

4000023271

236203



**ISOLATION, CHARACTERIZATION AND
MAPPING OF EXPRESSED SEQUENCE TAGS
(ESTs) FROM PINEAPPLE FRUIT cDNA
LIBRARY**

ONG WEN DEE

PERPUSTAKAAN
UNIVERSITI MALAYSIA SABAH

**THESIS SUBMITTED IN FULFILLMENT FOR
THE DEGREE OF MASTER OF SCIENCE**

**BIOTECHNOLOGY RESEARCH INSTITUTE
UNIVERSITI MALAYSIA SABAH
2011**



UMS
UNIVERSITI MALAYSIA SABAH

UNIVERSITI MALAYSIA SABAH

BORANG PENGESAHAN STATUS TESIS@

JUDUL: ISOLATION, CHARACTERIZATION, MAPPING OF THE EXPRESSED SEQUENCE TAGS FROM PINEAPPLE CDNA LIBRARY

IJAZAH: DEGREE OF MASTER OF SCIENCE

SAYA ONG WEN DEE

SESI PENGAJIAN: 2011

Mengaku membenarkan tesis Sarjana ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:-

1. Tesis adalah hakmilik Universiti Malaysia Sabah.
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institutsi pengajian tinggi.
4. Sila tandakan (/)

☐

SULIT

(Mengandungi maklumat yang berdarjah keselamatan atau Kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

☐

TERHAD

(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

☒

TIDAK TERHAD

(TANDATANGAN PENULIS)

Alamat Tetap: 12, Lorong 4,
Jalan Azizul Rahman, 34600
Kamunting, Taiping, Perak.

PERPUSTAKAAN
UNIVERSITI MALAYSIA SABAH

Disahkan Oleh
(TANDATANGAN PERPUSTAKAAN)

Prof. Madya Dr. Vijay Kumar
Nama Penyelea

Tarikh: 28 Julai 2011

CATATAN:-*Potong yang tidak berkenaan.

**Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh tesis ini perlu dikelaskan sebagai SULIT dan TERHAD.

@ Tesis ini dimaksudkan sebagai tesis bagi Ijazah Doktor Falsafah dan Sarjana secara penyelidikan atau sertai bagi pengajian secara kerja kursus dan Laporan Projek Sarjana Muda (LPSM).

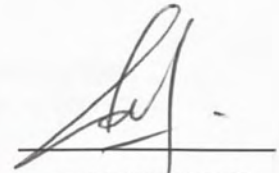


UMS
UNIVERSITI MALAYSIA SABAH

DECLARATION

I hereby declare that this dissertation is the result of my own research except for quotations and citations which have been duly acknowledged.

28 July 2011



ONG WEN DEE

PB2007-8432

CERTIFICATION

NAME : **ONG WEN DEE**

MATRIC NO. : **PB2007-8432**

TITLE : **ISOLATION, CHARACTERIZATION AND MAPPING OF
EXPRESSED SEQUENCE TAGS (ESTs) FROM PINEAPPLE
FRUIT cDNA LIBRARY**

DEGREE : **MASTER DEGREE OF SCIENCE**

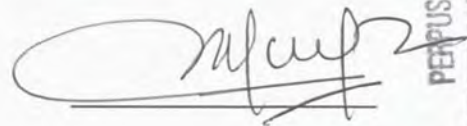
VIVA DATE : **4 JULY 2011**

DECLARED BY

Signature

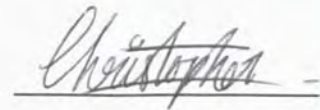
1. SUPERVISOR

Associate Professor Dr. Vijay Kumar



2. CO-SUPERVISOR

Dr. Christopher Voo Luk Yung



PERPUSTAKAAN
UNIVERSITI MALAYSIA SABAH

ABSTRACT

ISOLATION, CHARACTERIZATION AND MAPPING OF EXPRESSED SEQUENCE TAGS (ESTs) FROM PINEAPPLE FRUIT cDNA LIBRARY

Pineapple (*Ananas comosus* var. *comosus*), is an important tropical non-climacteric fruit with high commercial potential. Understanding the phenomena behind fruit ripening with a focus on improving fruit quality traits such as flavor, texture, appearance and sweetness may be possible through gene expression profiling of pineapple fruit transcriptome. As such, the objectives of this project are to, firstly, construct and sequenced mature green pineapple cDNA and *de novo* assembly of paired-end Solexa reads. Secondly, to characterize and functionally annotate the transcripts through similarity search and mapping against Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) database respectively. Finally, to develop a database of Expressed Sequence Tags containing Simple Sequence Repeats (EST-SSRs) using the newly obtained transcripts and/or through pineapple ESTs that are available in GenBank. The results show that both the unique transcripts (UT) assembled pineapple sequences and contigs from *de novo* assembly generated a total of 28,896 transcripts being generated with length ranges from 100 bp to 3.8 kb. A search for sequence similarity with NCBI's non-redundant database identified about 17,049 transcripts which were found to be associated with primary metabolisms, amino acid synthesis and processing, membrane and transport, cell division, cytoskeleton, cell wall and metabolism, RNA related gene expression, signal transduction, defense and stress related protein and also secondary metabolisms. Out of these transcripts, 71% returned GO terms with the distribution among the ontologies given as such: 35.8% in molecular function, 33.5% in cellular component and 30.7% in biological process. Annotation against the KEGG database pathways on the other hand, enabled the assignment of 542 enzyme commissions to 13,598 transcripts. The enzymes were further categorized into a total of 126 pathways with 122 pathways being involved in pineapple metabolism. The metabolic and cellular processes points out that there are tremendous changes in metabolic activities during pineapple fruit maturation as seen by the large numbers of the annotated transcripts. Data mining of the pineapple transcripts EST-SSRs showed that only 4% of the pineapple transcripts contained SSRs. Dinucleotide SSR (49.5%) was the most abundant followed by trinucleotide SSR (46.8%). The least abundant was tetranucleotide SSR (3.7%). Out of these, about 40% of the pineapple transcripts were found to have suitable flanking sites to enable the design of the upstream and downstream primers for future PCR amplification. This research cataloged the first pineapple fruit transcriptome. The transcripts will be subsequently useful to develop microarray chips for future gene expression studies among different plant tissue and development stages of the fruit. Further validation and/or relevant use of the EST-SSRs found will be useful in comparative mapping and genome mapping and gene tagging in pineapple.

ABSTRAK

Nenas, (Ananas comosus var. comosus) merupakan buah tropika yang mempunyai nilai komersial yang tinggi. Memahami fenomena disebalik pemasakan buah dengan tumpuan untuk memperbaharui nilai buah dari segi rasa, struktur, rupa dan manis buah boleh dicapai melalui analisa transkriptome ekspresi gene. Dengan itu, objektif kajian adalah pertamanya pembinaan perpustakaan jujukan saling melengkapi DNA dan pengelompokan pasangan hujung ke hujung jujukan Solexa. Keduanya, adalah menjelaskan transkript yang didapati melalui pencarian persamaan dan penentuan fungsi menggunakan pangkalan data ontologi serta 'Kyoto Encyclopedia of Genes and Genomes' (KEGG). Akhir sekali kajian ini akan membina satu pangkalan data transkript yang wujud dalam kawasan gen berkod dengan menggunakan transkript yang dihasilkan dan juga jujukan saling melengkapi DNA yang sedia ada dalam GenBank. Kesemua transcript unik (UT) dan contigs yang dihasilkan dapat dikelompokkan dalam lebih kurang 30, 000 transkript dengan panjang antara 100 bp ke 3.8 kb. Pencarian persamaan dengan pangkalan data "non-redundant" NCBI mengenalpasti sejumlah 17,049 transkript dengan penglibatan dalam metabolisme asas, penghasilan dan pemprosesan asid amino, dinding dan pengangkutan, pembahagian sel, metabolisme dan pembinaan struktur dinding, ekspresi gen berhubungkait dengan jujukan RNA, transduksi isyarat, protein berkait dengan pertahanan dan tekanan, dan juga metabolisme sekunder. Daripada jumlah ini, 71% mempunyai penanda ontologi dengan 35.8% dalam kumpulan fungsi molekul, 33.5% dalam komponen sel dan 30.7% dalam proses biologi. Penentuan fungsi menggunakan pangkalan data KEGG mendapati sebanyak 13,598 transkript mempunyai fungsi yang sama dengan sejumlah 542 kod enzim yang mana boleh dikelompokkan kepada 126 laluan. Daripada jumlah laluan ini 122 didapati berhubung kait dengan metabolisme nenas. Penentuan fungsi transcript mendapati kebanyakan transcript terlibat dalam metabolik and proses sel dinding. Ini menunjukkan semasa pemasakan buah nenas, aktiviti metabolik giat berlaku. Kajian rangkaian jujukan berulang dalam transkript nenas pula menunjukkan sebanyak 4% daripadanya mempunyai rangkaian jujukan berulang. Dua-nukleotid paling banyak dijumpai dengan sebanyak 676 (49.5%) jujukan penanda terungkap mengandungi rangkaian jujukan berulang. Ini diikuti dengan tiga-nukleotid dan empat-nukleotid dengan masing-masing sebanyak 639 (46.8%) dan 51 (3.7%). Daripada jumlah ini, 40% daripadanya dikenalpasti mempunyai rusuk yang sesuai untuk pencorakan "primers" bahagian depan dan belakang bagi kegunaan amplifikasi PCR pada masa akan datang. Kajian ini menghasilkan transkriptome buah nenas yang pertama. Jujukan penanda terungkap ini berguna untuk penghasilan cip microarray bagi kajian ekspresi gen dalam pelbagai tisu dan peringkat pembentukan buah. Analisa yang lebih terperinci dan/atau penggunaan jujukan penanda terungkap mengandungi rangkaian jujukan berulang boleh diaplikasi dalam pemetaan komparatif dan genome serta penandaan gene dalam nenas.

