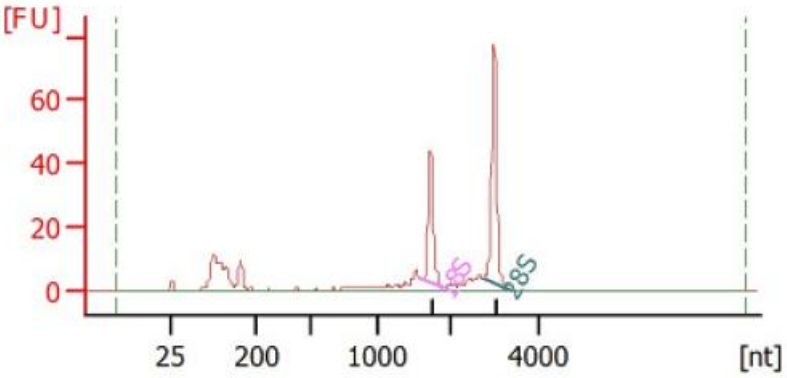
A – CTR



Overall Results for sample 1 : CTR

RNA Area:	188.4
RNA Concentration:	280 ng/µl
rRNA Ratio [28s / 18s]:	1.8
RNA Integrity Number (RIN):	9.2 (B.02.07)
Result Flagging Color:	
Result Flagging Label:	RIN: 9.20

B – TR100

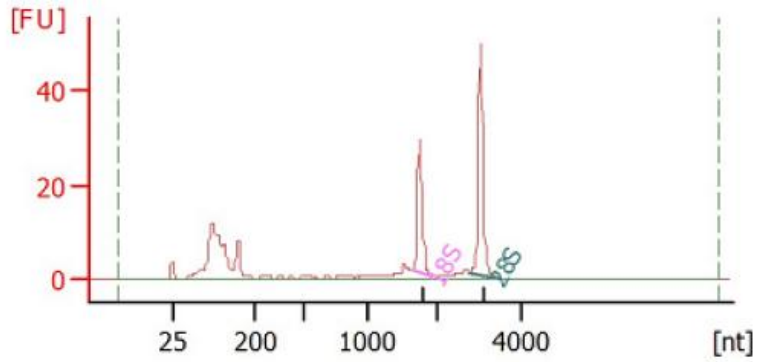


Overall Results for sample 2 : TR100

RNA Area:	271.0
RNA Concentration:	402 ng/µl
rRNA Ratio [28s / 18s]:	1.6
RNA Integrity Number (RIN):	8.9 (B.02.07)
Result Flagging Color:	
Result Flagging Label:	RIN: 8.90

Figure 4.18: Bioanalyzer (Agilent 2100) RNA analysis. A: RNA of CTR samples; B: RNA of TR100 samples; C: RNA of TR300 samples. CTR: untreated (0 mM NaCl) banana roots; TR100: 100 mM NaCl-treated banana roots; TR300: 300 mM NaCl-treated banana roots

C – TR300



Overall Results for sample 3 : TR300

RNA Area:	195.8
RNA Concentration:	291 ng/µl
rRNA Ratio [28s / 18s]:	1.7
RNA Integrity Number (RIN):	8.9 (B.02.07, Anomaly Threshold(s)
Result Flagging Color:	
Result Flagging Label:	RIN: 8.90

Figure 4.18: Continued

4.3.2.2 Sequence generation and processing

Sequencing of small RNA libraries constructed from CTR, TR100 and TR300 samples using Illumina GAxII sequencing platform generated more than 18 million tags in each library with 15-16 million tags being high quality (Table 4.16). After filtering adapter, polyA and other contaminant sequences, about 14.4, 14.7 and 13.6 million high quality clean reads ranged 17-35 nt were produced from CTR, TR100 and TR300, respectively (Table 4.16). After removal of redundancy, only about 3.44, 3.30 and 1.54 millions unique reads were present in sRNA datasets for CTR, TR100 and TR300, respectively (Table 4.16). The 21-nt sequence reads were dominant in all the three libraries, followed by 20-nt or 24-nt reads (Figure 4.19). Unlike CTR and TR100, TR300 had a higher percentage of 20-nt sRNA than 24-nt (Figure 4.19).

Table 4.16: Statistics of small RNA sequencing (sRNA-Seq) reads

Type	Read count		
	Control	**TR100**	**TR300**
Raw reads	18,189,390	18,396,868	19,641,136
High quality reads	15,911,422	16,039,100	15,190,814
Clean reads*	14,420,971	14,747,201	13,560,353
Reads after collapse**	3,438,498	3,302,845	1,539,695
Reads aligned to genome***	1,910,703	1,749,569	910,246

Control: untreated (0 mM NaCl) banana roots; TR100: 100 mM NaCl-treated banana roots; TR300: 300 mM NaCl-treated banana roots

*After filtering of contaminant sequences including sequencing adapters and poly(A)

**Non-redundant reads

***Reads mappable to the reference *Musa* genome (D'Hont et al., 2012)

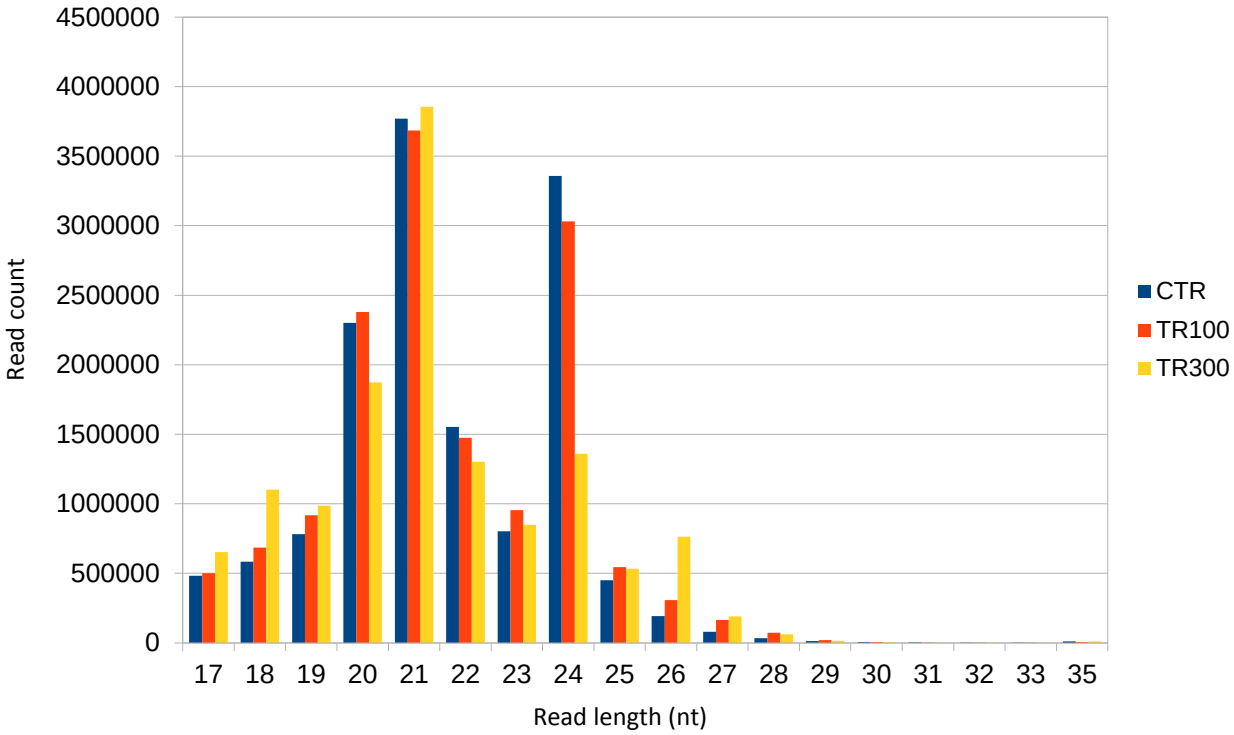


Figure 4.19: Length distribution of clean reads in small RNA libraries. CTR: untreated control (0 mM NaCl), TR100: 100 mM NaCl-treated banana roots; TR300: 300 mM NaCl-treated banana roots

4.3.2.3 MicroRNA annotation and mapping to *Musa* reference genome

After being annotated using Rfam and Plant MicroRNA Database (PMRD), the small RNA sequences were classified as shown in Figure 4.20. The vast majority (>88%) of the sequence reads in the banana root sRNA libraries are sequences not classified as an RNA with a known function. Among the reads annotated with an identity, miRNA sequences were the most abundant forming 6-10% of the sRNA libraries (Figure 4.20). The second most abundant small RNA with an identity was rRNA, forming 1-2% in the libraries (Figure 4.20). Each of the other known small RNAs (tRNA, snRNA and snoRNA) formed less than 1% of the sRNA libraries (Figure 4.20).

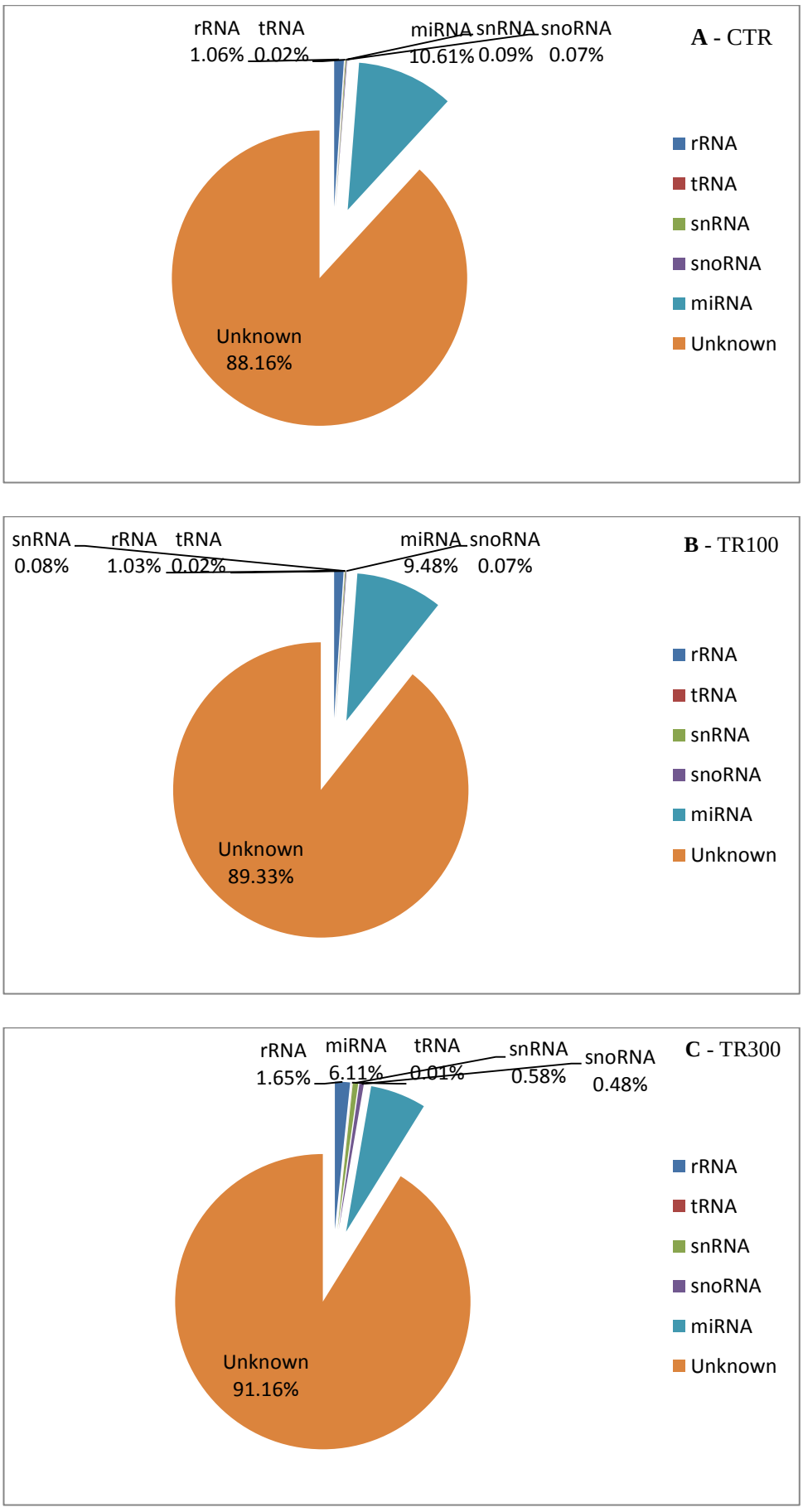


Figure 4.20: Classification of small RNA using PMRD and Rfam databases as reference

Using the Plant MicroRNA Database (http://bioinformatics.cau.edu.cn/PMRD/), we identified 153 (39 families), 149 (40 families) and 128 (35 families) orthologous miRNA sequences in CTR, TR100 and TR300 respectively (Table 4.17). All together, 181 miRNAs (Table S10) from 47 miRNA families were found in this study. A total number of 56 non-redundant putative *Musa*-specific miRNAs, which are not reported in species other than *Musa* to date, were identified from the banana root sRNA libraries in this study (Table 4.18).

About 1.91 (CTR), 1.75 (TR100) and 0.91 (TR300) million non-redundant sequence reads, could be mapped to the reference *Musa* genome (D'Hont et al., 2012) (Table 4.16). Of the mapped sequences, both the orthologus and *Musa*-specific miRNAs were distributed in all 11 chromosomes and the unrandom sequence of the reference banana genome as shown in Figure 4.6.

4.3.2.4 MicroRNA target prediction

A number of 110, 115 and 91 targets (RNA transcripts) could be found in our assembled unigenes for the orthologous miRNAs present in CTR, TR100 and TR300 respectively (Table 4.17). All together, our banana transcriptomes contain 247 unigenes (non-redundant) which are targets of the orthologous miRNAs in this study. From the 56 putative *Musa*-specific miRNAs, 120 targets could be found in the assembled unigenes in this study (Table 4.18).

Table 4.17: Annotation of orthologous miRNAs in banana root sRNAomes

	CTR	TR100	TR300	All (non-redundant)
Known miRNA	153	149	128	181
miRNA families	39	40	35	47
No. of miRNA targets	110	115	91	274

CTR: untreated control (0 mM NaCl); TR100: 100 mM NaCl-treated banana roots; TR300: 300 mM NaCl-treated banana roots

Table 4.18: Putative *Musa*-specific miRNAs in banana root sRNAomes

	CTR	TR100	TR300	All (non-redundant)
Novel miRNA	47	40	42	56
No. of miRNA targets	67	81	66	120

CTR: untreated control (0 mM NaCl); TR100: 100 mM NaCl-treated banana roots; TR300: 300 mM NaCl-treated banana roots

4.3.2.5 MicroRNA quantification and differential expression

A number of 43 orthologous miRNAs (belonging to 20 miRNA families) and 16 *Musa*-specific putative miRNAs were differentially expressed upon salt stress in banana roots in either 100 mM or 300 mM NaCl treatment, in comparison to the control (Figure 4.21). These miRNAs showed different expression patterns when different magnitudes of salt stress (100 mM and 300 mM NaCl) were applied (Figure 4.21).

Some of these miRNAs showed up-regulation in both 100 mM and 300 mM NaCl. These include mac-miR397a, mac-miR528, mac-miR827, mac-miR35, mac-miR38, mac-miR54, mac-miR60, mac-miR61, mac-miR62 and mac-miR66 (Figure 4.21). Whereas miRNAs including mac-miR156, mac-miR156k, mac-miR159c, mac-miR164h, mac-miR166c.4, mac-miR167g, mac-miR172a.2, mac-miR396a, mac-miR5139, mac-miR8, mac-miR14, mac-miR19 and mac-miR37 were among the miRNAs down-regulated in both 100 mM and 300 mM NaCl (Figure 4.21).

Some of these differentially expressed miRNAs were up-regulated in 100 mM NaCl but down-regulated in 300 mM. These include mac-miR156a, mac-miR56i, mac-miR156o, mac-miR156o.1, mac-miR156r, mac-miR157m, mac-miR159g, mac-miR162b, mac-miR162b.2, mac-miR167t, mac-miR169r, mac-miR171b, mac-miR396n, mac-miR529b, mac-miR12 and mac-miR49 (Figure 4.21). Whereas miRNAs, including mac-miR166b.3, mac-miR166e, mac-miR166i, mac-miR166m, mac-miR166t, mac-miR168, mac-miRf10192-akr, mac-miR5, mac-miR46 and mac-miR49, were down-regulated in 100 mM NaCl but up-regulated in 300 mM NaCl (Figure 4.21). Some of these miRNAs also showed other expression patterns, such as up-regulation (mac-miR156a.3 and mac-miR2910) or down-regulation (mac-miR162, mac-miR167m, mac-miR397 and mac-miR535) in TR300 but no change in expression in TR100.

A majority of the targets (found in the assembled unigenes) for these differentially expressed miRNAs are functionally related to binding (molecular function), cellular process (biological process), cell (cellular component) and membrane (cellular component) as shown in Figure 4.22.

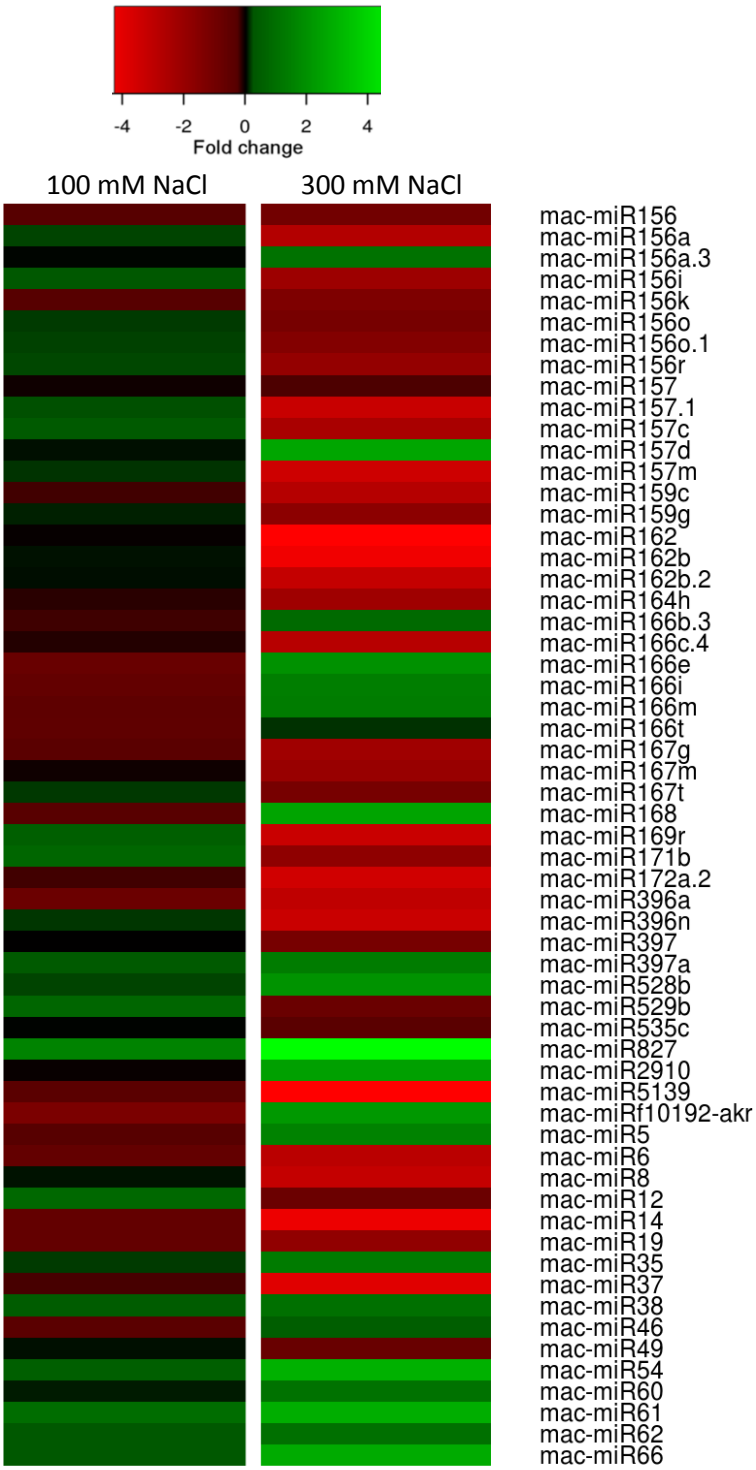


Figure 4.21: Orthologous and *Musa*-specific miRNAs differentially expressed in either 100 mM or 300 mM NaCl

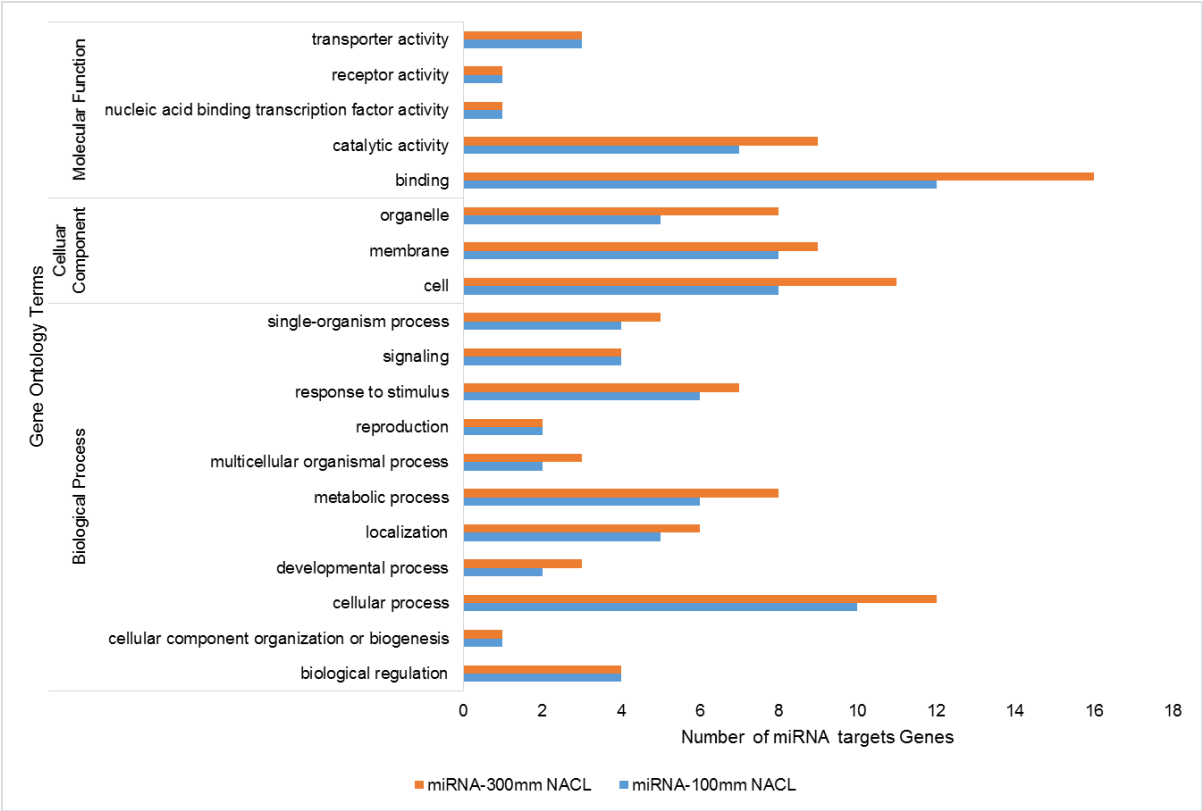


Figure 4.22: Gene Ontology (GO) assignment for targets of the differentially expressed miRNAs in salt-stressed banana roots

4.3.2.6 Differentially-expressed microRNAs and their targets

Forty three miRNAs (both orthologous and *Musa*-specific) were differentially expressed in 300 mM with at least a target in the RNA-Seq data (Figure 4.23). Most of these differentially expressed miRNAs showed an inverse expression pattern to at least one of their targets, with the exception of mac-miR166e, mac-miR169r, mac-miR396n, mac-miR827, mac-miR14 and mac-miR61 (Figure 4.23).

We found 12 *Musa*-specific miRNAs, which have not been reported yet in plants other than *Musa* species, differentially expressed in salt-stressed banana roots (300 mM NaCl) with at least one target found in the banana root transcriptomes (Figure 4.23). Among the 12 *Musa*-specific miRNAs, sequences of two miRNAs (mac-miR6 and mac-miR19) were previously reported in *Musa balbisiana* (Davey et al., 2013).

A Majority of the differentially-expressed miRNAs were down-regulated while

at least one of their predicted targets were up-regulated in the salt-stressed banana roots (TR300). These include mac-miR156, mac-miR156a,i,k,o,o.1,r, mac-miR157, mac-miR157.1, mac-miR157c,m, mac-miR159c,g, mac-miR162b.2, mac-miR164h, mac-miR166c.4, mac-miR171b, mac-miR172a.2, mac-miR172f, mac-miR390, mac-miR397, mac-miR529b, mac-miR535c, mac-miR6, mac-miR8, mac-miR12, mac-miR19, mac-miR37 and mac-miR49 (Figure 4.23).

Some differentially-expressed miRNAs were up-regulated (including mac-miR168, mac-miR397a, mac-miR528b, mac-miR35, mac-miR38, mac-miR62 and mac-miR66) and at least one of their targets were down-regulated (Figure 4.23). A small number of the differentially-expressed miRNAs showed the same expression pattern (either up- or down-regulation) as their targets, as shown in Figure 4.23. These miRNAs include mac-miR166e, mac-miR169r, mac-miR396n, mac-miR827, mac-miR2910, mac-miR14 and mac-miR61.

Among the predicted targets of the differentially expressed miRNAs are transcription factors, including AP2 domain-containing transcription factor (targeted by mac-miR172), GRAS family transcription factor (targeted by mac-miR166 and mac-miR171) and ethylene-responsive transcription factor (targeted by mac-miR172); transcripts coding for stress responsive proteins, such as, salt responsive protein (targetted by mac-miR159), osmotic stress-activated protein (targetted by mac-miR397), heat shock protein (targetted by mac-miR827); transcript coding for structural proteins such as protein root hair (targetted by mac-miR157), chloride channel (targetted by mac-miR37) and oligopeptide transporter 7-like (targetted by mac-miR66); enzymes involve in metabolism, such as, glutamate synthase (targetted by mac-miR169) and tropine dehydrogenase (targetted by mac-miR156), as well as transcripts coding for unknown or hypothetical proteins (Figure 4.23).

Among the unknown transcripts predicted to be targetted by the differentially-

expressed miRNAs, only one transcript, unigene CL1Contig328, was found with a domain annotation using InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan/). This unigene, which is targetted by mac-miR6, contains a dehydrin domain (InterProScan ID: IPR000167).

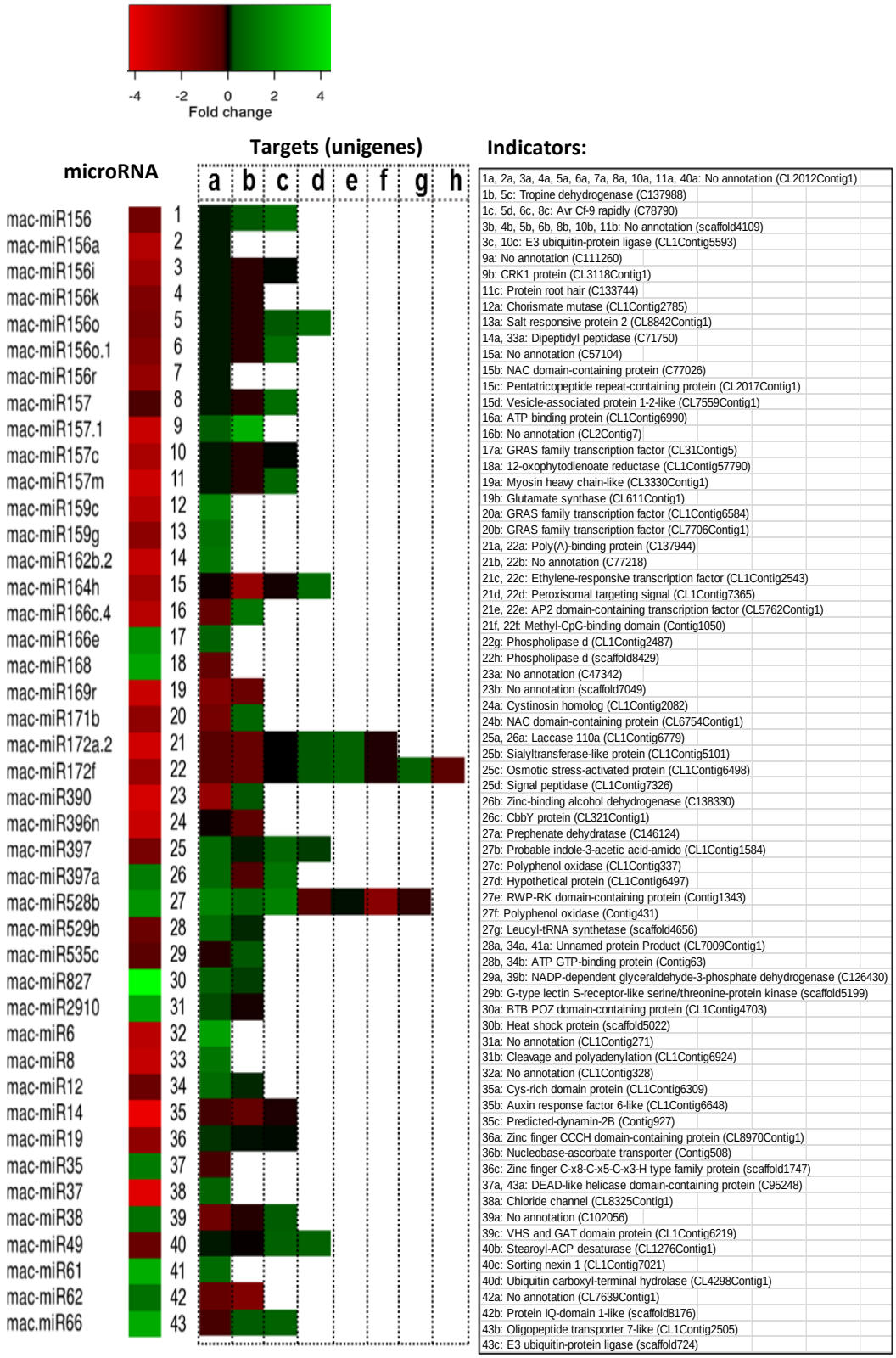


Figure 4.23: Differentially expressed miRNAs in 300 mM NaCl. Orthologous and *Musa*-specific miRNAs showed statistically significant expression change in banana roots upon 300 mM NaCl treatment, and their corresponding target (unigene) in the RNA-Seq data.

Indicators: 1a, 2a, 3a, 4a, 5a, 6a, 7a, 8a, 10a, 11a, 40a: No annotation (CL2012Contig1); 1b, 5c: Tropine dehydrogenase (C137988); 1c, 5d, 6c, 8c: Avr Cf-9 rapidly (C78790); 3b, 4b, 5b, 6b, 8b, 10b, 11b: No annotation (scaffold4109); 3c, 10c: E3 ubiquitin-protein ligase (CL1Contig5593); 9a: No annotation (C111260); 9b: CRK1 protein (CL3118Contig1); 11c: Protein root hair (C133744); 12a: Chorismate mutase (CL1Contig2785); 13a: Salt responsive protein 2 (CL8842Contig1); 14a, 33a: Dipeptidyl peptidase (C71750); 15a: No annotation (C57104); 15b: NAC domain-containing protein (C77026); 15c: Pentatricopeptide repeat-containing protein (CL2017Contig1); 15d: Vesicle-associated protein 1-2-like (CL7559Contig1); 16a: ATP-binding protein (CL1Contig6990); 16b: No annotation (CL2Contig7); 17a: GRAS family transcription factor (CL31Contig5); 18a: 12-oxophytodienoate reductase (CL1Contig57790); 19a: Myosin heavy chain-like (CL3330Contig1); 19b: Glutamate synthase (CL611Contig1); 20a: GRAS family transcription factor (CL1Contig6584); 20b: GRAS family transcription factor (CL7706Contig1); 21a, 22a: Poly(A)-binding protein (C137944); 21b, 22b: No annotation (C77218); 21c, 22c: Ethylene-responsive transcription factor (CL1Contig2543); 21d, 22d: Peroxisomal-targeting signal (CL1Contig7365); 21e, 22e: AP2 domain-containing transcription factor (CL5762Contig1); 21f, 22f: Methyl-CpG-binding domain (Contig1050); 22g: Phospholipase D (CL1Contig2487); 22h: Phospholipase D (scaffold8429); 23a: No annotation (C47342); 23b: No annotation (scaffold7049); 24a: Cystinosin homolog (CL1Contig2082); 24b: NAC domain-containing protein (CL6754Contig1); 25a, 26a: Laccase 110a (CL1Contig6779); 25b: Sialyltransferase-like protein (CL1Contig5101); 25c: Osmotic stress-activated protein (CL1Contig6498); 25d: Signal peptidase (CL1Contig7326); 26b: Zinc-binding alcohol dehydrogenase (C138330); 26c: CbbY protein (CL321Contig1); 27a: Prephenate dehydratase (C146124); 27b: Probable indole-3-acetic acid-amido (CL1Contig1584); 27c: Polyphenol oxidase (CL1Contig337); 27d: Hypothetical protein (CL1Contig6497); 27e: RWP-RK domain-containing protein (Contig1343); 27f: Polyphenol oxidase (Contig431); 27g: Leucyl-tRNA synthetase (scaffold4656); 28a, 34a, 41a: Unnamed protein product (CL7009Contig1); 28b, 34b: ATP/GTP-binding protein (Contig63); 29a, 39b: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (C126430); 29b: G-type lectin S-receptor-like serine/threonine-protein kinase (scaffold5199); 30a: BTB/POZ domain-containing protein (CL1Contig4703); 30b: Heat shock protein (scaffold5022); 31a: No annotation (CL1Contig271); 31b: Cleavage and polyadenylation (CL1Contig6924); 32a: No annotation (CL1Contig328); 35a: Cys-rich domain protein (CL1Contig6309); 35b: Auxin response factor 6-like (CL1Contig6648); 35c: Predicted-dynamin-2B (Contig927); 36a: Zinc finger CCCH domain-containing protein (CL8970Contig1); 36b: Nucleobase-ascorbate transporter (Contig508); 36c: Zinc finger C-x8-C-x5-C-x3-H type family protein (scaffold1747); 37a, 43a: DEAD-like helicase domain-containing protein (C95248); 38a: Chloride channel (CL8325Contig1); 39a: No annotation (C102056); 39c: VHS and GAT domain protein (CL1Contig6219); 40b: Stearoyl-ACP desaturase (CL1276Contig1); 40c: Sorting nexin 1 (CL1Contig7021); 40d: Ubiquitin carboxyl-terminal hydrolase (CL4298Contig1); 42a: No annotation (CL7639Contig1); 42b: Protein IQ-domain 1-like (scaffold8176); 43b: Oligopeptide transporter 7-like (CL1Contig2505); 43c: E3 ubiquitin-protein ligase (scaffold724)

CHAPTER 5: DISCUSSION

Salinity adversely affects productivity and quality of crops and the increase in arable land area being salinized poses a risk to global food security. Cultivated bananas have been demonstrated as salt sensitive. Therefore, understanding of the molecular basis of salt-stress sensitivity and tolerance is important in order to formulate strategies for improving salinity tolerance in banana as well as other economically important crops. A high throughput sequencing approach was used in this study to profile transcriptomes, including the functional small RNA, to reveal expressed genes, pathways and gene regulation in salt-stressed banana roots.

The tissue selected for this study was banana roots as these are the primary site of injury and signal perception for several types of water-limiting stress, including salinity and drought (Jiang & Deyholos, 2006; Zhao et al., 2013a). Therefore, it is of interest to profile transcripts in the banana roots to investigate the early responses, such as signal perception and transduction, transcriptional control and the effector mechanisms, in response to stress.

The RNA isolation method used in this study was modified from a CTAB-based method reported by Kiefer et al. (2000). We applied 1-2 extraction steps using phenol-chloroform-isoamylalcohol to improve the purity of the isolated nucleic acids as phenol is an effective reagent for removing contaminant proteins from the nucleic acids. RNA samples with high purity have A_{260nm}/A_{280nm} ratio between 1.8 and 2, a possible indicator of low protein contamination, and A_{260nm}/A_{230nm} ratio >1, indicating unlikely contamination of polyphenols and polysaccharides in the purified nucleic acids (Table 4.1 and Table 4.14). High RNA purity is important to ensure high success rate in cDNA library construction as the multiple enzymatic steps involved in the library construction are sensitive to protein, polyphenol and polysaccharide contaminants.

High RNA integrity (RIN > 8 and 28S/18S ratio approximately 2), determined by using Bioanalyzer (Figure 4.3 and Figure 4.18), is also very important in cDNA library construction for high throughput sequencing. Degraded and slightly degraded RNA can cause higher proportions of undesirable RNA sequences, including ribosomal RNA and tRNA, jeopardize the quality of the cDNA library and complicate *de novo* transcriptome assembly. Poor RNA integrity used for mRNA-Seq may prevent assembly of larger scaffolds (Johnson et al., 2012). Whereas the use of degraded and slightly degraded RNA for small RNA cDNA library can cause over-representation of degraded RNA species, including rRNA, tRNA and mRNA, which complicates the small RNA analysis, especially *in silico* transcript quantification.

High-throughput sequencing for plant tissue samples (especially roots) collected from *ex vitro* may experience contamination with microbial sequences. The presence of the microbial sequences may complicate the sequence assembly from short reads and additional filtering steps are required to mask the contaminant sequences (Schmieder & Edwards, 2011). Therefore, *in vitro* tissue materials were used in this study in order to avoid or minimize microbial nucleic acid sequences from the deeply sequenced banana root transcriptomes.

