



**Engineering yeast for the production of plant terpenoids and
characterization of their biosynthetic enzymes from *Salvia* sp.**

**Arabidopsis plants overexpressing select antioxidant genes exhibit resistance to oxidative stress and
bacterial infection**

by

Codruta Ignea

Heraklion, Greece

2009

Τροποποίηση του σακχαρομύκητα για την παραγωγή φυτικών τερπενίων και τον χαρακτηρισμό των βιοσυνθετικών του ενζύμων απο είδη *Salvia*

Φυτά *Arabidopsis* που υπερεκφράζουν επιλεγμένα αντιοξειδωτικά γονίδια παρουσιάζουν αντοχή στην οξειδωτική καταπόνηση και την βακτηριακή μόλυνση

Acknowledgments

My endless appreciation and deep gratitude to my supervisor Prof. Dr. Nicholas Panopoulos, who has encouraged me, and supported me in every single step during the work I have done.

My sincere gratitude and special thanks for my supervisor Dr. Antonios Makris, for his exceptional ideas, valuable comments, constructive criticism and tireless support that have kept me going.

Special thanks for Dr. Kalliopi Roubelakis, Dr. Despina Alexandraki, Dr. Kostas Tokatlidis, Dr. Kriton Kalantidis and Dr. Angelos Kanellis for their valuable guidance and constructive analysis of my scientific formation.

My deep gratitude for Dr. Sotiris Kampranis and Dr. Chris Johnson, for their valuable assistance and fruitful ideas they suggested.

Recognition to the whole staff of Food Quality Management Department - MAICH, for their huge help and special thanks to Panos Kefalas and Sofia Loupassaki.

Special thanks to Mr. Alkinoos Nikolaidis, Director of Mediterranean Agronomic Institute of Chania, MAICH, for his support during the entire time.

My colleagues in the lab, specially to Mustafa, Melania, Hadi, Remig, Sari, Peter. Thank you for your support and wish you all the best in your future career.

Special thanks to my friend, Stelios Hadzidakis, for his support and straight friendship.

To my parents who encouraged me all the time and to my daughter Silvana, all my love. God bless you.

Finally, to the memory of my brother and my grandparents who never left from my heart. I miss you.

TABLE OF CONTENT

Chapter I. Introduction

1. Yeast as a cell factory	1
1.1. Cell structure	2
1.1.1. General cellular characteristics of yeast.....	2
1.1.2. Yeast cell organelles and compartments.....	2
1.2. Molecular tools available for yeast genetic engineering	7
1.2.1. Mating.....	9
1.2.2. Homologous recombination.....	9
1.2.3. Mating – type switching.....	11
1.2.4. Mutagenesis and genetic screen.....	12
1.2.5. Yeast transformation.....	14
1.2.6. Shuttle vectors.....	15
1.2.7. Yeast Two-Hybrid System.....	16
1.3. Using yeast for the production of high added value compounds	17
1.3.1. Production of heterologous proteins in yeast.....	17
1.3.2. Production of terpenoids in yeast.....	18
1.3.3. Production of carotenoids in yeast.....	19
1.3.4. Production of flavonoids in yeast.....	20
1.3.5. Production of biofuels using yeast.....	22
1.4. Ergosterol biosynthetic pathway in yeast	24
1.4.1. Sterols in yeast.....	24
1.4.2. Branch point genes of sterol biosynthetic pathway.....	25
1.4.2.1. 3-Hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase.....	28
1.4.2.2. Farnesyl pyrophosphate synthase ERG20.....	29
1.4.2.3. Squalene synthase ERG9.....	30
1.4.2.4. Delta (24)-sterol C-methyltransferase ERG6.....	31
1.5. Yeast metabolic engineering – perspective	32

2. Plant Secondary Metabolism	33
2.1. Natural Products	33
2.1.1. Primary and secondary metabolism. Principal metabolic pathways.....	34
2.2. Terpenes in Lamiaceae, a family of medical and aromatic plants	38
2.2.1. Lamiaceae family. <i>Salvia</i> species.....	38
2.2.2. Terpenes. Term and significance.....	40
2.2.3. General structure, classification.....	41
2.2.4. Monoterpenes.....	42
2.2.5. Sesquiterpenes.....	44
2.2.6. Diterpenes.....	47
2.2.7. Other terpenes, carotenoids.....	48
2.3. Terpenes synthases, terpene modification enzymes	50
2.3.1. Mechanism, structure.....	51
2.3.2. Monoterpene synthases. Mechanism of action.....	53
2.3.3. Cineole synthase.....	56
2.3.4. Sesquiterpene synthases.....	58
2.4. Terpenoid biosynthesis	61
2.4.1. Cellular structures of synthesis.....	62
2.4.2. Biochemical pathways.....	65
2.5. Engineering of terpenoids production	69
2.5.1. Taxol.....	69
2.5.2. Lycopene.....	71
2.5.3. Artemisinin	73
2.6. Future directions in metabolic engineering of terpenoids	74
2.7. Natural Products Isolation	76
2.7.1. Solid Phase Extraction.....	76
2.7.2. Gas Chromatography – Mass Spectrometry.....	78
3. Oxidative Stress in plants	79
3.1. Oxidative stress	79
3.1.1. Peroxidoredoxins (Prx).....	80
3.1.2. Acrolein as a toxicant.....	81

3.2.	<i>Arabidopsis</i> sp. As a model system to study oxidative stress.....	81
3.2.1.	Molecular basis of <i>Agrobacterium</i> -mediated transformation.....	82
4.	Aim of study.....	83

Chapter II. Materials and Methods

1.	Strains.....	85
1.1.	<i>Escherichia coli</i> strains.....	85
1.2.	<i>Saccharomyces cerevisiae</i> strains.....	85
1.3.	<i>Agrobacterium tumefaciens</i> strain.....	87
1.4.	<i>Arabidopsis thaliana</i> strain.....	87
2.	Nutritional media.....	87
2.1.	Bacterial media.....	87
2.1.1.	Luria-Bertani (LB) medium.....	87
2.1.2.	YEP medium.....	87
2.1.3.	SOC medium.....	87
2.1.4.	Bacterial glycerol stocks.....	88
2.1.5.	Antibiotics.....	88
2.2.	Yeast media.....	88
2.2.1.	YPD medium.....	88
2.2.2.	Complete minimal (CM) dropout medium.....	89
2.2.3.	Alternative carbon source.....	89
2.2.4.	FOA medium.....	89
2.2.5.	Yeast glycerol stocks.....	90
2.3.	Plant media.....	90
2.3.1.	Murashige and Skoog (MS) medium.....	90
2.3.2.	Acrolein.....	91
3.	DNA vectors.....	91
3.1.	pCRII-TOPO.....	91
3.2.	pBluescript II SK.....	92
3.3.	pDNR-Lib.....	92
3.4.	pJG4-6.....	93