LIST OF CONTENTS

	Page
TITLE	i
DECLARATION	ii
CERTIFICATION	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
<i>ABSTRAK</i>	vi
LIST OF CONTENTS	vii
LIST OF TABLES	xii
LIST OF FIGURES	xiii
LIST OF ABBREVIATIONS	xvi
LIST OF SYMBOLS	xviii
LIST OF UNITS	xix
LIST OF EQUATIONS	xx
LIST OF APPENDICES	xxi
 CHAPTER 1: INTRODUCTION	
1.0 Introduction	1
1.1 The Objectives of the Study	3
 CHAPTER 2: LITERATURE REVIEW	
2.1 Pineapple	4
2.2 Pineapple Fruit Maturity and Ripening	5
2.3 Uses of Pineapple	7
2.3.1 Pineapple Fruit Processing	7
2.3.2 Application in the Meat Industry	7
2.3.3 Therapeutic Application	8
2.3.4 By-products	9
2.4 Problems in the Pineapple Industry	9
2.5 Application of Biotechnology for Crop Improvement	10
2.6 Transcriptomic Studies in Plant	12
2.7 Sanger Sequencing	13
2.8 Next Generation Sequencing (NGS)	15



2.9	Expressed Sequence Tags	18
2.10	Application of Expressed Sequence Tags	21
2.10.1	Hybridization Experiments	21
2.10.2	Gene Discovery	22
2.10.3	Simple Sequence Repeats (SSRs) Marker Development	24
2.11	ESTs in Non-climacteric Fruits	26
 CHAPTER 3: MATERIALS AND METHODS		
3.1	Overview of Methodology	29
3.2	Plant Materials and Total RNA Extraction	29
3.3	First Strand cDNA Synthesis	30
3.4	Confirmation of Successful of Reverse Transcription	31
3.5	Determination of cDNA Amplification Cycles	31
3.6	Normalization of cDNA Library	33
3.6.1	Hybridization	33
3.6.2	Duplex-specific Nuclease Treatment	33
3.7	Determination of Optimal Number of PCR Cycles for Normalized cDNA	34
3.8	Amplification of Normalized cDNA	35
3.9	Proteinase K and <i>Sfi</i> /Restriction Treatment	35
3.10	Size Fractionation	36
3.11	Ligation of cDNA to Vector and cDNA Library Generation	37
3.12	Screening of cDNA Library	38
3.12.1	Preparation of Host Bacteria	38
3.12.2	Mass Excision of λ TripIEx2 to pTripIEx2	38
3.13	Colony PCR	38
3.14	Plasmid Isolation by Alkaline Lysis	39
3.15	DNA Sequencing	40
3.16	Bioinformatic Analysis	40
3.16.1	Sequence Characterization and Gene Ontology (GO) Annotation of Transcripts	40
3.16.2	Functional Classification by Kyoto Encyclopedia of Genes and Genomes (KEGG)	41
3.16.3	Identification of EST-SSR Motifs and Flanking Primers	41

3.17	Solexa Sequencing	42
3.18	<i>De novo</i> Assembly of Solexa Sequencing Reads	42
3.18.1	Running Velvet	43
3.18.2	Running Velvetg	44
3.19	Counting of Contigs Size	44
 CHAPTER 4: RESULTS		
4.1	Overview of Result Presentation	47
4.2	Total RNA Extraction	47
4.3	Construction of cDNA Library	48
4.3.1	First and Second Strand cDNA Synthesis	48
4.3.2	Determination of Optimal PCR Cycles for Non-normalized and Normalized cDNA	48
4.3.3	<i>Sfi</i> I Enzyme Treatment, Size Fractionation and Vector Ligation of Normalized cDNA	51
4.3.4	Titrating of cDNA Library, Amplification of Insert and Plasmid Extraction of Positives Clones	52
4.3.5	Single Pass Sequencing of Partial cDNA Clones	54
4.3.6	Characterization of ESTs from Mature Green Fruit cDNA Library	56
4.4	Solexa Sequencing	61
4.4.1	Solexa Reads and <i>De novo</i> Assembly using Velvet Software	61
4.4.2	Distribution of Velvet <i>K</i> -mer 47 Contigs Size and Coverage (Expression Level)	64
4.4.3	Characterization of Velvet <i>K</i> -mer 47 Contigs	65
4.4.4	High Coverage Contigs	68
4.5	Functional Annotation	68
4.5.1	Gene Ontology (GO) Annotation	68
4.5.2	Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Assessment	73
4.5.3	Gene Encoding Important Traits in Pineapple Fruit	78
4.5.3.1	Fruit Ripening	78
4.5.3.2	Fragrance Biosynthesis	80
a)	Terpene Biosynthesis	80

b) Ester Biosynthesis	82
4.5.3.3 Flavor Biosynthesis	85
a) Organic Acid Synthesis	85
b) Starch and Sucrose Metabolism	85
4.5.3.4 Texture/structural Biosynthesis –Lignin Biosynthesis	89
4.5.3.5 Health-Related Compound Biosynthesis	89
a) Quinate Biosynthesis	89
b) Riboflavin Metabolism	91
c) Folate Biosynthesis	92
4.6 Detection of Simple Sequence Repeats (SSRs) in Pineapple ESTs	93
4.6.1 Pineapple EST-SSRs Generated from Sanger Sequencing	93
4.6.1.1 Occurrences of Different SSRs	93
4.6.1.2 Sizes of Pineapple SSRs	95
4.6.1.3 Distribution of EST-SSRs	96
4.6.1.4 Determination of Gene Identities of EST-SSRs	99
4.6.1.5 Identification of Flanking Sequences of Type I EST-SSRs	99
4.6.2 Pineapple EST-SSRs from Solexa Sequencing	117
4.6.2.1 Occurrences of Different SSRs in Pineapple Fruit Contigs	117
4.6.2.2 Sizes of Pineapple SSRs in Pineapple Fruit Contigs	117
4.6.2.3 Distribution of SSRs in Pineapple Fruit Contigs	119
4.6.2.4 Determination of Gene Identities of Contigs Containing SSRs	119
4.6.2.5 Identification of Flanking Sequences of Type I EST-SSRs Derived from Contigs	129

CHAPTER 5: DISCUSSION

5.1 Isolation of Total RNA from Pineapple Fruit Tissue	147
5.2 Factors Affecting the Construction of Mature Green Pineapple Fruit Library	148
5.3 Single Pass Sequencing of Green Mature Pineapple Clones	152
5.4 Assembly of Pineapple Transcripts	154
5.5 Characterization through Sequence Similarity Searches	157

5.6	Highly Expressed Transcripts	158
5.7	Functional Annotation of Pineapple Transcripts	163
5.8	Gene Encoding Important Traits in Pineapple Fruit	165
5.9	Pineapple Type I EST-SSRs	170
5.9.1	Distribution and Size	170
5.9.2	Distribution of EST-SSR Motifs	172
5.9.3	Identification of Gene Identities and Flanking Region of EST-SSRs	174
 CHAPTER 6: SUMMARY		
6.1	Summary and Conclusion	177
6.2	Future Prospects	179
 REFERENCES		181
APPENDICES		199

LIST OF TABLES

	Page
Table 2.1 Recent transgenic and mutant tomato and grape genotypes with relation to the nutrient or shelf-life change.	12
Table 2.2 Abundant transcripts in pineapple fruit identified by Moyle <i>et al.</i> (2005b).	28
Table 3.1 PCR cycling parameters for cDNA amplification.	32
Table 3.2 Dilution for DSN treatment.	34
Table 3.3 The various types of repeat motif screened for in the pineapple transcripts.	42
Table 4.1 Blast results of UTs against non-redundant NCBI database.	60
Table 4.2 Assembly and characterization of <i>k</i> -mer 47 contigs generated from <i>de novo</i> assembly.	66
Table 4.3 Contigs with significant identity and coverage over 500.	69
Table 4.4 Summary results of mapping and annotating of pineapple transcripts against Gene Ontology database.	71
Table 4.5 List of metabolism pathways in pineapple fruit transcripts.	74
Table 4.6 Summary of functional assessment of EST containing SSRs in pineapple ESTs.	99
Table 4.7 EST-SSRs with significant identities.	100
Table 4.8 EST-SSRs containing flanking primers.	108
Table 4.9 Summary of functional assessment of EST containing SSRs in pineapple fruit contigs.	123
Table 4.10 EST-SSRs in pineapple fruit contigs with significant identities.	123
Table 4.11 Pineapple fruit contigs EST-SSRs with flanking primers.	130

LIST OF FIGURES

	Page
Figure 2.1 The changes in the physiochemical properties of pineapple fruit from flowering to senescence.	6
Figure 2.2 Schematic workflow of paired-end tags methodology both using cloning based and cloning free procedures.	19
Figure 3.1 Perl scripts tracking of contigs size.	46
Figure 4.1 Total RNA extracted from pineapple fruit tissue using high salt concentration.	49
Figure 4.2 Amplification of metallothionein transcripts.	49
Figure 4.3 Amplification of double strand cDNA.	50
Figure 4.4 Amplification of cDNA with different number of PCR cycles.	50
Figure 4.5 Amplification of normalized cDNA treated with different DSN dilution.	51
Figure 4.6 Fractionation of normalized cDNA.	52
Figure 4.7 Colony PCR amplification of insert.	53
Figure 4.8 Plasmid extraction based on alkaline lysis of different positive clones.	53
Figure 4.9 Example of pineapple fruit cDNA sequence.	54
Figure 4.10 Chromatogram of single pass sequencing of mature green pineapple cDNA clones.	55
Figure 4.11 Contig generated from two pineapple sequences.	57
Figure 4.12 Length distribution of UTs assembled from green mature green pineapple sequences.	58
Figure 4.13 E-value distribution of mature green pineapple UTs with significant identities.	58
Figure 4.14 Blast hit species distribution of UTs of mature green pineapple sequences.	59
Figure 4.15 Example of forward and reverse Solexa reads.	62