In this study, salinity serves as an abiotic stress model for its ease and precise of application in laboratory settings. In addition to ionic stress, the application of NaCl in plants also represents other agronomically important stresses such as osmotic (similar to drought) and oxidative stresses (reviewed in Munns & Tester, 2008). 'Berangan' cultivar (triploid genome, AAA) was used in this study as it is a widely grown Malaysian local banana cultivar and its market value is about 60-120% higher than Cavendish (Figure S1), the most important cultivar in international trade.

5.1 Functional information from transcriptome sequencing of salt-stressed roots

Functional genomic data are valuable genetic resources for crop improvement programs. Over the past decade, the increase in sequencing capacity with lower cost using next-generation sequencing technologies has encouraged high-throughput production of gene expression data in various model and non-model plants, including the economically important crops. In this study, we produced about 3 Gbp of gene expression data from banana roots through Illumina sequencing of two transcriptomes (~2.1 Gbp) and three small RNA libraries (>900 Mbp).

Although a high-quality finished reference genome for banana is available (D'Hont et al., 2012), a *de novo* assembly approach was used in this study to include transcripts transcribed from the genome segments that are missing from the genome assembly (Martin & Wang, 2011). There are likely to be differences in genome constituent, mutation, chromosomal rearrangements and splice variants in 'Berangan' (triploid, AAA genome) cultivar used in this study, as compared to the wild banana, *Musa acuminata* ssp. *malaccensis* (diploid, AA genome) published reference genome.

A total of 31,390 unigenes were assembled from the two banana root transcriptomes and these assembled unigenes have a mean length of 517 bp, which is only 6.7% shorter than the banana root transcriptome assembled from 9.3 Gbp data generated using the same sequencing platform reported by Li et al. (2012a). The longer average length of assembled unigenes reported by Li et al. (2012a) is probably due to higher Illumina sequencing input (9.3 Gbp) compared to this present study (1.89 Gbp). More than 99.5% of the unigenes resulted from *de novo* transcriptome assembly in this study can be mapped to the reference A-genome (D'Hont et al., 2012). The length and high degree of matching indicates that these assembled unigenes are of high quality.

It has been demonstrated that a sequencing depth of 50 million of 50-bp single end reads, which is about 2.5 Gbp of sequence data resulted in near saturated coverage of expressed genes, reaching plateau with about 21,139 unique gene models in *Arabidopsis*. Further increase in number of sequence reads did not show significant increase in number of unique expressed gene models (Van Verk et al., 2013). In our study, we produced about 21 million reads or 1.89 Gbp of sequences from the banana root transcriptomes (CTR and TR300), with 45.33X and 38.75X coverage per base (CPB) per unigene in CTR and TR300 respectively, before clustering into 31,390 non-redundant unigenes. The number of putative proteins was 22,231 based on the matches found in the GenBank non-redundant protein database. This showed that our RNA-Seq is satisfactory in terms of sequencing depth and has probably captured most of the expressed transcripts in banana roots. Our sequencing coverage is also comparable to that reported in *Eucalyptus* tree transcriptome (37X) and *Ipomoea batatas* root transcriptome (48.36X) that have been sequenced and assembled using the same approaches but with higher amount of input sequences of about 6.90 Gbp (Mizrachi et al., 2010) and 4.4 Gbp (Wang et al., 2010e) respectively.

A number of 36,542 and 36,638 protein-coding gene models have been predicted in the reference *Musa acuminata* (A-genome) (D'Hont et al., 2012) and *Musa balbisiana* (B-genome) (Davey et al., 2013) respectively. In this study, we assembled 31,390 non-redundant unigenes from banana root transcriptomes including 22,231 unigenes (70.8%) with a qualifying match in the GenBank non-redundant protein database. The number of unigenes annotated with a putative protein identity was about 60% of the number of predicted protein-coding gene models in the reference *Musa* genomes.

In this study, putative protein functions were assigned for 70.8% or 22,231 of the assembled unigenes using the NCBI non-redundant (NR) protein database (Table

4.7), which the percentage of transcripts with an assigned protein identity is higher than the transcriptomes sequenced and assembled using similar methods, with only 69.99% (61,706 unigenes) unigenes showed hits in NR database in Cavendish banana roots (Li et al., 2012a), 48.54% (27,435 unigenes) in *Ipomoea batatas* roots (Wang, et al., 2010e) and 61.36% (34,684 unigenes) in *Lycium chinense* mixed tissues (Zhao et al., 2013b) but lower than that in 'Brazilian' banana root (85.36%) (Wang et al., 2012b) and sabaigrass (75.6% or 44,723 isogenes) (Zou et al., 2013). In addition, 47.5% of the assembled unigenes in this study have BLAST hits in the UniProt database, which the percentage is slightly higher than that observed in the recently reported *Lycium chinense* transcriptome (37.03%) (Zhao et al., 2013b). However, the previous studies have used a cut-off e-value $\leq 10^{-5}$ compared to the higher stringency with cut-off e-value $\leq 10^{-10}$ used for BLAST in this study.

BLAST search of the assembled unigenes against the NCBI NR database resulted in matches mostly from rice, followed by grapevine, soybean, poplar and Arabidopsis (Figure 4.7). The high number of matches with rice data was probably because of the availability of a larger number of annotated rice genes and the relatively close genetic distance from banana, as both are monocots, whereas grapevine, soybean, poplar and Arabidopsis are dicots.

Within the biological process category of Gene Ontology (GO), the largest proportion of the assembled unigenes in this present study was assigned in the cellular and metabolic processes (Figure 4.8). This observation is similar to that observed in transcriptomes reported in various plant species and tissues, including Cavendish banana roots, and banana leaf of Calcutta 4 and Grande Nain (Li et al., 2012a; Mizrachi, et al., 2010; Passos et al., 2013; Wang, et al., 2010e; Wang, et al., 2012b; Kaur et al., 2011; Zhao, et al., 2013b). These two processes are important in maintaining cellular activities with many genes investigated and functionally validated. In the molecular

function category, binding and catalytic activities are the most represented GO terms in the banana root transcriptome investigated in this study. This observation is also similar to those of many reported transcriptomes including banana roots (Li et al., 2012a; Wang, et al., 2012b), banana leaf (Passos, et al., 2013), *Ipomoea batatas* root (Wang et al., 2010e), *Eucalyptus* xylem (Mizrachi, et al., 2010) and *Lycium chinense* mixed tissues (Zhao, et al., 2013b). In the cellular component category, the most represented GO terms in banana root transcriptome are only similar to those reported for Cavendish banana roots (Li et al., 2012a; Wang et al., 2012b) and Calcutta 4 and Grande Nain banana leaf (Passos, et al., 2013) but not those for transcriptomes reported for other plant species and tissues (Wang et al., 2010e; Kaur et al., 2011; Zhao et al., 2013b). However, it is not conclusive that the dominance of functional groups of genes expressed in banana root is related specifically to particular GO terms as these well-represented GO terms were also observed to be the most represented GO terms in data reported for whitefly (Wang et al., 2010d). It is possible that there are many genes characterized and assigned in these GO terms and the possibility to get a match with these GO terms is high.

Similar to reports on transcriptomes by Wang et al. (2010b), Li et al. (2012a) and Zhao et al. (2013), 'general function prediction only' is the most represented Eukaryotic Orthologous Groups of Proteins (KOG) in the banana root in this present study (Figure 4.9). The second and third most represented KOGs in the banana root transcriptome of this study, the posttranslational modification, protein turnover and chaperones, and signal transduction mechanisms, however only ranked fourth and sixth in Cavendish banana root (Li et al., 2012a), fourth and fifth in *Ipomoea batatas* root transcriptome (Wang et al., 2010b) and third and seventh in *Lycium chinense* mixed tissues (Zhao et al., 2013b).

About 22.2% of the assembled unigenes in this study could be assigned to Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, compared to 31.14% unigenes in the *Ipomoea batatas* root transcriptome (Wang, et al., 2010e). Pathways for carbohydrate metabolism were well-represented in banana roots similar to that in the *Ipomoea batatas* root transcriptome (Wang, et al., 2010e). This indicates that root tissues may be actively involved in starch and sugar biosynthesis in banana. The High proportion of purine metabolism pathways expressed in banana roots may be related to active DNA replication and mitosis in banana roots, where purines are used as building blocks for nucleotide synthesis.

5.2 Salt stress-responsive genes in banana roots

Methods for *in silico* quantification of transcript expression in RNA-Seq data have been established and widely applied (Audic & Claverie, 1997; Cloonan et al., 2008; Mortazavi et al., 2008; Wilhelm & Landry, 2009). RNA-Seq is suitable for expression measurement and comparison of transcripts between libraries and it has been proven to perform as good as microarray approach, with strong correlation in expression fold change ($\log_2$ ratio) between the two methods (Malone & Oliver, 2011).

In Gene Ontology assignment for the differentially-expressed unigenes in the salt-stressed banana roots, the well-represented GO terms are similar to many reports as mentioned in section 5.1, which are catalytic activity, metabolic and cellular processes, cell, and binding (Figure 4.12). Therefore, this GO term profile is not specific to salt stress response in banana roots.

The differentially-expressed unigenes in the salt-stressed banana roots were assigned most abundantly in carbohydrate transport and metabolism of the Eukaryotic Orthologous Groups of Proteins (KOG) (Figure 4.13). This suggests that banana roots may actively respond to salt stress by altering metabolism and transport of sugars and

starch. Starch and sucrose metabolism was also the most represented KEGG pathway in salt-stressed banana roots (Figure 4.14). Sugar metabolism and mobilization are important for osmotic adjustment and stress adaptation in plants (Janz et al., 2010; Misic et al., 2012). It has been reported that salt tolerance is correlated with metabolism and allocation of carbohydrate in the leaves, stems and roots where hexoses accumulated and induced a feedback repression to photosynthesis in the salt-stressed perennial ryegrass (Hu et al., 2013). Phenyalanine metabolism and cysteine and methionine metabolism were the well-represented KEGG pathways in the differentially-expressed unigenes in the salt-stressed banana roots after the starch and sucrose metabolism. Phenylalanine, cysteine and methionine are important amino acids in plants. Phenylalanine serves as precursors for the synthesis of various vital metabolites in plants (Yoo et al., 2013), whereas cysteine is precursor for the biosynthesis of various important molecules, including vitamins, cofactors, antioxidants and compounds for defence (Alvarez et al., 2012). Methionine is an important precursor or donor for the metabolism of various other primary and secondary metabolites (Amir, 2008). The regulation of genes involved in these metabolic pathways may help banana roots to cope with salt stress through the production of a broad range of metabolites, as has previously been reported in various plants (reviewed in Ramakrishna & Ravishankar, 2011).

5.2.1 Up-regulation

In general, the strongly up-regulated genes in the salt-stressed stress banana roots are functionally related to signaling, transcriptional regulation and DNA repair, transport and cellular homeostasis, stress defence, cell wall, membrane and cytoskeleton structures, nutrient homeostasis and other functions not known to be associated with abiotic stresses (Table 5.1).

(a) Signal perception and transduction

Signal perception and transduction are important mechanisms in plants to sense the environmental changes and transduce the signals to activate a cascade of cellular responses. Receptor-like kinase (RLK) and protein phosphatase 2C are genes related to signaling in plants strongly regulated in the salt-stressed banana roots in this study. A putative RLK was up-regulated about 9-fold in the salt-stressed banana roots. Receptor-like protein kinases are a group of trans-membrane receptors with distinguished extracellular domains, likely to be involved in sensing and transducing extracellular signals (reviewed in Morris & Walker, 2003; Shiu & Bleecker, 2001). This putative RLK is probably specific for salinity stress signal sensing and transduction in the banana roots. A putative protein phosphatase 2C (PP2C) was up-regulated about 8-fold in the salt-stressed banana roots. Protein phosphatase 2C is a class of ubiquitous and evolutionarily conserved serine/threonine protein phosphatase that negatively regulates protein kinase pathways, thus, modulating signalling pathways that are involved in environmental stress responses and developmental processes (Sheen, 1998; Xue et al., 2008). The putative PP2C that was strongly up-regulated in the salt-stressed banana roots is probably a regulator to protein kinase pathways that are distinct to salt stress.

Plant lipid transfer proteins have been indicated in cutin and wax synthesis in leaf or adaptation to environmental changes in leaves and roots (Kader, 1997; Blilou et al., 2000; Hollenbach et al., 1997; Liu et al., 2001; Maghuly et al., 2009). The LTP has been shown to be differentially expressed in leaf and root in different stresses, and suggested to mediate signal transduction in *Prunus incisa* x *serrula* (Maghuly et al., 2009). A putative lipid transfer protein (LTP) was up-regulated about 7-fold in the salt-stressed banana roots, which probably played an important role in signalling pathway rather than cell wall modification in the banana roots.

(b) Transcriptional regulation and DNA repair

BTB/POZ domain-containing protein, H3 and H2A histones, homeobox-leucine zipper protein, DNA-3-methyladenine glycosylase I are among the genes involved in transcriptional regulation and DNA repair, and they were strongly up-regulated in the salt-stressed banana roots in this study (Table 5.1). Expression of a BTB/POZ domain-containing protein was increased about 8-fold in the salt-stressed banana roots. BTB/POZ (broad complex, tram track, bric-a-brac/POX virus and zinc finger) domain is a protein-protein interaction domain found in some zinc finger transcription factors and actin-binding proteins. This domain is in part responsible for chromatin remodelling and transcriptional control through oligomerization and interaction with its cofactors (Collins et al., 2001; reviewed in Weber & Hellmann, 2009). The strong up-regulation in expression suggests this gene may serve as a transcriptional modulator in banana roots under high salinity and further investigation is needed to reveal the specific functions of this gene in transcriptional regulation.

The transcripts of two histone proteins, H3 and H2A, were both increased about 9-fold in the salt-stressed banana roots. H3 and H2A are two of the five histone proteins in the eukaryotes involved in the packaging of DNA into nucleosome. Posttranslational modifications of histones regulate processes including transcription, DNA replication and DNA repair (reviewed in Kothapalli et al., 2005). Up-regulation of these histone-coding genes may help banana root cells to regulate gene expression to cope with salt stress, to replenish the injured or dead cells, and also to repair damaged DNA due to oxidative stress.

Homeobox-leucine zipper proteins belong to a transcription factor family that response to environmental cues and developmental signals (Deng et al., 2002; Rueda et al., 2005). Over-expression of an Arabidopsis homeobox-leucine zipper domain protein has been shown to confer salt tolerance in yeast (Shin et al., 2004). A putative

homeobox-leucine zipper protein was up-regulated about 7-fold in the salt-stressed banana roots in this present study. This putative homeobox-leucine zipper protein is probably a transcription factor regulating gene expression in response to salt stress in banana roots and further investigation is required to reveal its specific functions.

A putative DNA-3-methyladenine glycosylase I was up-regulated about 9-fold in the salt-stressed banana roots in this present study. DNA-3-methyladenine glycosylase has been shown as a DNA repair enzyme that removes cytotoxic alkylated bases in *E. coli* (Yamagata et al., 1996) and human (Hasplova et al., 2012). DNA-3-methyladenine glycosylase has also been reported as a DNA-binding protein and suggested to be involved in DNA repair in Arabidopsis (Malhotra & Sowdhamini, 2013). It is possible that this enzyme was up-regulated in banana roots to deal with DNA alkylation, such as DNA methylation, caused by salinity stress.

(c) Transport

Transport proteins are responsible for cellular homeostasis for plant adaptation in changing environments. Nodulin-like protein, inorganic phosphate transporter, phosphate transporter, auxin-induced protein 5NG4-like, VAMP protein SEC22 and cadmium/zinc-transporting ATPase 3-like are among the strongly up-regulated transport proteins found in the salt-stressed banana roots in this study (Table 5.1). Transcripts of a nodulin-like protein were up-regulated about 10 fold in the salt-stressed banana roots. Nodulin-like proteins have been shown as multifunctional transmembrane proteins in non-nodulating plants species important for transport of nutrients, solutes, amino acids or hormones. The nodulin-like proteins are involved in the major aspects of plant development and modulated by environmental cues (reviewed in Denance et al., 2014; Wallace et al., 2006).

A putative inorganic phosphate transporter and a putative phosphate transporter were up-regulated with 10- and 9-fold expression change in the salt-stressed banana roots. Phosphate transporters are membrane-localized proteins important for phosphate uptake, transportation/translocation and homeostasis (Jia et al., 2011). The phosphate transporters have been reported to be expressed in the roots of various plants, including Arabidopsis, tomato and rice, to increase phosphorous uptake capacity during phosphate starvation (Muchhal et al., 1996; Muchhal & Raghothama, 1999; Wang et al., 2014c). Activity of genes responsible for phosphate homeostasis was elevated in banana roots, which is a possible strategy deployed by banana roots to maintain nutrient availability when phosphate uptake is difficult due to osmotic and ionic imbalance.

Auxin-induced protein 5NG4-like is a membrane protein, possibly a nodulin-like or permease-like protein and functions as transporter for drugs and metabolites (UniProt, http://www.uniprot.org/uniprot/). Expression of a putative auxin-induced protein 5NG4-like was increased about 9-fold in the salt-stressed banana roots suggesting its possible role in transporting and accumulating metabolites for stress tolerance and removing harmful metabolites produced as a result of salt stress.

SEC22 has been described as a membrane trafficking protein involved in the secretory trafficking between ER and Golgi and contributes to growth and development in plants (Chatre et al., 2005; El-Kasmi et al., 2011). A putative VAMP (vesicle-associated membrane protein) protein SEC22 was up-regulated 7-fold in the salt-stressed banana roots. This putative SEC22 in banana root is possibly involved in the distribution of defence-related proteins and metabolites newly-synthesised in the ER and/or elimination of harmful metabolites after detoxification in the ER, through ER-Golgi transportation.

Cadmium/zinc-transporting ATPase has been demonstrated to regulate intracellular cadmium and zinc concentrations in human and bacteria (Noll & Lutsenko,

2000). It has also been shown to be involved in root-to-shoot translocation of Zn and Cd, sequestration of Cd into vacuoles in root cells and detoxification of zinc and cadmium in rice (Takahashi et al., 2012). Banana roots showed about 7-fold increase in expression of a putative cadmium/zinc-transporting ATPase 3-like. It is possible that zinc and cadmium level under normal circumstance can be lethal to the banana roots in salt stress condition and mechanism for metal translocation was activated.

(d) Stress defence

Some stress defence proteins were also strongly up-regulated in the salt-stressed banana roots in this study, which include peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase-like, peroxidase 47-like, peroxidase 3-like, late embryogenesis abundant (LEA) protein, xylanase inhibitor and glutamine synthetase (Table 5.1).

A peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase-like (PNGase) was up-regulated about 10-fold in the banana roots under 300 mM NaCl. PNGase has been reported as a component of the ER-associated degradation machinery and is involved in proper clearance of misfolded plant proteins by deglycosylating glycoproteins (Diepold et al., 2007; Masahara-Negishi et al., 2012; Suzuki, 2007). It is possible that proteins were improperly formed or modified when banana roots experienced high salinity (300 mM NaCl) and this protein degradation machinery is a part of the cell death process in highly stressed cells for recycling amino acids for new proteins.

A putative peroxidase 47-like and a putative peroxidase 3-like showed 8- and 7-fold up-regulation, respectively, in response to salt stress in banana roots. Peroxidases are anti-oxidant enzymes important in scavenging reactive oxygen species (ROS) produced as a result of environmental stresses (Davletova et al., 2005). High salinity

can cause ROS accumulation to the level which is deleterious to cell structures. Therefore, it is not surprising that peroxidase genes were strongly up-regulated in the salt-stressed banana roots.

A putative xylanase inhibitor was expressed 7-fold higher in the salt-stressed banana roots. Xylanase inhibitor has been demonstrated to inhibit xylanase activity of pathogenic *Fusarium graminearum* in durum wheat and confer resistance to *Rhizoctonia solani* in rice plant (Moscetti et al., 2013; Wu et al., 2013). In this present study, the relation between the increased expression of a putative xylanase inhibitor and salt stress in the banana roots is not clear. It is speculated that the increased expression of xylanase inhibitor may protect the root cell wall from degradation by endogenous xylanases.

Late embryogenesis abundant (LEA) protein is a large protein family that protects other proteins from aggregation, especially during desiccation or osmotic stresses in plants (Reddy et al., 2012). A putative LEA protein was elevated 7-fold in expression in the salt-stressed banana roots, which might confer protection to the banana roots from osmotic stress due to high salinity.

Glutamine synthetase is the key enzyme in nitrogen metabolism in plants and has been found to be induced by salinity in tomato roots (Debouba et al., 2006). Expression of a putative glutamate synthetase gene was increased about 7-fold in the salt-stressed banana roots. Modulation of nitrogen status by elevating nitrogen assimilation may be a strategy deployed by the banana roots to maintain nitrogen level or gain nitrogen source under salt stress where nutrient uptake is difficult due to osmotic and ionic imbalance.

(e) Cell wall modifications and cellular structures

The cell wall is the first cellular component confronting environmental stresses in plants. Modification of cell walls may improve the plant's tolerance against environmental stresses. Galactosyltransferase family protein, galacturonosyltransferase-like 3, fasciclin-like arabinogalactan, alpha-expansin 7, beta-expansin 3, cell-wall associated hydrolases and beta-glucanase are among the strongly- up-regulated cell wall-modifying enzymes in the salt-stressed banana roots.

A putative galactosyltransferase family protein and a putative galacturonosyltransferase-like 3 showed about 10- and 7-fold expression changes respectively, in the salt-stressed banana roots. Galactosyltransferase catalyses the transfer of galactose to a glycosyl acceptor to form complex carbohydrates or glycoproteins, such as proteoglycan, that is important for cell wall formation (Dilokpimol et al., 2014). Whereas, galacturonosyltransferase has been reported to be involved in the biosynthesis of pectin and/or xylans important for cell wall formation (Atmodjo et al., 2011). Up-regulation of the genes coding for cell wall proteins in the salt-stressed banana roots may confer protection to the root cells from salt stress and salt stress-related secondary stresses.

Fasciclin-like arabinogalactan proteins have been identified as cell surface adhesion proteins and are important in plant development and abiotic stress response (Johnson et al., 2003). The transcript of a putative fasciclin-like arabinogalactan was up-regulated about 9-fold in salt-stressed banana roots. The increased expression in this gene might be responsible for strengthening cell-to-cell adhesion in the root tissue to tolerate high salinity condition.

Two putative expansin proteins, alpha-expansin 7 and beta-expansin 3, were observed to be differentially expressed, with about 8-fold in expression change in the salt-stressed banana roots. Expansins are a group of cell wall proteins in plants,

responsible for wall-loosening and cell expansion that are involved in cell wall modifications for plant growth and development (Dal Santo et al., 2011; Gal et al., 2006; Lin et al., 2011). Overexpression of expansin has been found to alter leaf and root morphology, such as increased lateral number and fewer stomata, which confer drought and salt tolerance to Arabidopsis (Lu et al., 2013a). The strong up-regulation of the two putative expansins in the banana roots may modify root cells, such as root elongation and initiation of new roots, in response to high salinity.

Two putative cell wall-associated hydrolases were up-regulated about 7- and 8-fold respectively in the salt-stressed banana roots. Plant cell wall-associated hydrolases may be referred to as glycoside hydrolases that catalyse the hydrolysis of glycosidic bonds in cell wall polymers important for cell wall modifications, and are implicated in metabolism of secondary metabolites and phytohormone activation in response to abiotic stresses (reviewed in Sharma et al., 2013). Other plant cell wall-associated hydrolases include polygalacturonase, pectinmethylesterase and cellulase (Fischer & Bennett, 1991). Cell wall modifications and the production of secondary metabolites and phytohormones triggered by the increased expression of cell wall-associated hydrolases may serve as mechanisms in banana roots to survive high salinity.

Endo-beta-glucanase has been described as an enzyme mediating degradation of cell wall beta-glucans and is related to growth and development (UniProt, http://www.uniprot.org/). It is required for root cell elongation and cell division in rice (Zhang et al., 2012a). A putative beta-glucanase precursor was up-regulated about 7-fold in the salt-stressed banana roots. Up-regulation of the putative beta-glucanase precursor in the salt-stressed banana roots is probably responsible for the biosynthesis of beta-glucanases important for root cell modifications and root growth, in order to battle salt stress.

3-ketoacyl-CoA synthase is an enzyme involved in fatty acid synthesis by condensing acyl chain to form long-chain fatty acids (Chen et al., 2011c). It has been described in Arabidopsis to contribute to cuticular wax and root suberin biosynthesis and is differentially controlled under osmotic stress conditions (Lee et al., 2009). Expression of a putative 3-ketoacyl-CoA synthase 4-like was raised about 7-fold in the salt-stressed banana roots, which possibly helps in suberin deposition in the root cell wall as a selective barrier for certain molecules and stronger structural support to the root cells to contend with osmotic and oxidative stress in the high salinity condition.

Plants alter cell structures in order to cope with abiotic stress. Betu-tubulin is a component of the cytoskeleton in cells. It has been reported that the cytoskeleton is important for salt stress tolerance in Arabidopsis, where reduced levels of alpha- and beta-tubulin caused hypersensitivity to NaCl (Rodriguez-Milla & Salinas, 2009). A beta-tubulin gene was up-regulated about 8-fold in the salt-stressed banana roots. The increase in beta-tubulin expression is probably important for strengthening cell structure via the production of cytoskeleton protein in the salt-stressed banana roots.

Besides, some of the strongly up-regulated genes in the salt-stressed banana roots have no known stress-related functions as summarized in the Table 5.1. In addition, a number of unannotated genes (including CL1Contig7761, C91380, scaffold1398, CL1Contig501, CL9171Contig1, C85860, CL3948Contig1) and genes coding for hypothetical protein (including C114160, C134662. scaffold1645, CL1601Contig1, CL1Contig5300) were found to be strongly down-regulated in the salt-stressed banana roots in this study. These unknown genes, genes coding for hypothetical proteins and genes with unknown roles in abiotic stress response that were strongly up-regulated in the banana roots under high salinity (300 mM NaCl) are potential novel gene candidates for engineering salt stress tolerance in plants.

Table 5.1: Functions of the salt stress-responsive genes in banana roots

Function	Genes
Signaling	**Up-regulated:** Receptor-like kinase, protein phosphatase 2C, lipid transfer protein **Down-Regulated:** Heptahelical protein, calmodulin-binding proteins
Transcriptional regulation and DNA repair	**Up-regulated:** BTB/POZ domain-containing protein, H3 and H2A histones, homeobox-leucine zipper protein, DNA-3-methyladenine glycosylase I **Down-regulated:** AP2/ERF domain-containing transcription factor, ethylene response factor 2
Transport	**Up-regulated:** Nodulin-like protein, inorganic phosphate transporter, phosphate transporter, auxin-induced protein 5NG4-like, VAMP protein SEC22, cadmium/zinc-transporting ATPase 3-like **Down-regulated:** Hexose transporter, amino acid permease, amino acid transporter AAP4
Stress defence	**Up-regulated:** Peptide-N4-(N-acetyl-beta-glucosaminyl) asparagine amidase-like, peroxidase 47-like, peroxidase 3-like, late embryogenesis abundant (LEA) protein, xylanase inhibitor, glutamine synthetase **Down-regulated:** Heat shock protein 70, phenylalanine ammonia-lyase

Table 5.1: Continued

Function	Genes
Cell wall modifications and cellular structures	**Up-regulated:** Galactosyltransferase family protein, galacturonosyltransferase-like 3, fasciclin-like arabinogalactan, alpha-expansin 7, beta-expansin 3, cell-wall associated hydrolases, beta-glucanase, 3-ketoacyl-CoA synthase 4-like, beta-tubulin **Down-regulated:** Pectinesterase-1 precursor, pectinesterase/pectinesterase, dirigent-like protein, lignin-forming anionic peroxidase
Metabolism	**Down-regulated:** Alcohol dehydrogenase, pyruvate kinase, CBS domain-containing protein CBSX5-like
Other stress-related functions	**Down-regulated:** 70 kDa peptidyl-prolyl isomerase-like, gibberellin 2-oxidases
Unknown functions in abiotic stress response	**Up-regulated:** Nepenthesin, Ycf68, pollen Ole e 1 allergen, ORF16-lacZ **Down-regulated:** Pathogenesis-related protein 4, harpin-induced 1, protein kinase Xa21, GEM-like protein 5, putative acyl-acyl carrier protein (ACP) thioesterase type B, plant viral-response family, phytocyanin, WD-40 repeat family protein, WD-40 domain-containing F-box protein, stem-specific protein TSJT1-like

Partial list of genes with the highest expression fold change in the salt-stressed banana roots with 300 mM NaCl

5.2.2 Down-regulation

Like the up-regulated genes, the strongly down-regulated genes in the salt-stressed banana roots can functionally be classified into similar functional groups including signalling, transcriptional regulation, transport, cell wall modification, stress defence and unknown functions (Table 5.1).

Although coding for proteins advantageous for tolerating abiotic stresses as reported in other plant species, many of the strongly down-regulated genes were greatly down-regulated in the salt-stressed banana roots in this study (Table 5.1). This is probably due to the magnitude of salt stress above the threshold that banana roots can tolerate and some metabolic pathways and gene expression might be shut down. Besides, the expression of these genes may have elevated and peaked, and started to decrease at the time point when RNA was sampled in this present study. Moreover, stress tolerance conferred by these genes may be specific to particular stress factor, species and/or tissues only, but not in salt stress, root tissue and/or *Musa* species. Banana roots may deploy other strategies to overcome salt stress.

(a) Signaling

Heptahelical protein (HHP) is a transmembrane protein functions as receptors that are involved in signal transduction. Expression of HHP is required for plant tolerance to abiotic stresses (Chen et al., 2010a; Yadav & Tuteja, 2011). A HHP domain-containing protein in rice has been described as a transmembrane transducer and demonstrated to be responsive to NaCl and ABA treatments (Yadav & Tuteja, 2011). HHP has also been reported to express in regions in which the regulation of turgor pressure is crucial, such as roots and stomata, and serve as a negative regulator in abscisic acid (ABA) and osmotic signalling in Arabidopsis. A HHP-mutant showed decreased sensitivity to drought and ABA stress (Chen et al., 2010a). In this present

study, a putative HHP4 was decreased about 25-fold in expression in the salt-stressed banana roots. This indicates that this putative HHP probably functions as a signal transducer and negatively regulates salt stress signaling in banana roots.

Calmodulin-binding proteins (CaM-BPs) are proteins that bind to and mediate the action of a ubiquitous multifunctional calcium sensor, calmodulin (CaM), which in turn, regulate activity or function of many unrelated proteins (Golovkin & Reddy, 2003). There are many types of CaM-BP, including kinesin-like CaM-BP, CaM-binding protein kinase and a number of novel CaM-BPs, which are involved in the regulation of plant growth and development and responses to environmental clues (reviewed in Bouche et al., 2005; Wan et al., 2012; Zhou et al., 2012b; Yamamoto et al., 2013). A novel CaM-BP has been reported to be expressed in the guard cells and root cortex of Arabidopsis, and a CaM-BP-mutant Arabidopsis displayed a smaller stomatal aperture, a decreased water loss rate and a shorter primary root (Zhou et al., 2012b). A putative CaM-BP was down-regulated with about 18-fold in expression reduction in the salt-stressed banana roots. This putative CaM-BP probably mediates stress signaling in banana roots, which in turn, triggers morphological changes for salt stress adaptation in the banana roots.

(b) Transcriptional regulation

AP2/ERF domain-containing transcription factors are one of the largest and well-characterized transcription factor families in plants, and have been reported to positively regulate biotic and abiotic stresses (Mizoi et al., 2012; Shi et al., 2012; Sewelam et al., 2013;). However, a 44-fold reduction in expression was observed for a putative AP2/ERF domain-containing transcription factor in salt-stressed banana roots.

Ethylene is a class of plant growth regulator and stress hormone. Ethylene response factors (ERFs) are known to control transcription by binding to the GCC motif

found in the promoter region of the ethylene-regulated genes (Pirrello et al., 2006). Expression of ERFs has been indicated in improved plant tolerance to biotic and abiotic stresses (Anderson et al., 2010; Zhang et al., 2010a; Zhang et al., 2010b; Tian et al., 2011; Pan et al., 2012). In this study, however, a putative ethylene response factor 2 was down-regulated in the salt-stressed banana roots with about 15-fold in expression reduction.

(c) Transport

A putative hexose transporter showed expression reduction about 44-fold in the salt-stressed banana roots. Hexose transporters, which are found in the sink tissues, such as root, are involved in sugar-sensing and uptake of hexose, including glucose in plants (Schofield et al., 2009; Hayes et al., 2010). Although the uptake of hexose should be advantageous for roots to maintain high osmotic pressure under salt stress condition, banana roots in this present study, however, showed huge reduction of hexose transporter gene expression.

Amino acid permease, a membrane protein with high affinity to amino acid, is important for transportation of amino acid into the cells (Gotz et al., 2007; Svennerstam et al., 2008). Expression reduction about 34- and 23-fold for a putative amino acid permease and a putative amino acid transporter AAP4, respectively, was observed in salt-stressed banana roots. This reduction is probably because of the disruption of normal growth and metabolism in banana roots and energy was channeled to the mechanisms important for coping high salinity.

(d) Stress defence

Heat shock protein 70 (HSP70) is a conserved gene family in eukaryotes and is involved in a variety of cellular processes, including protein folding, as chaperones for

newly synthesized proteins and cytoprotection (Al-Whaibi, 2011; Jiang et al., 2014). However, a putative HSP70 was reduced about 27-fold in expression in the salt-stressed banana roots. Expression of HSP70 in *Fucus serratus* and *Lemna minor* has been shown to be increased to a peak and subsequently decreased as the levels of stressors (heat shock, osmotic stress and metal stress) increased (Elyse Ireland et al., 2004). This probably explains the decrease in the HSP70 expression in banana roots in 300 mM NaCl though HSP70 is a known stress defence protein in plants.

. Phenylalanine ammonia-lyase (PAL) is the key enzyme in the phenylpropanoid biosynthesis pathway leading to the production of phenylpropanoids such as lignins, flavonoids and coumarins (Schuster & Retey, 1995). It has been reported that expression of PAL is related to disease defence via production of defence-related secondary metabolites in the roots of rice, soybean and banana (Duan et al., 2014; Jain & Choudhary, 2014; Vaganan et al., 2014). A PAL gene has been demonstrated to be responsive to wounding, salicylic acid treatment and salinity stress in the seedlings of safflower, with expression peaked at 3-hour after treatment and levels off gradually afterwards. In this study, a putative was reduced in expression about 14-fold in the salt-stressed banana roots.

(e) Cell wall modifications

A putative pectinesterase-1 precursor (C125837) and a pectinesterase/pectinesterase inhibitor 41-like (CL1049Contig1) were strongly down-regulated with expression reduction about 68-fold and 15-fold respectively in the salt-stressed banana roots. Although the gene sequence and domain information (pectinesterase-1 precursor contains pectinesterase and pectin methylesterase inhibitor domains; pectinesterase/pectinesterase inhibitor 41-like contains pectinesterase 41 and pectinesterase inhibitor 41 domains) were deposited in the public protein databases,

their exact biological functions have not been fully elucidated. Pectinesterase probably acts by demethylesterification of cell wall pectin and subsequent degradation of cell wall (UniProt, http://www.uniprot.org/). In contrast, lignin-forming anionic peroxidases catalyze the formation of lignin polymer and form rigid cross-links between cellulose, pectin, lignin and extensin in the secondary plant cell wall (Lagrimini et al., 1987). A lignin-forming anionic peroxidase was also strongly down-regulated in the salt-stressed banana roots with 30-fold expression reduction. The strong down-regulation of the pectinesterase-related genes and lignin-forming anionic peroxidase in the banana roots may be involved in suspending cell wall extension and subsequent root growth to protect cell wall from degradation in the high saline condition.

Dirigent-like protein is involved in lignin/lignan biosynthesis and is related to lignin-based Casparian strip formation in Arabidopsis root (Hosmani et al., 2006), plant defence in spruce (Ralph et al., 2006) and abiotic stress response in sugarcane (Jin-Long et al., 2012). Although being a positive regulator of stress reported in other plants, a putative dirigent-like protein was decreased about 18-fold in expression in the salt-stressed banana roots.

(f) Metabolism

Alcohol dehydrogenase (Adh) is a well-characterized enzyme involved in anaerobic production of acetaldehyde and ethanol in higher plants, and has been reported to be induced in root tissues of cottonwood and Arabidopsis in response to anaerobic stresses (Chung & Ferl, 1999; Kimmerer, 1987). Expression of Adh gene has also been shown to be increased in the waterlogged roots of chrysanthemum and soybeans, which peaked at 2-hour and 6-hour respectively, and subsequently decreased (Komatsu et al., 2011; Yin et al., 2013a). In this present study, a putative alcohol

dehydrogenase 1 was reduced about 19-fold in expression when the salt-stressed banana roots were sampled at 48-hour.

Pyruvate kinase, an enzyme that is involved in glycolysis, produces a pyruvate molecule and energy in the form of ATP. It has been reported to be expressed in various cell types in rice including cortical parenchyma cells in roots and is important in plant development (Zhang et al., 2012c). Osmotic, ionic and oxidative stresses caused by high salinity may restrict the availability of oxygen to the root tissue and hamper the aerobic respiration process. Therefore, a strong expression reduction of about 20-fold in putative pyruvate kinase was observed in the salt-stressed banana roots.

Cystathionine β-synthase (CBS) domain binds adenosine-containing ligands, such as AMP, ATP or NADPH and helps regulating enzymatic activities of its adjacent domain (Baykov et al., 2011). CBS domain-containing proteins have been suggested to play regulatory roles for many enzymes and thus help in maintaining the cellular redox balance in Arabidopsis (Yoo et al., 2011) and function as sensor of cellular energy in human (Baykov et al., 2011). Expression of the CBS domain-containing proteins has been observed to be responsive to various stresses including salinity and drought (Kushwaha et al., 2009). A gene coding for a CBS domain-containing protein has been shown to be induced by high salinity, heavy metal, and oxidative stresses at seedling stage of rice, and overexpression of this gene enhanced tolerance in tobacco plants to abiotic stresses including high salinity (Singh et al., 2012). Although being a positive regulator to abiotic stresses in other plants, a putative CBS domain-containing protein, CBSX5-like was down-regulated in the banana roots in this study, with about 18-fold reduction in expression.