3.5.	pYES2 and variants	94
3.6.	pYX143 and variants	95
3.7.	pUG27	95
3.8.	pUG72	96
3.9.	pB227/GAL-<i>cre</i>	96
3.10.	M4366 (HO-hisG-URA3-poly-HO)	96
3.11.	<i>Agrobacterium</i> dual binary vector system	97
	3.11.1. pGREEN300:35S-nos.....	97
	3.11.2. pAL154.....	98
	3.11.3. pGREEN300:35S-nos variants.....	99
4.	DNA extraction	100
4.1.	Plasmid extraction	100
	4.1.1. Bacterial plasmid extraction.....	100
	4.1.2. Yeast plasmid extraction.....	100
4.2.	Genomic DNA extraction	102
	4.2.1. Yeast genomic DNA extraction.....	102
	4.2.2. Plant genomic DNA extraction.....	103
5.	Cloning protocols	104
5.1.	Agarose gel electrophoresis	104
5.2.	Restriction digestion of plasmid DNA	105
5.3.	Vector-insert ligation	106
6.	PCR amplification	107
6.1.	PCR reaction	107
6.2.	Colony PCR	108
6.3.	Generation of double stranded oligos	108
6.4.	Purification of the PCR product	109
	6.4.1. Purification by gel extraction.....	109
	6.4.2. Purification by ethanol:chloroform precipitation.....	110
6.5.	Cloning of the amplified PCR products in pCRII-TOPO vector	111
7.	Novel DNA recombinant vectors	111
7.1.	Generation of loxP-URA3-loxP-GALp-HA integration cassette	111

7.1.1.	pYES-HA.....	111
7.1.2.	pCOD2.....	112
7.2.	Generation of a loxP-URA3-loxP-ADH-myc integration cassette.....	114
7.2.1.	pCADH-myc.....	114
7.2.2.	pCOD3.....	115
7.3.	pDNR-GAL.....	116
8.	Subcloning.....	117
8.1.	Cloning of GPP synthase into pJG4-6 and pYES-myc plasmids.....	117
8.2.	Cloning of GPP synthase into pDNR-GAL.....	117
8.3.	Cloning of GALp-GPPS-CYC1 cassette into pUG27 vector.....	118
8.4.	Cloning of <i>SfCinS</i>(RC) and <i>SfCinS</i>(RR) into pJG4-6 and pYES plasmids.....	119
8.5.	Cloning of P330 into pYES2 plasmid.....	119
8.6.	Cloning of HSP90 into pYX143HA and pYES2 plasmids.....	120
8.7.	Cloning of 215R clone into pYESmyc plasmid.....	120
9.	Transformation protocols.....	121
9.1.	Quantification of nucleic acid concentration.....	121
9.2.	Measuring the optical density of cell cultures.....	121
9.3.	Bacterial transformation.....	122
9.3.1.	Preparation of chemical competent cells.....	122
9.3.2.	Bacterial transformation.....	123
9.4.	Yeast transformation.....	123
9.5.	<i>Arabidopsis</i> plant transformation.....	124
9.5.1.	Preparation of <i>Agrobacterium</i> electrocompetent cells.....	124
9.5.2.	<i>Agrobacterium</i> transformation by electroporation.....	125
9.5.3.	<i>In planta</i> transformation.....	125
10.	Molecular manipulation of yeast strains.....	126
10.1.	Diploid strain generation by mating.....	126
10.2.	URA3 selection marker recycling by Cre-lox recombination.....	126
10.3.	Chromosomal gene upregulation by promoter integration.....	127
10.3.1.	Genome insertion of a stabilized extra copy of HMG2 under the control	

of a Galactose promoter.....	127
10.3.2. Upregulation of ERG20 gene with a GAL promoter.....	129
10.3.3. Truncation and upregulation of HMG1 gene with ADH1 promoter.....	130
10.4. Gene disruption.....	131
10.4.1. Deletion of <i>erg6</i> gene.....	131
10.4.2. Deletion of one allele of <i>erg9</i> gene.....	131
11. Yeast protein protocol.....	132
11.1. Small scale of protein expression in yeast.....	132
11.1.1. Promoter induction.....	132
11.1.2. Terpene synthase induction.....	133
11.2. Protein extraction.....	133
11.2.1. Protein extraction using glass beads.....	133
11.2.2. Protein extraction by sonication.....	133
11.3. Determination of protein concentration.....	133
11.3.1. Spectrophotometric determination of protein concentration.....	134
11.3.2. Bradford assay.....	134
11.4. SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE).....	135
11.4.1. Materials and solutions.....	135
11.4.2. Method.....	137
11.5. Western Blotting of recombinant proteins.....	139
11.5.1. Materials.....	139
11.5.2. Solutions.....	139
11.5.3. Method.....	140
12. Terpene Synthases <i>in vitro</i> assay.....	142
12.1. Materials and Solutions.....	142
12.2. Method.....	142
13. Head Space – Solid Phase Microextraction (HS-SPME).....	143
13.1. Devices.....	144
13.2. Method.....	144
14. Gas Chromatography- Mass Spectrometry (GC-MS).....	145
15. Quantification of terpenes by SPME in Gas Chromatography.....	146