Figure 4.16	Results of <i>de novo</i> assembly of Solexa reads using different <i>k</i> -mers.	63
Figure 4.17	Length distribution of pineapple fruit contigs generated from <i>de novo</i> assembly.	64
Figure 4.18	Coverage distribution of contigs from <i>de novo</i> assembly.	65
Figure 4.19	E-value distribution of contigs with significant identities.	67
Figure 4.20	Blast hit species distribution of pineapple contigs.	67
Figure 4.21	Gene Ontology annotation of pineapple fruit transcripts (level 2).	72
Figure 4.22	List of biosynthesis involved in the pineapple fruit metabolism.	77
Figure 4.23	The ethylene synthesis and signal transduction.	79
Figure 4.24	The terpene biosynthesis.	81
Figure 4.25	The straight chain ester biosynthesis from fatty acids.	83
Figure 4.26	The branched chain ester biosynthesis.	84
Figure 4.27	The citrate acid cycle.	86
Figure 4.28	The sucrose metabolism.	87
Figure 4.29	The starch and sucrose metabolism.	88
Figure 4.30	The lignin biosynthesis.	90
Figure 4.31	The quinate biosynthesis.	91
Figure 4.32	The riboflavin (B2) biosynthesis.	92
Figure 4.33	The folate biosynthesis.	94
Figure 4.34	Distribution of Type I SSRs of single pass sequencing of pineapple ESTs.	95
Figure 4.35	The distribution of single pass sequencing SSRs in categories of repeats unit.	97
Figure 4.36	Distribution of single pass sequencing dinucleotide SSRs.	97
Figure 4.37	Distribution of single pass sequencing trinucleotide SSRs.	98
Figure 4.38	Distribution of single pass sequencing tetranucleotide SSRs.	98
Figure 4.39	Distribution of Type I simple sequence repeats in pineapple contigs.	118

Figure 4.40	The distribution of contigs SSRs in categories of repeats unit.	118
Figure 4.41	Distribution of dinucleotide SSRs in pineapple fruit contigs.	120
Figure 4.42	Distribution of trinucleotide SSRs in pineapple fruit contigs.	121
Figure 4.43	Distribution of tetranucleotide SSRs in pineapple fruit contigs.	122
Figure 5.1	Total RNA extracted from pineapple fruit using alkaline phenol:chloroform.	149
Figure 5.2	Colony PCR amplification of small insert cDNA clones.	152
Figure 5.3	Homopolymer slippage.	153

LIST OF ABBREVIATIONS

ACC	1-aminocyclopropane-1-carboxylate
AFLP	amplified fragment length polymorphism
CAD	cinnamyl alcohol dehydrogenase
CCD	charge coupled device
C-OMT	caffeate O-methyltransferase
CTAB	cetyltrimethylammonium bromide
DEPC	diethylpyrocarbonate
DSN	duplex-specific nuclease
DTT	dithiothreitol
cDNA	complementary DNA
dNTP	deoxynucleotide tri phosphate
ddNTP	dideoxynucleotide tri phosphate
ddATP	dideoxyadenine tri phosphat
ddTTP	dideoxythiamine tri phosphate
ddGTP	dideoxyguanine tri phosphate
ddCTP	dideoxycytosine tri phosphate
EB	extraction buffer
EC	enzyme commission
EDTA	ethylenediaminetetraacetic acid
EMBL	European Molecular Biology Laboratory
ESTs	expressed sequence tags
EtBr	ethidium bromide
EtOH	ethanol
FSH	ferulate 5-hydroxylase
Gb	gigabase
GO	Gene Ontology
HCl	hydrochloride
IPTG	isopropyl- β -D-thiogalactopyranoside
IP	internet protocol
KEGG	Kyoto Encyclopedia of Genes and Genomes
LB	luria brutani
MEP	methylerythritol

MDH	malate dehydrogenase
MgCl ₂	magnesium chloride
mRNA	messenger RNA
MgSO ₄	magnesium sulfate
NaCl	sodium chloride
NCBI	National Center for Biotechnology Information
NGS	next generation sequencing
nr	non-redundant
OD	optical density
PEPC	phosphoenolpyruvate carboxylase
PET	paired-end tags
PR	pathogenesis-related
PUFs	proteins of unknown functions
PTP	picotiterplate
PVP	polyvinylpyrrolidone
RAPD	random amplification of polymorphic DNA
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
SAM	S-adenosylmethione
SDS	sodium dodecyl sulfate
SNP	single nucleotide polymorphism
SOLiD	sequencing by oligo ligation and detection
SPS	sucrose phosphate synthase
SSRs	simple sequence repeats
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
THF	Tetrahydrofolate
UT	unique transcript
UTR	untranslated region
UV	ultraviolet
WinSCP	windows Secure copy
QTL	quantitative trait loci
β-ME	beta-mercaptoethanol

LIST OF SYMBOLS

%	percentage
>	more than
<	less than
$\leq$	less or equal to
=	equal to
$\approx$	approximately
/	per
A	absorbance
λ	lambda

LIST OF UNITS

bp	basepair
cm	centimeter
kg	kilogram
kb	kilobase
μ l	microliter
μ g	microgram
M	molar
m	meter
Mbp	megabasepair
min	minute
ml	mililiter
mM	milimolar
ng	nanogram
nm	nanometer
pfu	plaque-forming unit
rpm	rotation per minute
sec	second
v/v	volume per volume
μ M	micromolar
$^{\circ}$ C	degree celcius
w/v	weight per volume



LIST OF EQUATIONS

	Page
Equation 1 $N = X-7$	35
Equation 2 PCR cycles for normalized cDNA	35

LIST OF APPENDICES

	Page
Appendix A Sequences from single pass sequencing of green mature pineapple clones.	199
Appendix B Blast search of singletons assembled from green mature pineapple sequences with significant identity to non-redundant NCBI database.	203
Appendix C Contigs from <i>de novo</i> assembly of ripe yellow paired-end Solexa reads.	207

CHAPTER 1

INTRODUCTION

1.1 Introduction

The pineapple (*Ananas comosus* var. *comosus*), which is a member of the Bromeliaceae, is an economically important tropical fruit. Pineapple together with three other dominant tropical fruits (mango, papaya and avocado) are referred to as "major tropical fruits" as they account for the approximately 75% of global flesh tropical fruit production. The overall productions of pineapples fruit over the past few years has showed an increase and are expected grow in the global demand on the pineapple fruit flesh.

Pineapple fruit is mainly used in the processing industry to make canned pineapple and pineapple juice concentrate. Even though there is a very high demand for the fresh pineapple fruit, the short storage life of pineapple and the occurrence of blackheart disease disorder that is easily induced during storage, has hinder further export of pineapple fruits for direct consumption (Zhou *et al.*, 2003). As such, the export of pineapple is only limited to nearby countries.

Pineapple is a non-climacteric fruit where there is no increase in respiration and ethylene production upon ripening (Moyle *et al.*, 2005b). Therefore, the sweetness of the fruits relies on the time it is harvest. For climacteric fruits such as banana and tomato, the ripening process which follows the ethylene biosynthetic pathway is well characterized (Yang and Hoffman, 1984). In contrast, the mechanism of ripening in non-climacteric fruits such as pineapple, citrus and grape is totally unknown (Giovannoni, 2004).

Expressed Sequence Tags (ESTs) is a powerful tool for gene discovery, gene mapping, and for the analysis of quantitative traits. ESTs are partial sequencing of randomly picked cDNA clones generated by reverse transcription of mRNA. A large number of ESTs collections for various organisms representing



libraries of different tissue and development stages are available in the GenBank EST database, dbEST. As there is a need to sequence large numbers of clones to be able to isolate most if not all the transcripts in an organisms, sequencing of a single library has shifted to large scale sequencing generating EST libraries of more than 10,000 clones. These large scales sequencing has no doubt been able to identify a great number of transcripts but the overall library construction methodology is laborious, time consuming and expensive.

The emergence of next generation sequencing technology has brought molecular study to gain a deeper insight into the mechanisms regulating DNA and RNA level. Instead of a clone-by clone sequencing approach, the massively parallel sequencing, provide a better approach as this sequencing technology greatly reduces the costs, time, labour, errors associated by clone mishandling and also reduces bias associated with the type of vector used in during cloning (Weber, 2007). Aside from the capability to capture large amount of transcripts in a single sequencing reaction, the data generated were able to provide quantitative measurement of the levels of genes expression. This study attempts to both the gene discovery and the identification of up and down-regulated genes by comparison of the transcripts expression.

As of Feb 2011, the pineapple's EST in the publicly available NCBI database only account for approximately 6,000 sequences. Most of the sequences deposited were from pineapple nematode-infected gall cDNA library and root tips cDNA library. Only a small portion of the sequences were generated from pineapple fruit tissue. The limited number of pineapple fruit transcripts available hampers the understanding of the mechanism governing non-climacteric fruit, pineapple. This study applied both the Sanger sequencing and massively parallel sequencing using Solexa paired end sequencing to generate sequence data on the pineapple fruit transcriptome.

1.2 Objectives of the Study

The objectives are;

- a) To identify pineapple mRNA transcripts through the construction of a fruit flesh EST library and assembly of sequences generated from Solexa paired-end sequencing reads.
- b) To characterize and annotate the pineapple transcripts through similarities search against non-redundant NCBI GenBank database and against both GO and KEGG databases respectively.
- c) To identify Type I Simple Sequence Repeats (SSRs) in pineapple fruit transcripts through different motifs searches.