(g) Other stress-related functions

A transcript with sequence homology to 70 kDa peptidyl-prolyl isomerase-like was down-regulated about 30-fold in the salt-stressed banana roots. Peptidyl-prolyl isomerase (PPIase) superfamily consists of cyclophilins, FK506 binding proteins (FKBP) and parvulins and catalyzes the isomerization of peptide bonds important for protein folding and regulation of cellular processes such as cell signaling, biogenesis and activities of several receptors, plant growth and stress responses in plants (Edvardsson et al., 2003; Wang et al., 2014b; Yu et al., 2012b). PPIases have been reported to be induced by salinity and overexpression of PPIase improved salinity tolerance in plants (Trivedi et al., 2014; Wang, et al., 2014b). Although demonstrated as a positive regulator of salinity stress in plants, the putative peptidyl-prolyl isomerase-like gene was strongly down-regulated in the salt-stressed banana roots in this study.

Gibberellin 2-oxidases (GA2oxs) regulate GA homeostasis, plant growth and development, gravity response and stress tolerance by catalysing the degradation of endogenous bioactive gibberellins (GAs) (Lo et al., 2008). Overexpression of GA2ox has been shown to enhance salinity tolerance in rice (Shan et al., 2014). However, a putative gibberellin 2-oxidase showed reduction about 15-fold in expression in the salt-stressed banana roots.

Besides, a number of genes strongly down-regulated in the salt-stressed banana roots in this present study do not have a known abiotic stress-related function as shown in Table 5.1. This includes a putative PR-4, a putative protein kinase Xa21, a putative harpin-induced 1 (HIN1), a putative phytocyanin, GEM-like protein 5, WD-40 repeat family proteins, acyl-acyl carrier protein (ACP) thioesterase, a stem-specific protein TSJT1-like and a putative plant viral-response family gene (Table 5.1). In addition, a number of unannotated genes (including C124675, CL1Contig1647, CL1Contig3633, Contig1064, CL1Contig3838, CL1908Contig2, C93870, CL1Contig214, C78114,

CL1Contig1207, C50810, C36240, CL3071Contig1, CL1Contig5347 and CL1Contig3634) and genes coding for hypothetical protein (including CL1Contig2893, CL1Contig6581, Contig1187, CL1Contig3371, C143622, scaffold7670, CL2555Contig1) were found to be strongly down-regulated in the salt-stressed banana roots in this study. These novel genes are of interested for further salt-stress functional genomic investigations.

5.3 Functional information from small RNA sequencing of salt-stressed banana roots

Small RNA sequences surveyed using direct cloning approaches in this study yielded no to very low numbers of miRNA sequences, which is most probably due to the poor quality libraries constructed in this study. At the time this part of work was carried out, there was no facility available in our institution to analyze RNA integrity and RNA integrity was solely judged by the presence of intact 28S and 18 rRNA bands on agarose gel after electrophoresis. With the advent of advanced technology, such as Bioanalyzer™ (Agilent Technologies) and Experion™ (BioRad) to assess RNA integrity, it is advisable to carry out RNA integrity assay to ensure high quality RNA to be used for small RNA cDNA library construction. RNA Integrity Number (RIN) ranges from 1 to 10 with 1 indicating fully degraded RNA and 10 being intact RNA. RNA with RIN at least 8 is required for RNA-Seq (Schroeder et al., 2006). RNA samples with degradation or slight degradation can cause over-representation of degraded mRNA, rRNA, tRNA and other larger RNA species in the small RNA cDNA library. The cost of next generation sequencing became affordable and easier to access during this part of work. Therefore, we have shifted from direct cloning to high-throughput sequencing approach to study microRNA.

The three samples (CTR, TR100 & TR300) sequenced using Illumina high-throughput sequencing platform were shown to have similar length distribution of small RNA, with 21-nt being the most abundant sRNAs followed by 20-nt and 24-nt. The high abundance of 20-nt and 21-nt sRNAs in the small RNA libraries is probably due to the presence of miRNAs as most plant mature miRNAs are 20- to 21-nt long. It is also speculated that the abundance of 24-nt sRNAs is due to the presence of siRNAs.

Axtell & Bartel (2005), Zhang et al. (2006b) and Cuperus et al. (2011) have suggested that some miRNA families are evolutionarily conserved across all major plant lineages including mosses, gymnosperm, monocots and eudicots. Several individual miRNA regulatory circuits have remained intact throughout the evolution and diversification of plants (Axtell & Bartel, 2005). In this study, 181 distinct small RNA sequences found matches in the Plant MicroRNA Database (PMRD), indicating these miRNAs are evolutionary conserved. Using these 181 plant orthologous miRNA sequences as search queries, 247 miRNA targets were identified in the assembled unigenes in this study using psRNAtarget, a commonly used tool designed for plant miRNA target prediction (Dai & Zhao, 2011). These targets are either known (previously reported) to be targeted by miRNA, such as mRNAs coding for NAC domain-containing proteins (target of miR164) (Laufs et al., 2004; Baker et al., 2005; Guo et al., 2005), GRAS family transcription factors (target of miR171) (Branscheid et al., 2011), AP2/AP2-like transcription factors and laccases (target of miR397) (Lu et al., 2011; Ravet et al., 2011), or novel targets, which have not been observed as an miRNA target in other plant species.

Besides, 56 putative new miRNA sequences specific to *Musa* species, which have not been reported in other plant species and are probably unique in banana, were identified in this present study. The predicted targets for these new miRNAs in banana are also novel (not reported before) though some are mRNAs that have been reported to

be a target of a known miRNA, such as AUXIN RESPONSE FACTOR 6 (target of miR167) and DEAD-box helicases (target of miR164 and miR408). This shows that the transcriptome and gene regulation in plants are complex.

While some miRNAs were up-regulated or down-regulated in both TR100 and TR300, a number of miRNAs showed inverse expression pattern. This indicates that these miRNAs may be sensitive to a shift in stress magnitude. A raise from 100 mM to 300 mM NaCl diverted their expression pattern. Some miRNAs that showed no significant expression change in 100 mM NaCl (relative to control) were then up-regulated or down-regulated in 300 mM NaCl. This implies that some miRNAs may not be sensitive to lower salt stress (100 mM NaCl) but triggered by higher NaCl concentration (300 mM). MiRNAs respond to salt stress in a dose-dependent manner. In addition, the change in miRNA expression profiles from control to TR100 and TR300 may also be caused by the absence or under-representation of some data in one or other dataset due to the use of only single sRNA-seq dataset each in the control, TR100 and TR300. The focus of this present study is on high salinity (300 mM NaCl).

By comparing this present study to the previously reported high-throughput salt stress-related transcriptomic studies in plants, a number of miRNAs, such as miR156, miR159, miR167, miR168, miR171 and miR396, can be observed as common salt stress responsive miRNAs in plants (Table 5.2). A higher number of salt stress-responsive orthologous miRNAs is shared between the salt-stressed banana roots and the salt-stressed roots of *Zea mays* compared to other tissues and species, as shown in Table 5.2. Salt stress response of some plant orthologous miRNAs is probably unique to particular species, such as, miR158 and miR165 in *Arabidopsis* and miR399 in the roots of *Zea mays*. Plants express a different set of miRNAs in response to salt stress in different tissues/organs (Table 5.2).

Table 5.2: Salt stress-responsive plant orthologous miRNAs investigated using high-throughput miRNA expression profiling approaches

MiRNA	*Musa acuminata* roots[1]	*Zea mays* roots[2]	*Arabidopsis thaliana* seedlings[3]	*Populus euphratica* leaves[4]	*Saccharum sp.* shoots[5]
miR156	●	●	●	●	●
miR157	●				
miR158			●		
miR159	●	●	●		●
miR160		●		●	
miR162	●	●			
miR164	●	●		●	
miR165			●		
miR166	●	●			●
miR167	●	●	●		●
miR168	●	●	●		●
miR169	●		●		●
miR171	●	●	●	●	
miR172	●			●	
miR277				●	
miR319		●	●		
miR393			●	●	
miR394			●	●	
miR395		●		●	
miR396	●	●	●		●
miR397	●				●
miR398					●
miR399		●			
miR528	●				●
miR529	●				
miR530				●	
miR535	●				
miR827	●				
miR1444				●	
miR2910	●				
miR5139	●				
miRf10192-akr	●				

Salt-stress responsive miRNAs are the statistically significant differentially-expressed miRNAs
[1]Present study with *Musa acuminata* roots; 300 mM NaCl; sRNA-Seq
[2]*Zea mays* root; 200 mM NaCl; miRNA microarray (Ding et al., 2009)
[3]*Arabidopsis thaliana* seedlings; 300 mM NaCl; miRNA microarray (Liu et al., 2008)
[4]*Populus euphratica* shoots; 200 mM NaCl; sRNA-Seq (Li et al., 2013a)
[5] *Saccharum sp.* leaves; 170 mM NaCl; sRNA-Seq (Carnavale Bottino et al., 2013)

5.4 MicroRNAs were predicted to target important salt stress-responsive genes in banana roots

Most of the differentially-expressed miRNAs in this present study showed an inverse expression pattern (negatively correlated) with at least one of their predicted targets (Figure 4.23), indicating that these targets are possibly real, as mRNA cleavage is the major mode of posttranscriptional gene silencing and has been observed in many miRNAs in plants (Rhoades et al., 2002). Nevertheless, a minority of the miRNAs and their predicted targets did not show an inverse expression pattern (Figure 4.23). This is probably due to the use of different mode of action by miRNAs where translational inhibition rather than mRNA cleavage took place. It is also possible that these predicted targets are not true targets of the miRNAs in banana root. Below the various predicted targets and functions are described.

Seven distinct miR156 mature sequences (mac-miR156, mac-miR156a, mac-miR156i, mac-miR156k, mac-miR156o, mac-miR156o.1, mac-miR156r) were down-regulated in the salt-stressed banana roots, while their targets, tropine dehydrogenase, Avr/Cf-9 rapidly elicited gene and an unknown gene (CL2012Contig1) were up-regulated (Figure 4.23). Tropine dehydrogenase has been reported to be involved in the biosynthesis of tropane alkaloids and accumulated in *Datura stramonium* and *Hyoscyamus niger* roots (Hashimoto et al., 1992; Koelen & Gross, 1982). Tropane alkaloids have also been reported as metabolites for *Brugmansia suaveolens* defence against herbivorous insects (Arab et al., 2012). Whereas, Avr9/Cf-9 rapidly elicited gene is a resistance (R) gene involved in fungal and viral defence response in tomato and tobacco (Rowland et al., 2005; van den Burg et al., 2008). This suggests that the biotic defence mechanisms may also be deployed by the banana roots for salt stress defence, and there are possible connections between the biotic and abiotic stress responses. Therefore, miR156 is predicted to regulate stress defence in the banana roots.

Four distinct miR157 sequences (mac-miR157, mac-miR157.1, mac-miR157c, mac-miR157m) were found to be differentially expressed and down-regulated in the salt-stressed banana roots, and seven target genes can be found in the assembled banana transcriptome in this study (Figure 4.23). Among these targets, CL2012Contig1, Avr9/Cf-9 rapidly elicited gene, Scaffold4109 and E3 UBQ-protein ligase are also targets of the miR156 in this study. While unigene C111260, CRK1 gene and protein ROOT HAIR DEFECTIVE 3 (RHD3) are unique targets of miR157 and were up-regulated in the salt-stressed banana roots (Figure 4.23). CDPK-related kinase (CRK) has sequence similar to calcium-dependent protein kinase (CDPK) but with significant degenerate sequence in the calmodulin (CaM)-like domain. It is functionally a kinase but lacking calcium-binding activity (Furumoto et al., 1996; Leclercq et al., 2005) and has been shown to be involved in hormone signalling and root growth in Arabidopsis (Rigo et al., 2013). Whereas, ROOT HAIR DEFECTIVE 3 (RHD3) has been reported as a GTP-binding protein expressed ubiquitously in plants and it has been implied in root formation (Xu et al., 2012). Therefore, miR157 is predicted to regulate signalling and root formation in response to salt stress in the banana roots.

Two miR159 (mac-miR159c, mac-miR159g) mature sequences were down-regulated in the salt-stressed banana roots and they were predicted to target a putative chorismate mutase and a putative salt responsive protein 2 respectively (Figure 4.23). Both chorismate mutase and salt responsive protein 2 were up-regulated in the banana roots suggesting their important roles in the salt-stressed banana roots. Chorismate mutase catalyzes the early biosynthesis step of phenylalanine and tyrosine, the precursors to a number of secondary metabolites important for development and stress responses in Arabidopsis (Mobley et al., 1999). Salt responsive protein 2 has been described as an early salt stress response gene found in tomato root analyzed using suppression subtractive hybridization and microarray (Ouyang et al., 2007). However,

no further functional information is available for this gene. These showed the possible roles of miR159 in regulating secondary metabolism and early salt stress response in the banana roots.

In this study, a miR162 (mac-miR162b.2) was predicted to target a putative dipeptidyl peptidase. MiR162 was down-regulated in the salt-stressed banana root while its predicted target, dipeptidyl peptidase was up-regulated (Figure 4.23). Dipeptidyl peptidase has been suggested to degrade small proline-containing peptides and with physiological importance in barley (Davy et al., 2000). This suggests the possible role of miR162 in modulating salt stress by releasing proline, an osmoprotectant, from proline-containing peptides.

Mac-miR168 was up-regulated in the salt-stressed banana roots, while its only predicted target, 12-oxophytodienoate reductase was down-regulated (Figure 4.23). 12-oxophytodienoic acid reductase has been reported to be involved in biosynthesis of jasmonic acid, a stress signalling molecule in plants (Tani et al., 2008; Schaller & Stintzi, 2009). The 12-oxophytodienoic acid reductases have also been reported to be expressed predominantly in plant roots and induced by abiotic stresses including salinity and osmotic stress (Biesgen & Weiler, 1999; Zhang et al., 2007; Dong et al., 2013). Overexpression of 12-oxophytodienoic acid reductase has been shown to confer salinity tolerance in wheat (Dong, et al., 2013). Although being a positive regulator of abiotic stresses, 12-oxophytodienoic acid reductase, which was predicted to be targeted by miR168, was down-regulated in the salt-stressed banana roots.

SCARECROW-LIKE (SCL) is a member of the GRAS (GIBBERELLIN-INSENSITIVE, the REPRESSOR of ga1-3 and SCARECROW) family transcription factors and miR171 has been reported to target SCL transcription factors in plants (Curaba et al., 2013). GRAS family transcription factors are important transcriptional regulators for plant growth and development. The GRAS family transcription factor has

been reported to control lateral root development (Pysh et al., 1999; Tian et al., 2014). A miR171 (mac-miR171b) was down-regulated in the salt-stressed banana roots while its target, a putative GRAS family transcription factor, was down-regulated (Figure 4.23). This suggests a possible miR171-mediated down-regulation of GRAS family transcription factor in controlling banana root development in response to salt stress.

The two miR172 mature sequences differentially expressed in the salt-stressed banana roots were both down-regulated. Both mac-miR172a.2 and mac-miR172f were predicted to target the same targets, which are poly(A)-binding protein (down-regulated), C77218 (down-regulated), ethylene-responsive transcription factor (no significant expression change), peroxisomal targeting signal (up-regulated), AP2 domain-containing transcription factor (up-regulated) and methyl-CpG-binding domain-containing protein (down-regulated). MiR172f has two additional predicted targets, which are both putative phospholipase D, with one being down-regulated and the other being up-regulated (Figure 4.23). Peroxisomal-targeting signal is responsible for localization of protein to peroxisomes (Palma et al., 2009). Plant peroxisomes have been reported to be involved in various processes, including primary and secondary metabolism, development, and responses to abiotic and biotic stresses (reviewed in Hu et al., 2012). The salt-stressed banana roots may produce and accumulate ions, metabolites and free radicles which are toxic to the cells. The up-regulation of the peroxisomal-targeting signal in the banana roots may help increasing the peroxisomal-based detoxification activity in banana root under salt stress.

AP2 and AP2-like transcription factors are conserved targets of miR172 in plants (Aukerman & Sakai, 2003). AP2 domain-containing transcription factors have been reported to be a stress regulator with positive role in abiotic stress response in *Medicago sativa* and *Jatropha curcas* (Zhang et al., 2005b; Tang et al., 2007; Tang et al., 2011). An AP2 domain-containing transcription factor was up-regulated in the salt-

stressed banana roots and is predicted to be targeted by mac-miR172a.2 and mac-miR172f (Figure 4.23). This suggests a possible role of miR172 in regulating salt stress response in the banana roots.

Phospholipase D (PLD) is an enzyme that catalyses the hydrolysis of phospholipids into signalling molecule, phosphatidic acid (PA) (Gao et al., 2013). Phospholipase D is crucial for plant responses to stress and signal transduction and overexpression of a phospholipase has been shown to enhance abiotic stress tolerance in plant (Wang et al., 2014a). The expression a PLD was up-regulated in the salt-stressed banana roots. This suggests that miR172-mediated PLD expression regulates stress signalling and responses in the banana roots.

In this present study, mac-miR397 was down-regulated and all of its targets, a putative laccase 110a, a sialyl transferase-like protein, an osmotic stress-activated protein kinase and a signal peptidase were up-regulated (Figure 4.23). MiR397 has been described as a transcriptional regulator of laccase and members of laccase copper protein family (De Luis et al., 2012; Zhang et al., 2013c). Laccase has also been reported to be involved in lignin biosynthesis in poplar and is regulated by miRNA (Lu et al., 2013b) and laccase expression level in tomato (a dicot) and maize (a monocot) roots was elevated under salinity stress (Liang et al., 2006; Wei et al., 2000). A lac2 mutant has been reported to show reduced root elongation during dehydation compared to the wild type (Cai et al., 2006). The inverse expression pattern between mac-miR397 and its predicted target, a putative laccase 110a in the salt-stressed banana roots implies that miR397-mediated laccase regulation is important for modulating lignin synthesis and root growth in response to salt stress in banana.

Sialyl transferases (SiaTs), a type of glycosyltransferase, catalyze the transfer of sialic acid to oligosaccharide. This enzyme is commonly found in animals and is an important player in various biological processes. SiaT-like proteins have been described

to be involved in pollen tube growth and pollen germination in Arabidopsis and rice (Deng et al., 2010; Takashima et al., 2006). The function of the putative sialyl transferase-like protein in plant roots is not understood.

Osmotic stress-activated protein kinase has been reported to be involved in nitric oxide signaling and hyperosmotic stress response in tobacco (Wawer et al., 2010; Burza et al., 2006). This suggests the possible mac-miR397-mediated up-regulation of the putative osmotic stress-activated protein kinase in the salt-stressed banana roots and is probably important in stress signaling and coordinating osmotic stress response.

Signal peptidases are intramembrane proteases responsible for maturation of membrane proteins (including signal peptides) with developmental and physiological importance in Arabidopsis (Hoshi et al., 2013; Shipman-Roston et al., 2010). The expression of the putative signal peptidase in banana roots was predicted to be regulated by mac-miR397 and the up-regulation of this gene may be related to stress signaling and subsequent developmental and physiological stress adaptation.

Mac-miR528b was up-regulated in the salt-stressed banana roots. Three of its predicted targets were down-regulated – a hypothetical protein, a polyphenol oxidase, and a leucyl-tRNA synthetase (Figure 4.23). Polyphenol oxidases (PPOs) have been reported to be involved to environment adaptation in plants. Tissue browning following oxidation of ortho-diphenols to ortho-quinones might be important for defence in plants (Tran et al., 2012). In this study, however, a putative PPO was down-regulated in the salt-stressed banana roots. It is possible that the PPO-induced tissue browning was not triggered in the early stage of the salt stress in banana roots in this study and the low PPO expression was possibly modulated by mac-miR528b.

Leucyl-tRNA synthetase is an enzyme that catalyzes the transfer of L-leucine to tRNA (aminocylation of tRNA) and is important in protein translation (Chopra et al., 2013). The down-regulation of this gene, probably mediated by mac-miR528b, may be

a reaction to reduce protein translation process so that energy can be channeled to mechanisms for cellular homeostasis in the salt-stressed banana roots.

An miR529 (mac-miR529b) was down-regulated and both of its predicted targets, an unnamed protein product and an ATP/GTP-binding protein, were up-regulated in the salt-stressed banana roots. ATP/GTP binding proteins are group of proteins containing ATP/GTP binding motifs. These proteins include receptor-like kinases or R genes and heat shock proteins (Benedetti et al., 1998; Biselli et al., 2010; Tameling et al., 2002). Therefore, the functions of ATP/GTP binding proteins vary from signaling, defence, stress responses to growth and development in plants. This gene is predicted to be a target of mac-miR529b and up-regulation of this gene is probably important for stress signaling and response and the subsequent growth and developmental regulation in the salt-stressed banana roots.

A miR535 (mac-miR535c) was down-regulated in the salt-stressed banana roots while only one of its predicted targets, a putative G-type lectin S-receptor-like serine/threonine-protein kinase (GsSRK) was up-regulated. Receptor-like protein kinases (RLKs) are common signal sensing proteins. In a recent report, a GsSRK has been shown to be induced by salt stress and demonstrated as a positive regulator of plant tolerance to salt stress (Sun et al., 2013). This coincides with our finding which expression of a putative GsSRK was induced in the salt-stressed banana roots. In addition, this gene is predicted to be regulated by miR535 in this study.

A miR2910 (mac-miR2910) was up-regulated while one of its predicted targets, a cleavage and polyadenylation specificity factor (CPSF) was down-regulated. CPSF is a component of the mRNA 3'-end processing apparatus in eukaryotes and has been implicated in reproductive development (Xu et al., 2006) and immune response (Bruggeman et al., 2014) in Arabidopsis. The role of this mac-miR2910 and its predicted target, CPSF in the salt-stressed banana roots is not known.

Dehydrins (DHNs) are multi-family proteins induced by stresses that cause cellular dehydration in plants. These stress proteins are involved in plant protective reactions against environmental stresses (Yang et al., 2014). A *Musa*-specific miRNA, mac-miR6 was predicted to target CL1Contig328 that contains a dehydrin domain. Inverse relation was observed where mac-miR6 was down-regulated and its predicted target was up-regulated in the salt-stressed banana roots. The up-regulation of this dehydrin domain-containing gene was probably caused by salt stress that can dehydrate banana roots and may be responsible for cellular protection from dehydration.

Mac-miR19 was down-regulated in the salt-stressed banana root and only one of its targets, zinc finger CCCH domain-containing protein, was up-regulated. While the other two predicted targets, a putative nucleobase-ascorbate transporter and a putative zinc finger C-x8-C-x5-C-x3-H (CCCH) type family protein remained unchanged in expression. Cysteine3Histine (CCCH)-type zinc finger protein are RNA-binding protein with regulatory roles in mRNA processing. The CCCH-type or CCCH domain-containing proteins have been reported to regulate plant growth, development, and stress response via RNA regulation (reviewed in Bogamuwa & Jang, 2014; Peng et al., 2012; Wang et al., 2008). The miRNA-mediated regulation of a zinc finger CCCH domain-containing protein in the banana roots is possibly important in proper mRNA processing during salt stress.

Mac-miR35 was predicted to target a DEAD-like helicase domain-containing protein. Mac-miR35 was up-regulated and its target was down-regulated in the salt-stressed banana roots. Helicases, including DEAD-box RNA helicases, have been reported as one of the effector proteins in regulating salinity stress response (reviewed in Mahajan & Tuteja, 2005; Turan et al., 2012). Although known as an effector protein in salinity stress response, the transcript coding for a DEAD-like helicase domain-containing protein was down-regulated in the salt-stressed banana roots.

Mac-miR37 was down-regulated in the salt-stressed banana root while its predicted target, a chloride channel was up-regulated. Chloride channels (CLCs) are ubiquitous in prokaryote and eukaryotes, and function as ion channel and transporter in plants (Guo et al., 2014; Marmagne et al., 2007). The up-regulation of a putative CLC is possibly important for chloride transportation and homeostasis in banana roots during salt stress.

Mac-miR38 was up-regulated in the salt-stressed banana root and two predicted targets, unannotated gene C102056 and NADP-dependent glyceraldehyde-3-phophate dehydrogenase, were down-regulated. The other predicted target, a VHS and GAT domain protein was up-regulated. NADP-dependent glyceraldehyde-3-phophate dehydrogenase (NADP-GAPDH) is also a target of the mac-miR535c in this study. NADP-GAPDH catalyzes a glycolysis step in the presence of NADP, releasing energy and carbon molecules in the form of NADPH and 3-phospho-D-glycerate. Besides, it has also been shown as a part of the stress signaling pathway and plant protective systems, including intracellular detoxification and protection from reactive oxygen (Dizengremel et al., 2009; Holtgrefe et al., 2008). Although reported as a positive regulator of abiotic stress, transcript coding a putative NADP-GAPDH was down-regulated in the salt-stressed banana roots.

Mac-miR49 was down-regulated in the salt-stressed banana root while its targets CL2012Contigs (also a target of miR156 in this study), SORTING NEXIN 1 and ubiquitin carbonyl-terminal hydrolase were up-regulated. The other target, stearoyl-ACP desaturase did not show significant expression change. SORTING NEXIN 1 has been shown to be localized in the endosomal compartment of root cells, forming an endosome for transporting auxin efflux carrier PIN2 (Jaillais et al., 2006). The up-regulation of a putative SORTING NEXIN 1 in the salt-stressed banana roots may be important for elevating auxin transport activity for root morphology adaptation to salt

stress. Ubiquitin carbonyl-terminal hydrolase is a deubiquitinating enzyme involved in the processing of polyubiquitin precursors and ubiquitinated proteins (UniProt, http://www.uniprot.org/uniprot/). Up-regulation of this gene in the salt-stressed banana roots implies that the ubiquitin carbonyl-terminal hydrolase probably mediates post-translational modification and other protein regulation that is important for stress regulation in banana roots under salt stress.

Mac-miR62 was predicted to target an unannotated gene (CL7639) and a putative protein IQ-domain 1. This miRNA was up-regulated while both of its targets were down-regulated in the salt-stressed banana root. Protein IQ-domain belongs to the IQD family and contains isoleucine glutamine (IQ) domain(s). It may be involved in cooperative interactions with Ca^{2+} sensors, calmodulins (CaM) or CaM-like proteins, and associated with nucleic acids to regulate gene expression at the transcriptional or post-transcriptional level (Abel et al., 2013). Protein IQD1 in Arabidopsis has been reported to contain a putative nuclear localization signal and several motifs known to mediate calmodulin binding and stimulate glucosinolate metabolism and plant defence (Abel et al., 2005; Levy et al., 2005). Although reported as a positive regulator in stress response in plants, a putative IQD1 was down-regulated in the salt-stressed banana roots.

Mac-miR66 was up-regulated in the salt-stressed banana root and two predicted targets, oligopeptide transporter 7-like and E3 ubiquitin protein ligase were up-regulated. The other target, DEAD-like helicase, also a target of miR166 and mac-miR35 in this study, was down-regulated. Helicases have been reported as a group of effector proteins in regulating salinity stress response (reviewed in Mahajan & Tuteja, 2005; Turan et al., 2012). Similar to the observation in mac-miR35 in this study, the putative DEAD-like helicase in banana roots was down-regulated during stress though it functions as a positive regulator of stress.

In general, the targets that showed inverse expression pattern with their corresponding miRNAs differentially expressed in the 300 mM NaCl-stressed banana roots are involved in signaling, transcriptional and translations regulation, stress defence, transport, cellular homeostasis, metabolisms, as well as other stress-related and unknown functions (Table 5.3).

It is commonly known that a single miRNA may have multiple targets, and an mRNA may be targeted by multiple miRNAs. These were also observed in this study. However, validation of these miRNA-target interactions is required. Some transcripts, such as ATP/GTP-binding protein, NADP-GAPDH, CL2012Contig1 and CL7009Contig, are probably highly regulated in this present study as they were targeted by more than one miRNA family.

Some hypothetical or unnamed proteins and unannotated genes were predicted as miRNA targets and showed inverse expression pattern with their corresponding salt stress-responsive miRNAs in this present study. They are CL2012Contig (target of mac-miR156 and mac-miR49), C111260 (target of mac-miR157), CL1Contig6497 (target of miR528), CL7009Contig1 (target of miR529), C102056 (target of mac-miR38) and CL7639 (target of mac-miR62). These miRNA targets are probably novel transcripts or proteins expressed in the banana roots in response to salinity stress and are of interest for functional investigation.

Table 5.3: Functions of the predicted miRNA targets

Function	Genes
Signaling	CDPK-related kinase, 12-oxophytodienoic acid reductase, phospholipase, osmotic stress-activated kinase, signal peptidase, ATP/GTP-binding protein, G-type lectin S-receptor-like serine/threonine-protein kinase
Transcriptional and translational regulation	GRAS family transcription factor, AP2 domain-containing transcription factor, leucyl-tRNA synthetase, cleavage and polyadenylation specificity factor, Cysteine3Histine (CCCH)-type zinc finger protein, ubiquitin carbonyl-terminal hydrolase, protein IQ-domain,
Transport and cellular homeostasis	Chloride channel, SORTING NEXIN 1
Stress defence	Tropine dehydrogenase, Avr/cf-9 rapidly, dipeptidyl peptidase, peroxisomal-targeting signal, polyphenol oxidase, dehydrin domain-containing protein (CL1Contig328), DEAD-like helicase domain-containing protein, laccase 110a
Metabolisms	Chorismate mutase, NADP-dependent glyceraldehyde-3-phosphate dehydrogenase
Other stress-related functions	Root hair defective 3, salt responsive protein 2
Unknown functions in abiotic stress response	Sialyl transferase

List of predicted targets of the miRNAs differentially expressed in the salt-stressed banana roots with 300 mM NaCl that showed inverse expression pattern between the miRNAs and their corresponding targets.

5.5 Suggestions for future research

In this present study, high-throughput next-generation sequencing technology, a powerful tool in genomics and functional genomics studies, was applied to deeply sequence salt-stressed root transcriptomes and sRNAomes in order to reveal important gene sequences and regulation in banana.

More than 31,000 expressed transcript sequences were produced from banana roots with a majority being protein-coding gene sequences. Among the expressed transcripts, 6,469 SSR-containing unigenes were found. The frequency of SSRs at 21% reported in this present study is similar to that reported in the lentil (16%) (Kaur et al., 2011). However, frequency of SSR at 6.4% reported in sweet potato root (Wang et al., 2010e) is much lower than that in our present study although the same search criteria were used. Di- and tri-nucleotides repeat motifs are the most abundant ones (Wang et al. 2010e), which is similar to that observed in this present study. These SSR-containing unigenes are potentially a rich source of sequence data for the discovery of cDNA-derived SSR markers, important for genetic mapping and population genetic studies, as well as markers development for marker-assisted selection (Kaur et al., 2011; Wang, et al., 2010e).

The hypothetical or unnamed proteins and unigenes without functional annotation that showed high expression fold change in salt-stressed banana roots in this present study are potential novel stress responsive genes for biotechnological applications. Further investigations, including functional genetic studies, are required to reveal their functions by observing traits conferred by these genes through gene overexpression and/or silencing. In addition, mRNA alternative splicing is also another potential aspect to be investigated in order to reveal the other layer of gene regulation in banana, in response to salt stress (Xue & Loveridge, 2004).

The predicted miRNA-mRNA regulatory models in this present study may provide clues for further investigations that can potentially lead to crop genetic improvement for enhanced tolerance to abiotic stresses. There are reported uses of only miRNA sequences to produce desirable traits in plants. It has been shown that miR164 was down-regulated in order to produce more lateral roots and improve drought resistance in Arabidopsis (Guo et al., 2005). In tomato, overexpression of miR169 caused reduced stomatal opening, decreased transpiration rate, lowered leaf water loss and enhanced drought tolerance (Zhang et al., 2011b). It has been shown that improvement in biomass production in switchgrass when miR156 was over-expressed (Fu et al., 2012). MiR159 has been applied to control flowering time in an ornamental plant, gloxinia (Li et al., 2013e). In rice, overexpression of miR397 has been reported to increase grain size and panicle branching and increase the overall grain yield up to 25% in a field trial (Zhang, et al., 2013c). Therefore, overexpressing or blocking of the salt stress-responsive miRNAs or their targets found in this present study can potentially confer tolerance to abiotic stresses in bananas. The overexpression or silencing of miRNAs or miRNA targets can also be done without transgenesis/cisgenesis by exogenous application of siRNA to block mRNA or miRNA precursors, and the use of disarmed virus to overexpress gene for stress tolerance or silent genes for stress sensitivity *in planta* for improved tolerance. In addition to the plant orthologous and *Musa*-specific (‘novel’) miRNAs, the huge number of unannotated small RNA sequences in the three small RNA datasets is likely to include functional small RNAs such as siRNA with functions yet to be revealed.

Due to cost constraint, each transcriptomic library was only sequenced once with a depth of 1 Gbp for mRNA and at least 10-million reads for sRNA. Although gene expression profiling using RNA-Seq data is possible and proven realiable using proper statistical analyses (Malone & Oliver, 2011), the gene expression profiles are

more convincing if replication of RNA-Seq or further validation using RT-qPCR are conducted. The current findings would be more convicing if additional libraries as biological replicates are sequenced or RT-qPCR with sufficient biological replicates is carried out to validate the gene expression profiles inferred from RNA-Seq.

Whole genome sequencing and re-sequencing of two wild banana species ancestral to the modern cultivated bananas, *Musa acuminata* and *Musa balbisiana* have been reported and made publicly available. The sequenced genome of banana contains about 35,000 protein-coding gene models which have been predicted computationally (D'Hont et al., 2012; Davey et al., 2013). However, functions of many genes still remain unknown. In order to utilize these genomic information for crop genetic improvement, there is a need to elucidate the functions of the predicted gene models and understand how these genes are regulated in different tissues or growth conditions. The understanding of the molecular mechanisms in stress response is important for engineering stress tolerance in plants by enhancing or introducing stress tolerance mechanisms using molecular techniques (Mizoi & Yamaguchi-Shinozaki, 2013; Nakashima et al., 2014).

Unlike animals, plants have no or limited ability to move away from the environmental stresses. Plants have evolved with a number of mechanisms that provide them plasticity to response and adapt to the ever-changing environmental conditions. These mechanisms involve regulatory genes that code for transcription factors, protein kinases and other signalling proteins, and structural genes that code for enzymes and structural proteins with protective functions (Jamil et al., 2011) (Mizoi & Yamaguchi-Shinozaki, 2013). The stress responsive genes found in this present study are in agreement with the observations by Mizoi & Yamaguchi-Shinozaki (2013) though some potential new targets or novel genes for stress tolerance were also found in this present study.

It is unlikely only a single gene or mechanism is triggered and responsible for stress adaptation in plants. Like most other stress tolerance traits, salt tolerance is a highly complex trait and polygenic in nature controlled by multiple loci (Qin et al., 2011; Jamil et al., 2011). The tolerance or susceptibility to abiotic stress in plants is a coordinated action of multiple, estimated at hundreds, of stress responsive genes, which also cross talk with other components of stress signal transduction pathways (Carillo et al., 2011; Nakashima et al., 2014). Nevertheless, individual genes may contribute strongly to stress resistance and have been employed via transgenic approaches for crop improvement for example with a sodium ion transporter gene (TmHKT1;5) to relieve salinity stress in durum wheat (Munns et al. 2012); an ABA biosynthetic gene, 9-cis-epoxycarotenoid dioxygenase (NCED), to increase drought resistance in petunia (Estrada-Melo et al., 2015); a heat shock protein (HSP70) isolated from a wild relative of sugarcane for drought and salinity tolerance in sugarcane (Augustine et al. 2015); a rice heme activator protein gene (OsHAP2E) for resistance to salinity, drought and pathogens in rice (Alam et al. 2015) and a pea DNA helicase, PDH45, to improve the regulation of sodium, which in turn, increased salinity stress tolerance in both tobacco (dicot) and rice (monocot) (Nath et al. 2015).

Although a large number of salt responsive genes have been identified to date, there remain large gaps in complete comprehension of the salt tolerance/sensitivity trait to fully map plant responses to salt stress (Jamil et al., 2011). There are possibly more candidate targets or novel genes associated with stress tolerance/sensitivity and alternative means of how plants respond to salt stress. Therefore, more efforts are still required for further elucidation of stress response mechanisms (Qin et al., 2011; Jamil et al., 2011). In addition to the stress responsive protein-coding genes, this present study also described another layer of the stress regulation where the possible roles of miRNA in stress adaptation mechanisms were suggested. Therefore, the stress response genes

found in this present study need to be investigated for their functions individually, and together, their complex biological systems.

With the current knowledge and understanding of stress adaptation mechanisms in plant, researchers are currently putting effort in tailoring plants with abiotic stress tolerance by gene transfer or molecular breeding (Hirayama & Shinozaki, 2010; Nakashima et al., 2014). There have been very limited genes, pathways or mechanisms related to abiotic stresses investigated in *Musa* species (bananas). Some challenges need to be addressed in order to translate the findings in this current study into practical applications. The potential investigations that may lead to production of a promising genetically improved banana plant with salt stress tolerance are suggested as below:

1. Selecting genes that may be major players of stress adaptation from this present study for investigation in depth. Genes that are regulated by salt-responsive miRNAs may be good candidates. The presence of miRNA-regulated mechanisms during salt stress indicates that these genes are possibly important players and their expression needs to be modulated. The genes that are predicted to be regulated by miRNA in this present study include signal peptidase, serine/threonine-protein kinase and CCCH-type zinc finger protein which are signalling and regulation proteins, and protein root hair defective, chloride channel and chorismate mutase which are structural proteins.