15.1.	Cineole quantification.....	146
15.2.	Trans-caryophyllene quantification.....	148
16.	Yeast Two Hybrid System.....	149
16.1.	Plasmids used during yeast two hybrid approaches.....	150
16.1.1.	LexA fusion plasmids.....	150
16.1.1.1.	pEG202.....	150
16.1.1.2.	pGILDA.....	151
16.1.1.3.	pYES-LexA.....	152
16.1.2.	Activation domain fusion plasmid.....	154
16.1.2.1.	pJG4-5.....	154
16.1.2.2.	Yeast strain.....	155
16.1.2.3.	Screening libraries.....	155
16.2.	Bait protein characterization.....	155
16.2.1.	Leu requirement test.....	155
16.2.2.	Immunoblot analysis.....	156
16.3.	Interactor hunt.....	156
16.3.1.	Library transformation.....	156
16.3.2.	Primary transformants cell collection.....	156
16.3.3.	Replating efficiency determination.....	157
16.3.4.	Interacting proteins screening.....	157
16.3.5.	Plasmid isolation.....	158
16.3.6.	Positive colonies assessment.....	158
17.	Acrolein assay of transgenic plants.....	158
18.	Plant spray inoculation with <i>Pseudomonas syringae pv.tomato</i> DC3000.....	159

Chapter III. Production of plant terpenoids in yeast

1.	Expression of Cineole Synthase 1 in yeast cells.....	160
1.1.	N-terminus truncated monoterpenes are active and stable in yeast cells.....	164
2.	Establishment and optimization of an SPME-based assay for terpene detection....	169
2.1.	Influence of glycerol concentration on terpene production.....	171
3.	Assessing different yeast genetic backgrounds that produce the highest quantities	

of terpenes.....	172
4. Integration of GALp-(K6R)HMG2 construct into the HO locus of the EG60 strain.....	174
5. Terpenes Synthase genes are functional in the yeast.....	177
5.1. Monoterpene synthases are functional in yeast.....	181
5.1.1. Expression and analysis of enzyme activity of Sf 215 ORF from <i>Salvia fruticosa</i> in yeast.....	181
5.2. Sesquiterpene synthases are functional in yeast.....	187
5.2.1. Expression of clone P330 from <i>Salvia pomifera</i>	187
5.2.2. Expression of the putative terpene synthase clone 1025 from <i>Salvia fruticosa</i>	192
6. Construction of novel molecular tools for engineering yeast metabolic pathways.....	199
6.1. Generation of a recyclable cassette construct for integrating genes into chromosomal genome of yeast and expressing them under the control of the Galactose promoter.....	199
6.2. Generation of a recyclable cassette for promoter integration.....	202
6.2.1. Developing a GAL promoter-HA tag insertion cassette with a recyclable selection marker.....	202
6.2.2. Developing a ADH-myc tag insertion cassette with a recyclable selection marker.....	205
7. Yeast Strain Molecular Manipulation.....	207
7.1. Deletion of <i>erg6</i> gene in GALp-(K6R)HMG2 strain.....	208
7.2. Generation of diploid strain AM66.....	212
7.3. Upregulation of ERG20 gene by integrating a GAL promoter in one of the two alleles in a GALp-(K6R)HMG2 diploid strain.....	212
7.4. Construction of GALp-(K6R)HMG2, GALp-ERG20, $\Delta erg9$ haploinsufficient strain.....	214
7.5. Upregulation and stable truncation of HMG1 gene by integrating a ADH constitutive promoter in one of the two alleles in different diploid strains.....	217
8. Production of terpenes in modified yeast strains.....	219

8.1.	Production of monoterpenes in modified yeast strains.....	220
8.2.	Production of sesquiterpenes in modified yeast strains.....	222
8.2.1.	Expression and characterization of P330 enzyme in yeastcells.....	224
9.	Co-expression of monoterpene synthases with GPPsynthase in yeast.....	226
9.1.	Development of a recyclable cassette for recyclable integration of exogenous genes. Insertion of Galp- GPP synthase into the lox-URA pUG27 integration vector.....	227

Chapter IV. Identifying interacting proteins of Cineole Synthase

1.	Use of two hybrid system as a mean of finding interactors of Cineole Sythase full-length <i>SfCinS1</i>.....	230
1.1.	Generating the bait construct.....	230
1.2.	Screening for <i>SfCinS1</i> interacting proteins.....	231
1.3.	Essential control tests before library screening.....	231
1.3.1.	Testing the autoactivation potential of the bait (LEU2).....	232
1.3.2.	Testing expression of bait protein.....	232
1.4.	Preparation of master plates and selection based on leucine requirements..	233
1.5.	Plasmid curing.....	235
1.6.	Isolation and characterization of the library integrating clones.....	235
1.7.	Verification of the specificity of the inteactors.....	235
1.8.	cDNA clones identification.....	237
1.9.	Description of the identified interactors.....	237
1.9.1.	Chaperone Hsp90-2.....	237
1.9.2.	Putative phosphatase 2A inhibitor.....	238
1.9.3.	Glutathione S-transferase.....	239
2.	Yeast two hybrid screening of a <i>Salvia</i> library to identify interactors of RC truncated cineole synthase <i>SfCinS(RC)</i>.....	239
2.1.	Generation of CS (RC)-LexA bait plasmid.....	240
2.2.	The pilot scale of the screening.....	242
2.3.	cDNA clones identification.....	242
2.4.	Description of the identified interactors.....	243
2.4.1.	Luminal-binding protein 4 precursors (BiP4).....	243

2.4.2.	Heat shock protein 90-2.....	245
2.4.3.	SEC14 cytosolic factor family protein.....	245
3.	Assessment of <i>SfCinS1</i> activity in the presence of HSP90 interacting protein.....	246
3.1.	Cloning of Hsp90 into yeast vectors.....	246
Chapter V. Oxidative stress in plants		
1.	Expression of oxidative stress response genes in <i>Arabidopsis</i>.....	249
2.	Effect of acrolein on tomato thiredoxin peroxidase homologous.....	254
3.	Evaluation of the <i>Arabidopsis</i> transgenics' resistance to microbial pathogens.....	259
Chapter VI. Conclusions.....		
Appendix 1.....		270
Appendix 2.....		273
Appendix 3.....		287
References.....		290