REFERENCES

- Aggarwal, B.B. and Shishodia, S. 2006. Molecular targets of dietary agents for prevention and therapy of cancer. *Biochemical Pharmacology*. **71**:1397-1421.
- Akkaya, M.S. Bhagwat, A.A. and Cregan, P.B. 1992. Length polymorphisms of simple sequence repeat DNA in Soybean. *Genetics*. 1131-1139.
- Alba, R., Fei, Z., Payton, P., Liu, Y., Moore, S.L., Cohn, J., D'Ascenzo, M., Gordon, J.S., Rose, J.K.C., Martin, G., Tanksley, S.D., Bouzayen, M., Jahn, M.M. and Giovannoni, J. 2004. ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development. *The Plant Journal*. **39**:697-714.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990. Basic local alignment search tool. *Journal of Molecular Biology*. **215**:403-410.
- Anderson, J.R. and Lübberstedt, T. 2003. Functional markers in plants. *TRENDS in Plant Science*. **8**:554-560.
- Ansorge, W.J. 2009. Next-generation DNA sequencing techniques. *New Biotechnology*. **25**:195-203.
- Asopa, V.N. 2003. *Competitiveness in pineapple canning industry*. Hawaii International Conference on Business.
- Avallone, A., Guiraud, J-P., Briloeut, J-M. And Teisson, C. 2003. Enzymatic browning and biochemical alterations in black spots of pineapple [*Ananas comosus* (L.) Merr]. *Current Microbiology*. **47**:113-118.
- Bakarat, A., Bagniewska-Zadworna, A., Choi, A., Plakkat, U., DiLoreto, D.S., Yellanki, P. and Carlson, J.E. 2009. The cinnamyl alcohol dehydrogenase gene family in *Populus*: phylogeny, organization and expression. *BMC Plant Biology*. **9**:26.
- Bartholomew, D.P., Paull, R.E., Rohrbach, K.G. 2003. *The pineapple: botany, production and uses*. UK: CABI Publishing.
- Beldade, P., Rudd, S., Gruber, J.D. and Long, A.D. 2006. A wing expressed sequence tags resource for *Bicyclus anynana* butterflies, an evo-devo model. *BMC Genomics*. **7**:130-145.
- Bell, C.J. and Ecker, J.R. 1994. Assignment of 30 microsatellites loci to the linkage map of *Arabidopsis*. *Genomics*. **19**:137-144.
- Benedito, V.A., Dai, X., He, J., Zhao, P.X. and Udvardi, M.K. 2006. Functional genomics of plant transporters in legume nodules. *Functional Plant Biology*. **33**:731-736.

- Bevan, M. and Walsh, S. 2005. The *Arabidopsis* genome: A foundation for plant research. *Genome Research*. **15**:1632-1642.
- Birol, I., Jackman, S.D., Nielsen, C., Qian, J.Q., Varhol, R., Stazyk, G., Morin, R.D., Zhao, Y., Hirst, M., Schein, J.E., Horsman, D.E., Connors, J.M., Gascoyne, R.D., Marra, M.A. and Jones, S.J.M. 2009. *De novo* transcriptome assembly with AbySS. *Bioinformatics*. **25**:2872-2877.
- Blair, M.W., Hurtado, N., Chavarro, C.M., Munoz-Torres, M.C., Giraldo, M.C., Pedraza, F., Tomkins, J. and Wing, R. 2011. Gene-based SSR markers for common bean (*Phaseolus vulgaris* L.) derived from root and leaf tissue EST: and integration of the BMc series. *BMC Plant Biology*. **11**:50-59.
- Boerjan, A., Ralph, J. and Baucher, M. 2003. Lignin biosynthesis. *Annual Review Plant Biology*. **54**:519-546.
- Bombarely, A., Merchante, C., Csukasi, F., Cruz-Rus, E., Caballero, J.L., Medina-Escobar, N., Blanco-Portales, R., Botella, M.A., Munoz-Blanco, J., Sánchez-Sevilla, J.F. and Valpuesta, V. 2010. Generation and analysis of ESTs from strawberry (*Fragaria xananassa*) fruits and evaluation of their utility in genetic and molecular studies. *BMC Genomics*. **11**:503-520.
- Boo, K.H., Kim, D.W., Cho, S.K., Jin, S.B., Kim, J.H., Lee, H.Y. and Riu, K.Z. 2007. Construction and profiling of a cDNA library from young fruit of satsuma mandarin. *Journal of Plant Biology*. **50**:403-409.
- Boopathi, N.M. and Ravikesavan, R. 2009. Emerging trends in enhancement of cotton fiber productivity and quality using functional genomics tools. *Biotechnology and Molecular Biology Reviews*. **4**:11-28.
- Botella, J.R. and Fairbairn, D.J. 2005. Present and future potential of pineapple biotechnology. *Acta Horticulturae*. **666**:23-28.
- Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P. 2002. Leaf ESTs from *Stevia rebaudiana*: a resource for gene discovery in diterpene synthesis. *Plant Molecular Biology*. **50**:613-622.
- Burland, T.G. 1999. DNASTAR's Lasergene Sequence Analysis Software. *Methods in Molecular Biology*. **132**:71-91.
- Cara, B. And Giovannoni, J.J. 2008. Molecular biology of ethylene during tomato fruit development and maturation. *Plant Science*. **175**:106-113.
- Chan, Y.K. and Lee, H.K. 2000. Breeding for early fruiting in pineapple. *Acta Horticulturae*. **529**:139-146.
- Chang, S., Puryear, J. and Cairney, J. 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter*. **11**:113-116.

- Chen, H.J., Hou, W.C., Yang, C.Y., Huang, D.J., Liu, J.S. and Lin, Y.H. 2003. Molecular cloning of two metallothionein-like genes with differential expression patterns from sweet potato (*Ipomoea batatas*) leaves. *Journal of Plant Physiology*. **160**:547-555.
- Chen, C. and Paull, R.E. 1998. Changes in sugar content and activities of sugar metabolizing enzymes in pineapple fruit flesh during development. *ISHS Acta Horticulturae 529: III International Pineapple Symposium*.
- Cheung, F., Win, L., Lang, J.M., Hamilton, J., Vuong, H., Leach, J.E., Kamoun, S., Andre Levesque, C., Tisserat, N. and Buell, C.R. 2008. Analysis of the *Pythium ultimum* transcriptome using Sanger and Pyrosequencing approaches. *BMC Genomics*. **9**:542
- Chomczynski, P. and Sacchi, N. 2006. The single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nature Protocols*. **1**:581-585.
- Cho, Y.B., Jeon, H.J. and Hong, C.B. 2003. Cloning and functional annotation rare mRNA species from drought-stressed hot pepper (*Capsicum annuum*). *Journal of Plant Biology*. **46**:83-89.
- Collinge, D.B., Kragh, K.M., Mikkelsen, J.D., Nielsen, K.K., Rasmussen, U. and Vad, K. 1993. Plant chitinases. *The Plant Journal*. **3**:31-40.
- Collins, L.L., Biggs, P.J., Voelckel, C. and Joly, S. 2008. An approach to transcriptome analysis of non-model organisms using short-read sequences. *Genome Informatics*. **12**:3-14.
- Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M. and Robles, M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis of functional genomics research. *Bioinformatics*. **21**:3674-3676.
- Cordeiro, G.M., Casu, R., McIntyre, C.L., Manners, J.M., and Henry, R.J. 2001. Microsatellite markers from sugarcane ESTs cross transferable to erianthus and sorghum. *Plant Science*. **160**:1115-1123.
- Costa, G.G.L., Cardoso, K.C., Del Bem, L.E.V., Lima, A.C., Cunha, M.A.S., de Compos-Leite, L., Vicentini, R., Papes, F., Moreira, R.C., Yunes, J.A., Campos, F.A.P. and Da Silva, M.J. 2010. Transcriptome analysis of the oil-rich seed of the bioenergy crop *Jatropha curcas* L. *BMC Genomics*. **11**:462-470.
- Crawford, J.E., Guelbeogo, W.M., Sanou, A., Traoré, A., Vernick, K.D., Sagnon, N. and Lazzaro, B.P. 2010. De novo transcriptome sequencing in *Anopheles funestus* using Illumina RNA-seq technology. *PloS ONE*. **5**:e14202.
- Crowhurst, R.N., Gleave, A.P., MacRae, E.A., Ampomah-Dwamena, C., Atkinson, R.G., Beuning, L.L., Bulley, S.M., Chagne, D., Marsh, K.B., Matich, A.J., Montefiori, M., Newcomb, R.D., Schaffer, R.J., Usadel, B., Allan, A.C.,

- Bolding, H.L., Bowen, J.H., Davy, M.W., Eckloff, R., Ferguson, A.R., Fraser, L.G., Gera, E., Hellens, R.P., Janssen, B.J., Klages, K., Lo, K.R., MacDiarmid, R.M., Nain, B., McNeilage, M.A., Rassam, M., Richardson, A.C., Rikkerink, E.H.A., Ross, G.S., Schroder, R., Snowden, K.C., Souleyre, E.J.F., Templeton, M.D., Walton, E.F., Wang, D., Wang, M.Y., Wang, Y.Y., Wood, M., Wu, R., Yauk, Y-K. and Laing, W.A. 2008. Analysis of expressed sequence tags from Actinidia: applications of a cross species EST database for gene discovery in the areas of flavor, health, color and ripening. *BMC Genomics*. **9**:351.
- Davies, C. and Robinson, S.P. 2000. Differential screening indicates a dramatic change in mRNA profiles during grape berry ripening. Cloning and characterization of cDNAs encoding putative cell wall and stress response proteins. *Plant Physiology*. **122**:803-812.
- de Jong, A., Cordwener, J., Lo Schiavo, F., Terzi, M., Van-dekerckhove, J., Van Kammen, A. and De Vries, S.C. 1992. A carrot somatic embryo mutant is rescued by chitinase. *Plant Cell*. **4**:425-433.
- Deluc, L.G., Grimplet, J., Wheatley, M.D., Tillett, R.L., Quilici, D.R., Osborne, C., Schooley, D.A., Schlauch, K.A., Cushman, J.C. and Cramer, G.R. 2007. Transcriptomic and metabolite analyses of Cabernet Sauvignon grape berry development. *BMC Genomics*. **8**:429-471.
- Díaz de la Garza, R.I., Gregory III, J.F. and Hanson, A.D. 2007. Folate biofortification of tomato fruit. *PNAS*. **104**:4218-4222.
- Dudareva, N., Pichersky, E. and Gershenzon, J. 2004. Biochemistry of plant volatiles. *Plant Physiology*. **135**:1893-1902.
- Dull, G.G., Young, R.E. and Biale, J.B. 1967. Respiratory patterns in fruit of pineapple, *Ananas comosus*, detached at different stages of development. *Physiologia Plantarum*. **20**:1059-1065.
- Elss, S., Preston, C., Hertzog, F., Richling, E. And Schreier, P. 2005. Aroma profiles of pineapple fruit (*Ananas comosus* [L.] Merr.) and pineapple products. *LWT-Food Science and Technology*. **38**:263-274.
- Emrich, S.J., Barbazuk, W.B., Li, L. and Schnable, P.S. 2007. Gene discovery and annotation using LCM-454 transcriptome sequencing. *Genome Research*. **17**:69-73.
- Ezura, H. And Owino, W.O. 2008. Melon, an alternative model plant for elucidating fruit ripening. *Plant Science*. **175**:121-129.
- Fernandez-Silva, I., Eduardo, I., Blanca, J., Esteras, C., Picó, B., Nuez, F., Garcia-Mas, J. and Monforte, A.J. 2008. Bin mapping of genomic and EST-derived SSRs in melon (*Cucumis melo* L.). *TAG Theoretical and Applied Genetics*. **118**:139-150.