2. Functional annotation by homology-based searches alone is insufficient as a definitive argument to describe the precise activity or function of genes in banana. The predicted genes homologous to genes in other species still require thorough functional assessments through, for examples, gene overexpression and gene knock-out or knock-down in banana, or ectopic expression in model plants.

3. The gene functional investigation also includes study of gene structure, cis- and trans-regulatory elements, and the pathways and networks involved. It is also useful to include fine-scale time-point and cell-specific gene expression investigations using RT-

qPCR and *in situ* hybridisation assay.

4. It is also important to incorporate other *Musa* -omics data to determine crosstalk or interaction between the salt stress responsive mechanisms, pathways and networks using systems biology approaches.

5. The transcriptomic data produced in this present study can be used as molecular genetic resources for investigating genetic variation-associated function of salt tolerance genes between different *Musa* species and cultivars.

6. Stress-inducible promoter or tissue-specific promoter can be used to drive the expression of the candidate stress tolerance (protein-coding) genes in a spatiotemporal manner to avoid or minimize their possible undesirable side-effects to banana plants.

7. MiRNA genes that negatively regulate stress tolerance genes can be repressed or silenced using approaches such as miRNA sponges, and targetted gene mutation (including genome editing) to disrupt miRNA sequence or its promoter. Similarly, stress-inducible or tissue-specific promoter can be used to replace the original promoter of the miRNA genes.

8. Stacking of traits by gene pyramiding is useful for combining relevant genes or mechanisms in a plant. It has been suggested as a more promising approach in engineering stress tolerance in plants (Deinlein et al., 2014). This can be applied in banana by stacking genes found in this present studies. For examples, structural genes, chloride channel and hair root protein, and regulatory genes, signal peptidase and serine/threonine-protein kinase found in the present study can be introduced (overexpressed) together to provide multiple enhanced adaptative mechanisms to banana.

9. To date, limited success for commercial utilization of transgenic crops under field saline conditions as the test plants are not equipped with all necessary genes/mechanisms for successful growth under saline. They may require more genetic

determinants and physiological processes than those under controlled greenhouse (Jamil et al., 2011). Field trial of transgenic banana introduced with the candidate genes conferring salt tolerance should be carried out. Transgenic banana should be monitored for the levels of salt stress (as well as drought and oxidative stress) that it can stand. The transgenic plants must also be ensured that no alteration in agronomic traits and no off-target effects observed.

10. Engineered disarmed viral vector is an alternative approach to introduce and express candidate genes in plants. This approach is faster than stable transformation, and not transmittable nor heritable. This approach can be considered as an alternative way to overexpress or silence gene of interest in banana.

CHAPTER 6: CONCLUSION

In this study, gene expression of salinity-stressed banana root has successfully been profiled and characterised using Illumina high-throughput sequencing platforms. Two cDNA libraries (300 mM NaCl treatment and a control) and three small RNA cDNA libraries (100 mM and 300 mM NaCl treatments and a control) have successfully been constructed and sequenced, producing high throughput and quality gene expression data for banana roots.

A number of 31,390 transcript sequences were produced with 2,993 or 9.5% genes differentially expressed in the 300 mM NaCl-stressed banana roots. Genes or mechanisms with roles previously known in salinity stress were also observed in the salt-stressed banana roots in this study. These functions include signaling, transcriptional regulation, DNA repair, transport, stress defence and cell wall modifications. Besides, genes that have not been reported to be responsive to salt stress or abiotic stress were also found in this study. Putative novel salt-stress responsive genes, which include those differentially expressed unannotated genes and genes coding for hypothetical proteins, were also found in this study.

A combination of mRNA-seq and sRNA-Seq data in this study, as well as publicly available genetic resources, such as the *Musa* reference genome and the Plant MicroRNA Database (PMRD), was used to investigate miRNA-mediated gene regulation in the salt-stressed banana roots. A number of 181 plant orthologous miRNAs were found in the banana roots in this study. Plant orthologous miRNAs in the banana roots were observed to have conserved targets which have been reported in other plants species, as well as new targets which have not been reported elsewhere. Besides, 56 putative *Musa*-specific miRNAs, which sequences have not been reported in other species, were found the banana roots in this study. This indicates that banana roots

deploy both *Musa*-specific miRNA-mediated gene regulation mechanism and also the mechanism conserved in plant species.

In this study, banana roots expressed a unique set of miRNAs in response to salt stress and these miRNAs regulate diverse biological processes including stress signaling, transcriptional and translational regulations, stress defence, transport, cellular homeostasis, metabolisms and other stress-related functions. In addition, a number of unannotated genes and genes coding for unknown proteins that are responsive to salt stress was predicted to be regulated by miRNAs. New miRNA-target modules related to salt stress response in banana roots were proposed in this study.

In this study, high-throughput banana transcriptome data containing mRNA and sRNA sequences, which are valuable genetic resources for gene and marker discovery, were produced and made publicly available. To our knowledge, this is the first report of high-throughput genetic resources for an abiotic stress response in banana. This is also possibly the first reported high-throughput transcriptomic study for 'Berangan' cultivar (triploid genome, AAA), an important Malaysia banana cultivar. This present study may partly help in understanding gene regulation and abiotic stress response in banana roots and the high-throughput sequencing data generated in this present study may serve as important genetic resources for salt tolerance traits used for functional genomic studies and genetic improvement in banana.

REFERENCES

Abdel-Ghany, S. E., & Pilon, M. (2008). MicroRNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in Arabidopsis. *J Biol Chem, 283*(23), 15932-15945.

Abel, S., Burstenbinder, K., & Muller, J. (2013). The emerging function of IQD proteins as scaffolds in cellular signaling and trafficking. *Plant Signal Behav, 8*(6), e24369.

Abel, S., Savchenko, T., & Levy, M. (2005). Genome-wide comparative analysis of the IQD gene families in Arabidopsis thaliana and Oryza sativa. *BMC Evol Biol, 5*, 72.

Achard, P., Herr, A., Baulcombe, D. C., & Harberd, N. P. (2004). Modulation of floral development by a gibberellin-regulated microRNA. *Development, 131*(14), 3357-3365.

Adai, A., Johnson, C., Mlotshwa, S., Archer-Evans, S., Manocha, V., Vance, V., et al. (2005). Computational prediction of miRNAs in Arabidopsis thaliana. *Genome Res, 15*(1), 78-91.

Addo-Quaye, C., Eshoo, T. W., Bartel, D. P., & Axtell, M. J. (2008). Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. *Curr Biol, 18*(10), 758-762.

Aert, R., Sagi, L., & Volckaert, G. (2004). Gene content and density in banana (Musa acuminata) as revealed by genomic sequencing of BAC clones. *Theor Appl Genet, 109*(1), 129-139.

Alam, M. M., Tanaka, T., Nakamura, H., Ichikawa, H., Kobayashi, K., Yaeno, T., et al. (2015). Overexpression of a rice heme activator protein gene (OsHAP2E) confers resistance to pathogens, salinity and drought, and increases photosynthesis and tiller number. *Plant Biotechnol J, 13*, 85-96.

Allen, E., Xie, Z., Gustafson, A. M., & Carrington, J. C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell, 121*(2), 207-221.

Alonso-Peral, M. M., Li, J., Li, Y., Allen, R. S., Schnippenkoetter, W., Ohms, S., et al. (2010). The microRNA159-regulated GAMYB-like genes inhibit growth and promote programmed cell death in Arabidopsis. *Plant Physiol, 154*(2), 757-771.

Alvarez, C., Bermudez, M. A., Romero, L. C., Gotor, C., & Garcia, I. (2012). Cysteine homeostasis plays an essential role in plant immunity. *New Phytol, 193*(1), 165-177.

Al-Whaibi, M. H. (2011). Plant heat-shock proteins: a mini review. *J King Saud Univ Sci* 23, 139-150.

Amir, R. (2008). Towards improving methionine content in plants for enhanced nutritional quality. *Func Plant Sci Biotech 2*(1), 36-46.

Anderson, J. P., Lichtenzveig, J., Gleason, C., Oliver, R. P., & Singh, K. B. (2010). The B-3 ethylene response factor MtERF1-1 mediates resistance to a subset of root pathogens in Medicago truncatula without adversely affecting symbiosis with rhizobia. *Plant Physiol, 154*(2), 861-873.

APG III. (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot J Linn Soc 161*, 105–121.

Apse, M. P., & Blumwald, E. (2002). Engineering salt tolerance in plants. *Curr Opin Biotechnol, 13*(2), 146-150.

Arab, A., Alves, M. N., Sartoratto, A., Ogasawara, D. C., & Trigo, J. R. (2012). Methyl jasmonate increases the tropane alkaloid scopolamine and reduces natural herbivory in brugmansia suaveolens: is scopolamine responsible for plant resistance? *Neotrop Entomol, 41*(1), 2-8.

Arasu, P., Wightman, B., & Ruvkun, G. (1991). Temporal regulation of lin-14 by the

antagonistic action of two other heterochronic genes, lin-4 and lin-28. *Genes Dev, 5*(10), 1825-1833.

Arazi, T., Talmor-Neiman, M., Stav, R., Riese, M., Huijser, P., & Baulcombe, D. C. (2005). Cloning and characterization of micro-RNAs from moss. *Plant J, 43*(6), 837-848.

Arzanlou, M., Abeln, E. C., Kema, G. H., Waalwijk, C., Carlier, J., Vries, I., et al. (2007). Molecular diagnostics for the sigatoka disease complex of banana. *Phytopathology, 97*(9), 1112-1118.

Ashraf, M., & Foolad, M. R. (2007). Roles of glycine betaine and proline in improving plant abiotic stress resistance. *Environ Exp Bot 59*, 206-216.

Atmodjo, M. A., Sakuragi, Y., Zhu, X., Burrell, A. J., Mohanty, S. S., Atwood, J. A., 3rd, et al. (2011). Galacturonosyltransferase (GAUT)1 and GAUT7 are the core of a plant cell wall pectin biosynthetic homogalacturonan:galacturonosyltransferase complex. *Proc Natl Acad Sci U S A, 108*(50), 20225-20230.

Audic, S., & Claverie, J. M. (1997). The significance of digital gene expression profiles. *Genome Res, 7*(10), 986-995.

Augustine, S. M., Narayan, J. A., Syamaladevi, D. P., Appunu, C., Chakravarthi, M., Ravichandran, V., et al. (2015). Erianthus arundinaceus HSP70 (EaHSP70) overexpression increases drought and salinity tolerance in sugarcane (Saccharum spp. hybrid). *Plant Sci, 232*, 23-34.

Aukerman, M. J., & Sakai, H. (2003). Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell, 15*(11), 2730-2741.

Axtell, M. J. (2008). Evolution of microRNAs and their targets: are all microRNAs biologically relevant? *Biochim Biophys Acta, 1779*(11), 725-734.

Axtell, M. J. (2013). Classification and comparison of small RNAs from plants. *Annu Rev Plant Biol, 64*, 137-159.

Axtell, M. J., & Bartel, D. P. (2005). Antiquity of microRNAs and their targets in land plants. *Plant Cell, 17*(6), 1658-1673.

Baker, C. C., Sieber, P., Wellmer, F., & Meyerowitz, E. M. (2005). The early extra petals1 mutant uncovers a role for microRNA miR164c in regulating petal number in Arabidopsis. *Curr Biol, 15*(4), 303-315.

Barakat, A., Wall, K., Leebens-Mack, J., Wang, Y. J., Carlson, J. E., & Depamphilis, C. W. (2007). Large-scale identification of microRNAs from a basal eudicot (Eschscholzia californica) and conservation in flowering plants. *Plant J, 51*(6), 991-1003.

Bari, R., Datt Pant, B., Stitt, M., & Scheible, W. R. (2006). PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. *Plant Physiol, 141*(3), 988-999.

Barrera-Figueroa, B. E., Gao, L., Diop, N. N., Wu, Z., Ehlers, J. D., Roberts, P. A., et al. (2011). Identification and comparative analysis of drought-associated microRNAs in two cowpea genotypes. *BMC Plant Biol, 11*, 127.

Bartel, B. (2005). MicroRNAs directing siRNA biogenesis. *Nat Struct Mol Biol, 12*(7), 569-571.

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell, 116*(2), 281-297.

Baulcombe, D. C. (1996). RNA as a target and an initiator of post-transcriptional gene silencing in transgenic plants. *Plant Mol Biol, 32*(1-2), 79-88.

Baykov, A. A., Tuominen, H. K., & Lahti, R. (2011). The CBS domain: a protein module with an emerging prominent role in regulation. *ACS Chem Biol, 6*(11), 1156-1163.

Benedetti, C. E., Costa, C. L., Turcinelli, S. R., & Arruda, P. (1998). Differential

expression of a novel gene in response to coronatine, methyl jasmonate, and wounding in the Coi1 mutant of Arabidopsis. *Plant Physiol, 116*(3), 1037-1042.
Benes, V., & Castoldi, M. (2010). Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available. *Methods, 50*(4), 244-249.
Bhogale, S., Mahajan, A., Natarajan, B., Rajabhoj, M., Thulasiram, H. V., & Banerjee, A. K. (2013). miRNA156- A potential graft-transmissible microRNA that modulates plant architecture and tuberization in potato (Solanum tuberosum ssp andigena). *Plant Physiol*.
Biesgen, C., & Weiler, E. W. (1999). Structure and regulation of OPR1 and OPR2, two closely related genes encoding 12-oxophytodienoic acid-10,11-reductases from Arabidopsis thaliana. *Planta, 208*(2), 155-165.
Biselli, C., Urso, S., Bernardo, L., Tondelli, A., Tacconi, G., Martino, V., et al. (2010). Identification and mapping of the leaf stripe resistance gene Rdg1a in Hordeum spontaneum. *Theor Appl Genet, 120*(6), 1207-1218.
Blilou, I., Ocampo, J. A., & Garcia-Garrido, J. M. (2000). Induction of Ltp (lipid transfer protein) and Pal (phenylalanine ammonia-lyase) gene expression in rice roots colonized by the arbuscular mycorrhizal fungus Glomus mosseae. *J Exp Bot, 51*(353), 1969-1977.
Bogamuwa, S. P., & Jang, J. C. (2014). Tandem CCCH Zinc Finger Proteins in Plant Growth, Development, and Stress Response. *Plant Cell Physiol*.
Bonnet, E., Wuyts, J., Rouze, P., & Van de Peer, Y. (2004a). Detection of 91 potential conserved plant microRNAs in Arabidopsis thaliana and Oryza sativa identifies important target genes. *Proc Natl Acad Sci U S A, 101*(31), 11511-11516.
Bonnet, E., Wuyts, J., Rouze, P., & Van de Peer, Y. (2004b). Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. *Bioinformatics, 20*(17), 2911-2917.
Borsani, O., Zhu, J., Verslues, P. E., Sunkar, R., & Zhu, J. K. (2005). Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell, 123*(7), 1279-1291.
Bottcher, C., Keyzers, R. A., Boss, P. K., & Davies, C. (2010). Sequestration of auxin by the indole-3-acetic acid-amido synthetase GH3-1 in grape berry (Vitis vinifera L.) and the proposed role of auxin conjugation during ripening. *J Exp Bot, 61*(13), 3615-3625.
Boualem, A., Laporte, P., Jovanovic, M., Laffont, C., Plet, J., Combier, J. P., et al. (2008). MicroRNA166 controls root and nodule development in Medicago truncatula. *Plant J, 54*(5), 876-887.
Bouche, N., Yellin, A., Snedden, W. A., & Fromm, H. (2005). Plant-specific calmodulin-binding proteins. *Annu Rev Plant Biol, 56*, 435-466.
Branscheid, A., Devers, E. A., May, P., & Krajinski, F. (2011). Distribution pattern of small RNA and degradome reads provides information on miRNA gene structure and regulation. *Plant Signal Behav, 6*(10).
Bruggeman, Q., Garmier, M., de Bont, L., Soubigou-Taconnat, L., Mazubert, C., Benhamed, M., et al. (2014). The Polyadenylation Factor Subunit CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR30: A Key Factor of Programmed Cell Death and a Regulator of Immunity in Arabidopsis. *Plant Physiol, 165*(2), 732-746.
Burklew, C. E., Ashlock, J., Winfrey, W. B., & Zhang, B. (2012). Effects of aluminum oxide nanoparticles on the growth, development, and microRNA expression of tobacco (Nicotiana tabacum). *PLoS One, 7*(5), e34783.
Burza, A. M., Pekala, I., Sikora, J., Siedlecki, P., Malagocki, P., Bucholc, M., et al. (2006). Nicotiana tabacum osmotic stress-activated kinase is regulated by phosphorylation on Ser-154 and Ser-158 in the kinase activation loop. *J Biol*

Chem, 281(45), 34299-34311.
Cai, X., Davis, E. J., Ballif, J., Liang, M., Bushman, E., Haroldsen, V., et al. (2006). Mutant identification and characterization of the laccase gene family in Arabidopsis. *J Exp Bot, 57*(11), 2563-2569.
Capitao, C., Paiva, J. A., Santos, D. M., & Fevereiro, P. (2011). In Medicago truncatula, water deficit modulates the transcript accumulation of components of small RNA pathways. *BMC Plant Biol, 11*, 79.
Carillo, P., Annunziata M. G., Pontecorvo, G., Fuggi, A., & Woodrow, P. (2011). Salinity stress and salt tolerance. In A. Shanker (Ed.), *Abiotic Stress in Plants - Mechanisms and Adaptations* (pp. 21-38). Rijeka & Shanghai: InTech.
Carlsbecker, A., Lee, J. Y., Roberts, C. J., Dettmer, J., Lehesranta, S., Zhou, J., et al. (2010). Cell signalling by microRNA165/6 directs gene dose-dependent root cell fate. *Nature, 465*(7296), 316-321.
Carnavale Bottino, M., Rosario, S., Grativol, C., Thiebaut, F., Rojas, C. A., Farrineli, L., et al. (2013). High-throughput sequencing of small RNA transcriptome reveals salt stress regulated microRNAs in sugarcane. *PLoS One, 8*(3), e59423.
Casati, P. (2013). Analysis of UV-B regulated miRNAs and their targets in maize leaves. *Plant Signal Behav, 8*(10), doi: 10 4161/psb 26758.
Cellier, G., Remenant, B., Chiroleu, F., Lefeuvre, P., & Prior, P. (2012). Phylogeny and population structure of brown rot- and Moko disease-causing strains of Ralstonia solanacearum phylotype II. *Appl Environ Microbiol, 78*(7), 2367-2375.
Chatre, L., Brandizzi, F., Hocquellet, A., Hawes, C., & Moreau, P. (2005). Sec22 and Memb11 are v-SNAREs of the anterograde endoplasmic reticulum-Golgi pathway in tobacco leaf epidermal cells. *Plant Physiol, 139*(3), 1244-1254.
Chen, C. C., Liang, C. S., Kao, A. L., & Yang, C. C. (2010a). HHP1, a novel signalling component in the cross-talk between the cold and osmotic signalling pathways in Arabidopsis. *J Exp Bot, 61*(12), 3305-3320.
Chen, C., Ridzon, D. A., Broomer, A. J., Zhou, Z., Lee, D. H., Nguyen, J. T., et al. (2005a). Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res, 33*(20), e179.
Chen, C., Tan, R., Wong, L., Fekete, R., & Halsey, J. (2011a). Quantitation of microRNAs by real-time RT-qPCR. *Methods Mol Biol, 687*, 113-134.
Chen, L., Ren, Y., Zhang, Y., Xu, J., Zhang, Z., & Wang, Y. (2011b). Genome-wide profiling of novel and conserved Populus microRNAs involved in pathogen stress response by deep sequencing. *Planta, 235*(5), 873-883.
Chen, L., Wang, T., Zhao, M., & Zhang, W. (2012a). Ethylene-responsive miRNAs in roots of Medicago truncatula identified by high-throughput sequencing at whole genome level. *Plant Sci, 184*, 14-19.
Chen, L., Zhang, Y., Ren, Y., Xu, J., Zhang, Z., & Wang, Y. (2012b). Genome-wide identification of cold-responsive and new microRNAs in Populus tomentosa by high-throughput sequencing. *Biochem Biophys Res Commun, 417*(2), 892-896.
Chen, M., Meng, Y., Mao, C., Chen, D., & Wu, P. (2010b). Methodological framework for functional characterization of plant microRNAs. *J Exp Bot, 61*(9), 2271-2280.
Chen, X. (2005b). MicroRNA biogenesis and function in plants. *FEBS Lett, 579*(26), 5923-5931.
Chen, Y., Kelly, E. E., Masluk, R. P., Nelson, C. L., Cantu, D. C., & Reilly, P. J. (2011c). Structural classification and properties of ketoacyl synthases. *Protein Sci, 20*(10), 1659-1667.
Cheung, F., & Town, C. D. (2007). A BAC end view of the Musa acuminata genome. *BMC Plant Biol, 7*, 29.
Chiou, T. J. (2007). The role of microRNAs in sensing nutrient stress. *Plant Cell Environ, 30*(3), 323-332.

Chiou, T. J., Aung, K., Lin, S. I., Wu, C. C., Chiang, S. F., & Su, C. L. (2006). Regulation of phosphate homeostasis by MicroRNA in Arabidopsis. *Plant Cell, 18*(2), 412-421.

Chopra, S., Palencia, A., Virus, C., Tripathy, A., Temple, B. R., Velazquez-Campoy, A., et al. (2013). Plant tumour biocontrol agent employs a tRNA-dependent mechanism to inhibit leucyl-tRNA synthetase. *Nat Commun, 4*, 1417.

Christelova, P., Valarik, M., Hribova, E., De Langhe, E., & Dolezel, J. (2011). A multi gene sequence-based phylogeny of the Musaceae (banana) family. *BMC Evol Biol, 11*, 103.

Chuck, G., Meeley, R., Irish, E., Sakai, H., & Hake, S. (2007). The maize tasselseed4 microRNA controls sex determination and meristem cell fate by targeting Tasselseed6/indeterminate spikelet1. *Nat Genet, 39*(12), 1517-1521.

Chung, H. J., & Ferl, R. J. (1999). Arabidopsis alcohol dehydrogenase expression in both shoots and roots is conditioned by root growth environment. *Plant Physiol, 121*(2), 429-436.

Churchill, A. C. (2011). Mycosphaerella fijiensis, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development, and the challenges of control. *Mol Plant Pathol, 12*(4), 307-328.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogee, J., Allard, V., et al. (2005). Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature, 437*(7058), 529-533.

Clendennen, S. K., & May, G. D. (1997). Differential gene expression in ripening banana fruit. *Plant Physiol, 115*(2), 463-469.

Cloonan, N., Forrest, A. R., Kolle, G., Gardiner, B. B., Faulkner, G. J., Brown, M. K., et al. (2008). Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat Methods, 5*(7), 613-619.

Coemans, B., Matsumura, H., Terauchi, R., Remy, S., Swennen, R., & Sagi, L. (2005). SuperSAGE combined with PCR walking allows global gene expression profiling of banana (Musa acuminata), a non-model organism. *Theor Appl Genet, 111*(6), 1118-1126.

Cogoni, C., Irelan, J. T., Schumacher, M., Schmidhauser, T. J., Selker, E. U., & Macino, G. (1996). Transgene silencing of the al-1 gene in vegetative cells of Neurospora is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. *EMBO J, 15*(12), 3153-3163.

Collins, T., Stone, J. R., & Williams, A. J. (2001). All in the family: the BTB/POZ, KRAB, and SCAN domains. *Mol Cell Biol, 21*(11), 3609-3615.

Combier, J. P., Frugier, F., de Billy, F., Boualem, A., El-Yahyaoui, F., Moreau, S., et al. (2006). MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in Medicago truncatula. *Genes Dev, 20*(22), 3084-3088.

Conesa, A., & Gotz, S. (2008). Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics, 2008*, 619832.

Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., & Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics, 21*(18), 3674-3676.

Costa, V., Angelini, C., De Feis, I., & Ciccodicola, A. (2010). Uncovering the complexity of transcriptomes with RNA-Seq. *J Biomed Biotechnol, 2010*, 853916.

Cuperus, J. T., Fahlgren, N., & Carrington, J. C. (2011). Evolution and functional diversification of MIRNA genes. *Plant Cell, 23*(2), 431-442.

Curaba, J., Talbot, M., Li, Z., & Helliwell, C. (2013). Over-expression of microRNA171 affects phase transitions and floral meristem determinancy in barley. *BMC Plant*

Biol, 13, 6.
Dai, X., & Zhao, P. X. (2011). psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res, 39*(Web Server issue), W155-159.
Dal Santo, S., Fasoli, M., Cavallini, E., Tornielli, G. B., Pezzotti, M., & Zenoni, S. (2011). PhEXPA1, a Petunia hybrida expansin, is involved in cell wall metabolism and in plant architecture specification. *Plant Signal Behav, 6*(12), 2031-2034.
Daniells J., Jenny, C., Karamura, D. & Tomekpe, K. (2001a). *Musalogue: a catalogue of Musa germplasm. Diversity in the genus Musa* (E. Arnaud and S. Sharrock, compil.). Montpellier, France: International Network for the Improvement of Banana and Plantain.
Daniells, J. W., Geering, A. D. W., Bryde, N. J., & Thomas, J. E. (2001b). The effect of Banana streak virus on the growth and yield of dessert bananas in tropical Australia. *Ann of Appl Biol, 139*, 51-60.
Dasgan, H. Y., Aktas, H., Abak, K., & Cakmak, I. (2002). Determination of screening techniques to salinity tolerance in tomatoes and investigation of genotype responses. *Plant Sci, 163*, 695-703.
Davey, M. W., Graham, N. S., Vanholme, B., Swennen, R., May, S. T., & Keulemans, J. (2009). Heterologous oligonucleotide microarrays for transcriptomics in a non-model species; a proof-of-concept study of drought stress in Musa. *BMC Genomics, 10*, 436.
Davey, M. W., Gudimella, R., Harikrishna, J. A., Sin, L. W., Khalid, N., & Keulemans, J. (2013). "A draft Musa balbisiana genome sequence for molecular genetics in polyploid, inter- and intra-specific Musa hybrids". *BMC Genomics, 14*, 683.
Davletova, S., Rizhsky, L., Liang, H., Shengqiang, Z., Oliver, D. J., Coutu, J., et al. (2005). Cytosolic ascorbate peroxidase 1 is a central component of the reactive oxygen gene network of Arabidopsis. *Plant Cell, 17*(1), 268-281.
Davy, A., Thomsen, K. K., Juliano, M. A., Alves, L. C., Svendsen, I., & Simpson, D. J. (2000). Purification and characterization of barley dipeptidyl peptidase IV. *Plant Physiol, 122*(2), 425-432.
De Lapeyre de Bellaire, L., Foure´, E., Abadie, C. & Carlier, J. (2010). Black leaf streak disease is challenging the banana industry. *Fruits 65*, 327–342.
De Luis, A., Markmann, K., Cognat, V., Holt, D. B., Charpentier, M., Parniske, M., et al. (2012). Two microRNAs linked to nodule infection and nitrogen-fixing ability in the legume Lotus japonicus. *Plant Physiol, 160*(4), 2137-2154.
Debouba, M., Gouia, H., Valadier, M. H., Ghorbel, M. H., & Suzuki, A. (2006). Salinity-induced tissue-specific diurnal changes in nitrogen assimilatory enzymes in tomato seedlings grown under high or low nitrate medium. *Plant Physiol Biochem, 44*(5-6), 409-419.
Deinlein, U., Stephan, A. B., Horie, T., Luo, W., Xu, G., & Schroeder, J. I. (2014). Plant salt-tolerance mechanisms. *Trends Plant Sci, 19*(6), 371-379.
Denance, N., Szurek, B., & Noel, L. D. (2014). Emerging functions of nodulin-like proteins in non-nodulating plant species. *Plant Cell Physiol, 55*(3), 469-474.
Deng, X., Phillips, J., Meijer, A. H., Salamini, F., & Bartels, D. (2002). Characterization of five novel dehydration-responsive homeodomain leucine zipper genes from the resurrection plant Craterostigma plantagineum. *Plant Mol Biol, 49*(6), 601-610.
Deng, Y., Wang, W., Li, W. Q., Xia, C., Liao, H. Z., Zhang, X. Q., et al. (2010). MALE GAMETOPHYTE DEFECTIVE 2, encoding a sialyltransferase-like protein, is required for normal pollen germination and pollen tube growth in Arabidopsis. *J Integr Plant Biol, 52*(9), 829-843.
D'Hont, A., Denoeud, F., Aury, J. M., Baurens, F. C., Carreel, F., Garsmeur, O., et al.

(2012). The banana (Musa acuminata) genome and the evolution of monocotyledonous plants. *Nature*, *488*(7410), 213-217.

Diepold, A., Li, G., Lennarz, W. J., Nurnberger, T., & Brunner, F. (2007). The Arabidopsis AtPNG1 gene encodes a peptide: N-glycanase. *Plant J, 52*(1), 94-104.

Dilokpimol, A., Poulsen, C. P., Vereb, G., Kaneko, S., Schulz, A., & Geshi, N. (2014). Galactosyltransferases from Arabidopsis thaliana in the biosynthesis of type II arabinogalactan: molecular interaction enhances enzyme activity. *BMC Plant Biol, 14*, 90.

Ding, D., Li, W., Han, M., Wang, Y., Fu, Z., Wang, B., et al. (2013). Identification and characterisation of maize microRNAs involved in developing ears. *Plant Biol (Stuttg)*.

Ding, D., Wang, Y., Han, M., Fu, Z., Li, W., Liu, Z., et al. (2012a). MicroRNA transcriptomic analysis of heterosis during maize seed germination. *PLoS One, 7*(6), e39578.

Ding, D., Zhang, L., Wang, H., Liu, Z., Zhang, Z., & Zheng, Y. (2009). Differential expression of miRNAs in response to salt stress in maize roots. *Ann Bot, 103*(1), 29-38.

Ding, J., Zhou, S., & Guan, J. (2012b). Finding microRNA targets in plants: current status and perspectives. *Genomics Proteomics Bioinformatics, 10*(5), 264-275.

Ding, Y. F., & Zhu, C. (2009). The role of microRNAs in copper and cadmium homeostasis. *Biochem Biophys Res Commun, 386*(1), 6-10.

Ding, Y., Chen, Z., & Zhu, C. (2011). Microarray-based analysis of cadmium-responsive microRNAs in rice (Oryza sativa). *J Exp Bot, 62*(10), 3563-3573.

Dizengremel, P., Le Thiec, D., Hasenfratz-Sauder, M. P., Vaultier, M. N., Bagard, M., & Jolivet, Y. (2009). Metabolic-dependent changes in plant cell redox power after ozone exposure. *Plant Biol (Stuttg), 11 Suppl 1*, 35-42.

Dong, W., Wang, M., Xu, F., Quan, T., Peng, K., Xiao, L., et al. (2013). Wheat oxophytodienoate reductase gene TaOPR1 confers salinity tolerance via enhancement of abscisic acid signaling and reactive oxygen species scavenging. *Plant Physiol, 161*(3), 1217-1228.

Duan, L., Liu, H., Li, X., Xiao, J., & Wang, S. (2014). Multiple phytohormones and phytoalexins are involved in disease resistance to Magnaporthe oryzae invaded from roots in rice. *Physiol Plant*.

Dugas, D. V., & Bartel, B. (2004). MicroRNA regulation of gene expression in plants. *Curr Opin Plant Biol, 7*(5), 512-520.

Eamens, A., Wang, M. B., Smith, N. A., & Waterhouse, P. M. (2008). RNA silencing in plants: yesterday, today, and tomorrow. *Plant Physiol, 147*(2), 456-468.

Edvardsson, A., Eshaghi, S., Vener, A. V., & Andersson, B. (2003). The major peptidyl-prolyl isomerase activity in thylakoid lumen of plant chloroplasts belongs to a novel cyclophilin TLP20. *FEBS Lett, 542*(1-3), 137-141.

Egan, A. N., Schlueter, J., & Spooner, D. M. (2012). Applications of next-generation sequencing in plant biology. *Am J Bot, 99*(2), 175-185.

Elbashir, S. M., Lendeckel, W., & Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev, 15*(2), 188-200.

Eldem, V., Okay, S., & Ünver, T. (2013). Plant microRNAs: new players in functional genomics. *Turk J Agric For, 37*, 1-21.

El-Kasmi, F., Pacher, T., Strompen, G., Stierhof, Y. D., Muller, L. M., Koncz, C., et al. (2011). Arabidopsis SNARE protein SEC22 is essential for gametophyte development and maintenance of Golgi-stack integrity. *Plant J, 66*(2), 268-279.

Elyse Ireland, H., Harding, S. J., Bonwick, G. A., Jones, M., Smith, C. J., & Williams, J. H. (2004). Evaluation of heat shock protein 70 as a biomarker of environmental

stress in Fucus serratus and Lemna minor. *Biomarkers, 9*(2), 139-155.
English, J. J., Mueller, E., & Baulcombe, D. C. (1996). Suppression of Virus Accumulation in Transgenic Plants Exhibiting Silencing of Nuclear Genes. *Plant Cell, 8*(2), 179-188.
Estrada-Melo, A. C., Ma, C., Reid, M. S., & Jiang, C-Z (2015). Overexpression of an ABA biosynthesis gene using a stress-inducible promoter enhances drought resistance in petunia. *Hortic Res, 2*, 15013.
Fahlgren, N., Howell, M. D., Kasschau, K. D., Chapman, E. J., Sullivan, C. M., Cumbie, J. S., et al. (2007). High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes. *PLoS One, 2*(2), e219.
FAO. (2003). *The world banana economy 1985-2002*. Rome: Food and Agriculture Organization of the United Nations.
FAOSTAT. (2011). *FAOSTAT.* Accessed from http://faostat.fao.org/
Faure, S., Noyer, J. L., Carreel, F., Horry, J. P., Bakry, F., & Lanaud, C. (1994). Maternal inheritance of chloroplast genome and paternal inheritance of mitochondrial genome in bananas (Musa acuminata). *Curr Genet, 25*(3), 265-269.
Feng, H., Duan, X., Zhang, Q., Li, X., Wang, B., Huang, L., et al. (2013). The target gene of tae-miR164, a novel NAC transcription factor from the NAM subfamily, negatively regulates resistance of wheat to stripe rust. *Mol Plant Pathol*.
Feng, J., Liu, X., Lai, L., & Chen, J. (2011). Spatio-temporal expression of miRNAs in tomato tissues upon Cucumber mosaic virus and Tomato aspermy virus infections. *Acta Biochim Biophys Sin (Shanghai), 43*(4), 258-266.
Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature, 391*(6669), 806-811.
Fischer, R .L., & Bennett, A. B. (1991). Role of cell wall hydrolases in fruit ripening. *Annu Rev Plant Physiol Mol Biol 42*, 675-703.
Franchini, P., van der Merwe, M., & Roodt-Wilding, R. (2011). Transcriptome characterization of the South African abalone Haliotis midae using sequencing-by-synthesis. *BMC Res Notes, 4*, 59.
Frazier, T. P., Sun, G., Burklew, C. E., & Zhang, B. (2011). Salt and drought stresses induce the aberrant expression of microRNA genes in tobacco. *Mol Biotechnol, 49*(2), 159-165.
Friedlander, M. R., Mackowiak, S. D., Li, N., Chen, W., & Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res, 40*(1), 37-52.
Frison, E., & Sharrock, S. (1998). Introduction: The economic, social and nutritional importance of banana in the world. In C. Picq, E. Fouré, & E. A. Frison (Eds.), *Bananas and food security: Proceedings of an international symposium held in Douala, Cameroon, November 10-14, 1998* (pp. 21-35). Montpellier: International Network for the Improvement of Banana and Plantain.
Fu, C., Sunkar, R., Zhou, C., Shen, H., Zhang, J. Y., Matts, J., et al. (2012). Overexpression of miR156 in switchgrass (Panicum virgatum L.) results in various morphological alterations and leads to improved biomass production. *Plant Biotechnol J, 10*(4), 443-452.
Fu, H., Tie, Y., Xu, C., Zhang, Z., Zhu, J., Shi, Y., et al. (2005). Identification of human fetal liver miRNAs by a novel method. *FEBS Lett, 579*(17), 3849-3854.
Fujii, H., Chiou, T. J., Lin, S. I., Aung, K., & Zhu, J. K. (2005). A miRNA involved in phosphate-starvation response in Arabidopsis. *Curr Biol, 15*(22), 2038-2043.
Furumoto, T., Ogawa, N., Hata, S., & Izui, K. (1996). Plant calcium-dependent protein kinase-related kinases (CRKs) do not require calcium for their activities. *FEBS*

Lett, 396(2-3), 147-151.
Gal, T. Z., Aussenberg, E. R., Burdman, S., Kapulnik, Y., & Koltai, H. (2006). Expression of a plant expansin is involved in the establishment of root knot nematode parasitism in tomato. *Planta, 224*(1), 155-162.
Gandikota, M., Birkenbihl, R. P., Hohmann, S., Cardon, G. H., Saedler, H., & Huijser, P. (2007). The miRNA156/157 recognition element in the 3' UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. *Plant J, 49*(4), 683-693.
Gao, H. B., Chu, Y. J., & Xue, H. W. (2013). Phosphatidic acid (PA) binds PP2AA1 to regulate PP2A activity and PIN1 polar localization. *Mol Plant, 6*(5), 1692-1702.
Gao, P., Bai, X., Yang, L., Lv, D., Li, Y., Cai, H., et al. (2010). Over-expression of osa-MIR396c decreases salt and alkali stress tolerance. *Planta, 231*(5), 991-1001.
Gao, P., Bai, X., Yang, L., Lv, D., Pan, X., Li, Y., et al. (2011). osa-MIR393: a salinity- and alkaline stress-related microRNA gene. *Mol Biol Rep, 38*(1), 237-242.
Gardner, P. P., Daub, J., Tate, J., Moore, B. L., Osuch, I. H., Griffiths-Jones, S., et al. (2011). Rfam: Wikipedia, clans and the "decimal" release. *Nucleic Acids Res, 39*(Database issue), D141-145.
Garmire, L. X., & Subramaniam, S. (2012). Evaluation of normalization methods in mammalian microRNA-Seq data. *RNA, 18*(6), 1279-1288.
Gauggel, C. A., Sierra, F., & Arévalo, G. (2005). The problems of banana root deterioration and its impact on production: Latin America's experience. In D. W. Turner & F. E. Rosales (Eds.), *Banana root system: towards a better understanding for its productive management: Proceedings of an international symposium held in San José, Costa Rica, November 3-5, 2003* (pp. 13-22). Montpellier: International Network for the Improvement of Banana and Plantain.
Gayral, P., Blondin, L., Guidolin, O., Carreel, F., Hippolyte, I., Perrier, X., et al. (2010). Evolution of endogenous sequences of banana streak virus: what can we learn from banana (Musa sp.) evolution? *J Virol, 84*(14), 7346-7359.
Ge, Y., Li, Y., Zhu, Y. M., Bai, X., Lv, D. K., Guo, D., et al. (2010). Global transcriptome profiling of wild soybean (Glycine soja) roots under NaHCO3 treatment. *BMC Plant Biol, 10*, 153.
Gebelin, V., Argout, X., Engchuan, W., Pitollat, B., Duan, C., Montoro, P., et al. (2012). Identification of novel microRNAs in Hevea brasiliensis and computational prediction of their targets. *BMC Plant Biol, 12*, 18.
Gebelin, V., Leclercq, J., Hu, S., Tang, C., & Montoro, P. (2013). Regulation of MIR genes in response to abiotic stress in Hevea brasiliensis. *Int J Mol Sci, 14*(10), 19587-19604.
German, M. A., Luo, S., Schroth, G., Meyers, B. C., & Green, P. J. (2009). Construction of Parallel Analysis of RNA Ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. *Nat Protoc, 4*(3), 356-362.
German, M. A., Pillay, M., Jeong, D. H., Hetawal, A., Luo, S., Janardhanan, P., et al. (2008). Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat Biotechnol, 26*(8), 941-946.
Gifford, M. L., Dean, A., Gutierrez, R. A., Coruzzi, G. M., & Birnbaum, K. D. (2008). Cell-specific nitrogen responses mediate developmental plasticity. *Proc Natl Acad Sci U S A, 105*(2), 803-808.
Gleave, A. P., Ampomah-Dwamena, C., Berthold, S., Dejnoprat, S., Karunairetnam, S., Nain, B., Wang, Y. Y., Crowhurst, R. N., & MacDiarmid, R. M. (2008). Identification and characterisation of primary microRNAs from apple (Malus domestica cv. Royal Gala) expressed sequence tags. *Tree Genetics & Genomes 4*(2), 343-358
Golovkin, M., & Reddy, A. S. (2003). A calmodulin-binding protein from Arabidopsis

has an essential role in pollen germination. *Proc Natl Acad Sci U S A, 100*(18), 10558-10563.