LIST OF FIGURES

Figure I. 1.	Overview of diverse application areas of yeast biotechnology	1
Figure I. 2.	3-D visualization of a yeast cell	2
Figure I. 3.	Scheme of organelles and compartments in a yeast cell	3
Figure I. 4.	The yeast bud	4
Figure I. 5.	The general organization of a mitochondrion	6
Figure I. 6.	Scanning electron micrograph of budding yeast <i>Saccharomyces cerevisiae</i>	8
Figure I. 7.	Homologous recombination scheme	10
Figure I. 8.	Basics of gap-repair cloning in yeast	13
Figure I. 9.	The principle of the yeast two-hybrid system	16
Figure I. 10.	Carotenoid biosynthetic pathway	20
Figure I. 11.	Flavonoid biosynthetic pathway	21
Figure I. 12.	Route from sunlight to fuels	23
Figure I. 13.	Chemical structure of ergosterol	25
Figure I. 14.	Ergosterol biosynthetic pathway in yeast	26
Figure I. 15.	Function of Lys6 or Lys357 required the correct structure of the transmembrane region	29
Figure I. 16.	Biosynthesis of isoprenoids by mevalonate pathway	30
Figure I. 17.	Principal biosynthetic pathways of secondary metabolism	36
Figure I. 18.	<i>Salvia fruticosa</i>	38
Figure I. 19.	Representative members of various monoterpene subfamily structural types	43
Figure I. 20.	Representative sesquiterpene structure	45
Figure I. 21.	Representative diterpene structures	47
Figure I. 22.	Other representative terpene structures	49
Figure I. 23.	γ -Humulene synthase cyclization reaction mechanisms	52
Figure I. 24.	Scheme describing the reactions catalyzed by the <i>Salvia</i> monoterpene synthases and their mutants	55
Figure I. 25.	The structure of cineole synthase Sf-CinS1	56

Figure I. 26.	Amino acid sequence alignment of the C-terminal domain of monoterpene synthases from <i>Salvia</i> species	57
Figure I. 27.	The two main structural elements responsible for product specificity in Sf-CinS1	58
Figure I. 28.	The homology structural model for the γ -humulene synthase active site	59
Figure I. 29.	Overall structure of 5-epi-aristochelene	60
Figure I. 30.	Terminal cyclization steps of TEAS and HPS terpene synthases	61
Figure I. 31.	Isoprenoid biosynthesis	62
Figure I. 32.	Plant secretory structures	63
Figure I. 33.	Details on secretory gland of sage (<i>Salvia</i> sp.)	64
Figure I. 34.	Compartmentalized biosynthesis of IPP and DMAPP	66
Figure I. 35.	Overview of terpenoid biosynthesis and the generation of terpenoid starting blocks	68
Figure I. 36.	Taxol biosynthetic pathway	70
Figure I. 37.	Proposed mechanisms for the role of lycopene in preventing chronic diseases	72
Figure I. 38.	Pathway of artemisinin biosynthesis in <i>Artemisia annua</i>	73
Figure I. 39.	Unified reaction cycle of peroxiredoxines	80
Figure I. 40.	Schematic representation of a typical octopine-type Ti plasmid and the T-DNA region of a typical octopine-type Ti plasmid	82
Figure II. 1.	The pCRII-TOPO plasmid map	91
Figure II. 2.	The pBluescript II SK (+/-) plasmid map	92
Figure II. 3.	The pDNR-LIB plasmid map	92
Figure II. 4.	The pJG4-6 plasmid map	93
Figure II. 5.	The pYES2 plasmid map	94
Figure II. 6.	The pYES-myc plasmid map	94
Figure II. 7.	The pYX143 and the pYX143-HA plasmid maps	95
Figure II. 8.	The pUG27 plasmid map	95
Figure II. 9.	The pUG72 plasmid map	96
Figure II. 10.	The M4366 plasmid map	97
Figure II. 11.	The pGREEN300:35S-nos plasmid map	98

Figure II. 12.	Figure II. 12. The pAL154 plasmid map	99
Figure II. 13.	The pGREEN300:35S-nos/BI-GST and pGREEN300:35S-nos/LeTpx1 plasmid maps	99
Figure II. 14.	Schematic representation of the PCR standard reaction conditions	108
Figure II. 15.	Schematic representation of the PCR reaction conditions used to amplify HA oligo	109
Figure II. 16.	The pYES-HA plasmid map	112
Figure II. 17.	Cloning of loxP-URA3-loxP fragment into pCRII-TOPO/GALp-HA construct to generate pCOD2 integration vector	113
Figure II. 18.	The pCADH1-myc plasmid map	115
Figure II. 19.	The loxP-URA-loxP-ADH1-myc pCOD3 integration cassette	115
Figure II. 20.	The pDNR-GAL-HA plasmid map	116
Figure II. 21.	The pJG4-6/GPPS and pYES2-myc/GPPS plasmid maps	117
Figure II. 22.	The pDNR-GAL/GPPS plasmid map	118
Figure II. 23.	The pUG27-GAL-GPPS integration vector map	119
Figure II. 24.	The HO-hisG-URA3-hisG-GALp-HA plasmid map	127
Figure II. 25.	The HO-hisG-URA3-hisG-GALp-K6R HMG2-HO integration cassette	128
Figure II. 26.	The SPME fiber-holder	144
Figure II. 27.	The introduction of the SPME fiber into the head-space	145
Figure II. 28.	The graph of cineole quantification	147
Figure II. 29.	The graph of trans-caryophyllene quantification	149
Figure II. 30.	The pEG202 plasmid map	150
Figure II. 31.	The pGILDA vector map	151
Figure II. 32.	The pYES-CLexA plasmid map	152
Figure II. 33.	The pYES-CinS-CLexA plasmid map	153
Figure II. 34.	The pJG4-5 vector map	154
Figure III. 1.	Volatiles produced by yeast cells were sampled by solid-phase micro-extraction (SPME) and analyzed by Gas Chromatography	161
Figure III. 2.	Volatiles produced by yeast cells expressing Sf-CinS1(RC) were sampled by solid-phase micro-extraction (SPME) and analyzed by Gas Chromatography/ Mass Spectrometry	163