- Fernie, A.R. and Willmitzer, L. 2001. Molecular and biochemical triggers of potato tuber development. *Plant physiology*. **127**:1459-1465.
- Fleige, S. and Pfaffl, M.W. 2006. RNA integrity and the effect on the real-time qRT-PCR performance. *Molecular Aspects of Medicine*. **27**:126-139.
- Franca, L.T.C., Carrilho, E. and Kist, T.B.L. 2002. A review of DNA sequencing techniques. *Quarterly Review of Biophysics*. **35**:169-200.
- Food and Agriculture Organization of the United Nations. 2009. *A case study of tropical fruits in asia, with special reference to mangoes and pineapples*.
- Fullwood, M.J., Wei, C., Liu, E.T. and Ruan, Y. 2009. Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. *Genome Research*. **19**:521-532.
- Gao, L., Tang, J., Li, H. and Jia, J. 2003. Analysis of microsatellites in major crops assessed by computational and experimental approaches. *Molecular Breeding*. **12**:245-261.
- Giovannoni, J. 2001. Molecular biology of fruit maturation and ripening. *Annual Review of Plant Physiology*. **52**:725-749.
- Giovannoni, J.J. 2004. Genetic Regulation of Fruit Development and Ripening. *The Plant Cell*. **16**:S170-S180.
- Giovannoni, J.J. 2007. Fruit ripening mutants yield insights into ripening control. *Current Opinion in Plant Biology*. **10**:283-289.
- Gollery, M., Harper, J., Crushman, J., Mittler, T., Girke, T., Zhu, J.K., Bailey-Serres, J. and Mittler, R. 2007. POFs: what we don't know can hurt us. Trends in *Plant Science*. **12**:492-496.
- Goossens, A., Häkkinen, S.T., Laakso, I., Seppänen-Laakso, T., Biondi, S., De Sutter, V., Lammertyn, F., Nuutila, A.M., Söderlund, H., Zabeau, M., Inzé, D. and Oksman-Caldentey, K-M. 2003. A functional genomics approach toward the understanding of secondary metabolism in plant cells. *PNAS*. **100**:8595-8600.
- Gotz, S., Garcia-Gomez, J.M., Terol, J., Williams, T.D. Nagaraj, S.H., Nueda, M.J. Robles, M., Talon, M., Dopazo, J. And Conesa, A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acid Reseach* **36**:3420-3435.
- Grima-Pettinati, J., Campargue, C., Boudet, A. And Boudet, A.M. 1994. Purification and characterization of cinnamyl alcohol dehydrogenase isoforms from *Phaseolus vulgaris*. *Pytochemistry*. **37**:941-947.

- Guadix, A, Guadix, E.M., P´aez-Due˜nas, M.P., Gonz´alez-Tello, P., Camacho, F. 2000. Technological processes and methods of control in the hydrolysis of proteins. *Ars Pharmaceutica*. **41**:79–89.
- Guillon, F., Philippe, S., Bouchet, B., Devaux, M-F., Frasse, P., Jones, B., Bouzayen, M. and Lahaye, M. 2008. Down-regulation of an auxin response factor in the tomato induces modification of fine pectin structure and tissue architecture. *Journal of Experimental Botany*. **59**:273-288.
- Gupta, P.K., Balyan, H.S., Sharma, P.C. and Ramesh, B. 1996. Microsatellites in plants: A new class of molecular markers. *Current Science*. **70**:45-53.
- Hadfield, K.A. and Bennett, A.B. 1998. Polygalacturonases: Many genes in search of a function. *Plant Physiology*. **117**:337-343.
- Hall, N. 2007. Advances sequencing and their widest impact in microbiology. The *Journal of Experimental Biology*. **209**:1518-1525. A unique vascular processing enzyme responsible for conversion of several propotein precursors into the mature forms. *FEBS Letter*. **294**:89-93.
- Hanson, A.D. and Gregory III, J.F. 2002. Synthesis and turnover of folates in plants. *Current Opinion in Plant Biology*. **5**:244-249.
- Hara-Nishimura, I., Inoue, K. and Nishimura, M. 1991. A unique vacuolar processing enzyme responsible for conversion of several proprotein precursors into the mature forms. *FEBS Letters*. **294**:89-93.
- Harrison, M.D., Jones, C.E. and Dameron, C.T. 1999. Copper chaperones: function, structure and copper-binding properties. *JBIC*. **4**:145-153.
- Hemberger, M., Cross, J.C., Ropers, H-H., Lehrach, H., Fundele, R. and Himmelbauer, H. 2001. UniGene cDNA array-based monitoring of transcriptome changes during mouse placental development. *PNAS*. **98**:13126-13131.
- Horan, K., Jang, C., Bailey-Serres, J., Mittler, R., Shelton, C., Harper, J.F., Zhu, J-K., Cushman, J.C., Gollery, M. and Girke, T. 2008. Annotating genes of known and unknown function by large-scale co-expression analysis. *Bioinformatics*. **147**:41-47.
- Iandolino, A. Nobuta, K., da Silva, F.G., Cook, D.R. and Meyers, B.C. 2008. Comparative expression profiling in grape (*Vitis vinifera*) berries derived from frequency analysis of ESTs and MPSS signatures. *BMC Plant Biology*. **8**:53-68.
- International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. *Nature*. **409**:860-921.

- Jabrín, S., Ravanel, S., Gambonnet, B., Douce, R. and Rébeillé, F. 2003. One-carbon metabolism in Plants. Regulation of tetrahydrofolate synthesis during germination and seedling development. *Plant Physiology*. **131**:1431-1439.
- Jiang, D., Zhong, G.Y. and Hong, Q.B. 2006. Analysis of microsatellites in citrus unigenes. *Acta Genetica Sinica*. **33**:345-353.
- Jones, W.J. 2009. High-throughput sequencing and metagenomics. *Estuaries and Coasts*.
- Ju, Z., Wells, M.C., Al Martinez, Hazlewood, L and Walter, B.R. 2005. An silico mining for simple sequence repeats from expressed sequence tags of zebrafish, merdeka, *Fundulus*, and *Xiphophorus*. *In Silico Biology*. **5**:41.
- Kantety, R.V., La Rota, M., Matthews, D.E. and Sorrells, M.E. 2002. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology*. **48**:501-510.
- Kato, H., Sutoh, K. and Minamikawa, T. 2003. Identification, cDNA cloning and possible roles of seed-specific rice asparaginyl endopeptidase, REP-2. *Planta*. **217**:676-685.
- Kay, R.M. 1982. Dietary fiber. *The Journal of Lipid Research*. **23**:221-242.
- Kinsuat, M.J. and Kumar, S.V. 2007. Polymorphic microsatellites and cryptic simple repeat sequence markers in pineapples (*Ananas comosus* var. *comosus*). *Molecular Ecology*. **7**:1032-1035.
- Kisiel, A. and Podkowiński, J. 2005. Expressed Sequence Tags and their application for plant research. *Acta Physiologiae Plantarum*. **27**:157-161.
- Korbel, J.O., Urban, A.E., Affourtit, J.P., Godwin, B., Grubert, F., Simons, J.F., Kim, P.M., Palejev, D., Carriero, N.J., Du, L., Taillon, B.E., Chen, Z., Tanzer, A., Saunders, A.C.E., Chi, J., Yang, F., Carter, N.P., Hurles, M.E., Weissman, S.M., Harkins, T.T., Gerstein, M.B., Egholm, M. and Snyder, M. 2007. Paired-end mapping reveals extensive structural variation in the human genome. *Science*. **318**:420-426.
- Kristiansson, E., Asker, N., Forlin, L. And Larsson, D.G. 2009. Characterization of the *Zoarces viviparus* liver transcriptome using massively parallel pyrosequencing. *BMC Genomics*. **10**:345-355.
- Lagercrantz, U., Ellegren, H. And Andersson, L. 1993. The abundance of various polymorphic microsatellites motifs differs between plants and vertebrates. *Nucleic Acid Research*. **21**:1111-1115.
- La Rota, M., Kantety, R.V., Yu, J-K. and Sorrells, M.E. 2005. Nonrandom distribution and frequencies of genomic and EST-derived microsatellite markers in rice, wheat and barley. *BMC Genomics*. **6**:1-12.