Gotz, K. P., Staroske, N., Radchuk, R., Emery, R. J., Wutzke, K. D., Herzog, H., et al. (2007). Uptake and allocation of carbon and nitrogen in Vicia narbonensis plants with increased seed sink strength achieved by seed-specific expression of an amino acid permease. *J Exp Bot, 58*(12), 3183-3195.

Gou, J. Y., Felippes, F. F., Liu, C. J., Weigel, D., & Wang, J. W. (2011). Negative regulation of anthocyanin biosynthesis in Arabidopsis by a miR156-targeted SPL transcription factor. *Plant Cell, 23*(4), 1512-1522.

Gowen, S. R. (1995). Pest. In S. Gowen (Ed.), *Bananas and plantains* (pp. 382-402). London: Chapman & Hall.

Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A., & Enright, A. J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res, 34*(Database issue), D140-144.

Griffiths-Jones, S., Moxon, S., Marshall, M., Khanna, A., Eddy, S. R., & Bateman, A. (2005). Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res, 33*(Database issue), D121-124.

Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., et al. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell, 106*(1), 23-34.

Guan, Q., Lu, X., Zeng, H., Zhang, Y., & Zhu, J. (2013). Heat stress induction of miR398 triggers a regulatory loop that is critical for thermotolerance in Arabidopsis. *Plant J, 74*(5), 840-851.

Guo, H. S., Xie, Q., Fei, J. F., & Chua, N. H. (2005). MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for arabidopsis lateral root development. *Plant Cell, 17*(5), 1376-1386.

Guo, N., Ye, W. W., Wu, X. L., Shen, D. Y., Wang, Y. C., Xing, H., et al. (2011). Microarray profiling reveals microRNAs involving soybean resistance to Phytophthora sojae. *Genome, 54*(11), 954-958.

Guo, Q., Xiang, A., Yang, Q., & Yang, Z. (2007). Bioinformatic identification of microRNAs and their target genes from Solanum tuberosum expressed sequence tags. *Chinese Science Bulletin, 52*(17), 2380-2389.

Guo, W., Zuo, Z., Cheng, X., Sun, J., Li, H., Li, L., et al. (2014). The chloride channel family gene CLCd negatively regulates pathogen-associated molecular pattern (PAMP)-triggered immunity in Arabidopsis. *J Exp Bot, 65*(4), 1205-1215.

Gutierrez, L., Bussell, J. D., Pacurar, D. I., Schwambach, J., Pacurar, M., & Bellini, C. (2009). Phenotypic plasticity of adventitious rooting in Arabidopsis is controlled by complex regulation of AUXIN RESPONSE FACTOR transcripts and microRNA abundance. *Plant Cell, 21*(10), 3119-3132.

Häkkinen, M. (2013). Reappraisal of sectional taxonomy in *Musa* (*Musaceae*). *Taxon 62*(4), 809-813.

Häkkinen, M., & Sharrock, S. (2002). Diversity in the genus Musa – Focus on Rhodochlamys. In Author, *INIBAP annual report 2001* (pp. 16-23). Montpellier: INIBAP.

Han, Y., Zhu, B., Luan, F., Zhu, H., Shao, Y., Chen, A., Lu, & C., Luo, Y. (2010). Conserved miRNAs and their targets identified in lettuce (Lactuca) by EST analysis. *Gene, 463*(1-2), 1-7.

Harbers, M., & Carninci, P. (2005). Tag-based approaches for transcriptome research and genome annotation. *Nat Methods, 2*(7), 495-502.

Harper, G., Osuji, J. O., Heslop-Harrison, J. S., & Hull, R. (1999). Integration of banana streak badnavirus into the Musa genome: molecular and cytogenetic evidence.

Virology, 255(2), 207-213.
Harris, M. A., Clark, J., Ireland, A., Lomax, J., Ashburner, M., Foulger, R., et al. (2004). The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res, 32*(Database issue), D258-261.
Harrison, J., Moore, K. A., Paszkiewicz, K., Jones, T., Grant, M .R., Ambacheew, D., et al. (2014). A draft genome sequence for Ensete ventricosum, the drought-tolerant "tree against hunger". *Agronomy, 4*, 13-33.
Hashimoto, T., Nakajima, K., Ongena, G., & Yamada, Y. (1992). Two Tropinone Reductases with Distinct Stereospecificities from Cultured Roots of Hyoscyamus niger. *Plant Physiol, 100*(2), 836-845.
Hasplova, K., Hudecova, A., Magdolenova, Z., Bjoras, M., Galova, E., Miadokova, E., et al. (2012). DNA alkylation lesions and their repair in human cells: modification of the comet assay with 3-methyladenine DNA glycosylase (AlkD). *Toxicol Lett, 208*(1), 76-81.
Hayes, M. A., Feechan, A., & Dry, I. B. (2010). Involvement of abscisic acid in the coordinated regulation of a stress-inducible hexose transporter (VvHT5) and a cell wall invertase in grapevine in response to biotrophic fungal infection. *Plant Physiol, 153*(1), 211-221.
Heslop-Harrison, J. S., & Schwarzacher, T. (2007). Domestication, genomics and the future for banana. *Ann Bot, 100*(5), 1073-1084.
Hirayama, T., & Shinozaki, K. (2010). Research on plant abiotic stress responses in the post-genome era: past, present and future. *Plant J, 61*(6), 1041-1052.
Hirimburegama, W. K., Dias, W. K. G., & Hirimburegama, K. (2004). Banana improvement through gamma irradiation and testing for banana bract mosaic virus in Sri Lanka. In S. M. Jain & R. Swennen (Eds.), *Banana improvement: Cellular, molecular biology, and induced mutations.* Plymouth: Science Publishers, Inc.
Ho, T., Pallett, D., Rusholme, R., Dalmay, T., & Wang, H. (2006). A simplified method for cloning of short interfering RNAs from Brassica juncea infected with Turnip mosaic potyvirus and Turnip crinkle carmovirus. *J Virol Methods, 136*(1-2), 217-223.
Hofacker, I. L. (2003). Vienna RNA secondary structure server. *Nucleic Acids Res, 31*(13), 3429-3431.
Hollenbach, B., Schreiber, L., Hartung, W., & Dietz, K. J. (1997). Cadmium leads to stimulated expression of the lipid transfer protein genes in barley: implications for the involvement of lipid transfer proteins in wax assembly. *Planta, 203*(1), 9-19.
Holtgrefe, S., Gohlke, J., Starmann, J., Druce, S., Klocke, S., Altmann, B., et al. (2008). Regulation of plant cytosolic glyceraldehyde 3-phosphate dehydrogenase isoforms by thiol modifications. *Physiol Plant, 133*(2), 211-228.
Hooks, C. R. R., Wright, M. G., Kabasawa, D. S., Manandhar, R., & Almeida, R. P. P. (2008). Effect of banana bunchy top virus infection on morphology and growth characteristics of banana. *Ann of Appl Biol, 153*, 1-9.
Hoshi, M., Ohki, Y., Ito, K., Tomita, T., Iwatsubo, T., Ishimaru, Y., et al. (2013). Experimental detection of proteolytic activity in a signal peptide peptidase of Arabidopsis thaliana. *BMC Biochem, 14*, 16.
Hosmani, P. S., Kamiya, T., Danku, J., Naseer, S., Geldner, N., Guerinot, M. L., et al. (2013). Dirigent domain-containing protein is part of the machinery required for formation of the lignin-based Casparian strip in the root. *Proc Natl Acad Sci U S A, 110*(35), 14498-14503.
Hribova, E., Neumann, P., Matsumoto, T., Roux, N., Macas, J., & Dolezel, J. (2010). Repetitive part of the banana (Musa acuminata) genome investigated by low-

depth 454 sequencing. *BMC Plant Biol, 10*, 204.
Hu, J., Baker, A., Bartel, B., Linka, N., Mullen, R. T., Reumann, S., et al. (2012). Plant peroxisomes: biogenesis and function. *Plant Cell, 24*(6), 2279-2303.
Hu, T., Hu, L., Zhang, X., Zhang, P., Zhao, Z., & Fu, J. (2013). Differential responses of CO2 assimilation, carbohydrate allocation and gene expression to NaCl stress in perennial ryegrass with different salt tolerance. *PLoS One, 8*(6), e66090.
Huang, H. H., Xu, L. L., Tong, Z. K., Lin, E. P., Liu, Q. P., Cheng, L. J., et al. (2012). De novo characterization of the Chinese fir (Cunninghamia lanceolata) transcriptome and analysis of candidate genes involved in cellulose and lignin biosynthesis. *BMC Genomics, 13*, 648.
Huang, S. Q., Peng, J., Qiu, C. X., & Yang, Z. M. (2009). Heavy metal-regulated new microRNAs from rice. *J Inorg Biochem, 103*(2), 282-287.
Huang, S. Q., Xiang, A. L., Che, L. L., Chen, S., Li, H., Song, J. B., et al. (2010). A set of miRNAs from Brassica napus in response to sulphate deficiency and cadmium stress. *Plant Biotechnol J, 8*(8), 887-899.
INIBAP. (2006). *INIBAP annual report 2005*. Montpellier: International Network for the Improvement of Banana and Plantain.
IPGRI–INIBAP/CIRAD. (1996). *Descriptors for banana (Musa spp.).* Rome: IPGRI; Montpellier: INIBAP; Montpellier: CIRAD.
Israeli, Y., Lahav, E., & Nameri, N. (1986). The effect of salinity and sodium adsorption ratio in the irrigation water, on growth and productivity of bananas under irrigation conditions. *Fruits, 41*, 297-302.
Iyer, N. J., Jia, X., Sunkar, R., Tang, G., & Mahalingam, R. (2012). microRNAs responsive to ozone-induced oxidative stress in Arabidopsis thaliana. *Plant Signal Behav, 7*(4), 484-491.
Jagadeeswaran, G., Saini, A., & Sunkar, R. (2009). Biotic and abiotic stress down-regulate miR398 expression in Arabidopsis. *Planta, 229*(4), 1009-1014.
Jaillais, Y., Fobis-Loisy, I., Miege, C., Rollin, C., & Gaude, T. (2006). AtSNX1 defines an endosome for auxin-carrier trafficking in Arabidopsis. *Nature, 443*(7107), 106-109.
Jain, S., & Choudhary, D. K. (2014). Induced defense-related proteins in soybean (Glycine max L. Merrill) plants by Carnobacterium sp. SJ-5 upon challenge inoculation of Fusarium oxysporum. *Planta, 239*(5), 1027-1040.
Jamil, A., Riaz, S., Ashraf, M., & Foolad, M. R. (2011). Gene expression profling of plants under salt stress. *Crit Rev in Plant Sci, 30*(5), 435-458.
Janz, D., Behnke, K., Schnitzler, J. P., Kanawati, B., Schmitt-Kopplin, P., & Polle, A. (2010). Pathway analysis of the transcriptome and metabolome of salt sensitive and tolerant poplar species reveals evolutionary adaption of stress tolerance mechanisms. *BMC Plant Biol, 10*, 150.
Jeger, M. J., Eden-Green, S., Thresh, J. M., Johanson, A., Waller, J. M., & Brown, A. E. (1995). Banana diseases. In S. Gowen (Ed.), *Bananas and plantains* (pp. 317-381). London: Chapman & Hall.
Jeong, D. H., Park, S., Zhai, J., Gurazada, S. G., De Paoli, E., Meyers, B. C., et al. (2011). Massive analysis of rice small RNAs: mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage. *Plant Cell, 23*(12), 4185-4207.
Ji, L., Liu, X., Yan, J., Wang, W., Yumul, R. E., Kim, Y. J., et al. (2011). ARGONAUTE10 and ARGONAUTE1 regulate the termination of floral stem cells through two microRNAs in Arabidopsis. *PLoS Genet, 7*(3), e1001358.
Jia, H., Ren, H., Gu, M., Zhao, J., Sun, S., Zhang, X., et al. (2011). The phosphate transporter gene OsPht1;8 is involved in phosphate homeostasis in rice. *Plant Physiol, 156*(3), 1164-1175.

Jia, X., Ren, L., Chen, Q. J., Li, R., & Tang, G. (2009a). UV-B-responsive microRNAs in Populus tremula. *J Plant Physiol, 166*(18), 2046-2057.

Jia, X., Wang, W. X., Ren, L., Chen, Q. J., Mendu, V., Willcut, B., et al. (2009b). Differential and dynamic regulation of miR398 in response to ABA and salt stress in Populus tremula and Arabidopsis thaliana. *Plant Mol Biol, 71*(1-2), 51-59.

Jian, X., Zhang, L., Li, G., Wang, X., Cao, X., Fang, X., et al. (2010). Identification of novel stress-regulated microRNAs from Oryza sativa L. *Genomics, 95*(1), 47-55.

Jiang, S., Lu, Y., Li, K., Lin, L., Zheng, H., Yan, F., et al. (2014). Heat shock protein 70 is necessary for Rice stripe virus infection in plants. *Mol Plant Pathol*.

Jiang, Y., & Deyholos, M. K. (2006). Comprehensive transcriptional profiling of NaCl-stressed Arabidopsis roots reveals novel classes of responsive genes. *BMC Plant Biol, 6*, 25.

Jin, W., Li, N., Zhang, B., Wu, F., Li, W., Guo, A., et al. (2008). Identification and verification of microRNA in wheat (Triticum aestivum). *J Plant Res, 121*(3), 351-355.

Jin-Long, G., Li-Ping, X., Jing-Ping, F., Ya-Chun, S., Hua-Ying, F., You-Xiong, Q., et al. (2012). A novel dirigent protein gene with highly stem-specific expression from sugarcane, response to drought, salt and oxidative stresses. *Plant Cell Rep, 31*(10), 1801-1812.

Johnson, K. L., Jones, B. J., Bacic, A., & Schultz, C. J. (2003). The fasciclin-like arabinogalactan proteins of Arabidopsis. A multigene family of putative cell adhesion molecules. *Plant Physiol, 133*(4), 1911-1925.

Johnson, M. T., Carpenter, E. J., Tian, Z., Bruskiewich, R., Burris, J. N., Carrigan, C. T., et al. (2012). Evaluating methods for isolating total RNA and predicting the success of sequencing phylogenetically diverse plant transcriptomes. *PLoS One, 7*(11), e50226.

Jones-Rhoades, M. W. (2012). Conservation and divergence in plant microRNAs. *Plant Mol Biol, 80*(1), 3-16.

Jones-Rhoades, M. W., & Bartel, D. P. (2004). Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol Cell, 14*(6), 787-799.

Juarez, M. T., Kui, J. S., Thomas, J., Heller, B. A., & Timmermans, M. C. (2004). microRNA-mediated repression of rolled leaf1 specifies maize leaf polarity. *Nature, 428*(6978), 84-88.

Jung, H. J., & Kang, H. (2007). Expression and functional analyses of microRNA417 in Arabidopsis thaliana under stress conditions. *Plant Physiol Biochem, 45*(10-11), 805-811.

Jung, J. H., Seo, P. J., Kang, S. K., & Park, C. M. (2011). miR172 signals are incorporated into the miR156 signaling pathway at the SPL3/4/5 genes in Arabidopsis developmental transitions. *Plant Mol Biol, 76*(1-2), 35-45.

Kader, J.-C. (1997). Lipid-transfer proteins: a puzzling family of plant proteins. *Trends Plant Sci, 2*(2), 66-70.

Kanehisa, M. (2002). The KEGG database. *Novartis Found Symp, 247*, 91-101; discussion 101-103, 119-128, 244-152.

Kang, G. Z., Wang, Z. X., Xia, K. F., & Sun, G. C. (2007). Protection of ultrastructure in chilling-stressed banana leaves by salicylic acid. *J Zhejiang Univ Sci B, 8*(4), 277-282.

Kant, S., Peng, M., & Rothstein, S. J. (2011). Genetic regulation by NLA and microRNA827 for maintaining nitrate-dependent phosphate homeostasis in arabidopsis. *PLoS Genet, 7*(3), e1002021.

Kantar, M., Unver, T., & Budak, H. (2010). Regulation of barley miRNAs upon

dehydration stress correlated with target gene expression. *Funct Integr Genomics, 10*(4), 493-507.

Kaur, S., Cogan, N. O., Pembleton, L. W., Shinozuka, M., Savin, K. W., Materne, M., et al. (2011). Transcriptome sequencing of lentil based on second-generation technology permits large-scale unigene assembly and SSR marker discovery. *BMC Genomics, 12*, 265.

Khraiwesh, B., Zhu, J. K., & Zhu, J. (2012). Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim Biophys Acta, 1819*(2), 137-148.

Kiefer, E., Heller, W., & Ernst, D. (2000). A simple and efficient protocol for isolation of functional RNA from plant tissues rich in secondary metabolites. *Plant Mol Biol Report, 18*(1), 33-39.

Kim, B., Yu, H. J., Park, S. G., Shin, J. Y., Oh, M., Kim, N., et al. (2012a). Identification and profiling of novel microRNAs in the Brassica rapa genome based on small RNA deep sequencing. *BMC Plant Biol, 12*, 218.

Kim, J. J., Lee, J. H., Kim, W., Jung, H. S., Huijser, P., & Ahn, J. H. (2012b). The microRNA156-SQUAMOSA PROMOTER BINDING PROTEIN-LIKE3 module regulates ambient temperature-responsive flowering via FLOWERING LOCUS T in Arabidopsis. *Plant Physiol, 159*(1), 461-478.

Kim, J. Y., Kwak, K. J., Jung, H. J., Lee, H. J., & Kang, H. (2010a). MicroRNA402 affects seed germination of Arabidopsis thaliana under stress conditions via targeting DEMETER-LIKE Protein3 mRNA. *Plant Cell Physiol, 51*(6), 1079-1083.

Kim, J. Y., Lee, H. J., Jung, H. J., Maruyama, K., Suzuki, N., & Kang, H. (2010b). Overexpression of microRNA395c or 395e affects differently the seed germination of Arabidopsis thaliana under stress conditions. *Planta, 232*(6), 1447-1454.

Kim, J., Jung, J. H., Reyes, J. L., Kim, Y. S., Kim, S. Y., Chung, K. S., et al. (2005). microRNA-directed cleavage of ATHB15 mRNA regulates vascular development in Arabidopsis inflorescence stems. *Plant J, 42*(1), 84-94.

Kim, S. W., Li, Z., Moore, P. S., Monaghan, A. P., Chang, Y., Nichols, M., et al. (2010c). A sensitive non-radioactive northern blot method to detect small RNAs. *Nucleic Acids Res, 38*(7), e98.

Kim, W., Ahn, H. J., Chiou, T. J., & Ahn, J. H. (2011). The role of the miR399-PHO2 module in the regulation of flowering time in response to different ambient temperatures in Arabidopsis thaliana. *Mol Cells, 32*(1), 83-88.

Kimmerer, T. W. (1987). Alcohol Dehydrogenase and Pyruvate Decarboxylase Activity in Leaves and Roots of Eastern Cottonwood (Populus deltoides Bartr.) and Soybean (Glycine max L.). *Plant Physiol, 84*(4), 1210-1213.

Koelen, K. J., & Gross, G. G. (1982). Partial purification and properties of tropine dehydrogenase from root cultures of Datura stramonium. *Planta Med, 44*(4), 227-230.

Komatsu, S., Deschamps, T., Hiraga, S., Kato, M., Chiba, M., Hashiguchi, A., et al. (2011). Characterization of a novel flooding stress-responsive alcohol dehydrogenase expressed in soybean roots. *Plant Mol Biol, 77*(3), 309-322.

Kong, W. W., & Yang, Z. M. (2010). Identification of iron-deficiency responsive microRNA genes and cis-elements in Arabidopsis. *Plant Physiol Biochem, 48*(2-3), 153-159.

Kothapalli, N., Camporeale, G., Kueh, A., Chew, Y. C., Oommen, A. M., Griffin, J. B., et al. (2005). Biological functions of biotinylated histones. *J Nutr Biochem, 16*(7), 446-448.

Kulcheski, F. R., de Oliveira, L. F., Molina, L. G., Almerao, M. P., Rodrigues, F. A., Marcolino, J., et al. (2011). Identification of novel soybean microRNAs involved

in abiotic and biotic stresses. *BMC Genomics, 12*, 307.
Kurihara, Y., & Watanabe, Y. (2004). Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. *Proc Natl Acad Sci U S A, 101*(34), 12753-12758.
Kushwaha, H. R., Singh, A. K., Sopory, S. K., Singla-Pareek, S. L., & Pareek, A. (2009). Genome wide expression analysis of CBS domain containing proteins in Arabidopsis thaliana (L.) Heynh and Oryza sativa L. reveals their developmental and stress regulation. *BMC Genomics, 10*, 200.
Kutter, C., Schob, H., Stadler, M., Meins, F., Jr., & Si-Ammour, A. (2007). MicroRNA-mediated regulation of stomatal development in Arabidopsis. *Plant Cell, 19*(8), 2417-2429.
Lagos-Quintana, M., Rauhut, R., Lendeckel, W., & Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science, 294*(5543), 853-858.
Lagrimini, L. M., Burkhart, W., Moyer, M., & Rothstein, S. (1987). Molecular cloning of complementary DNA encoding the lignin-forming peroxidase from tobacco: Molecular analysis and tissue-specific expression. *Proc Natl Acad Sci U S A, 84*(21), 7542-7546.
Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol, 10*(3), R25.
Lau, N. C., Lim, L. P., Weinstein, E. G., & Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science, 294*(5543), 858-862.
Laufs, P., Peaucelle, A., Morin, H., & Traas, J. (2004). MicroRNA regulation of the CUC genes is required for boundary size control in Arabidopsis meristems. *Development, 131*(17), 4311-4322.
Lauter, N., Kampani, A., Carlson, S., Goebel, M., & Moose, S. P. (2005). microRNA172 down-regulates glossy15 to promote vegetative phase change in maize. *Proc Natl Acad Sci U S A, 102*(26), 9412-9417.
Leclercq, J., Ranty, B., Sanchez-Ballesta, M. T., Li, Z., Jones, B., Jauneau, A., et al. (2005). Molecular and biochemical characterization of LeCRK1, a ripening-associated tomato CDPK-related kinase. *J Exp Bot, 56*(409), 25-35.
Lee, R. C., & Ambros, V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. *Science, 294*(5543), 862-864.
Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell, 75*(5), 843-854.
Lee, S. B., Jung, S. J., Go, Y. S., Kim, H. U., Kim, J. K., Cho, H. J., et al. (2009). Two Arabidopsis 3-ketoacyl CoA synthase genes, KCS20 and KCS2/DAISY, are functionally redundant in cuticular wax and root suberin biosynthesis, but differentially controlled by osmotic stress. *Plant J, 60*(3), 462-475.
Lee, W. S., Leong, P. L., Ho, C. L., & Harikrishna, J. A. (2011). Identification of microRNA precursors in Bruguiera spp. *Bot Mar, 54*, 313-324.
Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H., et al. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J, 23*(20), 4051-4060.
Lelandais-Briere, C., Naya, L., Sallet, E., Calenge, F., Frugier, F., Hartmann, C., et al. (2009). Genome-wide Medicago truncatula small RNA analysis revealed novel microRNAs and isoforms differentially regulated in roots and nodules. *Plant Cell, 21*(9), 2780-2796.
Lescot, M., Piffanelli, P., Ciampi, A. Y., Ruiz, M., Blanc, G., Leebens-Mack, J., et al. (2008). Insights into the Musa genome: syntenic relationships to rice and between Musa species. *BMC Genomics, 9*, 58.
Lescot, T. (2011). The genetic diversity of banana in figures. *FruiTrop 189*, 58–62.

Levy, M., Wang, Q., Kaspi, R., Parrella, M. P., & Abel, S. (2005). Arabidopsis IQD1, a novel calmodulin-binding nuclear protein, stimulates glucosinolate accumulation and plant defense. *Plant J, 43*(1), 79-96.

Li, B., Duan, H., Li, J., Deng, X. W., Yin, W., & Xia, X. (2013a). Global identification of miRNAs and targets in Populus euphratica under salt stress. *Plant Mol Biol, 81*(6), 525-539.

Li, B., Qin, Y., Duan, H., Yin, W., & Xia, X. (2011a). Genome-wide characterization of new and drought stress responsive microRNAs in Populus euphratica. *J Exp Bot, 62*(11), 3765-3779.

Li, C. Y., Deng, G. M., Yang, J., Viljoen, A., Jin, Y., Kuang, R. B., et al. (2012a). Transcriptome profiling of resistant and susceptible Cavendish banana roots following inoculation with Fusarium oxysporum f. sp. cubense tropical race 4. *BMC Genomics, 13*, 374.

Li, C., Shao, J., Wang, Y., Li, W., Guo, D., Yan, B., et al. (2013b). Analysis of banana transcriptome and global gene expression profiles in banana roots in response to infection by race 1 and tropical race 4 of Fusarium oxysporum f. sp. cubense. *BMC Genomics, 14*, 851.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics, 25*(14), 1754-1760.

Li, H., Deng, Y., Wu, T., Subramanian, S., & Yu, O. (2010a). Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation. *Plant Physiol, 153*(4), 1759-1770.

Li, H., Dong, Y., Yin, H., Wang, N., Yang, J., Liu, X., et al. (2011b). Characterization of the stress associated microRNAs in Glycine max by deep sequencing. *BMC Plant Biol, 11*, 170.

Li, J., Gao, G., Zhang, T., & Wu, X. (2013c). The putative phytocyanin genes in Chinese cabbage (Brassica rapa L.): genome-wide identification, classification and expression analysis. *Mol Genet Genomics, 288*(1-2), 1-20.

Li, J., Han, Y., Zhao, Q., Li, C., Xie, Q., Chong, K., et al. (2013d). The E3 ligase AtRDUF1 positively regulates salt stress responses in Arabidopsis thaliana. *PLoS One, 8*(8), e71078.

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., et al. (2010b). De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res, 20*(2), 265-272.

Li, T., Chen, J., Qiu, S., Zhang, Y., Wang, P., Yang, L., et al. (2012b). Deep sequencing and microarray hybridization identify conserved and species-specific microRNAs during somatic embryogenesis in hybrid yellow poplar. *PLoS One, 7*(8), e43451.

Li, T., Li, H., Zhang, Y. X., & Liu, J. Y. (2011c). Identification and analysis of seven HO-responsive miRNAs and 32 new miRNAs in the seedlings of rice (Oryza sativa L. ssp. indica). *Nucleic Acids Res, 39*(7), 2821-2833.

Li, W. X., Oono, Y., Zhu, J., He, X. J., Wu, J. M., Iida, K., et al. (2008). The Arabidopsis NFYA5 transcription factor is regulated transcriptionally and posttranscriptionally to promote drought resistance. *Plant Cell, 20*(8), 2238-2251.

Li, X., Bian, H., Song, D., Ma, S., Han, N., Wang, J., et al. (2013e). Flowering time control in ornamental gloxinia (Sinningia speciosa) by manipulation of miR159 expression. *Ann Bot, 111*(5), 791-799.

Li, Y. F., Zheng, Y., Addo-Quaye, C., Zhang, L., Saini, A., Jagadeeswaran, G., et al. (2010c). Transcriptome-wide identification of microRNA targets in rice. *Plant J, 62*(5), 742-759.

Liang, G., & Yu, D. (2010). Reciprocal regulation among miR395, APS and SULTR2;1

in Arabidopsis thaliana. *Plant Signal Behav, 5*(10), 1257-1259.
Liang, G., He, H., & Yu, D. (2012). Identification of nitrogen starvation-responsive microRNAs in Arabidopsis thaliana. *PLoS One, 7*(11), e48951.
Liang, G., Yang, F., & Yu, D. (2010). MicroRNA395 mediates regulation of sulfate accumulation and allocation in Arabidopsis thaliana. *Plant J, 62*(6), 1046-1057.
Liang, M., Haroldsen, V., Cai, X., & Wu, Y. (2006). Expression of a putative laccase gene, ZmLAC1, in maize primary roots under stress. *Plant Cell Environ, 29*(5), 746-753.
Liang, R. Q., Li, W., Li, Y., Tan, C. Y., Li, J. X., Jin, Y. X., et al. (2005). An oligonucleotide microarray for microRNA expression analysis based on labeling RNA with quantum dot and nanogold probe. *Nucleic Acids Res, 33*(2), e17.
Lima, J. C., Arenhart, R. A., Margis-Pinheiro, M., & Margis, R. (2011). Aluminum triggers broad changes in microRNA expression in rice roots. *Genet Mol Res, 10*(4), 2817-2832.
Lin, C., Choi, H. S., & Cho, H. T. (2011). Root hair-specific EXPANSIN A7 is required for root hair elongation in Arabidopsis. *Mol Cells, 31*(4), 393-397.
Lin, J. S., Lin, C. C., Lin, H. H., Chen, Y. C., & Jeng, S. T. (2012). MicroR828 regulates lignin and H2O2 accumulation in sweet potato on wounding. *New Phytol, 196*(2), 427-440.
Lindberg, J., & Lundeberg, J. (2010). The plasticity of the mammalian transcriptome. *Genomics, 95*(1), 1-6.
Liu, B., Li, P., Li, X., Liu, C., Cao, S., Chu, C., et al. (2005). Loss of function of OsDCL1 affects microRNA accumulation and causes developmental defects in rice. *Plant Physiol, 139*(1), 296-305.
Liu, H. H., Tian, X., Li, Y. J., Wu, C. A., & Zheng, C. C. (2008). Microarray-based analysis of stress-regulated microRNAs in Arabidopsis thaliana. *RNA, 14*(5), 836-843.
Liu, H., Trieu, A. T., Blaylock, L. A., & Harrison, M. J. (1998). Cloning and characterization of two phosphate transporters from Medicago truncatula roots: regulation in response to phosphate and to colonization by arbuscular mycorrhizal (AM) fungi. *Mol Plant Microbe Interact, 11*(1), 14-22.
Liu, H., Xue, L., Li, C., Zhang, R., & Ling, Q. (2001). Calmodulin-binding protein BP-10, a probable new member of plant nonspecific lipid transfer protein superfamily. *Biochem Biophys Res Commun, 285*(3), 633-638.
Liu, P. P., Montgomery, T. A., Fahlgren, N., Kasschau, K. D., Nonogaki, H., & Carrington, J. C. (2007). Repression of AUXIN RESPONSE FACTOR10 by microRNA160 is critical for seed germination and post-germination stages. *Plant J, 52*(1), 133-146.
Liu, Q., & Zhang, H. (2012). Molecular identification and analysis of arsenite stress-responsive miRNAs in rice. *J Agric Food Chem, 60*(26), 6524-6536.
Liu, X., Huang, J., Wang, Y., Khanna, K., Xie, Z., Owen, H. A., et al. (2010). The role of floral organs in carpels, an Arabidopsis loss-of-function mutation in MicroRNA160a, in organogenesis and the mechanism regulating its expression. *Plant J, 62*(3), 416-428.
Llave, C., Kasschau, K. D., Rector, M. A., & Carrington, J. C. (2002a). Endogenous and silencing-associated small RNAs in plants. *Plant Cell, 14*(7), 1605-1619.
Llave, C., Xie, Z., Kasschau, K. D., & Carrington, J. C. (2002b). Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science, 297*(5589), 2053-2056.
Lo, S. F., Yang, S. Y., Chen, K. T., Hsing, Y. I., Zeevaart, J. A., Chen, L. J., et al. (2008). A novel class of gibberellin 2-oxidases control semidwarfism, tillering, and root development in rice. *Plant Cell, 20*(10), 2603-2618.

Lu, C., Kulkarni, K., Souret, F. F., MuthuValliappan, R., Tej, S. S., Poethig, R. S., et al. (2006). MicroRNAs and other small RNAs enriched in the Arabidopsis RNA-dependent RNA polymerase-2 mutant. *Genome Res, 16*(10), 1276-1288.

Lu, C., Tej, S. S., Luo, S., Haudenschild, C. D., Meyers, B. C., & Green, P. J. (2005a). Elucidation of the small RNA component of the transcriptome. *Science, 309*(5740), 1567-1569.

Lu, P., Kang, M., Jiang, X., Dai, F., Gao, J., & Zhang, C. (2013a). RhEXPA4, a rose expansin gene, modulates leaf growth and confers drought and salt tolerance to Arabidopsis. *Planta, 237*(6), 1547-1559.

Lu, S., Li, Q., Wei, H., Chang, M. J., Tunlaya-Anukit, S., Kim, H., et al. (2013b). Ptr-miR397a is a negative regulator of laccase genes affecting lignin content in Populus trichocarpa. *Proc Natl Acad Sci U S A, 110*(26), 10848-10853.

Lu, S., Sun, Y. H., & Chiang, V. L. (2008). Stress-responsive microRNAs in Populus. *Plant J, 55*(1), 131-151.

Lu, S., Sun, Y. H., Shi, R., Clark, C., Li, L., & Chiang, V. L. (2005b). Novel and mechanical stress-responsive MicroRNAs in Populus trichocarpa that are absent from Arabidopsis. *Plant Cell, 17*(8), 2186-2203.

Lu, S., Yang, C., & Chiang, V. L. (2011). Conservation and Diversity of MicroRNA-associated Copper-regulatory Networks in Populus trichocarpa(F). *J Integr Plant Biol, 53*(11), 879-891.

Lu, X., Guan, Q., & Zhu, J. (2013c). Downregulation of CSD2 by a heat-inducible miR398 is required for thermotolerance in Arabidopsis. *Plant Signal Behav, 8*(8).

Luo, Y. C., Zhou, H., Li, Y., Chen, J. Y., Yang, J. H., Chen, Y. Q., et al. (2006). Rice embryogenic calli express a unique set of microRNAs, suggesting regulatory roles of microRNAs in plant post-embryogenic development. *FEBS Lett, 580*(21), 5111-5116.

Lv, D. K., Bai, X., Li, Y., Ding, X. D., Ge, Y., Cai, H., et al. (2010). Profiling of cold-stress-responsive miRNAs in rice by microarrays. *Gene, 459*(1-2), 39-47.

Macovei, A., & Tuteja, N. (2012). microRNAs targeting DEAD-box helicases are involved in salinity stress response in rice (Oryza sativa L.). *BMC Plant Biol, 12*, 183.

Maghuly, F., Borroto-Fernandez, E. G., Khan, M. A., Herndl, A., Marzban, G., & Laimer, M. (2009). Expression of calmodulin and lipid transfer protein genes in Prunus incisa x serrula under different stress conditions. *Tree Physiol, 29*(3), 437-444.

Mahajan, S., & Tuteja, N. (2005). Cold, salinity and drought stresses: an overview. *Arch Biochem Biophys, 444*(2), 139-158.

Malhotra, S., & Sowdhamini, R. (2013). Genome-wide survey of DNA-binding proteins in Arabidopsis thaliana: analysis of distribution and functions. *Nucleic Acids Res, 41*(15), 7212-7219.

Mallory, A. C., & Vaucheret, H. (2006). Functions of microRNAs and related small RNAs in plants. *Nat Genet, 38 Suppl*, S31-36.

Mallory, A. C., Bartel, D. P., & Bartel, B. (2005). MicroRNA-directed regulation of Arabidopsis AUXIN RESPONSE FACTOR17 is essential for proper development and modulates expression of early auxin response genes. *Plant Cell, 17*(5), 1360-1375.

Mallory, A. C., Reinhart, B. J., Jones-Rhoades, M. W., Tang, G., Zamore, P. D., Barton, M. K., et al. (2004). MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. *EMBO J, 23*(16), 3356-3364.