Figure III. 3.	The volatiles produced by yeast cells expressing empty vector were sampled by solid-phase micro-extraction (SPME) and analyzed by Gas Chromatography/ Mass Spectrometry	163
Figure III. 4.	1, 8- cineole synthase predicted protein diagramme	164
Figure III. 5.	Alignment of the deduced amino acid sequences of the transit peptide of the truncated <i>Salvia fruticosa</i> 1,8-cineole synthase full length <i>SfCinS1</i> (fl) and the two truncated versions <i>SfCinS1</i> (RC) and <i>SfCinS1</i> (RR)	165
Figure III. 6.	Cloning of <i>SfCinS1</i> (RR) and <i>Sf CinS1</i> (RC) fragments into yeast vector pJG4-6	165
Figure III. 7.	Immunodetection of <i>Sf-CinS1</i> (RR) and <i>Sf-CinS1</i> (RC) proteins by HA tag	166
Figure III. 8.	Cineole production by yeast strain AM63 transformed with two truncated cineole synthases <i>Sf-CinS1</i> (RR) and <i>Sf-CinS1</i> (RC) for a period of 13 days	167
Figure III. 9.	The GC-MS analysis of yeast cells expressing cineole synthase <i>Sf-CinS1</i> (RR) and evaluated by conventional enzymatic assay	168
Figure III. 10.	The GC-MS analysis of yeast cells expressing cineole synthase <i>Sf-CinS1</i> (RC) and evaluated by conventional enzymatic assay	168
Figure III. 11.	Agarose gel electrophoresis of pYES/ <i>Sf-CinS1</i> (RC)	169
Figure III. 12.	Suppression of non-specific volatile compounds from terpene profil by resuspending in enzymatic reaction buffer	170
Figure III. 13.	The average amount of cineole produced per hour by cineole synthase in EG60 yeast cell for a period of time of 14 days	171
Figure III. 14.	The influence of glycerol concentration of cineole production in EG60 yeast cell	172
Figure III. 15.	Cineole production in two different yeast strain, BY4741 and EG60	173
Figure III. 16.	Sterol biosynthetic pathway	174
Figure III. 17.	The HO-hisG-URA3-hisG-GALp-K6R HMG2-HO integration cassette	175
Figure III. 18.	Cineole production in modified yeast strain	176
Figure III. 19.	Phylogenetic tree resulting from the multiple protein alignment of the isolated cDNA clones predicted protein sequences expressed in yeast cells with other terpene synthases previously reported	178
Figure III. 20.	Alignment of deduced amino acid sequences of the terpene synthases	

	clones employed in the current study with several reported terpene synthases	180
Figure III. 21.	Phylogenetic tree resulting from the multiple protein alignment of the isolated 215 clone predicted protein sequences expressed in yeast cells with other monoterpene synthases previously reported	182
Figure III. 22.	Alignment of deduced amino acid sequences of the 215 monoterpene synthase clone from <i>Salvia fruticosa</i> with several reported monoterpene synthases	183
Figure III. 23.	Cloning of 215R into pYESmyc vector	184
Figure III. 24.	The Gas Chromatography / Mass Spectrometry analysis of volatiles produced by AM63 yeast cells expressing 215R monoterpene synthase	186
Figure III. 25.	Phylogenetic tree resulting from the multiple protein alignment of the isolated P330 clone predicted protein sequences expressed in yeast cells with other sesquiterpene synthases previously reported	187
Figure III. 26.	Alignment of deduced amino acid sequences of the P330 sesquiterpene synthase clone from <i>Salvia pomifera</i> with several reported sesquiterpene synthases	189
Figure III. 27.	Cloning of P330 into yeast vector pYES2	189
Figure III. 28.	GS-MS profile of the sesquiterpenes produced by AM63 yeast strain expressing P330 clone	190
Figure III. 29.	Mass spectra of target products	191
Figure III. 30.	Phylogenetic tree resulting from the multiple protein alignment of the isolated 1025 clone predicted protein sequences expressed in yeast cells with other sesquiterpene synthases previously reported	192
Figure III. 31.	Alignment of deduced amino acid sequences of the 1025 sesquiterpene synthase clone from <i>Salvia fruticosa</i> with several reported sesquiterpene synthases	194
Figure III. 32.	The GC-MS profile of the sesquiterpene produced by AM63 strain expressing 1025 clone after three days of incubation	195
Figure III. 33.	The GC-MS analysis of yeast cells expressing 1025 and evaluated by conventional enzymatic assay	196

Figure III. 34.	The GC-MS profile of the sesquiterpene produced by 1025 clone in AM74 and AM75 strains	197
Figure III. 35.	Mass spectra of target products detected in AM75 strain expressing 1025 protein	198
Figure III. 36.	The pDNR-GAL-HA plasmid map	200
Figure III. 37.	Agarose gel electrophoresis of pYES-HA vector	201
Figure III. 38.	PCR product of GAL-HA-cyc1 amplification on agarose gel electrophoresis	201
Figure III. 39.	Cloning of GALp-HA-cyc1 cassette into pDNR vector	202
Figure III. 40.	LoxP-URA-loxP-GALp-HA pCOD2 integration cassette	203
Figure III. 41.	PCR product of GALp-HA amplification on agarose gel electrophoresis	203
Figure III. 42.	Cloning of GALp-HA fragment into pCRII-TOPO vector	204
Figure III. 43.	Cloning of GALp-HA-cyc1 cassette into pUG27 vector	204
Figure III. 44.	The loxP-URA-loxP-ADH1-myc pCOD3 integration cassette	205
Figure III. 45.	The pCADH1-myc plasmid map	206
Figure III. 46.	Cloning of ADH fragment into pYESmyc vector	206
Figure III. 47.	Cloning of ADHmyc fragment into pCOD2 vector	207
Figure III. 48.	Deletion scheme of <i>ERG6</i> gene	209
Figure III. 49.	PCR amplification of pUG27 disruption cassette	210
Figure III. 50.	Cineole production in modified yeast strain	211
Figure III. 51.	PCR product of pCOD2 resolved on agarose gel electrophoresis	212
Figure III. 52.	Immunodetection of ERG20 protein by HA tag	213
Figure III. 53.	PCR confirmation of GALp-HA integration upstream of <i>erg20</i> gene in AM67 strain	213
Figure III. 54.	Deletion scheme of ERG9 gene	215
Figure III. 55.	PCR product of pUG27 resolved on agarose gel electrophoresis	216
Figure III. 56.	PCR confirmation of <i>erg9</i> gene disruption in AM70 strain	217
Figure III. 57.	The loxP-URA3-loxP-ADH1-myc-HMG1 truncated linear map	218
Figure III. 58.	PCR product of pCOD3 resolved on agarose gel electrophoresis	219
Figure III. 59.	Metabolic engineering of sterol biosynthetic pathway	220
Figure III. 60.	Cineole production in modified yeast strains	221