- Lawson, M.J. and Zhang, L. 2006. Distinct patterns of SSR distribution in the *Arabidopsis thaliana* and rice genomes. *Genome Biology*. **7**:R14.
- Leide, J., Hildebrandt, U., Reussing, K., Riederer, M. and Vogg, G. 2007. The developmental pattern of tomato fruit wax accumulation and its impact on cuticular transpiration barrier properties: effects of a deficiency in a β -ketoacyl-Coenzyme A synthase (LeCER6). *Plant Physiology*. **144**:1667-1679.
- Levi, A., Davis, A., Hernandez, A., Wechter, P., Thimmapuram, J., Trebitsh, T., Tadmor, Y., Katzir, N., Portnoy, V. and King, S. 2006. Genes expressed during the development and ripening of watermelon fruit. *Plant Cell Reports*. **25**:1233-1245.
- Li, J.H., Tang, C.H., Song, C.Y., Chen, M.J., Feng, Z.Y. And Pan, Y.J. 2006. A simple, rapid and effective method for total RNA extraction from *Lentinula edodes*. *Biotechnol. Letter*. **26**:1193-1197.
- Li, R., Li, Y., Kristiansen, K. and Wang, J. 2008. SOAP: short oligonucleotide alignment program. *Bioinformatics*. **24**:713-714.
- Li, Y.C., Korol, A.B., Fahima, T. and Nevo, E. 2004. Microsatellites within genes: structure, function and evolution. *Molecular Biology and Evolution*. **21**:991-1006.
- Liu, Q., Zhu, A., Chai, L., Zhou, W., Yu, K., Ding, J., Xu, J. and Deng, X. 2009. Transcriptome analysis of a spontaneous mutant in sweet orange [*Citrus sinensis* (L.) Osbeck] during fruit development. *Journal of Experimental Botany*. **60**:801-813.
- Livak, K.J. and Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(-Delta Delta C(T)). *Method*. **25**:402-408.
- Lokanathan, Y., Mohd-Adnan, A., Wan, K-L. and Nathan, S. 2010. Transcriptome analysis of the *Cryptocaryon irritans* tomont stage identifies potential genes for the detection and control of cryptocaryonosis. *BMC Genomics*. **11**:76-90.
- Lutz, R.E. 2007. Trinucleotide repeats disorder. *Seminars in Pediatric Neurology*. **14**:26-33.
- Mackey, M. 2002. The application of biotechnology to nutrition: an overview. *Journal of the American College of Nutrition*. **21**:157S-160S.
- Maiti, A.K., Ahlawat, S.S., Sharma, D.P. and Khanna, N. 2008. Application of natural tenderizers in meat – A review. *Agricultural Reviews*. **29**:0253-1496.
- Manthey, K., Krajinski, F., Hohnjec, N., Firnhaber, C., Puhler, A., Perlick, A.M. and Kuster, H. 2004. Transcriptome profiling in root nodules and arbuscular mycorrhiza identifies a collection of novel genes induced during *Medicago*

truncatula root endosymbioses. *Molecular Plant-Microbe Interactions*. **17**:1063-1077.

- Mardis, E.R. 2007. Next-generation DNA sequencing methods. *Annual Review and Genomics and Human Genetics*. **9**:387-402.
- Margis-Pinheiro, M., Martin, C., Didierjean, L. and Burkard, G. 1993. Differential expression of bean chitinase genes by virus infection, chemical treatment and UV irradiation. *Plant Molecular Biology*. **22**:659-668.
- Marioni, J.M., Mason, C.E., Mane, S.M., Stephens, M. and Gilad, Y., 2008. RNA-seq: An assessment of technical reproducibility and comparison with the gene expression arrays. *Genome Research*. **18**:1509-1517.
- Matas, A.J., Gapper, N.E., Chung, M-Y., Giovannoni, J.J. and Rose, J.K.C. 2009. Biology and genetic engineering of fruit maturation for enhanced quality and shelf-life. *Current Opinion in Biotechnology*. **20**:197-203.
- Maurer, H.R. 2001. Bromelian: biochemistry, pharmacology and medical use. *Cellular and Molecular Life Sciences*. **58**:1234-1245.
- McKillop, D.J., Pentieva, K., Daly, D., McPartlin, J.M., Hughes, J., Strain, J.J., Scott, J.M. and McNulty, H. 2002. The effect of different cooking methods on folate retention in various foods that are amongst the major contributors to folate intake in the UK diet. *British Journal of Nutrition*. **88**:681-688.
- Metzgar, D., Bytof, J. and Wills, C. 2000. Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Research*. **10**:72-80.
- Meyer, E., Aglyamova, G.V., Wang, S., Buchanan-Carter, J., Abrego, D., Colbourne, J.K. Willis, B.L. and Matz, M. 2009. Sequencing and *de novo* analysis of a coral larval transcriptome using 454 GSFlx. *BMC Genomics*. **10**:219-35.
- Michal, J., 2008. Next-generation Genome Sequencing: Towards personalized medicine. Wiley-Blackwell.
- Minoru, K. and Susumu, G. 2000. KEGG:Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*. **28**:27-30.
- Majeed Mohammed. 1994. Post-harvest physiology of pineapples. *Proceedings of Regional Workshop on Pineapple in Martinique*. March, 22-24, Martinique. 47-49.
- Morgante, M. and Olivieri, A.M. 1993. PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal*. **3**:175-182.
- Moriguchi, T., Kita, M., Hisada, S., Endo-Inagaki, T. and Omura, M. 1998. Characterization of gene repertoires at mature stage of citrus fruits through random sequencing and analysis of redundant metallothionein-like genes expressed during fruit development. *Gene*. **211**:211-227.

- Moyle, R.L., Crowre, M.L., Ripi-Koia, J., Fairbairn, D.J. and Botella, J.R. 2005a. PineappleDB: An online pineapple bioinformatics resource. *BMC Plant Biology*. **5**:21-25.
- Moyle, R.L., Fairbairn, D.J., Ripi, J., Crowre, M. and Botella, J.R. 2005b. Developing pineapple fruit has a small transcriptome dominated by metallothionein. *Journal of Experimental Botany*. **56**:101-112.
- Müntz, K. and Shutov, A.D. 2002. Legumains and their function in plants. *TRENDS in Plant Science*. **7**:340-344.
- Nagaraj, S.H., Gasser, R.B. and Ranganathan, S. 2006. A hitchhiker's guide to expressed sequence tag (EST) analysis. *Briefings in Bioinformatics*. **8**:6-21.
- Nam, Y-W., Tichit, L., Leperlier, M., Cuerq, B., Marty, I. and Lelièvre, J-M. 1999. Isolation and characterization of mRNAs differentially expressed during ripening of wild strawberry (*Fragaria vesca* L.) fruits. *Plant Molecular Biology*. **39**:629-636.
- Negri, A.S., Prinsi, B., Rossoni, M., Failla, O., Scienza, A., Cocucci, M. and Espen, L. 2008. Proteome changes in the skin of the grape cultivar Barbera among different stages of ripening. *BMC Genomics*. **9**:378-397.
- Newcomb, R.D., Crowhurst, R.N., Gleave, A.P., Rikkerink, E.H.A., Allan, A.C., Beuning, L.L., Bowen, J.H., Gera, E., Jamieson, K.R., Janssen, B.J., Laing, W.A., McCartney, S., Nain, B., Ross, G.S., Snowden, K.C., Souleyre, E.J.F., Walton, E.F. and Yauk, Y-K. 2006. Analyses of Expressed Sequence Tags from Apple. *Plant Physiology*. **141**:147-166.
- Nigam, J.N. 2000. Continuous ethanol production from pineapple cannery waste using immobilized yeast cells. *Journal of Biotechnology*. **80**: 189-193.
- Ohlrogge, J. and Benning, C. 2000. Unravelling plant metabolism by EST analysis. *Current Opinion in Plant Biology*. **3**:224-228.
- Oshlack, A. and Wakefield. 2009. Transcript length bias in RNA-seq data confounds systems biology. *Biology Direct*. **4**:14-23.
- Palmieri, D.A., Novelli, V.M., Bastianel, M., Cristofani-Yaly, M., Astúa-Monge, G., Carlos, E.F., de Oliveira, A.C. and Machamirido, M.A. 2007. Frequency and distribution of microsatellites from ESTs of citrus. *Genetics and Molecular Biology*. **30**:1009-1018.
- Pan, Y.W., Chou, H.H., You, E.M. and Yu, H.T. 2004. Isolation and characterization of 23 polymorphic microsatellite markers for diversity and stock analysis in tiger shrimp (*Panaeus monodon*). *Molecular Ecology Notes*. **4**:345-347.
- Parchman, T.L., Geist, K.S., Grahnen, J.A., Benkman, C.W. and Buerkle, C.A. 2010. Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. *BMC Genomics*. **11**:180-195.

- Park, S., Sugimoto, N., Larson, M.D., Beaudry, R. and van Nocker, S. 2006. Identification of genes with potential roles in apple fruit development and biochemistry through large-scale statistical analysis of expressed sequence tags. *Genome Analysis*. **141**:811-824.
- Pavy, N., Laroche, J., Bousquet, J. and Mackay, J. 2005. Large-scale statistical analysis of secondary xylem ESTs in pine. *Plant Molecular Biology*. **5**:203-224.
- Payne, T.C. 2009. Enzymes. In Tarte (ed.). *Ingredients in Meat Products: Properties, Functionality and Applications*, pp. 107-183. Springer Science.
- Payton, P., Kottapalli, K.R., Rowland, D., Faircloth, W., Guo, B., Burow, M., Puppala, N. and Gallo, M. 2009. Gene expression profiling in peanut using high density oligonucleotide microarrays. *BMC Genomics*. **10**:265-276.
- Pei, Z-M., Ghassemian, M., Kwak, C.M., McCourt, P. and Schoeder, J.I. 1998. Role of farnesyltransferase in ABA regulation of guard cell anion channels and plant water loss. *Science*. **282**:287-290.
- Pereira, A. 1999. Plant genomics is revolutionizing agricultural research. *Biotechnology and Development Monitor*. **40**:2-7.
- Pereira, J.F., Zhou, G., Delhaize, E., Richardson, T., Zhou, M. and Ryan, P.R. 2010. Engineering greater aluminium resistance in wheat by over-expressing *TaALMT1*. *Annals of Botany*. **106**:205-214.
- Perez, F., Ortiz, J., Zhinaula, M., Gonzabay, C., Calderon, J. and Volckaert, F.A. 2005. Development of EST-SSR markers by data mining in three species of shrimp, *Litopenaeus vannamei*, *Litopenaeus stylirostris* and *Trachypenaeus birdy*. *Marine Biotechnology*. **7**:554-569.
- Pevsner, J. 2009. *Bioinformatics and functional genomics*. Canada: Wiley-Blackwell.
- Pilati, S., Perazzolli, M., Malossini, A., Cestaro, A., Dematte, L., Fontana, P., Ri, A.D., Viola, R., Velasco, R and Moser, C. 2007. Genome-wide transcriptional analysis of grapevine berry ripening reveals a set of genes similarly modulated during three seasons and the occurrence of an oxidative burst at véraison. *BMC Genomics*. **8**:428-450.
- Platt, A.R., Woodhall, R.W. and George, A.L. 2007. Improved DNA sequencing quality and efficiency using an optimized fast cycle sequencing protocol. *Biotechniques*. **43**:58-62.
- Pop, M. and Salzberg, S.L. 2008. Bioinformatics challenges of new sequencing technology. *Trends in Genetics*. **24**:142-149.
- Pracaya, I. 1982. *Bertanam Nenas*. Jakarta: Pt Penebar Swadaya.