Malone, J. H., & Oliver, B. (2011). Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol, 9*, 34.

Manrique-Trujillo, S. M., Ramirez-Lopez, A. C., Ibarra-Laclette, E., & Gomez-Lim, M.

A. (2007). Identification of genes differentially expressed during ripening of banana. *J Plant Physiol, 164*(8), 1037-1050.

Marin, E., Jouannet, V., Herz, A., Lokerse, A. S., Weijers, D., Vaucheret, H., et al. (2010). miR390, Arabidopsis TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. *Plant Cell, 22*(4), 1104-1117.

Marmagne, A., Vinauger-Douard, M., Monachello, D., de Longevialle, A. F., Charon, C., Allot, M., et al. (2007). Two members of the Arabidopsis CLC (chloride channel) family, AtCLCe and AtCLCf, are associated with thylakoid and Golgi membranes, respectively. *J Exp Bot, 58*(12), 3385-3393.

Martin, J. A., & Wang, Z. (2011). Next-generation transcriptome assembly. *Nat Rev Genet, 12*(10), 671-682.

Masahara-Negishi, Y., Hosomi, A., Della Mea, M., Serafini-Fracassini, D., & Suzuki, T. (2012). A plant peptide: N-glycanase orthologue facilitates glycoprotein ER-associated degradation in yeast. *Biochim Biophys Acta, 1820*(10), 1457-1462.

Mathieu, J., Yant, L. J., Murdter, F., Kuttner, F., & Schmid, M. (2009). Repression of flowering by the miR172 target SMZ. *PLoS Biol, 7*(7), e1000148.

May, P., Liao, W., Wu, Y., Shuai, B., McCombie, W. R., Zhang, M. Q., et al. (2013). The effects of carbon dioxide and temperature on microRNA expression in Arabidopsis development. *Nat Commun, 4*, 2145.

McCormick, K. P., Willmann, M. R., & Meyers, B. C. (2011). Experimental design, preprocessing, normalization and differential expression analysis of small RNA sequencing experiments. *Silence, 2*(1), 2.

Meng, F., Liu, H., Wang, K., Liu, L., Wang, S., Zhao, Y., et al. (2013). Development-associated microRNAs in grains of wheat (Triticum aestivum L.). *BMC Plant Biol, 13*, 140.

Meng, Y., Shao, C., Wang, H., & Chen, M. (2011). The regulatory activities of plant microRNAs: a more dynamic perspective. *Plant Physiol, 157*(4), 1583-1595.

Metpally, R. P., Nasser, S., Malenica, I., Courtright, A., Carlson, E., Ghaffari, L., et al. (2013). Comparison of Analysis Tools for miRNA High Throughput Sequencing Using Nerve Crush as a Model. *Front Genet, 4*, 20.

Meyer, S. U., Pfaffl, M. W., & Ulbrich, S. E. (2010). Normalization strategies for microRNA profiling experiments: a 'normal' way to a hidden layer of complexity? *Biotechnol Lett, 32*(12), 1777-1788.

Meyers, B. C., Axtell, M. J., Bartel, B., Bartel, D. P., Baulcombe, D., Bowman, J. L., et al. (2008). Criteria for annotation of plant MicroRNAs. *Plant Cell, 20*(12), 3186-3190.

Miller, J. R., Koren, S., & Sutton, G. (2010). Assembly algorithms for next-generation sequencing data. *Genomics, 95*(6), 315-327.

Miri, S. M., Mousavi, A., Naghavi, M. R., Mirzali, M., Talaei, A. R., & Khiabani, B. N. (2009). Analysis of induced mutants of salinity resistant banana (Musa acuminata cv. Dwarf Cavendish) using morphological and molecular markers. *Iran J Biotech, 7*(2), 86-92.

Misic, D., Dragicevic, M., Siler, B., Nestorovic Zivkovic, J., Maksimovic, V., Momcilovic, I., et al. (2012). Sugars and acid invertase mediate the physiological response of Schenkia spicata root cultures to salt stress. *J Plant Physiol, 169*(13), 1281-1289.

Miyashima, S., Koi, S., Hashimoto, T., & Nakajima, K. (2011). Non-cell-autonomous microRNA165 acts in a dose-dependent manner to regulate multiple differentiation status in the Arabidopsis root. *Development, 138*(11), 2303-2313.

Mizoi, J., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2012). AP2/ERF family transcription factors in plant abiotic stress responses. *Biochim Biophys Acta,*

1819(2), 86-96.
Mizoi, J., & Yamaguchi-Shinozaki, K. (2013). Molecular approaches to improve rice abiotic stress tolerance. *Methods Mol Biol, 956*, 269-283.
Mizrachi, E., Hefer, C. A., Ranik, M., Joubert, F., & Myburg, A. A. (2010). De novo assembled expressed gene catalog of a fast-growing Eucalyptus tree produced by Illumina mRNA-Seq. *BMC Genomics, 11*, 681.
Mobley, E. M., Kunkel, B. N., & Keith, B. (1999). Identification, characterization and comparative analysis of a novel chorismate mutase gene in Arabidopsis thaliana. *Gene, 240*(1), 115-123.
Molnar, A., Schwach, F., Studholme, D. J., Thuenemann, E. C., & Baulcombe, D. C. (2007). miRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii. *Nature, 447*(7148), 1126-1129.
Morin, R. D., Aksay, G., Dolgosheina, E., Ebhardt, H. A., Magrini, V., Mardis, E. R., et al. (2008). Comparative analysis of the small RNA transcriptomes of Pinus contorta and Oryza sativa. *Genome Res, 18*(4), 571-584.
Morris, E. R., & Walker, J. C. (2003). Receptor-like protein kinases: the keys to response. *Curr Opin Plant Biol, 6*(4), 339-342.
Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods, 5*(7), 621-628.
Moscetti, I., Tundo, S., Janni, M., Sella, L., Gazzetti, K., Tauzin, A., et al. (2013). Constitutive expression of the xylanase inhibitor TAXI-III delays Fusarium head blight symptoms in durum wheat transgenic plants. *Mol Plant Microbe Interact, 26*(12), 1464-1472.
Mou, G., Wang, K., Xu, D., & Zhou, G. (2013). Evaluation of Three RT-qPCR-Based miRNA Detection Methods Using Seven Rice miRNAs. *Biosci Biotechnol Biochem, 77*(6), 1349-1353
Moxon, S., Jing, R., Szittya, G., Schwach, F., Rusholme Pilcher, R. L., Moulton, V., et al. (2008). Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. *Genome Res, 18*(10), 1602-1609.
Muchhal, U. S., & Raghothama, K. G. (1999). Transcriptional regulation of plant phosphate transporters. *Proc Natl Acad Sci U S A, 96*(10), 5868-5872.
Muchhal, U. S., Pardo, J. M., & Raghothama, K. G. (1996). Phosphate transporters from the higher plant Arabidopsis thaliana. *Proc Natl Acad Sci U S A, 93*(19), 10519-10523.
Munns, R., James, R. A., Xu, B., Athman, A., Conn, S. J., Jordans, C., et al. (2012). Wheat grain yield on saline soils is improved by an ancestral Na(+) transporter gene. *Nat Biotechnol, 30*, 360-364.
Munns, R., & Tester, M. (2008). Mechanisms of salinity tolerance. *Annu Rev Plant Biol, 59*, 651-681.
Murashige, T, & Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco cultures. *Physiol Plant, 15*, 473-497.
Nag, A., King, S., & Jack, T. (2009). miR319a targeting of TCP4 is critical for petal growth and development in Arabidopsis. *Proc Natl Acad Sci U S A, 106*(52), 22534-22539.
Nakashima, K., Yamaguchi-Shinozaki, K., & Shinozaki, K. (2014). The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress responses including drought, cold, and heat. *Front Plant Sci, 5*, 170.
Nasaruddin, N. M., Harikrishna, K., Othman, R. Y., Lim, S. H., & Harikrishna, J. A. (2007). Computational prediction of microRNAs from oil palm (Elaeis guineensis Jacq.) expressed sequence tags. *AsPac J Mol Biol Biotechnol, 15*(3), 107-113.

Nath, M., Garg, B., Sahoo, R. K., & Tuteja, N. (2015). PDH45 overexpressing transgenic tobacco and rice plants provide salinity stress tolerance via less sodium accumulation. *Plant Signal Behav, 10*, e992289

Navarro, L., Dunoyer, P., Jay, F., Arnold, B., Dharmasiri, N., Estelle, M., et al. (2006). A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science, 312*(5772), 436-439.

Ni, Z., Hu, Z., Jiang, Q., & Zhang, H. (2012). Overexpression of gma-MIR394a confers tolerance to drought in transgenic Arabidopsis thaliana. *Biochem Biophys Res Commun, 427*(2), 330-335.

Ni, Z., Hu, Z., Jiang, Q., & Zhang, H. (2013). GmNFYA3, a target gene of miR169, is a positive regulator of plant tolerance to drought stress. *Plant Mol Biol, 82*(1-2), 113-129.

Nischal, L., Mohsin, M., Khan, I., Kardam, H., Wadhwa, A., Abrol, Y. P., et al. (2012). Identification and comparative analysis of microRNAs associated with low-N tolerance in rice genotypes. *PLoS One, 7*(12), e50261.

Noll, M., & Lutsenko, S. (2000). Expression of ZntA, a zinc-transporting P1-type ATPase, is specifically regulated by zinc and cadmium. *IUBMB Life, 49*(4), 297-302.

Oh, T. J., Wartell, R. M., Cairney, J., & Pullman, G. S. (2008). Evidence for stage-specific modulation of specific microRNAs (miRNAs) and miRNA processing components in zygotic embryo and female gametophyte of loblolly pine (Pinus taeda). *New Phytol, 179*(1), 67-80.

Omar, S., Jamaluddin, S. H., Tahir, M., & Bahari, U. M. (2012). *Collection of bananas in Malaysia*. Serdang: Malaysian Agricultural Research and Development Institute (MARDI).

Ori, N., Cohen, A. R., Etzioni, A., Brand, A., Yanai, O., Shleizer, S., et al. (2007). Regulation of LANCEOLATE by miR319 is required for compound-leaf development in tomato. *Nat Genet, 39*(6), 787-791.

Osborn, T. (2002). *Banana research and development in the Pacific: Status report 2002*. INIBAP-BAPNET, Montpellier: INIBAP; Los *Baños, Laguna*: BAPNET.

Ouyang, B., Yang, T., Li, H., Zhang, L., Zhang, Y., Zhang, J., et al. (2007). Identification of early salt stress response genes in tomato root by suppression subtractive hybridization and microarray analysis. *J Exp Bot, 58*(3), 507-520.

Ozsolak, F., & Milos, P. M. (2011). RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet, 12*(2), 87-98.

Palatnik, J. F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J. C., et al. (2003). Control of leaf morphogenesis by microRNAs. *Nature, 425*(6955), 257-263.

Pall, G. S., Codony-Servat, C., Byrne, J., Ritchie, L., & Hamilton, A. (2007). Carbodiimide-mediated cross-linking of RNA to nylon membranes improves the detection of siRNA, miRNA and piRNA by northern blot. *Nucleic Acids Res, 35*(8), e60.

Palma, J. M., Corpas, F. J., & del Rio, L. A. (2009). Proteome of plant peroxisomes: new perspectives on the role of these organelles in cell biology. *Proteomics, 9*(9), 2301-2312.

Pan, Y., Seymour, G. B., Lu, C., Hu, Z., Chen, X., & Chen, G. (2012). An ethylene response factor (ERF5) promoting adaptation to drought and salt tolerance in tomato. *Plant Cell Rep, 31*(2), 349-360.

Pandey, S. P., Shahi, P., Gase, K., & Baldwin, I. T. (2008). Herbivory-induced changes in the small-RNA transcriptome and phytohormone signaling in Nicotiana attenuata. *Proc Natl Acad Sci U S A, 105*(12), 4559-4564.

Pang, M., Woodward, A. W., Agarwal, V., Guan, X., Ha, M., Ramachandran, V., et al.

(2009). Genome-wide analysis reveals rapid and dynamic changes in miRNA and siRNA sequence and expression during ovule and fiber development in allotetraploid cotton (Gossypium hirsutum L.). *Genome Biol, 10*(11), R122.

Pant, B. D., Buhtz, A., Kehr, J., & Scheible, W. R. (2008). MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. *Plant J, 53*(5), 731-738.

Pantaleo, V., Szittya, G., Moxon, S., Miozzi, L., Moulton, V., Dalmay, T., et al. (2010). Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis. *Plant J, 62*(6), 960-976.

Papp, I., Mette, M. F., Aufsatz, W., Daxinger, L., Schauer, S. E., Ray, A., et al. (2003). Evidence for nuclear processing of plant micro RNA and short interfering RNA precursors. *Plant Physiol, 132*(3), 1382-1390.

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., et al. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature, 408*(6808), 86-89.

Passos, M. A., de Cruz, V. O., Emediato, F. L., de Teixeira, C. C., Azevedo, V. C., Brasileiro, A. C., et al. (2013). Analysis of the leaf transcriptome of Musa acuminata during interaction with Mycosphaerella musicola: gene assembly, annotation and marker development. *BMC Genomics, 14*, 78.

Passos, M. A., de Oliveira Cruz, V., Emediato, F. L., de Camargo Teixeira, C., Souza, M. T., Jr., Matsumoto, T., et al. (2012). Development of expressed sequence tag and expressed sequence tag-simple sequence repeat marker resources for Musa acuminata. *AoB Plants, 2012*, pls030.

Peng, X., Zhao, Y., Cao, J., Zhang, W., Jiang, H., Li, X., et al. (2012). CCCH-type zinc finger family in maize: genome-wide identification, classification and expression profiling under abscisic acid and drought treatments. *PLoS One, 7*(7), e40120.

Perez-Quintero, A. L., Quintero, A., Urrego, O., Vanegas, P., & Lopez, C. (2012). Bioinformatic identification of cassava miRNAs differentially expressed in response to infection by Xanthomonas axonopodis pv. manihotis. *BMC Plant Biol, 12*, 29.

Perrier, X., De Langhe, E., Donohue, M., Lentfer, C., Vrydaghs, L., Bakry, F., et al. (2011). Multidisciplinary perspectives on banana (Musa spp.) domestication. *Proc Natl Acad Sci U S A, 108*(28), 11311-11318.

Personal communication. Dr. Hugo Alfried Volkaert, Kasetsart University, Thailand.

Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S., et al. (2003). TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics, 19*(5), 651-652.

Pestana, R. K., Amorim, E. P., Ferreira, C. F., Amorim, V. B., Oliveira, L. S., Ledo, C. A., et al. (2011). Genetic dissimilarity of putative gamma-ray-induced 'Preciosa-AAAB-Pome type' banana (Musa sp) mutants based on multivariate statistical analysis. *Genet Mol Res, 10*(4), 3976-3986.

Pillay, M., & Tripathi, L. (2007). Banana. In C. Kole (Ed.), *Genome mapping and molecular breeding in plants* (Vol. 4, pp. 281-297). Berlin Heidelberg: Springer-Verlag.

Pirrello, J., Jaimes-Miranda, F., Sanchez-Ballesta, M. T., Tournier, B., Khalil-Ahmad, Q., Regad, F., et al. (2006). Sl-ERF2, a tomato ethylene response factor involved in ethylene response and seed germination. *Plant Cell Physiol, 47*(9), 1195-1205.

Powell, J. (2000). SAGE. The serial analysis of gene expression. *Methods Mol Biol, 99*, 297-319.

Pysh, L. D., Wysocka-Diller, J. W., Camilleri, C., Bouchez, D., & Benfey, P. N. (1999). The GRAS gene family in Arabidopsis: sequence characterization and basic

expression analysis of the SCARECROW-LIKE genes. *Plant J, 18*(1), 111-119.
Qin, F., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2011). Achievements and challenges in understanding plant abiotic stress responses and tolerance. *Plant Cell Physiol, 52*(9), 1569-1582.
Qiu, C. X., Xie, F. L., Zhu, Y. Y., Guo, K., Huang, S. Q., Nie, L., et al. (2007). Computational identification of microRNAs and their targets in Gossypium hirsutum expressed sequence tags. *Gene, 395*(1-2), 49-61.
Ralph, S., Park, J. Y., Bohlmann, J., & Mansfield, S. D. (2006). Dirigent proteins in conifer defense: gene discovery, phylogeny, and differential wound- and insect-induced expression of a family of DIR and DIR-like genes in spruce (Picea spp.). *Plant Mol Biol, 60*(1), 21-40.
Ramakrishna, A., & Ravishankar, G. A. (2011). Influence of abiotic stress signals on secondary metabolites in plants. *Plant Signal Behav, 6*(11), 1720-1731.
Raman, S., Greb, T., Peaucelle, A., Blein, T., Laufs, P., & Theres, K. (2008). Interplay of miR164, CUP-SHAPED COTYLEDON genes and LATERAL SUPPRESSOR controls axillary meristem formation in Arabidopsis thaliana. *Plant J, 55*(1), 65-76.
Ramkissoon, S. H., Mainwaring, L. A., Sloand, E. M., Young, N. S., & Kajigaya, S. (2006). Nonisotopic detection of microRNA using digoxigenin labeled RNA probes. *Mol Cell Probes, 20*(1), 1-4.
Ravet, K., Danford, F. L., Dihle, A., Pittarello, M., & Pilon, M. (2011). Spatiotemporal Analysis of Copper Homeostasis in Populus trichocarpa Reveals an Integrated Molecular Remodeling for a Preferential Allocation of Copper to Plastocyanin in the Chloroplasts of Developing Leaves. *Plant Physiol, 157*(3), 1300-1312.
Ravi, I., Uma, S., Vaganan, M. M., & Mustaffa, M. M. (2013). Phenotyping bananas for drought resistance. *Front Physiol, 4*, 9.
Raymond, C. K., Roberts, B. S., Garrett-Engele, P., Lim, L. P., & Johnson, J. M. (2005). Simple, quantitative primer-extension PCR assay for direct monitoring of microRNAs and short-interfering RNAs. *RNA, 11*(11), 1737-1744.
Reddy, P. S., Reddy, G. M., Pandey, P., Chandrasekhar, K., & Reddy, M. K. (2012). Cloning and molecular characterization of a gene encoding late embryogenesis abundant protein from Pennisetum glaucum: protection against abiotic stresses. *Mol Biol Rep, 39*(6), 7163-7174.
Reichel, H., Martínez, A. K., Arroyave, J. A., Sedano, R., Morales, F. J., Duterme, O., et al. (2003). First Report of Banana mild mosaic virus Isolated from Plantains (Musa AAB) in Colombia. *Plant Disease, 87*(9), 1150-1150.
Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., et al. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature, 403*(6772), 901-906.
Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B., & Bartel, D. P. (2002). MicroRNAs in plants. *Genes Dev, 16*(13), 1616-1626.
Reyes, J. L., & Chua, N. H. (2007). ABA induction of miR159 controls transcript levels of two MYB factors during Arabidopsis seed germination. *Plant J, 49*(4), 592-606.
Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., & Bartel, D. P. (2002). Prediction of plant microRNA targets. *Cell, 110*(4), 513-520.
Rigo, G., Ayaydin, F., Tietz, O., Zsigmond, L., Kovacs, H., Pay, A., et al. (2013). Inactivation of plasma membrane-localized CDPK-RELATED KINASE5 decelerates PIN2 exocytosis and root gravitropic response in Arabidopsis. *Plant Cell, 25*(5), 1592-1608.
Rodriguez, R. E., Mecchia, M. A., Debernardi, J. M., Schommer, C., Weigel, D., & Palatnik, J. F. (2010). Control of cell proliferation in Arabidopsis thaliana by

microRNA miR396. *Development, 137*(1), 103-112.
Rodriguez-Milla, M. A., & Salinas, J. (2009). Prefoldins 3 and 5 play an essential role in Arabidopsis tolerance to salt stress. *Mol Plant, 2*(3), 526-534.
Roux, N., Baurens, F.-C., Doleˇzel, J., Hˇribová, E., Heslop-Harrison, P., Town, C., et al. (2008). Genomics of banana and plantain (Musa spp.), major staple crops in the tropics. In P. H. Moore, R. Ming (Eds.), *Genomics of Tropical Crop Plants* (pp. 83-111). New York: Springer.
Rowland, O., Ludwig, A. A., Merrick, C. J., Baillieul, F., Tracy, F. E., Durrant, W. E., et al. (2005). Functional analysis of Avr9/Cf-9 rapidly elicited genes identifies a protein kinase, ACIK1, that is essential for full Cf-9-dependent disease resistance in tomato. *Plant Cell, 17*(1), 295-310.
Ru, P., Xu, L., Ma, H., & Huang, H. (2006). Plant fertility defects induced by the enhanced expression of microRNA167. *Cell Res, 16*(5), 457-465.
Rubio-Somoza, I., & Weigel, D. (2011). MicroRNA networks and developmental plasticity in plants. *Trends Plant Sci, 16*(5), 258-264.
Rueda, E. C., Dezar, C. A., Gonzalez, D. H., & Chan, R. L. (2005). Hahb-10, a sunflower homeobox-leucine zipper gene, is regulated by light quality and quantity, and promotes early flowering when expressed in Arabidopsis. *Plant Cell Physiol, 46*(12), 1954-1963.
Rumble, S. M., Lacroute, P., Dalca, A. V., Fiume, M., Sidow, A., & Brudno, M. (2009). SHRiMP: accurate mapping of short color-space reads. *PLoS Comput Biol, 5*(5), e1000386.
Sablok, G., Luo, C., Lee, W. S., Rahman, F., Tatarinova, T. V., Harikrishna, J. A., & Luo Z. (2011). Bioinformatic analysis of fruits-specific expressed sequence tag libraries of Diospyros kaki Thunb.: view at the transcriptome at different developmental stages. 3 Biotech 1, 35-45.
Safar, J., Noa-Carrazana, J. C., Vrana, J., Bartos, J., Alkhimova, O., Sabau, X., et al. (2004). Creation of a BAC resource to study the structure and evolution of the banana (Musa balbisiana) genome. *Genome, 47*(6), 1182-1191.
Saleh, O., Arazi, T., & Frank, W. (2011). MicroRNA-mediated establishment of transcription factor gradients controlling developmental phase transitions. *Plant Signal Behav, 6*(6), 873-877.
Santos, C. M., Martins, N. F., Horberg, H. M., de Almeida, E. R., Coelho, M. C., Togawa, R. C., et al. (2005). Analysis of expressed sequence tags from Musa acuminata ssp. burmannicoides, var. Calcutta 4 (AA) leaves submitted to temperature stresses. *Theor Appl Genet, 110*(8), 1517-1522.
Sattar, S., Song, Y., Anstead, J. A., Sunkar, R., & Thompson, G. A. (2012). Cucumis melo microRNA expression profile during aphid herbivory in a resistant and susceptible interaction. *Mol Plant Microbe Interact, 25*(6), 839-848.
Schaller, A., & Stintzi, A. (2009). Enzymes in jasmonate biosynthesis - structure, function, regulation. *Phytochemistry, 70*(13-14), 1532-1538.
Schmid, M., Uhlenhaut, N. H., Godard, F., Demar, M., Bressan, R., Weigel, D., et al. (2003). Dissection of floral induction pathways using global expression analysis. *Development, 130*(24), 6001-6012.
Schmieder, R., & Edwards, R. (2011). Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PLoS One, 6*(3), e17288.
Schmittgen, T. D., Jiang, J., Liu, Q., & Yang, L. (2004). A high-throughput method to monitor the expression of microRNA precursors. *Nucleic Acids Res, 32*(4), e43.
Schofield, R. A., Bi, Y. M., Kant, S., & Rothstein, S. J. (2009). Over-expression of STP13, a hexose transporter, improves plant growth and nitrogen use in Arabidopsis thaliana seedlings. *Plant Cell Environ, 32*(3), 271-285.

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., et al. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol, 7*, 3.

Schuster, B., & Retey, J. (1995). The mechanism of action of phenylalanine ammonia-lyase: the role of prosthetic dehydroalanine. *Proc Natl Acad Sci U S A, 92*(18), 8433-8437.

Sewelam, N., Kazan, K., Thomas-Hall, S. R., Kidd, B. N., Manners, J. M., & Schenk, P. M. (2013). Ethylene response factor 6 is a regulator of reactive oxygen species signaling in Arabidopsis. *PLoS One, 8*(8), e70289.

Shan, C., Mei, Z., Duan, J., Chen, H., Feng, H., & Cai, W. (2014). OsGA2ox5, a gibberellin metabolism enzyme, is involved in plant growth, the root gravity response and salt stress. *PLoS One, 9*(1), e87110.

Shapira, O., Khadka, S., Israeli, Y., Shani, U., & Schwartz, A. (2009). Functional anatomy controls ion distribution in banana leaves: significance of Na+ seclusion at the leaf margins. *Plant Cell Environ, 32*(5), 476-485.

Sharma, R., Cao, P., Jung, K. H., Sharma, M. K., & Ronald, P. C. (2013). Construction of a rice glycoside hydrolase phylogenomic database and identification of targets for biofuel research. *Front Plant Sci, 4*, 330.

Sheen, J. (1998). Mutational analysis of protein phosphatase 2C involved in abscisic acid signal transduction in higher plants. *Proc Natl Acad Sci U S A, 95*(3), 975-980.

Shen, J., Xie, K., & Xiong, L. (2010). Global expression profiling of rice microRNAs by one-tube stem-loop reverse transcription quantitative PCR revealed important roles of microRNAs in abiotic stress responses. *Mol Genet Genomics, 284*(6), 477-488.

Shi, X., Gupta, S., & Rashotte, A. M. (2012). Solanum lycopersicum cytokinin response factor (SlCRF) genes: characterization of CRF domain-containing ERF genes in tomato. *J Exp Bot, 63*(2), 973-982.

Shin, D., Koo, Y. D., Lee, J., Lee, H. J., Baek, D., Lee, S., et al. (2004). Athb-12, a homeobox-leucine zipper domain protein from Arabidopsis thaliana, increases salt tolerance in yeast by regulating sodium exclusion. *Biochem Biophys Res Commun, 323*(2), 534-540.

Shipman-Roston, R. L., Ruppel, N. J., Damoc, C., Phinney, B. S., & Inoue, K. (2010). The significance of protein maturation by plastidic type I signal peptidase 1 for thylakoid development in Arabidopsis chloroplasts. *Plant Physiol, 152*(3), 1297-1308.

Shiu, S. H., & Bleecker, A. B. (2001). Plant receptor-like kinase gene family: diversity, function, and signaling. *Sci STKE, 2001*(113), re22.

Simmonds, N. W. (1962). *Evolution of the bananas.* London: Longmans, Green & Co.

Simmonds, N. W. (1966). *Bananas*. London: Longmans.

Singh, A. K., Kumar, R., Pareek, A., Sopory, S. K., & Singla-Pareek, S. L. (2012). Overexpression of rice CBS domain containing protein improves salinity, oxidative, and heavy metal tolerance in transgenic tobacco. *Mol Biotechnol, 52*(3), 205-216.

Singh, Z., Jones, R. A. C., & Jones, M. G. K. (1995). Identification of Cucumber Mosaic Virus subgroup I isolates from banana plants infected by infectious chlorosis disease using RT-PCR. *Plant Disease, 79* (7), 713-716.

Smith, M. R., Willmann, M. R., Wu, G., Berardini, T. Z., Moller, B., Weijers, D., et al. (2009). Cyclophilin 40 is required for microRNA activity in Arabidopsis. *Proc Natl Acad Sci U S A, 106*(13), 5424-5429.

Sreedharan, S., Shekhawat, U. K., & Ganapathi, T. R. (2013). Transgenic banana plants overexpressing a native plasma membrane aquaporin MusaPIP1;2 display high

tolerance levels to different abiotic stresses. *Plant Biotechnol J, 11*(8), 942-952.
Srivastava, A., Raj, S. K., Haq, Q. M., Srivastava, K. M., Singh, B. P., & Sane, P. V. (1995). Association of a cucumber mosaic virus strain with mosaic disease of banana, Musa paradisiaca--an evidence using immuno/nucleic acid probe. *Indian J Exp Biol, 33*(12), 986-988.
Srivastava, S., Srivastava, A. K., Suprasanna, P., & D'Souza, S. F. (2013). Identification and profiling of arsenic stress-induced microRNAs in Brassica juncea. *J Exp Bot, 64*(1), 303-315.
Stover, R. H., & Simmonds, N. W. (1987). *Bananas (3rd edition).* New York: Longman Scientific & Technical.
Strickler, S. R., Bombarely, A., & Mueller, L. A. (2012). Designing a transcriptome next-generation sequencing project for a nonmodel plant species. *Am J Bot, 99*(2), 257-266.
Sun, G. (2012). MicroRNAs and their diverse functions in plants. *Plant Mol Biol, 80*(1), 17-36.
Sun, X. L., Yu, Q. Y., Tang, L. L., Ji, W., Bai, X., Cai, H., et al. (2013). GsSRK, a G-type lectin S-receptor-like serine/threonine protein kinase, is a positive regulator of plant tolerance to salt stress. *J Plant Physiol, 170*(5), 505-515.
Sunkar, R., & Zhu, J. K. (2004). Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. *Plant Cell, 16*(8), 2001-2019.
Sunkar, R., Girke, T., Jain, P. K., & Zhu, J. K. (2005). Cloning and characterization of microRNAs from rice. *Plant Cell, 17*(5), 1397-1411.
Sunkar, R., Kapoor, A., & Zhu, J. K. (2006). Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in Arabidopsis is mediated by downregulation of miR398 and important for oxidative stress tolerance. *Plant Cell, 18*(8), 2051-2065.
Sunkar, R., Li, Y. F., & Jagadeeswaran, G. (2012). Functions of microRNAs in plant stress responses. *Trends Plant Sci, 17*(4), 196-203.
Sunkar, R., Zhou, X., Zheng, Y., Zhang, W., & Zhu, J. K. (2008). Identification of novel and candidate miRNAs in rice by high throughput sequencing. *BMC Plant Biol, 8*, 25.
Suzuki, T. (2007). Cytoplasmic peptide:N-glycanase and catabolic pathway for free N-glycans in the cytosol. *Semin Cell Dev Biol, 18*(6), 762-769.
Svennerstam, H., Ganeteg, U., & Nasholm, T. (2008). Root uptake of cationic amino acids by Arabidopsis depends on functional expression of amino acid permease 5. *New Phytol, 180*(3), 620-630.
Tabara, H., Yigit, E., Siomi, H., & Mello, C. C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in C. elegans. *Cell, 109*(7), 861-871.
Takahashi, R., Bashir, K., Ishimaru, Y., Nishizawa, N. K., & Nakanishi, H. (2012). The role of heavy-metal ATPases, HMAs, in zinc and cadmium transport in rice. *Plant Signal Behav, 7*(12), 1605-1607.
Takashima, S., Abe, T., Yoshida, S., Kawahigashi, H., Saito, T., Tsuji, S., et al. (2006). Analysis of sialyltransferase-like proteins from Oryza sativa. *J Biochem, 139*(2), 279-287.
Tameling, W. I., Elzinga, S. D., Darmin, P. S., Vossen, J. H., Takken, F. L., Haring, M. A., et al. (2002). The tomato R gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity. *Plant Cell, 14*(11), 2929-2939.
Tang, M., Liu, X., Deng, H., & Shen, S. (2011). Over-expression of JcDREB, a putative AP2/EREBP domain-containing transcription factor gene in woody biodiesel plant Jatropha curcas, enhances salt and freezing tolerance in transgenic Arabidopsis thaliana. *Plant Sci, 181*(6), 623-631.

Tang, M., Sun, J., Liu, Y., Chen, F., & Shen, S. (2007). Isolation and functional characterization of the JcERF gene, a putative AP2/EREBP domain-containing transcription factor, in the woody oil plant Jatropha curcas. *Plant Mol Biol, 63*(3), 419-428.

Tani, T., Sobajima, H., Okada, K., Chujo, T., Arimura, S., Tsutsumi, N., et al. (2008). Identification of the OsOPR7 gene encoding 12-oxophytodienoate reductase involved in the biosynthesis of jasmonic acid in rice. *Planta, 227*(3), 517-526.

Tan Yew Seong. (2010). Computational identification of micro RNA in Musa sp. EST sequences. M. Biotech. Dissertation, University of Malaya, Kuala Lumpur, Malaysia.

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., et al. (2003). The COG database: an updated version includes eukaryotes. *BMC Bioinformatics, 4*, 41.

Tester, M., & Davenport, R. (2003). Na^+ tolerance and Na^+ transport in higher plants. *Ann Bot, 91*(5), 503-527.

The Global Musa Genomics Consortium. (2002). *A strategy for the Global Musa Genomics Consortium. Report of a meeting held in Arlington, USA, 17-20 July 2001*. Montpellier: The International Network for the Improvement of Banana and Plantain.

Thiebaut, F., Rojas, C. A., Almeida, K. L., Grativol, C., Domiciano, G. C., Lamb, C. R., et al. (2011). Regulation of miR319 during cold stress in sugarcane. *Plant Cell Environ., 35*(3), 502-512.

Tian, H., Jia, Y., Niu, T., Yu, Q., & Ding, Z. (2014). The key players of the primary root growth and development also function in lateral roots in Arabidopsis. *Plant Cell Rep, 33*(5), 745-753.

Tian, Y., Zhang, H., Pan, X., Chen, X., Zhang, Z., Lu, X., et al. (2011). Overexpression of ethylene response factor TERF2 confers cold tolerance in rice seedlings. *Transgenic Res, 20*(4), 857-866.

Tran, L. T., Taylor, J. S., & Constabel, C. P. (2012). The polyphenol oxidase gene family in land plants: Lineage-specific duplication and expansion. *BMC Genomics, 13*, 395.

Trindade, I., Capitao, C., Dalmay, T., Fevereiro, M. P., & Santos, D. M. (2010). miR398 and miR408 are up-regulated in response to water deficit in Medicago truncatula. *Planta, 231*(3), 705-716.

Trivedi, D. K., Ansari, M. W., Bhavesh, N. S., Johri, A. K., & Tuteja, N. (2014). Response of tobacco T2 transgenic matured plant to potential tolerance to salinity stress. *Plant Signal Behav, 9*(1).

Turan, S., Cornish, K., & Kumar, S. (2012). Salinity tolerance in plants: Breeding and genetic engineering. *Aust J Crop Sci 6*(9), 1337-1348.

Unver, T., Namuth-Covert, D. M., & Budak, H. (2009). Review of current methodological approaches for characterizing microRNAs in plants. *Int J Plant Genomics, 2009*, 262463.

Unver, T., Parmaksiz, I., & Dundar, E. (2010). Identification of conserved micro-RNAs and their target transcripts in opium poppy (Papaver somniferum L.). *Plant Cell Rep, 29*(7), 757-769.

Vaganan, M. M., Ravi, I., Nandakumar, A., Sarumathi, S., Sundararaju, P., & Mustaffa, M. M. (2014). Phenylpropanoid enzymes, phenolic polymers and metabolites as chemical defenses to infection of Pratylenchus coffeae in roots of resistant and susceptible bananas (Musa spp.). *Indian J Exp Biol, 52*(3), 252-260.