Figure III. 61.	Production of caryophyllene in modified yeast strains	223
Figure III. 62.	The activity of P330 sesquiterpene synthase in modified yeast strains	224
Figure III. 63.	Effect of GPP synthase expression and Sf-Cin1 (RC) co-expression on cineole production in AM63 cells	227
Figure III. 64.	Cloning of GPPS fragment into pUG27 integration vector	228
Figure III. 65.	Maps of GAL-GPPS-cyc1 cassette	229
Figure IV. 1.	Subcloning of <i>SfCinS1</i> into pGILDA	231
Figure IV. 2.	The self activation assay of the bait proteins	232
Figure IV. 3.	Expression of LexA- <i>SfCinS1</i> recombinant protein in yeast cells. Western blot from yeast protein extract using anti-LexA antibody recognized the N-terminal fusion of the <i>SfCinS1</i> protein	233
Figure IV. 4.	Comparison between colonies grown on solid medium containing galactose and raffinose (where the gene coding for leucine gene was triggered) and glucose plates (where growth on leucine lacking plates was the result of non-specific activation of the reporter gene)	234
Figure IV. 5.	Separation of interacting clones after double digestion with <i>EcoRI</i> and <i>XhoI</i> by agarose gel electrophoresis	235
Figure IV. 6.	Interaction-dependent phenotype of 4 positive cDNA Tomato library clones and <i>SfCinS1</i>	236
Figure IV. 7.	Cloning of LexA fragment into pYES2 vector	240
Figure IV. 8.	Cloning of Sf-CinS1 fragment into pYES-LexA vector	241
Figure IV. 9.	The volatiles produced by yeast cells expressing Sf-CinS1(RC) subcloned in pYES-LexA vector were sampled by solid-phase micro-extraction (SPME) and analyzed by Gas Chromatography	241
Figure IV. 10.	Interaction-dependent phenotype of 5 positive cDNA Salvia library clones and <i>SfCinS(RC)</i>	242
Figure IV. 11.	Cloning of AtHSP90-2 into pYES and pYX143HA yeast vectors	247
Figure IV. 12.	Cineole yield by yeast strain AM63 engineered with <i>Sf-CinS1(RC)</i> and HSP90, empty plasmid	248
Figure V. 1.	Colony PCR products of <i>Agrobacterium</i> cells	250
Figure V. 2.	Basta selection process	251

Figure V. 3.	<i>Arabidopsis thaliana</i> genomic DNA isolated by agarose gel electrophoresis	252
Figure V. 4.	PCR amplification of the bacterial bialophos resistance gene (BAR) using <i>Arabidopsis thaliana</i> genomic DNAs as a template	253
Figure V. 5.	Third generation of <i>Arabidopsis thaliana</i> transgenic plants	254
Figure V. 6.	<i>Arabidopsis thaliana</i> wild type seedlings on the 10 th day of growth on MS media supplemented with different concentration of acrolein	255
Figure V. 7.	Germination rate of <i>Arabidopsis thaliana</i> wild type, <i>Arabidopsis thaliana</i> LeTpx-wt, and <i>Arabidopsis thaliana</i> LeTpx-76CS on the 5 th day of growth on MS media supplemented with different concentration of oxidants	256
Figure V. 8.	The effect of acrolein on the rate of germination of <i>Arabidopsis thaliana</i> wild type Col-0, and LeTpx1 transgenic <i>Arabidopsis</i> seedlings of 5 days	257
Figure V. 9.	The effect of acrolein on transgenic <i>Arabidopsis</i> seedlings of 12 days, overexpressing a thioredoxin peroxidase (LeTpx1) and its mutant on 76Cys residue	258
Figure V. 10.	<i>Arabidopsis</i> plants at leaf fully expanded, prior infection with <i>Pseudomonas syringae</i>	260
Figure V. 11.	<i>Arabidopsis thaliana</i> plants infected with <i>Pseudomonas syringae</i> pv. <i>tomato</i> DC3000, 8 days after inoculation	261
Figure V. 12.	Expression analysis of <i>Arabidopsis</i> plants infected with <i>Pseudomonas</i>	262

LIST OF TABLES

Table I. 1.	Principal classes of secondary compounds	37
Table I. 2.	Ranges of the quantity of the main compounds in <i>Salvia</i> species upper leaves (% of total oil)	39
Table I. 3.	Ranges of the main essential oil components (% of the total oil) of the sage species in Greece	40
Table I. 4.	Parent hydrocarbons of terpenes (isoprenoids)	41
Table II. 1.	Yeast strains used in the present project	86
Table III. 1.	Range of variation of the quantity of the main monoterpene compounds in <i>Salvia fruticosa</i> comparing with bacterial and yeast expression product profile of Sf-CinS(RC)	162
Table III. 2.	The size of the truncated 1, 8-cineole synthase variants	166
Table III. 3.	Ranges of the quantity of monoterpene compounds in <i>Salvia</i> species upper leaves (% of total oil)	177
Table III. 4.	Ranges of the quantity of sesquiterpene compounds in <i>Salvia</i> species upper leaves (% of total oil)	178
Table III. 5.	Monoterpene profile of 215 monoterpene synthase versus SfCinS1 cineole synthase	185
Table III. 6.	Sesquiterpene profile of P330 sesquiterpene synthase	192
Table III. 7.	Sesquiterpene profile of 1025 sesquiterpene synthase (3 rd day of incubation on buffer)	196
Table III. 8.	Comparison of sesquiterpene profile of 1025 enzyme in vivo expressed in AM63 and AM74	198
Table III. 9.	Relative ratio of sesquiterpenes produced by each engineered yeast strains	225
Table III.10.	Relative yields of sesquiterpenes produced by modified yeast strains (mg/L)	225
Table IV. 1.	Cineole synthase Tomato interactors	237
Table IV. 2.	Cineole synthase <i>Salvia</i> interactors	243