- Primrose, S.B. 1998. Principle of genome analysis. United Kingdom: Blackwell Science.
- Prober, J.M., Trainor, G.L., Dam, R.J., Hobbs, F.W., Robertson, C.W., Zagursky, R.J., Cocuzza, A.J., Jensen, M.A. and Baumeister, K. 1987. A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science*. **238**:336-341.
- Puig, S. and Thiele, D.J. 2002. Molecular mechanisms of copper uptake and distribution. *Current Opinion in Chemical Biology*. **6**:171-180.
- Punja, Z.K. and Zhang, Y-Y. 1993. Plant chitinases and their roles in resistance to fungal diseases. *The journal of Nematology*. **25**:526-540.
- Purseglove, J.W. *Tropical crops:monocotyledons*. pp.74-91. London:Longman.
- Quilang, J., Wang, S., Li, P., Abernathy, J., Peatman, E., Wang, Y., Wang, L., Shi, Y., Wallace, R., Guo, X. and Liu, Z. 2007. Generation and analysis of ESTs from the eastern oyster, *Crassostrea virginica* Gmelin and identification of microsatellites and SNP markers. *BMC Genomics*. **8**:157-178.
- Qureshi, S.N., Saha, S., Kantety, R.V. and Jenkins, J.N. 2004. EST-SSR: A new class of genetic markers in cotton. *The Journal of Cotton Science*. **8**:112-123.
- Rani, D.S. and Nand, K. 2004. Ensilage of pineapple processing waste for methane generation. *Waste Management*. **24**:523-528.
- Rébeillé, F., Ravanel, S., Jabrin, S., Douce, R., Storozhenko, S. and Van Der Straeten, D. 2006. Folates in plants: biosynthesis, distribution and enhancement. *Physiologia Plantarum*. **126**:330-342.
- Richmond, T. and Somerville, S. 2000. Chasing the dream: plant EST microarrays. *Current Opinion in Plant Biology*. **3**:108:116.
- Rohrer, G.A., Fahrenkrug, S.C., Nonneman, D., Tao, N. and Warren, W.C. 2002. Mapping microsatellites markers identified in porcine EST sequences. *Animal Genetics*. **33**:372-376.
- Ronning, C.M., Stegalkina, S.S., Ascenzi, R.A., Bougri, O., hart, A.L., Utterbach, T.R., Vanaken, S.E., Riedmuller, S.B., White, J.A., Cho, J., Perte, G.M., Lee, Y., Karamycheva, S., Sultana, R., Tsai, J., Quackenbush, J., Griffiths, H.M., Restrepo, S., Smart, C.D., Fry, W.E., van der Hoeven, R., Tanksley, S., Zhang, P., Jin, H., Yamamoto, M.L., Baker, B.J. and Buell, R. 2003. Comparative analyses of potato expressed sequence tag libraries. *Genome Analysis*. **131**:419-429.
- Rogers, Y-H. and Venter, C. 2005. Massively parallel sequencing. *Nature*. **437**:326-327.

- Ronaghi, M. 2001. Pyrosequencing sheds light on DNA sequencing. *Genome Research*. **11**:3-11.
- Rose, J.K.C., Hadfield, K.A., Labavitch, J.M. and Bennett, A.B. 1998. Temporal sequence of cell wall disassembly in rapidly ripening melon fruit. *Plant Physiology*. **117**:345-361.
- Rowan, A.D., Buttle, D.J. and Barrett, A.J. 1990. The cysteine proteinases of the pineapple plant. *Biochemistry Journal*. **266**:869-875.
- Rounsley, S., Lin, X. and Ketchum, K.A. 1998. Large-scale sequencing of plant genomes. *Current Opinion in Plant Biology*. **1**:136-141.
- Rozen, R. and Skaletsky, H.J. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. pp 365-386. Totowa, Humana Press.
- Ruiz Martinez, M.C., Berka, J., Belenkii, A., Foret, F., Miller, A.W. and Karger, B.L. 1993. DNA sequencing by capillary electrophoresis with replaceable linear polyacrylamide and laser-induced fluorescence detection. *Analytical Chemistry*. **65**:2851-2858.
- Saha, S., Sparks, A.B., Rago, C., Akmaev, V., Wang, C.J., Vogelstein, B., Kinzler, K.W. and Velculescu, V.E. 2002. Using the transcriptome to annotate the genome. *Nature Biotechnology*. **20**:508-512.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA Polymerase. *Science*. **239**:487-491.
- Salem, M., Rexroad III, G.E., Wang, J., Thorgaard, G.H. and Yao, J. 2010. Characterization of the rainbow trout transcriptome using Sanger and 454-pyrosequencing approaches. *BMC Genomics*. **11**:564-588.
- Sanewski, C. and Scott, C. 2000. The Australian Pineapple Industry. *Acta Horticulturae*. **529**: 53-56.
- Sanger, F. 1981. Determination of nucleotide sequences in DNAs. *Bioscience Reports*. **24**:239-253.
- Sanger, F., Nicklen S. and Coulson, R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci*. **74**:5463-5467.
- Sawano, Y., Muramatsu, T., Hatano, K., Nagata, K. and Tanokura, M. 2002. Characterization of genomic sequence coding for bromelain inhibitors in pineapple and expression of its recombinant isoform. *The Journal of Biological Chemistry*. **277**:28222-28227.

- Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. 1995. Quantitative monitoring gene expression patterns with a complementary DNA microarray. *Science*. **270**:467-470.
- Schieber, A., Stintzing, F.C. and Carle, R. 2001. By products of plant food processing as a source of functional compounds-recent developments. *Trends in Food Science and Technology*. **12**:401-413.
- Schlötterer, C. and Tautz, D. 1992. Slippage synthesis of simple sequence DNA. *Nucleic Acids Research*. **20**:211-215.
- Schneiderbauer, A., Sandermann, H., Jr. and Ernst, D. 1991. Isolation of functional RNA from plant tissues rich in phenolic compounds. *Analytical Biochemistry*. **197**:91-95.
- Scott, K.D., Eggler, P., Seaton, G., Rossetto, M., Ablett, E.M., Lee, L.S., Henry, R.J. 2000. Analysis of SSRs derived from grape ESTs. *Theoretical and Applied Genetics*. **100**:723-726.
- Serapion, J., Kucuktas, H., Feng, J. and Liu, Z. 2004. Bioinformatic mining of Type I microsatellites from expressed sequence tags of channel catfish (*Ictalurus punctatus*). *Marine Biotechnology*. **6**:364-377.
- Sharma, A.D., Gill, P.K. and Singh, P. 2003. RNA isolation from plant tissues rich polysaccharides. *Analytical Biochemistry*. **314**:319-321.
- Sharma, H.C., Crouch, J.H., Sharma, K.K., Seetharama, N. and Hash, C.T. 2002. Applications of biotechnology for crop improvement: prospects and constraints. *Plant Science*. **163**: 381-395.
- Shendure, J. and Ji, H. 2008. Next-generation DNA sequencing. *Nature Biotechnology*. **26**:1135-1145.
- Simko, I. 2009. Development of EST-SSR markers for the study of population structure in lettuce (*Lactuca sativa* L.). *Journal of Heridity*. **100**:256-262.
- Smith, A.G., Croft, M.T., Moulin, M and Webb, M.E. 2007. Plants need their vitamins too. *Current Opinion in Plant Biology*. **10**:266-275.
- Sinden, R.R. 2001. Origins of instability. *Nature*. **411**:757-758.
- Stallings, R.L., Torney, D.C., Hildebrand, C.E., Longmire, J.L., Deaven, L.L., Tett, J.H, Doggett, N.A. and Moyzis, R.K. 1990. Physical mapping of human chromosomes by repetitive sequence fingerprinting. *Proceedings of the National Academy of Science*. **87**:6218-6222.
- Stea, T.H., Johansson, M., Jägerstad, M. and Frølich, W. 2006. Retention of folates in cooked, stored and reheated peas, broccoli and potatoes for use in modern large-scale service systems. *Food Chemistry*. **101**:1095-1107.