Valdes-Lopez, O., Arenas-Huertero, C., Ramirez, M., Girard, L., Sanchez, F., Vance, C. P., et al. (2008). Essential role of MYB transcription factor: PvPHR1 and microRNA: PvmiR399 in phosphorus-deficiency signalling in common bean

roots. *Plant Cell Environ, 31*(12), 1834-1843.
Valdes-Lopez, O., Yang, S. S., Aparicio-Fabre, R., Graham, P. H., Reyes, J. L., Vance, C. P., et al. (2010). MicroRNA expression profile in common bean (Phaseolus vulgaris) under nutrient deficiency stresses and manganese toxicity. *New Phytol, 187*(3), 805-818.
Valmayor, R. V., Jamaluddin, S. H., Silayoi, B., Kusumo, S., Danh, L. D., Pascua, O. C., et al. (2000). Banana cultivar names and synonyms in Southeast Asia. Los Baños, Laguna: International Network for the Improvement of Banana and plantain.
Valoczi, A., Hornyik, C., Varga, N., Burgyan, J., Kauppinen, S., & Havelda, Z. (2004). Sensitive and specific detection of microRNAs by northern blot analysis using LNA-modified oligonucleotide probes. *Nucleic Acids Res, 32*(22), e175.
van den Burg, H. A., Tsitsigiannis, D. I., Rowland, O., Lo, J., Rallapalli, G., Maclean, D., et al. (2008). The F-box protein ACRE189/ACIF1 regulates cell death and defense responses activated during pathogen recognition in tobacco and tomato. *Plant Cell, 20*(3), 697-719.
Van Verk, M. C., Hickman, R., Pieterse, C. M., & Van Wees, S. C. (2013). RNA-Seq: revelation of the messengers. *Trends Plant Sci, 18*(4), 175-179.
Vaucheret, H., Vazquez, F., Crete, P., & Bartel, D. P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev, 18*(10), 1187-1197.
Vega-Sánchez, M. E., Gowda, M., & Wang, G.-L. (2007). Tag-based approaches for deep transcriptome analysis in plants. *Plant Sci, 173*, 317-380.
Velculescu, V. E., Vogelstein, B., & Kinzler, K. W. (2000). Analysing uncharted transcriptomes with SAGE. *Trends Genet, 16*(10), 423-425.
Vidal, E. A., Araus, V., Lu, C., Parry, G., Green, P. J., Coruzzi, G. M., et al. (2010). Nitrate-responsive miR393/AFB3 regulatory module controls root system architecture in Arabidopsis thaliana. *Proc Natl Acad Sci U S A, 107*(9), 4477-4482.
Vilarinhos, A. D., Piffanelli, P., Lagoda, P., Thibivilliers, S., Sabau, X., Carreel, F., et al. (2003). Construction and characterization of a bacterial artificial chromosome library of banana (Musa acuminata Colla). *Theor Appl Genet, 106*(6), 1102-1106.
Vinocur, B., & Altman, A. (2005). Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. *Curr Opin Biotechnol, 16*(2), 123-132.
Wallace, I. S., Choi, W. G., & Roberts, D. M. (2006). The structure, function and regulation of the nodulin 26-like intrinsic protein family of plant aquaglyceroporins. *Biochim Biophys Acta, 1758*(8), 1165-1175.
Wan, D., Li, R., Zou, B., Zhang, X., Cong, J., Wang, R., et al. (2012). Calmodulin-binding protein CBP60g is a positive regulator of both disease resistance and drought tolerance in Arabidopsis. *Plant Cell Rep, 31*(7), 1269-1281.
Wang, B., Sun, Y. F., Song, N., Wang, X. J., Feng, H., Huang, L. L., et al. (2013a). Identification of UV-B-induced microRNAs in wheat. *Genet Mol Res, 12*(4), 4213-4221.
Wang, J. F., Zhou, H., Chen, Y. Q., Luo, Q. J., & Qu, L. H. (2004a). Identification of 20 microRNAs from Oryza sativa. *Nucleic Acids Res, 32*(5), 1688-1695.
Wang, J. W., Czech, B., & Weigel, D. (2009a). miR156-regulated SPL transcription factors define an endogenous flowering pathway in Arabidopsis thaliana. *Cell, 138*(4), 738-749.
Wang, J. W., Park, M. Y., Wang, L. J., Koo, Y., Chen, X. Y., Weigel, D., et al. (2011a). miRNA control of vegetative phase change in trees. *PLoS Genet, 7*(2), e1002012.
Wang, J. W., Wang, L. J., Mao, Y. B., Cai, W. J., Xue, H. W., & Chen, X. Y. (2005). Control of root cap formation by MicroRNA-targeted auxin response factors in

Arabidopsis. *Plant Cell, 17*(8), 2204-2216.
Wang, J., Ding, B., Guo, Y., Li, M., Chen, S., Huang, G., et al. (2014a). Overexpression of a wheat phospholipase D gene, TaPLDalpha, enhances tolerance to drought and osmotic stress in Arabidopsis thaliana. *Planta*.
Wang, L., Feng, Z., Wang, X., & Zhang, X. (2010a). DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics, 26*(1), 136-138.
Wang, L., Gu, X., Xu, D., Wang, W., Wang, H., Zeng, M., et al. (2011b). miR396-targeted AtGRF transcription factors are required for coordination of cell division and differentiation during leaf development in Arabidopsis. *J Exp Bot, 62*(2), 761-773.
Wang, L., Gu, X., Xu, D., Wang, W., Wang, H., Zeng, M., et al. (2011c). miR396-targeted AtGRF transcription factors are required for coordination of cell division and differentiation during leaf development in Arabidopsis. *J Exp Bot, 62*(2), 761-773.
Wang, L., Li, P., & Brutnell, T. P. (2010b). Exploring plant transcriptomes using ultra high-throughput sequencing. *Brief Funct Genomics, 9*(2), 118-128.
Wang, L., Liu, H., Li, D., & Chen, H. (2011d). Identification and characterization of maize microRNAs involved in the very early stage of seed germination. *BMC Genomics, 12*, 154.
Wang, L., Mai, Y. X., Zhang, Y. C., Luo, Q., & Yang, H. Q. (2010c). MicroRNA171c-targeted SCL6-II, SCL6-III, and SCL6-IV genes regulate shoot branching in Arabidopsis. *Mol Plant, 3*(5), 794-806.
Wang, L., Xu, Y., Zhang, C., Ma, Q., Joo, S. H., Kim, S. K., et al. (2008). OsLIC, a Novel CCCH-Type Zinc Finger Protein with Transcription Activation, Mediates Rice Architecture via Brassinosteroids Signaling. *PLoS One, 3*(10), e3521.
Wang, M., Wang, Q., & Zhang, B. (2013b). Response of miRNAs and their targets to salt and drought stresses in cotton (Gossypium hirsutum L.). *Gene, 530*(1), 26-32.
Wang, P., Li, X. Z., Cui, H. R., Feng, Y. G., & Wang, X. Y. (2014b). Identification and functional analysis of a novel parvulin-type peptidyl-prolyl isomerase from Gossypium hirsutum. *Plant Physiol Biochem, 76*, 58-66.
Wang, T., Chen, L., Zhao, M., Tian, Q., & Zhang, W. H. (2011e). Identification of drought-responsive microRNAs in Medicago truncatula by genome-wide high-throughput sequencing. *BMC Genomics, 12*, 367.
Wang, W., Vinocur, B., & Altman, A. (2003). Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. *Planta, 218*(1), 1-14.
Wang, X. J., Reyes, J. L., Chua, N. H., & Gaasterland, T. (2004b). Prediction and identification of Arabidopsis thaliana microRNAs and their mRNA targets. *Genome Biol, 5*(9), R65.
Wang, X. W., Luan, J. B., Li, J. M., Bao, Y. Y., Zhang, C. X., & Liu, S. S. (2010d). De novo characterization of a whitefly transcriptome and analysis of its gene expression during development. *BMC Genomics, 11*, 400.
Wang, X., Wang, Y., Pineros, M. A., Wang, Z., Wang, W., Li, C., et al. (2014c). Phosphate transporters OsPHT1;9 and OsPHT1;10 are involved in phosphate uptake in rice. *Plant Cell Environ, 37*(5), 1159-1170.
Wang, Y., Sun, F., Cao, H., Peng, H., Ni, Z., Sun, Q., et al. (2012a). TamiR159 directed wheat TaGAMYB cleavage and its involvement in anther development and heat response. *PLoS One, 7*(11), e48445.
Wang, Z., Fang, B., Chen, J., Zhang, X., Luo, Z., Huang, L., et al. (2010e). De novo assembly and characterization of root transcriptome using Illumina paired-end

sequencing and development of cSSR markers in sweet potato (Ipomoea batatas). *BMC Genomics, 11*, 726.

Wang, Z., Gerstein, M., & Snyder, M. (2009b). RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet, 10*(1), 57-63.

Wang, Z., Zhang, J., Jia, C., Liu, J., Li, Y., Yin, X., et al. (2012b). De novo characterization of the banana root transcriptome and analysis of gene expression under Fusarium oxysporum f. sp. Cubense tropical race 4 infection. *BMC Genomics, 13*, 650.

Ward, J. A., Ponnala, L., & Weber, C. A. (2012). Strategies for transcriptome analysis in nonmodel plants. *Am J Bot, 99*(2), 267-276.

Waters, B. M., McInturf, S. A., & Stein, R. J. (2012). Rosette iron deficiency transcript and microRNA profiling reveals links between copper and iron homeostasis in Arabidopsis thaliana. *J Exp Bot, 63*(16), 5903-5918.

Wawer, I., Bucholc, M., Astier, J., Anielska-Mazur, A., Dahan, J., Kulik, A., et al. (2010). Regulation of Nicotiana tabacum osmotic stress-activated protein kinase and its cellular partner GAPDH by nitric oxide in response to salinity. *Biochem J, 429*(1), 73-83.

Weber, H., & Hellmann, H. (2009). Arabidopsis thaliana BTB/ POZ-MATH proteins interact with members of the ERF/AP2 transcription factor family. *FEBS J, 276*(22), 6624-6635.

Wei, J., Tirajoh, A., Effendy, J., & Plant, A. L. (2000). Characterization of salt-induced changes in gene expression in tomato (Lycopersicon esculentum) roots and the role played by abscisic acid. *Plant Sci, 159*(1), 135-148.

Wei, L. Q., Yan, L. F., & Wang, T. (2011). Deep sequencing on genome-wide scale reveals the unique composition and expression patterns of microRNAs in developing pollen of Oryza sativa. *Genome Biol, 12*(6), R53.

Wei, M., Wei, H., Wu, M., Song, M., Zhang, J., Yu, J., et al. (2013). Comparative expression profiling of miRNA during anther development in genetic male sterile and wild type cotton. *BMC Plant Biol, 13*, 66.

Wightman, B., Ha, I., & Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell, 75*(5), 855-862.

Wilhelm, B. T., & Landry, J. R. (2009). RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing. *Methods, 48*(3), 249-257.

Williams, L., Grigg, S. P., Xie, M., Christensen, S., & Fletcher, J. C. (2005). Regulation of Arabidopsis shoot apical meristem and lateral organ formation by microRNA miR166g and its AtHD-ZIP target genes. *Development, 132*(16), 3657-3668.

Wong, C. E., Zhao, Y. T., Wang, X. J., Croft, L., Wang, Z. H., Haerizadeh, F., et al. (2011). MicroRNAs in the shoot apical meristem of soybean. *J Exp Bot, 62*(8), 2495-2506.

Wu, G., & Poethig, R. S. (2006). Temporal regulation of shoot development in Arabidopsis thaliana by miR156 and its target SPL3. *Development, 133*(18), 3539-3547.

Wu, J., Wang, Y., Kim, S. T., Kim, S. G., & Kang, K. Y. (2013). Characterization of a newly identified rice chitinase-like protein (OsCLP) homologous to xylanase inhibitor. *BMC Biotechnol, 13*, 4.

Wu, M. F., Tian, Q., & Reed, J. W. (2006). Arabidopsis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. *Development, 133*(21), 4211-4218.

Xie, F. L., Huang, S. Q., Guo, K., Xiang, A. L., Zhu, Y. Y., Nie, L., et al. (2007). Computational identification of novel microRNAs and targets in Brassica napus.

FEBS Lett, 581(7), 1464-1474.
Xie, F., Stewart, C. N., Jr., Taki, F. A., He, Q., Liu, H., & Zhang, B. (2013). High-throughput deep sequencing shows that microRNAs play important roles in switchgrass responses to drought and salinity stress. *Plant Biotechnol J*.
Xie, K., Wu, C., & Xiong, L. (2006). Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice. *Plant Physiol, 142*(1), 280-293.
Xie, Z., Kasschau, K. D., & Carrington, J. C. (2003). Negative feedback regulation of Dicer-Like1 in Arabidopsis by microRNA-guided mRNA degradation. *Curr Biol, 13*(9), 784-789.
Xin, M., Wang, Y., Yao, Y., Xie, C., Peng, H., Ni, Z., et al. (2010). Diverse set of microRNAs are responsive to powdery mildew infection and heat stress in wheat (Triticum aestivum L.). *BMC Plant Biol, 10*, 123.
Xing, S., Salinas, M., Hohmann, S., Berndtgen, R., & Huijser, P. (2010). miR156-targeted and nontargeted SBP-box transcription factors act in concert to secure male fertility in Arabidopsis. *Plant Cell, 22*(12), 3935-3950.
Xu, L., Wang, Y., Xu, Y., Wang, L., Zhai, L., Zhu, X., et al. (2013a). Identification and characterization of novel and conserved microRNAs in radish (Raphanus sativus L.) using high-throughput sequencing. *Plant Sci, 201-202*, 108-114.
Xu, L., Wang, Y., Zhai, L., Xu, Y., Wang, L., Zhu, X., et al. (2013b). Genome-wide identification and characterization of cadmium-responsive microRNAs and their target genes in radish (Raphanus sativus L.) roots. *J Exp Bot, 64*(14), 4271-4287.
Xu, M., Xie, W., & Huang, M. (2012). Overexpression of PeRHD3 alters the root architecture in Populus. *Biochem Biophys Res Commun, 424*(2), 239-244.
Xu, Q., Liu, Y., Zhu, A., Wu, X., Ye, J., Yu, K., et al. (2010). Discovery and comparative profiling of microRNAs in a sweet orange red-flesh mutant and its wild type. *BMC Genomics, 11*, 246.
Xu, R., Zhao, H., Dinkins, R. D., Cheng, X., Carberry, G., & Li, Q. Q. (2006). The 73 kD subunit of the cleavage and polyadenylation specificity factor (CPSF) complex affects reproductive development in Arabidopsis. *Plant Mol Biol, 61*(4-5), 799-815.
Xue, G. P., & Loveridge, C. W. (2004). HvDRF1 is involved in abscisic acid-mediated gene regulation in barley and produces two forms of AP2 transcriptional activators, interacting preferably with a CT-rich element. *Plant J, 37*(3), 326-339.
Xue, L. J., Zhang, J. J., & Xue, H. W. (2009). Characterization and expression profiles of miRNAs in rice seeds. *Nucleic Acids Res, 37*(3), 916-930.
Xue, T., Wang, D., Zhang, S., Ehlting, J., Ni, F., Jakab, S., et al. (2008). Genome-wide and expression analysis of protein phosphatase 2C in rice and Arabidopsis. *BMC Genomics, 9*, 550.
Yadav, D. K., & Tuteja, N. (2011). Rice G-protein coupled receptor (GPCR): in silico analysis and transcription regulation under abiotic stress. *Plant Signal Behav, 6*(8), 1079-1086.
Yamagata, Y., Kato, M., Odawara, K., Tokuno, Y., Nakashima, Y., Matsushima, N., et al. (1996). Three-dimensional structure of a DNA repair enzyme, 3-methyladenine DNA glycosylase II, from Escherichia coli. *Cell, 86*(2), 311-319.
Yamamoto, M., Shitsukawa, N., Yamada, M., Kato, K., Takumi, S., Kawaura, K., et al. (2013). Identification of a novel homolog for a calmodulin-binding protein that is upregulated in alloplasmic wheat showing pistillody. *Planta, 237*(4), 1001-1013.
Yamasaki, H., Abdel-Ghany, S. E., Cohu, C. M., Kobayashi, Y., Shikanai, T., & Pilon, M. (2007). Regulation of copper homeostasis by micro-RNA in Arabidopsis. *J Biol Chem, 282*(22), 16369-16378.

Yan, Z., Hossain, M. S., Wang, J., Valdes-Lopez, O., Liang, Y., Libault, M., et al. (2013). miR172 regulates soybean nodulation. *Mol Plant Microbe Interact, 26*(12), 1371-1377.

Yang, J., Liu, X., Xu, B., Zhao, N., Yang, X., & Zhang, M. (2013). Identification of miRNAs and their targets using high-throughput sequencing and degradome analysis in cytoplasmic male-sterile and its maintainer fertile lines of Brassica juncea. *BMC Genomics, 14*, 9.

Yang, L., Conway, S. R., & Poethig, R. S. (2011). Vegetative phase change is mediated by a leaf-derived signal that represses the transcription of miR156. *Development, 138*(2), 245-249.

Yang, Y., Sun, X., Yang, S., & Li, X. (2014). Molecular cloning and characterization of a novel SK3-type dehydrin gene from Stipa purpurea. *Biochem Biophys Res Commun, 448*(2), 145-150.

Yano-Melo, A. M., Saggin, O. J., & Maia, L. C. (2003). Tolerance of mycorrhized banana (*Musa* sp. cv. Pacovan) plantlets to saline stress. *Agric Ecosyst Environ, 95*, 343-348.

Yao, X., Wang, H., Li, H., Yuan, Z., Li, F., Yang, L., et al. (2009). Two types of cis-acting elements control the abaxial epidermis-specific transcription of the MIR165a and MIR166a genes. *FEBS Lett, 583*(22), 3711-3717.

Yao, Y., Guo, G., Ni, Z., Sunkar, R., Du, J., Zhu, J. K., et al. (2007). Cloning and characterization of microRNAs from wheat (Triticum aestivum L.). *Genome Biol, 8*(6), R96.

Yi, F., Xie, S., Liu, Y., Qi, X., & Yu, J. (2013). Genome-wide characterization of microRNA in foxtail millet (Setaria italica). *BMC Plant Biol, 13*, 212.

Yin, D., Ni, D., Song, L., & Zhang, Z. (2013a). Isolation of an alcohol dehydrogenase cDNA from and characterization of its expression in chrysanthemum under waterlogging. *Plant Sci, 212*, 48-54.

Yin, X., Wang, J., Cheng, H., Wang, X., & Yu, D. (2013b). Detection and evolutionary analysis of soybean miRNAs responsive to soybean mosaic virus. *Planta, 237*(5), 1213-1225.

Yin, Z., Li, Y., Yu, J., Liu, Y., Li, C., Han, X., et al. (2012). Difference in miRNA expression profiles between two cotton cultivars with distinct salt sensitivity. *Mol Biol Rep, 39*(4), 4961-4970.

Yoo, H., Widhalm, J. R., Qian, Y., Maeda, H., Cooper, B. R., Jannasch, A. S., et al. (2013). An alternative pathway contributes to phenylalanine biosynthesis in plants via a cytosolic tyrosine:phenylpyruvate aminotransferase. *Nat Commun, 4*, 2833.

Yoo, K. S., Ok, S. H., Jeong, B. C., Jung, K. W., Cui, M. H., Hyoung, S., et al. (2011). Single cystathionine beta-synthase domain-containing proteins modulate development by regulating the thioredoxin system in Arabidopsis. *Plant Cell, 23*(10), 3577-3594.

Yoon, E. K., Yang, J. H., Lim, J., Kim, S. H., Kim, S. K., & Lee, W. S. (2010). Auxin regulation of the microRNA390-dependent transacting small interfering RNA pathway in Arabidopsis lateral root development. *Nucleic Acids Res, 38*(4), 1382-1391.

Yu, H., Song, C., Jia, Q., Wang, C., Li, F., Nicholas, K. K., et al. (2011). Computational identification of microRNAs in apple expressed sequence tags and validation of their precise sequences by miR-RACE. *Physiol Plant, 141*(1), 56-70.

Yu, X., Wang, H., Lu, Y., de Ruiter, M., Cariaso, M., Prins, M., et al. (2012a). Identification of conserved and novel microRNAs that are responsive to heat stress in Brassica rapa. *J Exp Bot, 63*(2), 1025-1038.

Yu, Y., Zhang, H., Li, W., Mu, C., Zhang, F., Wang, L., et al. (2012b). Genome-wide

analysis and environmental response profiling of the FK506-binding protein gene family in maize (Zea mays L.). *Gene, 498*(2), 212-222.
Zeng, H. Q., Zhu, Y. Y., Huang, S. Q., & Yang, Z. M. (2010). Analysis of phosphorus-deficient responsive miRNAs and cis-elements from soybean (Glycine max L.). *J Plant Physiol, 167*(15), 1289-1297.
Zeng, Q. Y., Yang, C. Y., Ma, Q. B., Li, X. P., Dong, W. W., & Nian, H. (2012). Identification of wild soybean miRNAs and their target genes responsive to aluminum stress. *BMC Plant Biol, 12*, 182.
Zhang, B. H., Pan, X. P., Wang, Q. L., Cobb, G. P., & Anderson, T. A. (2005a). Identification and characterization of new plant microRNAs using EST analysis. *Cell Res, 15*(5), 336-360.
Zhang, B., Pan, X., & Anderson, T. A. (2006a). Identification of 188 conserved maize microRNAs and their targets. *FEBS Lett, 580*(15), 3753-3762.
Zhang, B., Pan, X., Cannon, C. H., Cobb, G. P., & Anderson, T. A. (2006b). Conservation and divergence of plant microRNA genes. *Plant J, 46*(2), 243-259.
Zhang, B., Pan, X., Cobb, G. P., & Anderson, T. A. (2006c). Plant microRNA: a small regulatory molecule with big impact. *Dev Biol, 289*(1), 3-16.
Zhang, H., Huang, Z., Xie, B., Chen, Q., Tian, X., Zhang, X., et al. (2004). The ethylene-, jasmonate-, abscisic acid- and NaCl-responsive tomato transcription factor JERF1 modulates expression of GCC box-containing genes and salt tolerance in tobacco. *Planta, 220*(2), 262-270.
Zhang, H., Liu, W., Wan, L., Li, F., Dai, L., Li, D., et al. (2010a). Functional analyses of ethylene response factor JERF3 with the aim of improving tolerance to drought and osmotic stress in transgenic rice. *Transgenic Res, 19*(5), 809-818.
Zhang, J. P., Liu, T. S., Zheng, J., Jin, Z., Zhu, Y., Guo, J. F., et al. (2007). Cloning and characterization of a putative 12-oxophytodienoic acid reductase cDNA induced by osmotic stress in roots of foxtail millet. *DNA Seq, 18*(2), 138-144.
Zhang, J. W., Xu, L., Wu, Y. R., Chen, X. A., Liu, Y., Zhu, S. H., et al. (2012a). OsGLU3, a putative membrane-bound endo-1,4-beta-glucanase, is required for root cell elongation and division in rice (Oryza sativa L.). *Mol Plant, 5*(1), 176-186.
Zhang, J. Y., Broeckling, C. D., Blancaflor, E. B., Sledge, M. K., Sumner, L. W., & Wang, Z. Y. (2005b). Overexpression of WXP1, a putative Medicago truncatula AP2 domain-containing transcription factor gene, increases cuticular wax accumulation and enhances drought tolerance in transgenic alfalfa (Medicago sativa). *Plant J, 42*(5), 689-707.
Zhang, J. Z., Ai, X. Y., Guo, W. W., Peng, S. A., Deng, X. X., & Hu, C. G. (2012b). Identification of miRNAs and their target genes using deep sequencing and degradome analysis in trifoliate orange [Poncirus trifoliata L. Raf] [corrected]. *Mol Biotechnol, 51*(1), 44-57.
Zhang, J., Xu, Y., Huan, Q., & Chong, K. (2009a). Deep sequencing of Brachypodium small RNAs at the global genome level identifies microRNAs involved in cold stress response. *BMC Genomics, 10*, 449.
Zhang, J., Zeng, R., Chen, J., Liu, X., & Liao, Q. (2008a). Identification of conserved microRNAs and their targets from Solanum lycopersicum Mill. *Gene, 423*(1), 1-7.
Zhang, L. W., Song, J. B., Shu, X. X., Zhang, Y., & Yang, Z. M. (2013a). miR395 is involved in detoxification of cadmium in Brassica napus. *J Hazard Mater, 250-251*, 204-211.
Zhang, L., Chia, J. M., Kumari, S., Stein, J. C., Liu, Z., Narechania, A., et al. (2009b). A genome-wide characterization of microRNA genes in maize. *PLoS Genet, 5*(11), e1000716.

Zhang, Q., Zhao, C., Li, M., Sun, W., Liu, Y., Xia, H., et al. (2013b). Genome-wide identification of Thellungiella salsuginea microRNAs with putative roles in the salt stress response. *BMC Plant Biol, 13*, 180.

Zhang, W., Gao, S., Zhou, X., Chellappan, P., Chen, Z., Zhang, X., et al. (2011a). Bacteria-responsive microRNAs regulate plant innate immunity by modulating plant hormone networks. *Plant Mol Biol, 75*(1-2), 93-105.

Zhang, X., Zou, Z., Gong, P., Zhang, J., Ziaf, K., Li, H., et al. (2011b). Over-expression of microRNA169 confers enhanced drought tolerance to tomato. *Biotechnol Lett, 33*(2), 403-409.

Zhang, Y. C., Yu, Y., Wang, C. Y., Li, Z. Y., Liu, Q., Xu, J., et al. (2013c). Overexpression of microRNA OsmiR397 improves rice yield by increasing grain size and promoting panicle branching. *Nat Biotechnol, 31*(9), 848-852.

Zhang, Y., Xiao, W., Luo, L., Pang, J., Rong, W., & He, C. (2012c). Downregulation of OsPK1, a cytosolic pyruvate kinase, by T-DNA insertion causes dwarfism and panicle enclosure in rice. *Planta, 235*(1), 25-38.

Zhang, Z., Li, F., Li, D., Zhang, H., & Huang, R. (2010b). Expression of ethylene response factor JERF1 in rice improves tolerance to drought. *Planta, 232*(3), 765-774.

Zhang, Z., Wei, L., Zou, X., Tao, Y., Liu, Z., & Zheng, Y. (2008b). Submergence-responsive MicroRNAs are potentially involved in the regulation of morphological and metabolic adaptations in maize root cells. *Ann Bot, 102*(4), 509-519.

Zhang, Z., Yu, J., Li, D., Liu, F., Zhou, X., Wang, T., et al. (2010c). PMRD: plant microRNA database. *Nucleic Acids Res, 38*(Database issue), D806-813.

Zhao, B., Ge, L., Liang, R., Li, W., Ruan, K., Lin, H., et al. (2009). Members of miR-169 family are induced by high salinity and transiently inhibit the NF-YA transcription factor. *BMC Mol Biol, 10*, 29.

Zhao, B., Liang, R., Ge, L., Li, W., Xiao, H., Lin, H., et al. (2007a). Identification of drought-induced microRNAs in rice. *Biochem Biophys Res Commun, 354*(2), 585-590.

Zhao, C. Z., Xia, H., Frazier, T. P., Yao, Y. Y., Bi, Y. P., Li, A. Q., et al. (2010). Deep sequencing identifies novel and conserved microRNAs in peanuts (Arachis hypogaea L.). *BMC Plant Biol, 10*, 3.

Zhao, J. P., Jiang, X. L., Zhang, B. Y., & Su, X. H. (2012). Involvement of microRNA-mediated gene expression regulation in the pathological development of stem canker disease in Populus trichocarpa. *PLoS One, 7*(9), e44968.

Zhao, L., Kim, Y., Dinh, T. T., & Chen, X. (2007b). miR172 regulates stem cell fate and defines the inner boundary of APETALA3 and PISTILLATA expression domain in Arabidopsis floral meristems. *Plant J, 51*(5), 840-849.

Zhao, M., Ding, H., Zhu, J. K., Zhang, F., & Li, W. X. (2011). Involvement of miR169 in the nitrogen-starvation responses in Arabidopsis. *New Phytol, 190*(4), 906-915.

Zhao, Q., Zhang, H., Wang, T., Chen, S., & Dai, S. (2013a). Proteomics-based investigation of salt-responsive mechanisms in plant roots. *J Proteomics, 82*, 230-253.

Zhao, S., Tuan, P. A., Li, X., Kim, Y. B., Kim, H., Park, C. G., et al. (2013b). Identification of phenylpropanoid biosynthetic genes and phenylpropanoid accumulation by transcriptome analysis of Lycium chinense. *BMC Genomics, 14*, 802.

Zhao, Y., Xu, Z., Mo, Q., Zou, C., Li, W., Xu, Y., et al. (2013c). Combined small RNA and degradome sequencing reveals novel miRNAs and their targets in response to low nitrate availability in maize. *Ann Bot, 112*(3), 633-642.

Zhou, G. K., Kubo, M., Zhong, R., Demura, T., & Ye, Z. H. (2007a). Overexpression of miR165 affects apical meristem formation, organ polarity establishment and vascular development in Arabidopsis. *Plant Cell Physiol, 48*(3), 391-404.

Zhou, J., Liu, M., Jiang, J., Qiao, G., Lin, S., Li, H., et al. (2012a). Expression profile of miRNAs in Populus cathayana L. and Salix matsudana Koidz under salt stress. *Mol Biol Rep, 39*(9), 8645-8654.

Zhou, L., Liu, Y., Liu, Z., Kong, D., Duan, M., & Luo, L. (2010). Genome-wide identification and analysis of drought-responsive microRNAs in Oryza sativa. *J Exp Bot, 61*(15), 4157-4168.

Zhou, M., Li, D., Li, Z., Hu, Q., Yang, C., Zhu, L., et al. (2013a). Constitutive expression of a miR319 gene alters plant development and enhances salt and drought tolerance in transgenic creeping bentgrass. *Plant Physiol, 161*(3), 1375-1391.

Zhou, X., Oshlack, A., & Robinson, M. D. (2013b). miRNA-Seq normalization comparisons need improvement. *RNA, 19*(6), 733-734.

Zhou, X., Wang, G., & Zhang, W. (2007b). UV-B responsive microRNA genes in Arabidopsis thaliana. *Mol Syst Biol, 3*, 103.

Zhou, X., Wang, G., Sutoh, K., Zhu, J. K., & Zhang, W. (2008). Identification of cold-inducible microRNAs in plants by transcriptome analysis. *Biochim Biophys Acta, 1779*(11), 780-788.

Zhou, Y. P., Duan, J., Fujibe, T., Yamamoto, K. T., & Tian, C. E. (2012b). AtIQM1, a novel calmodulin-binding protein, is involved in stomatal movement in Arabidopsis. *Plant Mol Biol, 79*(4-5), 333-346.

Zhou, Z. S., Song, J. B., & Yang, Z. M. (2012c). Genome-wide identification of Brassica napus microRNAs and their targets in response to cadmium. *J Exp Bot, 63*(12), 4597-4613.

Zhou, Z. S., Zeng, H. Q., Liu, Z. P., & Yang, Z. M. (2012d). Genome-wide identification of Medicago truncatula microRNAs and their targets reveals their differential regulation by heavy metal. *Plant Cell Environ, 35*(1), 86-99.

Zhu, C., Ding, Y., & Liu, H. (2011). MiR398 and plant stress responses. *Physiol Plant, 143*(1), 1-9.

Zhu, H., Xia, R., Zhao, B., An, Y. Q., Dardick, C. D., Callahan, A. M., et al. (2012). Unique expression, processing regulation, and regulatory network of peach (Prunus persica) miRNAs. *BMC Plant Biol, 12*, 149.

Zhu, Q. H., Spriggs, A., Matthew, L., Fan, L., Kennedy, G., Gubler, F., et al. (2008). A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. *Genome Res, 18*(9), 1456-1465.

Zhu, Q. H., Upadhyaya, N. M., Gubler, F., & Helliwell, C. A. (2009). Over-expression of miR172 causes loss of spikelet determinacy and floral organ abnormalities in rice (Oryza sativa). *BMC Plant Biol, 9*, 149.

Zou, D., & Chen, X. (2013). Sequencing, de novo assembly, annotation and SSR and SNP detection of sabaigrass (Eulaliopsis binata) transcriptome. *Genomics, 102*(1), 57-62.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res 31*, 3406-3415.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

Publications:

Wan Sin Lee, Ranganath Gudimella, Gwo Rong Wong, Martti Tapani Tammi, Norzulaani Khalid, Jennifer Ann Harikrishna (2015). Transcripts and microRNAs responding to salt stress in Musa acuminata Colla (AAA Group) cv. Berangan roots. PLoS ONE (*In Press*) (*ISI-indexed publication*)

Jennifer Ann Harikrishna, **Wan-Sin Lee**, Ranganath Gudimella & Mark W. Davey (2015). The banana salt stress transcriptome: Large and small RNA expression in stressed banana. Acta Horticulturae (*In Press*) (*Scopus-indexed publication*)

Mark W. Davey, Ranganath Gudimella, Jennifer Ann Harikrishna, **Lee Wan Sin**, Norzulaani Khalid and Johan Keulemans (2013). "A draft Musa balbisiana genome sequence for molecular genetics in polyploid, inter- and intra-specific Musa hybrids". BMC Genomics, 14, 683 (*ISI-indexed publication*).

Lee Wan Sin, Leong Pei Li, Ho Chai Ling and Jennifer Ann Harikrishna (2011). Identification of microRNA precursors in Bruguiera spp. Botanica Marina, 54, 313-324 (*ISI-indexed publication*).

Gaurav Sablok, Luo Chun, **Lee Wan Sin**, Farzana Rahman, Tatania V. Tatarinova, Jennifer Ann Harikrishna and Luo Zhengrong (2011). Bioinformatic analysis of fruits-specific expressed sequence tag libraries of Diospyros kaki Thunb.: view at the transcriptome at different developmental stages. 3 Biotech 1, 35-45.

Conferences:

Lee Wan Sin, Ranganath Gudimella, Martti Tammi, Norzulaani Khalid and Jennifer Ann Harikrishna. Novel transcripts and microRNAs regulating salt-stress response in banana roots. Oral presentation at The 2nd Asian Regional Conference on Systems Biology, 8th-9th October 2013, Hotel Equatorial Bangi-Putrajaya, Malaysia.

Lee Wan Sin, Ranganath Gudimella, Martti Tammi, Norzulaani Khalid and Jennifer Ann Harikrishna. Transcriptome and sRNAome: Gene expression dynamics in abiotic-stressed banana plants. Poster presentation at The University of Malaya's Researchers' Conference 2013, 19th-20th November 2013, Research Management and Innovation Complex, University of Malaya, Kuala Lumpur, Malaysia.

Lee Wan Sin, Ranganath Gudimella, Martti Tammi, Norzulaani Khalid and Jennifer Ann Harikrishna. Conserved and novel microRNA expressed in salt-stressed banana roots revealed by high-throughput sequencing of small RNA. Poster presentation at The University of Malaya's Researchers' Conference 2012, 23rd-24th April 2012, Research Management and Innovation Complex, University of Malaya, Kuala Lumpur, Malaysia.

Lee Wan Sin, Norzulaani Khalid and Jennifer Ann Harikrishna. Expression profiling of microRNAs in salt stressed banana roots. Poster presentation at The Asian Regional Conference on Systems Biology, 29th November – 1st December 2010, Royale Chulan, Kuala Lumpur, Malaysia.

Lee Wan Sin, Tan Yew Seong, Norzulaani Khalid and Jennifer Ann Harikrishna. Discovery of microRNA Genes in Banana (*Musa acuminata*). Poster presentation at

The 18th Malaysian Society of Molecular Biology & Biotechnology (MSMBB) Scientific Meeting, August 18-20, 2009, Saujana Kuala Lumpur Hotel

Poster Session

Lee Wan Sin, Tan Yew Seong, Norzulaani Khalid and Jennifer Ann Harikrishna. Investigation study of abiotic stress-related microRNAs in *Musa acuminata* var. Berangan (AAA genome). Oral presentation at The 14th Biological Sciences Graduate Congress, December 10-12, 2009, Chulalongkorn University, Bangkok, Thailand.

Lee Wan Sin. Study of microRNA genes in banana (*Musa* species). Oral presentation at The Plant Biotechnology Postgraduate Symposium 2009, 17th December, 2009, University of Putra Malaysia, Serdang, Selangor, Malaysia.