ABBREVIATIONS

5-FOA	5-fluoroorotic acid
ATP	adenosine triphosphate
cDNA	complementary DNA
DMAPP	Dimethylallyl diphosphate
DXP	1-deoxyxyulose-5-phosphate
DXS, DOXP	1-deoxy-D-xylulose-5-phosphate synthase
DXR, DOXP	1-deoxy-D-xylulose-5-phosphate reductoisomerase
DTT	dithiothreitol
ERG6	delta(24)-sterol C-methyltransferase
ERG9	squalene synthase
ERG20	farnesyl pyrophosphate synthase
FPP	farnesyl-pyrophosphate
FID	flame ionisation detector
GGPP	geranylgeranyl diphosphate
GPP	geranyl diphosphate;
GC-MS	Gas Chromatography-Mass Spectrometry
GC	Gas Chromatography
GSH	glutathione
GST	glutathione S-transferase
HMG-CoA	3-hydroxy-3-methylglutaryl-CoA;
HMGR	3-hydroxy-3-methylglutaryl-CoA reductase;
HPS	henbane prenaspirodiene synthase
HR	hypersensitive response
HS-SPME	Head Space – Solid Phase Microextraction
IPP	isopentenyl diphosphate
kDa	kilo dalton
MEP	2-C-methyl-D-erythritol-phosphate
MS	Murashige and Skoog

MVA	mevalonate; The mevalonate-dependent pathway
PPi	pyro-phosphate
PR	pathogenesis-related proteins
ROS	reactive oxygen species
SfCinS1	<i>Salvia fruticosa</i> 1,8-cineole synthase
SPE	Solid Phase Extraction
SPME	Solid Phase Microextraction
sssDNA	sonicated single-stranded carrier DNA
TEAS	5-epi-aristolochene synthase
TPS	terpene synthase

ABSTRACT

Plants produce an enormous variety of low molecular weight compounds called secondary metabolites, through various biosynthetic pathways. Terpenoids and isoprenoids contribute more than 50,000 compounds to this chemical diversity which include numerous commercial flavors, fragrances and medicines. Artemisinin and taxol are such terpene-based drug compounds. In general, most useful terpenoids are produced in small quantities in plants, which has slowed considerably their commercial utilization. Research has recently targeted the development of microbial fermentative processes as an alternative approach. *Saccharomyces cerevisiae* is an amenable organism for metabolic engineering and the diversion of the metabolic machinery towards the overproduction of terpenoids compounds.

Terpene synthases, the enzymes synthesizing terpenes utilize geranyl pyrophosphate (GPP), farnesyl pyrophosphate (FPP) or geranyl-geranyl pyrophosphate (GGPP) as substrates. In yeast they are synthesized by the sterol biosynthetic pathway. Although, the whole genetic pathway is present in yeast, it is tightly regulated and the precursors are present in limited quantities. To overcome this problem, a yeast strain producing high sterol levels was identified and tested for its capacity to produce a monoterpene cineole. This was achieved by stable transformation of a plasmid carrying the cineole synthase gene, a monoterpene synthase from *Salvia fruticosa*, under the control of an inducible promoter, and facile detection of the terpene products as volatiles by Head Space - Solid Phase Microextraction (HS-SPME) coupled with Gas Chromatography/ Mass Spectrometry (GC/MS) analysis.

The selected strain was targeted for further modifications. A mutant stabilized version (K6R) of HMG2 under the control of the inducible Galactose promoter was stably integrated into the HO locus to generate strain AM63. The modified strain produces on average 1.5 fold more cineole than the parental strain and exhibited reduced background volatile metabolites when transformed with Sf-CinSyn (RC).

Further molecular engineering of AM63 yeast strain aimed to maximize terpene productivity without sacrificing cell viability which could hamper the biofermentation process. Two different modifications targeted the upregulation of FPP synthase encoded by ERG20 gene by GAL promoter integration into the yeast chromosome of AM68 strain, and deletion of one allele of *erg9* gene encoding for a squalene synthase in the diploid AM70 strain resulting in

further enhancement in terpene production. AM68 strains produces 3 fold more cineole, while AM70 cells produced 3.5 fold more sesquiterpenes than the parental strain. Additional modifications targeted the HMG1 gene by truncation of the N-terminus and expression of the one allele under a stable constitutive promoter.

The increased yield of terpenes in yeast enabled the identification of several novel terpene synthases isolated from *Salvia fruticosa* (Greek sage) and *Salvia pomifera*. Four of them failed to yield any products when tested as bacterially expressed proteins. One clone was a monoterpene synthase producing mostly cineole, distinct from the previously identified canonical *Salvia* cineole synthase, while another one was a naturally truncated form which failed to yield any products. The two other genes encoded for sesquiterpene synthases, the first producing beta-farnesene and nerolidol and the second was a multiproduct enzyme synthesizing alpha-cubebene, alpha-copaene, trans-caryophyllene and delta-cadinene.

Focusing on the terpene synthase molecule we tested whether additional modifications in the N-terminus of CS in the chloroplastic targeting sequence could further enhance product yield. The truncated *SfCinS1*(RR) and *SfCinS1*(RC) catalyzed the formation of multiple monoterpenes using the endogenous GPP pool as substrate with a significant peak of 1,8-cineole. However, *SfCinS1*(RC) was found to be more stable, efficient, with high activity during long incubation times. Additionally, using the two-hybrid system we screened a *Salvia fruticosa* glandular trichome library to identify interacting proteins to the Cineole monoterpene synthase. One of the interactors an, HSP90 when co-expressed with cineole synthase (CS) reproducibly increased product yield by 20%.

Parallel work in the context of the PENED funded project requirements focused on the *in vivo* characterization of two recently isolated enzymes involved in secondary metabolism and plant defense by expressing them in *Arabidopsis* transgenic plants. When exposed to acrolein-induced oxidative stress *Arabidopsis* plants overexpressing a thioredoxin-peroxidase transgene, LeTpx1, exhibited less sensitivity than wild type plants. Correspondingly, the BI-GST, a plant GST-like protein inhibiting Bax lethality in yeast cells, and the LeTpx1 transgenes significantly increased plant resistance to the microbial pathogen, *Pseudomonas syringe pv.tomato* DC3000.