- Stover, P.J. 2004. Physiology of folate and vitamin B₁₂ in health and disease. *Nutrition Reviews*. **62**:S3-S12.
- Subramanian, S., Madgula, V.M., George, R., Mishra, R.K., Pandit, M.W., Kumar, C.S. and Singh, L. 2003. Triplet repeats in human genome: distribution and their association with genes and other genomic regions. *Bioinformatics*. **19**:549-552.
- Suzuki, Y., Kawazu, T. and Koyama, H. 2004. RNA isolation from siliques, dry seeds, and other tissues of *Arabidopsis thaliana*. *Biotechniques*. **37**:542-544.
- Takasuga, A., Hirotsune, S., Itoh, R., Jitohzono, A., Suzuki, H., Aso, H. and Sugimoto, Y. 2001. Establishment of a high throughput EST sequencing system using Poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs. *Nucleic Acid Research*. **29**:e108.
- Tanaka, K. Hilary, Z.D. and Ishizaki, A. 1999. Investigation of the utility of pineapple juice and Pineapple Waste as low-cost substrate for ethanol fermentation by *Zymomonas mobilis*. *Journal of Bioscience and Bioengineering*. **87**:642-646.
- Taussig, S.T. and Batkin, S. 1987. Bromelian, the enzyme complex of pineapple (*Ananas comosus* and its clinical application. An update. *Journal of Ethnopharmacology*. **22**:191-203.
- Tassanakajon, A., Klinbunga, S., Paunglarp, N., Rimphanitchayakit, V., Udomkit, A., Jitrapakdee, S., Sritunyalucksana, K., Phongdara, A., Pongsomboon, S., Supungul, P., Tang, S., Kuphanumart, K., Pichyangkura, R. and Lursinsap, C. 2006. *Panaeus monodon* gene discovery project: The generation of an EST collection and establishment of a database. *Gene*. **384**:104-112.
- Terrier, N., Ageorges, A., Abbal, P. Romieu, C. 2001. Generation of ESTs from grape berry at various development stages. *Journal of Plant Physiology*. **158**:1575-1583.
- Terol, J., Conesa, A., Colmenero, J.M., Cercos, M., Tadeo, F., Agusti, J., Alós, E., Soler, G., Brumos, J., Iglesias, D.J., Götz, S., Legaz, F., Argout, X., Courtois, B., Ollitrault, P., Dossat, C., Wincker, P., Morillon, R. and Talon, M. 2007. Analysis of 13 000 unique Citrus clusters associated with fruit quality, production and salinity tolerance. *BMC Genomics*. **8**:31-53.
- Tesfaye, M., Temple, S.J., Allan, D.L., Vance, C.P. and Samac, D.A. 2001. Overexpression of malate dehydrogenase in transgenic alfalfa enhances organic acid synthesis and confers tolerance to alluminium. *Plant Physiology*. **127**:1836-1844.
- Thiel, T., Michalek, W., Varshney, R.K. and Graner, A. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-

- markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics*. **106**:411-422.
- Tochi, B.N., Wang, Z., Xu, S-Y. and Zhang W. 2008. Therapeutic application of pineapple protease (Bromelian): A Review. *Pakistan Journal of Nutrition*. **7**:513-520.
- Tóth, G., Gáspári, Z. and Jurka, J. 2000. microsatellites in different eukaryotes genomes: survey and analysis. *Genome Research*. **10**:967-981.
- Udall, J.A., Flagel, F.E., Cheung, F., Woodward, A.W., Hovav, R., Rarr, R.A., Swanson, J.M., Jinsuk, J.L., Gingle, A.R., Nettleton, D., Town, C.D., Chen, Z.J. and Wendel, J.F. 2007. Spotted cotton oligonucleotide microarrays for gene expression analysis. *BMC Genomics*. **8**:81-89.
- Udall, J.A., Swanson, J.M., Haller, K., Rapp, R.A., Sparks, M.E., Hatfield, J., Yu, Y., Wu, Y., Dowd, C., Arpat, A.B., Sickler, B.A., Wilkins, T.A., Guo, J.Y., Chen, X.Y., Scheffer, J., Taliercio, E., Turley, R., McFadden, H., Payton, P., Klueva, N., Allen, R., Zhang, D., Haigler, C., Wilkerson, C., Suo, J., Schulze, S.R., Pierce, M.L., Essenberg, M., Kim, H., Llewellyn, D.J., Dennis, E.S., Kudrna, D., Wing, R., Paterson, A.H., Soderlund, C. and Wendel, J.F. 2006. A global assembly of cotton ESTs. *Genome Research*. **16**:441-450.
- Vahteristo, L., Lehtikoinen, K., Ollilainen, V. Varo, P. 1997. Application of an HPLC assay for the determination of folate derivatives in some vegetables, fruits and berries consumed in Finland. *Food Chemistry*. **59**:589-597.
- van Hal, N.L.W., Vorst, O. van Houwelingen, A.M.M.L., Kok, E.J., Peijnenburg, A., Aharoni, A., van Tunen, A.J. and Keijer, J. 2000. The application of DNA microarrays in gene expression analysis. *Journal of Biotechnology*. **78**:271-280.
- Van Hoof, N.A., Hassinen, V.H., Hakvoort, H.W., Ballintijn, K.F., Schat, H., Verkleij, J.A., Ernst, W.H., Karenlampi, S.O. and Tervahauta, A.I. 2001. Enhanced copper tolerance in *Silene vulgaris* (Moench) Garcke populations from copper mines is associated with increase transcripts levels of a 2b-type metallothionein gene. *Plant Physiology*. **126**:1519-1526.
- Varshney, R.K., Graner, A. and Sorrells, M.E. 2005a. Genic microsatellite markers in plants: features and applications. *TRENDS in Biotechnology*. **23**:48-55.
- Varshney, R.K., Sigmund, R., Börner, A., Korzun, V., Stein, N., Sorrells, M.E., Langridge, P. and Graner, A. 2005b. Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. *Plant Science*. **168**:195-202.
- Varshney, R.K., Thiel, T., Stein, N., Langridge, P., and Graner, A. 2002. In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species. *Cellular and Molecular Biology Letters*. **7**:537-546.

- Vera, J.C., Wheat, C.W., Fescemyer, H.W., Frilander, M.J., Crawford, D.L., Hanski, I. and Marden, J.H. 2008. Rapid transcriptome characterization for a non-model organism using 454 pyrosequencing. *Molecular Ecology*. **17**:1636-1647.
- Wang, Z., Fang, B., Chen, J., Zhang, X., Luo, Z., Huang, L., Chen, X. and Li, Y. 2010. *De novo* assembly and characterization of root transcriptome using Illumina paired-end sequencing and development of cSSR markers in sweetpotato (*Ipomoea batatas*). *BMC Genomics*. **11**:726-755.
- Webb, M. 1987. Toxicological significance of metallothionein. *Experientia Supplementum* **52**:109-134.
- Weber, A.P.M., Weber, K.L., Carr, K., Wilkerson, C., Ohlrogge, J.B. 2007. Sampling the Arabidopsis transcriptome by massively parallel pyrosequencing. *Plant Physiology*. **144**: 32-42.
- Whetten, R. and Sederoff, R. 1995. Lignin Biosynthesis. *The Plant Cell*. **7**:1001-1013.
- Whetten, R.W., MacKay, J.J. and Sederoff, R. 1998. Recent advances in understanding lignin biosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology*. **49**:585-609.
- White, J.A., Todd, J., Newman, T., Focks, N., Girke, T., de Ilárduya, O.M., Jaworski, J.G., Ohlrogge, J.B. and Benning, C. 2000. A new set of Arabidopsis Expressed Sequence Tags from developing seeds. The metabolic pathway from carbohydrates to seed oil. *Plant Physiology*. **124**:1582-1594.
- White, P.J. 2002. Recent advances in fruit development and ripening: an overview. 2002. *Journal of Experimental Botany*. **53**:1995-2000.
- Wintz, H., Fox, T. and Vulpe, C. 2002. Responses of plants to iron, zinc and copper deficiencies. *Biochemical Society Transactions*. **30**:766-768.
- Wintz, H. and Vulpe, C. 2002. Plant copper chaperones. *Biochemical Society Transactions*. **30**:732-735.
- Wong, L.Y., Hafeman, A. Boyd, V.L., Bodeau, J., Lazaruk, K.D., Liew, S.N., Casey, Belonogoff, V., Bit, S., Sumner, C., Bredo, A., Ho, N., Chu, E., Olson, S., Rabkin, S., Maltchenko, S., Spier, G., Gilbert, D. and Baumhueter, S. 2003. Assessing gene expression variation in normal human tissues using GeneTag™, a novel, global, sensitive profiling method. *Journal of Biotechnology*. **101**:199-217.
- Yamamoto, K. and Sasaki, T. 1997. Large-scale EST sequencing in rice. *Plant Molecular Biology*. **35**:135-144.

- Yan, Q., Zhang, Y., Li, H., Wei, C., Niu, L., Guan, S., Li, S. and Du, L. 2008. Identification of microsatellites in cattle unigenes. *Journal of Genetics and Genomics*. **35**:261-266.
- Yang, S.F. and Hoffman, N.E. 1984. Ethylene biosynthesis and its regulation in higher plants. *Annual Review of Plant Physiology*. **35**:155-189.
- Yosodha, R., Sumathi, R., Chezhian, P., Kavitha, S. and Ghosh, M. 2008. *Eucalyptus* microsatellites mined *in silico*: survey and evaluation. *Journal of Genetics*. **87**:21-25.
- Yu, J.-K., La Rota, M., Kantety, R.V. and Sorrells, M.E. 2004. EST derived SSR markers for comparative mapping in wheat and rice. *Molecular Genetics and Genomics*. **271**:742-751.
- Zerbino, D.R. and Birney, E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Research*. **18**:821-829.
- Zhang, L., Yuan, D., Yu, S., Li, Z., Cao, Y., Miao, Z., Qian, H. and Tang, K. 2004. Preference of simple sequence repeats in coding and non-coding regions of *Arabidopsis thaliana*. *Bioinformatics*. **20**:1081-1086.
- Zhang, X.M., Dou, M.A., Yao, Y.L., Du, L.Q., Li, J.G. and Sun, G.M. 2011. Dynamic of sugar metabolism in different harvest seasons of pineapple (*Ananas comosus* L. (Merr.)). *African Journal of Biotechnology*. **10**: 2716-2723.
- Zhou, Y., Dahler, J.M., Underhill, J.R. and Wills, R.B.H. 2003. Enzyme associated with blackheart development in pineapple fruit. *Food Chemistry*. **80**:565-572.