APPENDIX A: SUPPLEMENTARY DATA

Table S1: MS media (Murashige & Skoog, 1962)

Components		Concentration (mg/l)
Macronutrients	$CaCl_2 \cdot 2H_2O$	440
	NH_4NO_3	1,650
	KNO_3	1,900
	KH_2PO_4	170
	$MgSO_4 \cdot 7H_2O$	370
Micronutrients	KI	0.83
	$CoCl \cdot 6H_2O$	0.025
	H_3BO_3	6.2
	$Na_2MoO_4 \cdot 2H_2O$	0.25
	$MnSO_4 \cdot 4H_2O$	22.3
	$CuSO_4 \cdot 5H_2O$	0.025
	$ZnSO_4 \cdot 7H_2O$	8.6
FeEDTA	$FeSO_4 \cdot 7H_2O$	27.85
	$Na_2EDTA \cdot 2H_2O$	37.25
Vitamins	Glycine	2
	Nicotinic acid	0.5
	Pyridoxine	0.5
	Thiamine HCl	0.1
	Myo-inositol	100

Table S2: Length distribution of contigs

Length	**CTR-Contigs**	**Percentage**	**TR300-Contigs**	**Percentage**
100-200	33,197	47.81	34,942	46.89
201-400	25,747	37.08	28,251	37.91
401-600	6,294	9.06	6,562	8.81
601-800	2,269	3.27	2,468	3.31
801-1000	952	1.37	1,108	1.49
1001-1200	461	0.66	568	0.76
1201-1400	247	0.36	259	0.35
1401-1600	130	0.19	171	0.23
1601-1800	66	0.10	77	0.10
1801-2000	29	0.04	45	0.06
2001-2200	16	0.02	31	0.04
2201-2400	7	0.01	14	0.02
2401-2600	8	0.01	7	0.01
2601-2800	8	0.01	4	0.01
2801-3000	4	0.005	4	0.005
>3000	6	0.008	14	0.018
Total reads	69,441		74,525	
Mean length	263		265	
Maximum length	4340		4924	
Minimum length	52		52	

CTR: untreated control (0 mM NaCl); TR300: 300 mM NaCl-treated banana roots

Table S3: Length distribution of scaffolds

Length	**CTR-Scaffold**	**Percentage**	**TR300-Scaffold**	**Percentage**
100-200	17,788	35.88	20,990	37.10
201-400	19,010	38.35	22,294	39.41
401-600	6,172	12.45	6,636	11.73
601-800	2,915	5.88	2,928	5.18
801-1000	1,544	3.11	1,586	2.80
1001-1200	891	1.80	858	1.52
1201-1400	498	1.00	475	0.84
1401-1600	311	0.63	327	0.58
1601-1800	177	0.36	177	0.31
1801-2000	100	0.20	101	0.18
2001-2200	51	0.10	74	0.13
2201-2400	40	0.08	45	0.08
2401-2600	29	0.06	26	0.05
2601-2800	17	0.03	20	0.04
2801-3000	11	0.02	6	0.01
>3000	22	0.04	39	0.05
Total reads	49,576		56,572	
Mean length	349		334	
Maximum length	4,542		4,924	
Minimum length	100		100	

CTR: untreated control (0 mM NaCl); TR300: 300 mM NaCl-treated banana roots

Table S4: Length distribution of unigenes

Length	CTR	Percentage (%)	TR300	Percentage (%)	All unigene	Percentage (%)
100-200	17,788	35.88	20,990	37.10	4,433	14.12
201-400	18,987	38.30	22,285	39.39	12,504	39.83
401-600	6,169	12.44	6,645	11.75	5,945	18.94
601-800	2,900	5.85	2,944	5.20	3,204	10.21
801-1000	1,556	3.14	1,587	2.81	1,923	6.13
1001-1200	899	1.81	851	1.50	1,169	3.72
1201-1400	504	1.02	470	0.83	722	2.30
1401-1600	315	0.64	327	0.58	537	1.71
1601-1800	183	0.37	174	0.31	323	1.03
1801-2000	100	0.20	100	0.18	220	0.70
2001-2200	54	0.11	76	0.13	145	0.46
2201-2400	38	0.08	42	0.07	82	0.26
2401-2600	32	0.06	26	0.05	60	0.19
2601-2800	18	0.04	19	0.03	33	0.11
2801-3000	11	0.02	7	0.012	32	0.10
>3000	22	0.04	29	0.05	55	0.17
Total reads	49,576		56,572		31,390	
Mean length	350		334		517	
Maximum length	4,542		4,924		5,995	
Minimum length	100		100		100	

CTR: untreated control (0 mM NaCl); TR300: 300 mM NaCl-treated banana roots

Table S5: Gap distribution (N/size) of the scaffolds and unigenes

Gap distribution (N/Length) %	CTR_unigene	Percentage (%)	TR300_unigene	Percentage (%)	CTR_scaffold	Percentage (%)	TR300_scaffold	Percentage (%)	All-unigene	Percentage (%)
0	42,826	86.35	50,782	89.74	24,912	86.78	49,584	87.37	26,164	83.31
1-5	6,447	13.00	5,465	9.65	1,340	4.66	905	1.59	4,950	15.76
6-10	51	0.102	58	0.102	300	1.045	547	0.96	45	0.14
11-15	53	0.106	53	0.093	276	0.96	578	1.018	50	0.15
16-20	44	0.088	45	0.079	188	0.65	547	0.96	33	0.105
21-25	37	0.074	34	0.06	246	0.85	594	1.04	43	0.13
25-30	43	0.086	33	0.058	183	0.63	576	1.01	38	0.12
>30	91	0.18	113	0.19	1259	4.38	3417	6.02	80	0.25
Total	49592	100	56583	100	28704	100	56748	100	31403	100

Note: (N/length) % = percentage of ambiguous nucleotide ('N') in a sequence

CTR: untreated control (0 mM NaCl); TR300: 300 mM NaCl-treated banana roots

Table S6: Nucleotide content of the banana transcriptomes

	CTR-Unigene	**Percentage**	**TR300-Unigene**	**Percentage**	**All-Unigene**	**Percentage**
A	4,309,606	24.83	4,750,275	25.13	4,068,171	25.04
C	4,369,065	25.18	4,702,783	24.88	4,068,150	25.03
G	4,360,913	25.13	4,680,642	24.76	4,049,806	24.92
T	4,296,493	24.76	4,751,549	25.14	4,048,727	24.92
N	13,969	0.080	13,892	0.073	11,817	0.072
Total	17,350,046	100	18,899,141	100	16,246,671	100

CTR: untreated control (0 mM NaCl); TR300: 300 mM NaCl-treated banana roots

Table S7: cDNA sequences directly cloned and sequenced from banana mixed tissue samples using method Ho et al. (2006)

Clones	cDNA sequence	Length (nt)	Putative identity
A1	CCAGGCACCCAG	12	Unknown
A2	TACGCGATGGGGCATTGTAAGTGGTAGAGTGGCC	34	26S ribosomal RNA
A2	TGCACATGGGTTAGCCGATCC	21	26S ribosomal RNA
A4	GCTCCTGGGGATTAGTGGCGAACGGG	26	16S ribosomal RNA
A5	GAGAGGACTGGGTTGGGTCAACCTATGGTG	30	Mitochondrial sequence
B1	CCTCGGGAAAGTACTTGTTTGCAGGAGTGGTTGATGGGAGGAACATCTGGG	51	mRNA
B3	N/I	-	-
B4	N/I	-	-
B5	GACAGAGAGGTGCAACCGCCTGGG	24	*Musa acuminata* BAC clone
B6	TGCACTGCCTCTTCCCTGGC	20	miR408, positive strand
B7	GCCACGGCTGGTGCACCA	18	mRNA
B8	GGAGGGGGATGCCGAAGGCAGGGCTAGTGACTGGAGTGA	39	16S ribosomal RNA
B9	CCAGCCCTGCGTCGCACGGATTCG	24	26S ribosomal RNA
B11	GGTACGCGAGCTGGG	15	Chloroplast sequence
B11	TGGGGCGCGGCCAAGCGGTAAGGCAGCGGG	30	Chloroplast sequence
C4	N/I	-	-
C23	N/I	-	-
D1	N/I	-	-
D3	N/I	-	-
D4	N/I	-	-
D5	TATGACTACCGCAATCGAAAAACATAAAATTGCTACATGGCTTTCTGGG	49	Ty3-gypsy-like retrotransposon
D6	CGTGCGAGGTATTGTCAGCT	20	mRNA
D6	AGCCCCAGCAATGGGA	16	Unknown
E9	TGGGGCGTGGCCAAGCGGTAAGGCAGCGGG	30	Chloroplast sequence
E11	N/I	-	-
F1	CACCCCGCAATGGGC	15	Unknown
F3	N/I	-	-
F4	GCCGGCCAGGGGACGGACTGGGA	23	26S ribosomal RNA
F5	TGGGTATAGGGGCGAAAGACTAATCGAACCATCTAGTAGCTGG	43	28S ribosomal RNA

Table S7: continued

Clones	cDNA sequence	Length (nt)	Putative identity
F6	TGGGGCGTGGCCAAGCGGTAAGGCAGCGGG	30	Chloroplast sequence
F6	GTCGGAGGTAGGGTCCAGCGGC	22	26S ribosomal RNA
F7	N/I	-	-
F8	TACGCGATGGGGCATTGTAAGTGGTAGAGTGGCC	34	26S ribosomal RNA
F8	CACCCGGCCGAGGGCACGCCTGCCTGGGCGTCACG	35	5.8S ribosomal RNA
F9	CCCGTCGTTCGCCTCCCGACCCACAGTAGGGGCC	34	26S ribosomal RNA
F10	GCCGGCCGGGGGACGGCCTGGG	22	26S ribosomal RNA
F10	CACCCACTAGCATCC	15	Unknown
F11	GCCGGCCGGGGGATGGACCGGG	22	mRNA
F12	CGAGGGCACGTCTGCCTGGGCGTCACG	27	18S ribosomal RNA
F13	CAAGGGTTCCAGGGCCAGG	19	Unknown
		Average = 24	

Note: Some clones contained two inserts
N/I: No insert
Unknown: Unknown identity

Table S8: cDNA sequences directly cloned and sequenced from banana mixed tissue samples using method Fu et al. (2005)

Clones	cDNA sequence	Length (nt)	Putative identity
P1	CTGGAGACCGGGGTTCGACCCCCCGTATCGGAGC	34	tRNA
P2	TGAGAATAGATGGGTTAATTCCTATTAAGGACGC	34	tRNA
P2	GGCGTAAGTCATCGGTTCAAATCCGATAAGGGGCTC	36	tRNA
P3	ATTCGTATTTCATAGTCAGAGGTGAAATTCTTGGATTTATGAAAGACGAACCACTGCGAAAGCATTTGCCA AGGATGTTTTCATTAATCAAGAA	94	18S ribosomal RNA
P4	TCGAAGGGTTGCAGGTTCAATTCCTGTCCGTTTCACG	37	tRNA
P5	GACCTTATCGTCTAATGGTTACGACATCACCTCATCATGTTGATAATATCGGTTCGATTCCGATTAAGGTTA CCG	75	tRNA
P6	CTCACTGGGGTCGC	14	mRNA
P6	CGTGGAGACCGGGGTTCGACTCCCCGTATCGGAG	34	Unknown
P7	TGTGGAGATCATCGGTTCAAATCCGATTGGAAGCAC	36	tRNA
P8	GTTCAACCCTCACTGGGGTCG	21	tRNA
P9	ATTCGTATTTCATAGCTACTAGGTGAAATTCTTGGATTTATGAAAGACGAACCACTGCGAAAGCATTTGCC AAGGATGTTTTCATTAATCAAGAA	95	tRNA
P11	CGTGGAGACCGGGGTTCGACTCCCCGTATCGGAACCG	37	18S ribosomal RNA
P14	TGTGGAGATCATCGGTTCAAATCCGATTGGAGGCAC	36	tRNA
P15	TCGCAAGATCGTGAGTTCAACCCTCACTGGGGTCG	35	tRNA
P16	CGTGGAGATCATCGCTTATAATCCGATTGGAAGCAC	36	tRNA
P17	CTCAATGACTTAGGTCTTCATAGGTTCAATTCCTATTCCCTTCAC	45	tRNA
P18	TCGAAAGGTTTGGGGTTCAAATCCCTAATATAACAC	36	Mitochondrial sequence
P19	TCGAAAGGTTTGGGGTTCAAATCCCTAATATAACAC	36	Mitochondrial sequence
P21	GACCTTATCGTCTAATGGTTACGACATCACCTCATCATGTTGATAATATCGGTTCGATTCCGATTAAGGTTA C	73	Mitochondrial sequence
P23	ATGTGGAGATCATCGGTTCAAATCCGATTGGAAGCAC	37	tRNA
PD5	TGTGGAGATCATCGGTTCAAATCCGATTGGAAGCAC	36	tRNA
PD11	ATCTTTTGGGTTTTGCCCGACCCCGTTCGAGTCCTGCAGTTGTCGC	46	tRNA
PD13	GGGTTATATTAGGGATTTGAACCCCAAACCTTTCGA	36	tRNA
PD15	TCAGAAGATTATGGGTTCGACCCCCATCGTGAGTGC	36	tRNA
PD16	ATGTGGAGATCATCGGTTCAAATCCGATTGGAAGCAC	37	tRNA
PD17	CCAGAAGATTATGGGTTCGACCCCCATCGTGAGTG	35	tRNA
PD19	TCGGAAGATCGTGAGTTCAACCCTCACTGGGGT	33	tRNA

Table S8: Continued

Clones	cDNA sequence	Length (nt)	Putative identity
PD1	TCTGTTGGACGGTTGTCCGCGCGAGTTCGAACCTCGCATCCTTCAC	46	tRNA
PD1	TGTGGAGATCATCGGTTCAAATCCGATTCCAAGCCC	36	tRNA
PD2	ACGTGGAGATCATCGGTTCAAATCCGATTGGAAGCACCAGGCATTGGGTTTTACCTGCGCAGGTTCGAATC CTGTCTGTGACGC	84	tRNA
PD4	CTTGATAATATCGGTTCGATTCCGATTAAGGTTAC	35	tRNA
PD7	CTTGATAATATCGGTTCGATTCCGATTAAGGTTAC	35	tRNA
PD12	AATTCCTGTCCGTTTCAC	18	tRNA
PD13	TCGAAAGGTTTGGGGTTCAAATCCCTAATATAACCC	36	tRNA
PD14	TCTTTTGGGCTTTGCCCGCGCAGGTTCGAGTCCCGCAGTTGTCG	44	tRNA
PD23	TCGAAAGGTTTGGGGTTCAAATCCCTAATATAACAC	36	tRNA
PD24	TCTTTGGGGTTTTGCCCGCGCAGTTTAGAGTCCTACAGTTGTCG	44	tRNA
		Average = 40	

Note: Some clones contained two inserts
Unknown: Unknown identity

Table S9: List of banana orthologous miRNAs named in this study and their matches in Plant MicroRNA Database (PMRD)

No.	*Musa acuminata*	PMRD
1	mac-miR156	ssp-miR156
2	mac-miR156a	pts-miR156a
3	mac-miR156a.3	vun-miR156a.3
4	mac-miR156c	smo-miR156c
5	mac-miR156e	vvi-miR156e
6	mac-miR156h	vvi-miR156h
7	mac-miR156h.1	ath-miR156h
8	mac-miR156i	vvi-miR156i
9	mac-miR156j	zma-miR156j
10	mac-miR156k	zma-miR156k
11	mac-miR156o	zma-miR156o
12	mac-miR156o.1	osa-miR156o
13	mac-miR156p	osa-miR156p
14	mac-miR156q	zma-miR156q
15	mac-miR156q.1	osa-miR156q
16	mac-miR156r	zma-miR156r
17	mac-miR157	ghr-miR157
18	mac-miR157.1	mtr-miR157
19	mac-miR157.2	sbi-miR157
20	mac-miR157c	sly-miR157c
21	mac-miR157d	ath-miR157d
22	mac-miR157d*	aly-miR157d*
23	mac-miR157m	zma-miR157m
24	mac-miR159	smo-miR159
25	mac-miR159.1	psi-miR159
26	mac-miR159.2	aqc-miR159
27	mac-miR159.3	bdi-miR159
28	mac-miR159a	pta-miR159a
29	mac-miR159b	ath-miR159b
30	mac-miR159c	vvi-miR159c
31	mac-miR159c.1	ath-miR159c
32	mac-miR159f	osa-miR159f
33	mac-miR159g	zma-miR159g
34	mac-miR159m	zma-miR159m
35	mac-miR160b	ttu-miR160b
36	mac-miR160f	vun-miR160f
37	mac-miR160g	ppt-miR160g
38	mac-miR160h	ptc-miR160h
39	mac-miR160m	zma-miR160m
40	mac-miR162	zma-miR162
41	mac-miR162b	zma-miR162b
42	mac-miR162b.2	vun-miR162b.2
43	mac-miR164	gar-miR164
44	mac-miR164b	sbi-miR164b
45	mac-miR164c	osa-miR164c
46	mac-miR164c.1	ath-miR164c
47	mac-miR164e	osa-miR164e

Table S9: Continued

48	mac-miR164h	zma-miR164h
49	mac-miR165	hce-miR165
50	mac-miR165b	aly-miR165b
51	mac-miR166	ctr-miR166
52	mac-miR166b	vvi-miR166b
53	mac-miR166b.1	crt-miR166b
54	mac-miR166b.3	vun-miR166b.3
55	mac-miR166c*.4	vun-miR166c*.4
56	mac-miR166e	bdi-miR166e
57	mac-miR166e*	aly-miR166e*
58	mac-miR166i	zma-miR166i
59	mac-miR166k	zma-miR166k
60	mac-miR166k.1	sbi-miR166k
61	mac-miR166m	zma-miR166m
62	mac-miR166m.1	osa-miR166m
63	mac-miR166q	ptc-miR166q
64	mac-miR166q.1	gma-miR166q
65	mac-miR166t	zma-miR166t
66	mac-miR167	aqc-miR167
67	mac-miR167b	bna-miR167b
68	mac-miR167c	vvi-miR167c
69	mac-miR167c.1	tcc-miR167c
70	mac-miR167c.2	rco-miR167c
71	mac-miR167f	vun-miR167f
72	mac-miR167g	Vun-miR167g
73	mac-miR167m	zma-miR167m
74	mac-miR167n	gma-miR167n
75	mac-miR167t	zma-miR167t
76	mac-miR168	vvi-miR168
77	mac-miR168.1	stu-miR168
78	mac-miR168m	zma-miR168m
79	mac-miR169f	tcc-miR169f
80	mac-miR169i	sbi-miR169i
81	mac-miR169k	zma-miR169k
82	mac-miR169l	mtr-miR169l
83	mac-miR169m	zma-miR169m
84	mac-miR169q	zma-miR169q
85	mac-miR169r	zma-miR169r
86	mac-miR169u	zma-miR169u
87	mac-miR171	hce-miR171
88	mac-miR171.1	ctr-miR171
89	mac-miR171.2	vun-miR171.2
90	mac-miR171a	smo-miR171a
91	mac-miR171b	tae-miR171b
92	mac-miR171c	tcc-miR171c
93	mac-miR171c.1	mtr-miR171c
94	mac-miR171f	vvi-miR171f
95	mac-miR171j	zma-miR171j
96	mac-miR171n*	zma-miR171n*
97	mac-miR172a	vvi-miR172a

Table S9: Continued

98	mac-miR172a.2	vun-miR172a.2
99	mac-miR172b	gra-miR172b
100	mac-miR172c	vvi-miR172c
101	mac-miR172d	zma-miR172d
102	mac-miR172d.1	vvi-miR172d
103	mac-miR172d.2	ath-miR172d
104	mac-miR172e	sbi-miR172e
105	mac-miR172e.1	aly-miR172e
106	mac-miR172f	zma-miR172f
107	mac-miR172n	zma-miR172n
108	mac-miR319	pta-miR319
109	mac-miR319a.2	vun-miR319a.2
110	mac-miR319b	ppt-miR319b
111	mac-miR319b.1	mtr-miR319b
112	mac-miR319e	ppt-miR319e
113	mac-miR319f	vvi-miR319f
114	mac-miR319g	vvi-miR319g
115	mac-miR319m	zma-miR319m
116	mac-miR390	zma-miR390
117	mac-miR393b	vun-miR393b
118	mac-miR393g	zma-miR393g
119	mac-miR394a.3	vun-miR394a.3
120	mac-miR394b	zma-miR394b
121	mac-miR395h.3	vun-miR395h.3
122	mac-miR395m	vvi-miR395m
123	mac-miR395p	zma-miR395p
124	mac-miR396	smo-miR396
125	mac-miR396a	vvi-miR396a
126	mac-miR396a*	aly-miR396a*
127	mac-miR396b	vvi-miR396b
128	mac-miR396b*	vun-miR396b*
129	mac-miR396d	tcc-miR396d
130	mac-miR396d.1	sbi-miR396d
131	mac-miR396e	gma-miR396e
132	mac-miR396f	ptc-miR396f
133	mac-miR396f.1	osa-miR396f
134	mac-miR396g	zma-miR396g
135	mac-miR396l	zma-miR396l
136	mac-miR396n	zma-miR396n
137	mac-miR397	sly-miR397
138	mac-miR397.1	pab-miR397
139	mac-miR397a	vvi-miR397a
140	mac-miR397b	bna-miR397b
141	mac-miR397b.1	bdi-miR397b
142	mac-miR398c	zma-miR398c
143	mac-miR398c.1	vvi-miR398c
144	mac-miR399	tae-miR399
145	mac-miR399f	sbi-miR399f
146	mac-miR408	zma-miR408
147	mac-miR408.1	vvi-miR408

Table S9: Continued

148	mac-miR408.2	smo-miR408
149	mac-miR408b	ppt-miR408b
150	mac-miR444b	ssp-miR444b
151	mac-miR528b	zma-miR528b
152	mac-miR529	zma-miR529
153	mac-miR529.1	sbi-miR529
154	mac-miR529.2	far-miR529
155	mac-miR529*	zma-miR529*
156	mac-miR529b	osa-miR529b
157	mac-miR535	aqc-miR535
158	mac-miR535c	vvi-miR535c
159	mac-miR5538	osa-miR5538
160	mac-miR827	zma-miR827
161	mac-miR829.1	ath-miR829.1
162	mac-miR1318	osa-miR1318
163	mac-miR1511	gma-miR1511
164	mac-miR2005	tae-miR2005
165	mac-miR2018	tae-miR2018
166	mac-miR2029	tae-miR2029
167	mac-miR2910	peu-miR2910
168	mac-miR2914	peu-miR2914
169	mac-miR2915	peu-miR2915
170	mac-miR4995	gma-miR4995
171	mac-miR5083	osa-miR5083
172	mac-miR5139	rgl-miR5139
173	mac-miR5179	bdi-miR5179
174	mac-miRf10192-akr	ptc-miRf10192-akr
175	mac-miRf10238-akr	ptc-miRf10238-akr
176	mac-miRf10271-akr	ptc-miRf10271-akr
177	mac-miRf10467-akr	ptc-miRf10467-akr
178	mac-miRf10514-akr	ptc-miRf10514-akr
179	mac-miRf10514-akr.1	ptc-miRf10514-akr.1
180	mac-miRf10629-akr	ptc-miRf10629-akr
181	mac-miRf12052-akr	ptc-miRf12052-akr

Table S10: List of putative *Musa*-specific miRNA genes

miRNA[§]	Sequence	Genome Coordinates:Strand[‡]
mac-miR1	GGCGAUGAUGAUUGGUGAAU	chr1:4999003..4999188:+
mac-miR2	CCGCAGGAGAGAUGAUGCCGCUA	chr5:6134793..6134884:-
mac-miR3	AAAAGAGCGCAACGAUGA	chr1:6510183..6510232:+
mac-miR4	UAUGCUUUGAAGACAAAAUU	chr1:12623066..12623180:-
mac-miR5	UCGGACCAGGCUUCAUUCCUCUC	chr3:4807845..4807933:-
mac-miR6	UAGGAGAGAUGACACCGGCUU	chr7:1377251..1377323:-
mac-miR8	UCGAUAAACCUCUGCAUCCGG	chr3:9020608..9020718:-
mac-miR10	GCGUAGGGCUCUGACUUGACGUGC	chr4:5771259..5771321:+
mac-miR11	UAGAUCGGCUUAGAAAAGAG	chr4:29437465..29437677:-
mac-miR12	AGAAGAGAGAGAGUACAGCUU	chr8:17200596..17200680:-
mac-miR13	UAGCCAAGAAUGACUUGCCUG	chr7:24974039..24974130:+
mac-miR14	UCAAGCUGCCAGCAUGAUCUGA	chr7:1098212..1098297:-
mac-miR15	UGCUUGGAUUGAAGGGAGC	chr10:27333936..27334101:-
mac-miR16	AAUGACGCUAGAAGAAGAGCUG	chr7:24392147..24392444:-
mac-miR17	CUGGAGAAGCAGGGCACGU	chr9:9325871..9325942:+
mac-miR18	GACAGGAAGAGAAGUGAGCA	chr9:27488940..27489028:+
mac-miR19	UCCAGGAGAGAUGACACCAAC	chr4:22573796..22573893:+
mac-miR20a	GAGGCGAUGAUGAUUGGUGAAUGU	chr1:4999001..4999190:+
mac-miR29	UGGAGAAGCAGGGUACGUGC	chr10:4251952..4252039:-
mac-miR30	UGGAGAAGCAGGUCACGUGCAAA	chr11:5005384..5005461:+
mac-miR31	GGAGAAGCAGGUCACGUGCAG	chr11:23428280..23428351:+
mac-miR32	AAACUUUUGUUGGAGAGGAACC	chr11:10286593..10286657:-
mac-miR33	UCGCAGGAGAGAUGAUGCCG	chrUn_random:74512625..74512684:+
mac-miR34	AGGUAUUUCCCGGAUGAUGCCCCCU	chrUn_random:114945177..114945234:+
mac-miR35	AGCGGGGUAGAGGAAUUGGUC	chrUn_random:42177693..42177785:-
mac-miR36	GUUCAAUAAAGCUGUGGGAAA	chrUn_random:47515512..47515588:-
mac-miR37	UAAAGCUGCCAGCAUGAUCUGA	chr1:3895783..3895851:+
mac-miR38	AGGCGAUGAUGAUUGGUGAAUG	chr1:4999002..4999189:+
mac-miR39	UUCCACAGCUUUCUUGAACGG	chr1:7522673..7522749:-
mac-miR40	CGCUGCGCCGGUGGAGACAGC	chr1:13054622..13054786:-
mac-miR41	AGAUGUAGGUUUUUUUUA	chr3:5115827..5115939:+
mac-miR42	UGGUGGAGGGUCGAUGAAGAGGC	chr3:27575488..27575548:+
mac-miR43	AACAGUUGUACGAGCCUAUAGAACA	chr6:32944524..32944583:+
mac-miR44	GAAGAGAGAGAGUACAGC	chr6:28252452..28252519:-
mac-miR45	AGAGAAUAUAUAGAGGAAGAGGAU	chr7:16671938..16672021:+
mac-miR46	UGGAGAAGCAGGGCACGUG	chr9:9325872..9325942:+
mac-miR47	CGGACAAGAGACGGGGAUGG	chr9:11418951..11419032:+
mac-miR48	GACAGGAAGAGAAGUGAGCAC	chr9:21664919..21665003:-
mac-miR49	CUGACAGGAGAGAGUGAGCAC	chr9:27488936..27489027:-
mac-miR50	GAGAAGCAGGGUACGUGCAUU	chr10:4251954..4252037:-

Table S10: Continued

miRNA[§]	Sequence	Genome Coordinates:Strand[ǂ]
mac-miR51	UGGAGAAGCAGGUCACGUG	chr11:5005384..5005461:+
mac-miR52	UUAAAUCUGUGUUAUGGUAA	chr11:8969900..8969968:-
mac-miR53	UCGCAGGAGAGAUGAUGCC	chrUn_random:74512626..74512683:+
mac-miR54	GCGGGGUAGAGGAAUUGGU	chrUn_random:42177694..42177784:-
mac-miR55	AAAGCAUACUAAAAAGGGGCGUA	chr1:12623079..12623172:+
mac-miR56	GAGGAAGAGGAGGUAGAGUGGUG	chr2:18260095..18260208:-
mac-miR57	AUUAAUUUUUUUUGGGGG	chr3:1660129..1660193:-
mac-miR58	AGUCAAACAAUAAUCAGCUGA	chr4:8986260..8986337:+
mac-miR59	CAUCAACCCAGCCUGACCUUUA	chr4:10109433..10109591:+
mac-miR60	CAGGAGAGAUGACACCAAC	chr4:22573798..22573891:+
mac-miR61	GAAGAGAGAGAGUACAGCUU	chr8:17200596..17200679:-
mac-miR62	AGAAGAGAGAGAGUACAGCCU	chr6:28252452..28252520:-
mac-miR63	AAGCUGCCAGCAUGAUCUG	chr7:1098211..1098295:-
mac-miR64	GGAGAAGCAGGUCACGUGCA	chr11:5005385..5005460:+
mac-miR65	AUGGAUCAGGAUCGUCGAGG	chrUn_random:75942924..75943004:+
mac-miR66	GCGGGGUAGAGGAAUUGGUC	chrUn_random:42177693..42177785:-

[§] Putative *Musa*-specific miRNA genes reported in this present study. mac-miR7, mac-miR9, mac-miR20 and mac-miR21 to 28 have been reported in Davey et al. (2013). The mac-miR20a in this study is one nucleotide different from mac-miR20a reported in Davey et al. (2013)

[ǂ] Location of putative *Musa*-specific miRNA precursor in the reference *Musa* A-genome (D'Hont et al., 2012)

(a)

(b)

Figure S1: Market price of Berangan and Cavendish banana in Ringgit Malaysia. Survey on (a) 3rd April 2013 at Tesco, Mutiara Damansara, Selangor, Malaysia and (b) 29th May 2014 at Aeon Big, Tropicana City Mall, Petaling Jaya, Selangor, Malaysia. Berangan price was about 60% (a) and 120% (b) higher than Cavendish

APPENDIX B: R script for transcripts normalization and calculation of differential gene expression

```
# Usage: Rscript DEGseq-script.R
# Input: Tab delimited file of raw counts (Geneid,sample1,sample2)
# Output: Tab delimited file i.e output_score.txt and output.html
#####################################################################
library(DEGseq)

####### Read the tab delimited raw counts file ####################
Musa.RNA<-read.delim("filepath",row.names=1,header=TRUE)


Musa.RNA.mat <- as.matrix(Musa.RNA)                                    #
Convert the dataframe into matrix
CTR.norm <- Musa.RNA.mat[,1]*1000000/sum(Musa.RNA.mat[,1])             #
Calculate transcripts per million (TPM) for normalization
TR300.norm <- Musa.RNA.mat[,2]*1000000/sum(Musa.RNA.mat[,2])
Musa.RNA.norm <- data.frame(CTR.norm,TR300.norm)                       #
Convert into dataframe

############## Differential expression calculation #################

geneExpFile <- Musa.RNA.norm                                          #
read the datafram into DEGseq program
CTR <- readGeneExp(file=geneExpFile, geneCol=1, valCol=2)
TR300 <- readGeneExp(file=geneExpFile, geneCol=1, valCol=3)
mapResultBatch1 <- c(CTR)
mapResultBatch2 <- c(TR300)
DEGexp( geneExpMatrix1=CTR, geneCol1=1, expCol1=2, groupLabel1="CTR",
        geneExpMatrix2=TR300, geneCol2=1, expCol2=2,
groupLabel2="TR300",method="MARS",qValue=0.05,thresholdKind=3,foldChan
ge = 2, rawCount= FALSE,outputDir='path of directory')
```

APPENDIX C: Python script for mapping small RNA reads

```
# Usage : map_reads_script.py input file > output.fasta
# input file = Length,abundance and sequence tab delimited format
# Output file = miRNA matches in fasta format
#
#########################################################################
#
from Bio import SeqIO
import csv
import sys

readFile = sys.argv[1]             # Read input from command line

it = SeqIO.parse(open('Non-redundant miRBase/PMRD miRNA in fasta
format'),'fasta') #

nrD = {}
for rec in it:
    nrD[rec.seq.tostring()] = rec.id              # Creating ids for
matching reads and identifies 5' and 3' pairing.
    nrD[rec.seq.tostring()[1:]] = rec.id+'_5pdg'
    nrD[rec.seq.tostring()[:-1]] = rec.id+'_3pdg'

linkedMature = ' '.join(nrD.keys())
print linkedMature

reader = csv.reader(open(readFile,'rU'), delimiter='\t')         #
Read input file

for line in reader:
    length,  count,  seq = line
    seq = seq.replace('T', 'U')
    if seq in linkedMature:

        try:
            print '>%s %s'%(nrD[seq], count)                     #
print match sequence in fasta format
            print seq
        except KeyError:
            pass
```

APPENDIX D: A brief report on the use of an alternative approach for the genome search for microRNA precursors in banana prior to the release of the *Musa* reference genome

Introduction

Prior to the release of the *Musa* reference genome by D'Hont et al. (2012), genetic resources for banana are limited, with only some BACs, EST libraries and GSS sequences available. It has been reported that miRNA precursors can be found from banana EST sequences as reported by Tan, Y. S. (2010). However, the number of miRNA that can be found was very little, which only 9 miRNA candidates predicted from 42,978 ESTs, which is about 1 miRNA candidate found per 5,000 ESTs. Therefore, we investigated an alternative approach to clone miRNA precursors from banana genome.

A PCR-based approach was used to search for miRNA precursors from the banana genome as some miRNAs have high conservation within plant kingdom and both miRNA and miRNA* strands have good complementarities (Figure D1). Based on these characteristics, primers were designed to amplify several miRNAs of interest (which have been reported as abiotic stress-responsive miRNAs in other plants) from *Musa acuminata* cultivar 'Berangan' (AAA genome).

Figure D1: microRNA mature and its corresponding microRNA* on a miRNA precursor strand. miRNA and the miRNA* have good complementary so that a stem-loop precursor can form

Results and discussion

PCR products with ~100 bp as shown in Figure D2 were cloned and sequenced and two different clones, which sized 96 bp and 144 bp respectively, were obtained. Multiple PCR amplicons were also observed in primers designed for amplifying other miRNA precursors (Figure D3 and other data not shown). These indicated that there were possibly different miRNA precursors (paralogs) of the same miRNA family in banana. After sequencing, the cloned sequences were predicted for their secondary structures using mfold (Zuker, 2003). Two characteristic miR166 precursors (stem-loop structures) and three miR171 were obtained as shown in Figure D4 and Figure D5.

Although putative ortologous miRNA precursors can be obtained by simple primer design and PCR, this approach can only be applied for highly conserved miRNAs as the primers used were designed based on the conservation of miRNA mature and miRNA* sequences.

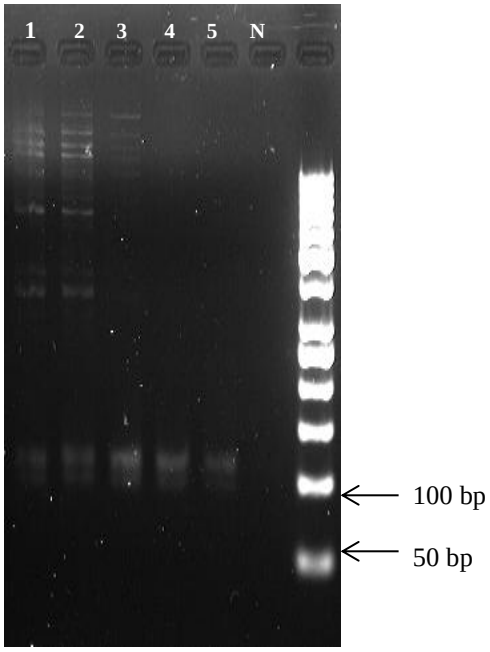


Figure D2: PCR amplification using primers flanking miR166. 1:5: annealing temperature gradient from 40 to 60°C; L: 50-bp DNA marker

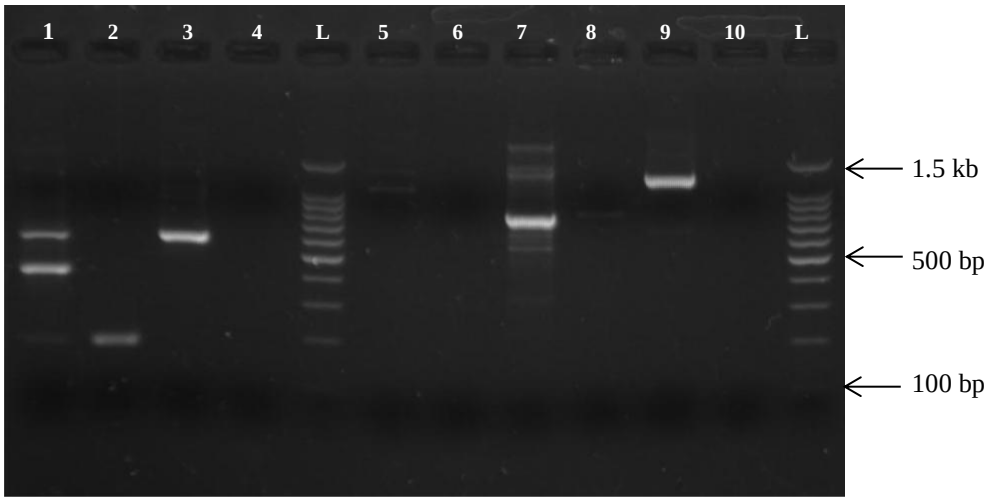


Figure D3: PCR amplification with annealing temperature at 50 (well 1, 3, 5, 7 and 9) and 60°C (well 2, 4, 6, 8 and 10). Primers used for PCR shown in this gel picture flanking (1-2) miR171a, (3-4) miR171b, (5-6) miR393b, (7-8) miR397a and (9-10) miR402. L: 100bp DNA marker

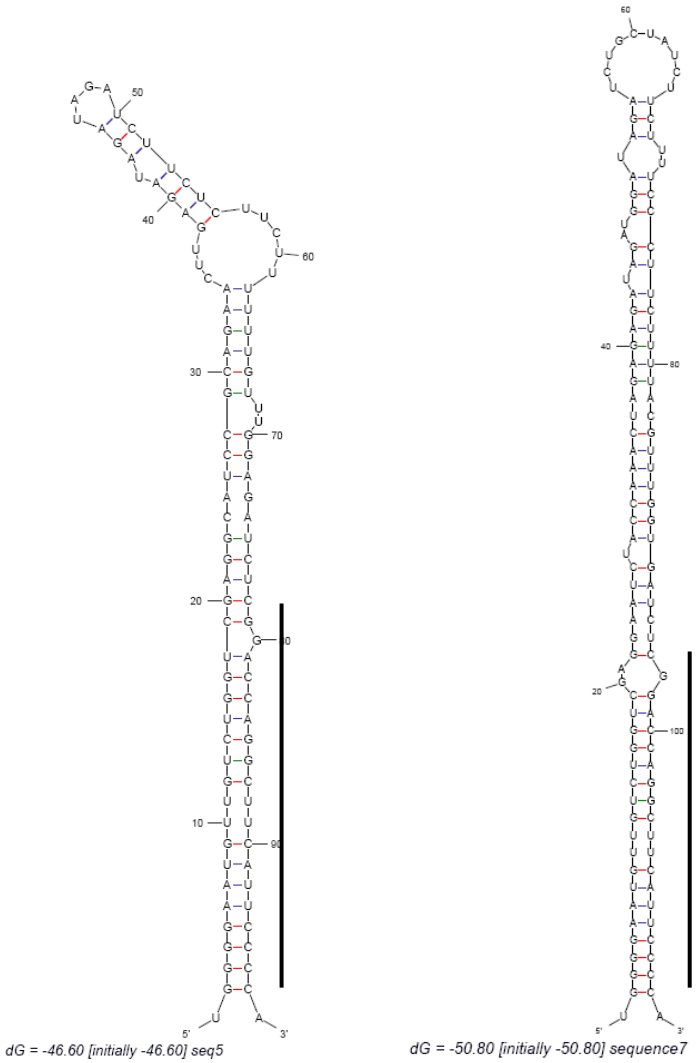


Figure D4: Putative miRNA166 precursors, from banana. Two possible paralogs with the length of 96-nt and 144-nt were found

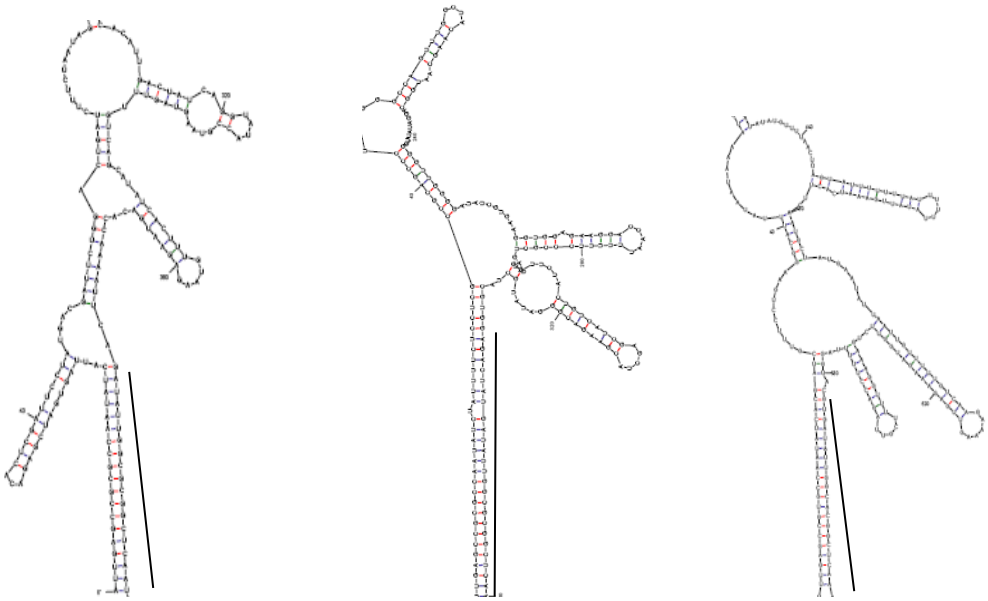


Figure D5: Possible miRNA171 paralogs (putative precursors) from banana