Περίληψη

Μια σημαντική κατηγορία φυσικών προϊόντων, τα τερπενοϊδή και ισοπρενοϊδή συνισφέρουν περισσότερες απο 50,000 ουσίες στην χημική ποικιλότητα των φυτών. Πολλά τερπένια έχουν προσελκύσει εμπορικό ενδιαφέρον για τις βιομηχανικές και ιατρικές τους χρήσεις. Ένα σημαντικό εμπόδιο στην επέκταση της χρήσης τους, είναι οι περιορισμένες ποσότητες που παράγουν τα φυτά. Ο *Saccharomyces cerevisiae* είναι οργανισμός που επιτρέπει την μεταβολική μηχανική με στόχο την παραγωγή φυτικών τερπενών. Οι συνθάσες τερπενίων είναι τα ένζυμα που συνθέτουν τερπένια αξιοποιώντας το διφωσφοπικό γερανύλιο (GPP), το διφωσφορικό φαρνεσύλιο (FPP) ή το διφωσφορικό γερανυλ-γερανύλιο (GGPP) ως υποστρώματα. Στη ζύμη συνθέτονται από το μονοπάτι βιοσύνθεσης στερολών. Ένα στέλεχος ζύμης που βρέθηκε να παράγει υψηλές στερόλες χρησιμοποιήθηκε για περαιτέρω τροποποιήσεις. Ένα μετάλλαγμα του γονιδίου HMG2 (K6R), με αυξημένη σταθερότητα, ενσωματώθηκε στο HO τόπο του γονιδιώματος κάτω από τον έλεγχο του επαγόμενου εκκινητή γαλακτόζης (στέλεχος AM63).

Μια σειρά χαρακτηρισμένων και νέων συνθασών τερπενίων που απομονώθηκαν απο τη *Salvia fruticosa* (φασκόμηλο) εκφράστηκαν στο στέλεχος AM63 και το πατρικό του. Το τροποποιημένο στέλεχος παράγει κατά μέσο όρο 1,5 φορές περισσότερο από το πατρικό του και εκκλύει μειωμένους πτητικούς μη-ειδικούς μεταβολίτες. Αυτό επιτρέπει τον χαρακτηρισμό τριών νέων συνθασών τερπενίων που στο παρελθόν δεν είχαν καταλυτική δράση όταν εκφράστηκαν ως βακτηριακές πρωτεΐνες. Μία απο αυτές ήταν μια συνθάση μονοτερπενίων που παράγει κύρια σινεόλη, αλλά διαφέρει απο την ορθόλογη συνθάση σινεόλης. Δύο άλλες κωδικοποιούν για σεσκιτερπενικές συνθάσες, η πρώτη παράγει β-φαρνεσίνη και νερολιδόλη ενώ η δεύτερη δ-καδινένιο, α-κουμπεμπένιο, α-κοπαένιο και trans-καρυοφυλένιο.

Περαιτέρω μοριακή μηχανική του στελέχους AM63 στόχευσε στη μεγιστοποίηση της παραγωγικότητας σε τερπένια χωρίς να θυσιαστεί η βιωσιμότητα του οργανισμού κατά την διαδικασία της βιοζύμωσης. Δύο διαφορετικές τροποποιήσεις στόχευσαν στην υπερέκφραση της FPP συνθάσης που κωδικοποιείται από το γονίδιο ERG20 με ενσωμάτωση του GAL εκκινητή στο χρωμόσωμα του AM68 στελέχους και την αδρανοποίηση ενός αλληλόμορφου του *erg9* που κωδικοποιεί την συνθάση σκουαλενίου (στέλεχος AM70) με αποτέλεσμα την περαιτέρω αυξημένη παραγωγή τερπενίων. Το AM68 παράγει 3 φορές περισσότερη σινεόλη, ενώ το AM70

παράγει 3,5 φορές περισσότερα σεσκιτερπένια από το πατρικό τους. Περαιτέρω τροποποιήσεις στόχευσαν το HMG1 γονίδιο αποκόπτοντας το N-terminus και υπερεκφράζοντας το αλληλόμορφο με ισχυρό σταθερό εκκινητή.

Εστιάζοντας στην μονοτερπενική συνθάση ελέγξαμε κατά πόσο οι περαιτέρω τροποποιήσεις στο N-τελικό άκρο της ΣΣ στις αλληλουχίες χλωροπλαστικής στόχευσης μπορούν να επιρέασουν περαιτέρω την παραγωγή. Επιπρόσθετα χρησιμοποιώντας μια βιβλιοθήκη δύο υβριδίων απομονώσαμε πρωτεΐνες που αλληλεπιδρούν με τη συνθάση σινεόλης. Μια από αυτές η HSP90 όταν εκφράζεται μαζί με την συνθάση σινεόλης οδηγεί σε αύξηση της παραγωγής κατά 30%.

Παράλληλη δουλειά στα πλαίσια των απαιτήσεων του προγράμματος ΠΕΝΕΔ εστιάστηκε στον in-vivo χαρακτηρισμό δύο πρόσφατα απομονωθέντων ενζύμων που συμμετέχουν στο δευτερογενή μεταβολισμό και την άμυνα, σε φυτά *Arabidopsis*. Έκθεση στην ακρολεΐνη επάγει οξειδωτικό στρές. Τα φυτά *Arabidopsis* που υπερεκφράζουν την υπεροξειδάση θειορεδοξίνης LeTPX1 παρουσίασαν μικρότερη ευαισθησία από τα αγρίου τύπου φυτά. Αντίστοιχα η BI-GST μια τρανσφεράση της γλουταθειόνης και η LeTPX1 αύξησαν σημαντικά την φυτική αντοχή στον παθογόνο μικροοργανισμό *Pseudomonas syringae* pv. *tomato* DC3000.

CHAPTER I

INTRODUCTION

1. YEAST AS A FACTORY CELL

Yeasts have a long history as domesticated and cultivated organisms. The brewing of beer and wine and the leavening of bread dough are well known ‘artisanal’ applications of yeast. These early examples of yeast biotechnology have clearly contributed to the acceptance of yeasts both, as biotechnological workhorses and as a model system for detailed understanding of eukaryotic molecular cell biology and genetics (Fig. I.1). Research using yeasts has offered generic fundamental insights into processes, such as cell cycle, vesicular transport, and gene expression. Yeast has further established its ‘super model organism’ status (Barr, 2003) as an ideal platform for the development, validation and application of post-genomic technologies, such as those used in large-scale gene knockout genetics and functional genomics studies (Kumar and Snyder, 2001), large-scale analysis of the yeast transcriptome, proteome and metabolome (appropriate references are cited in subsequent sections of this Chapter).

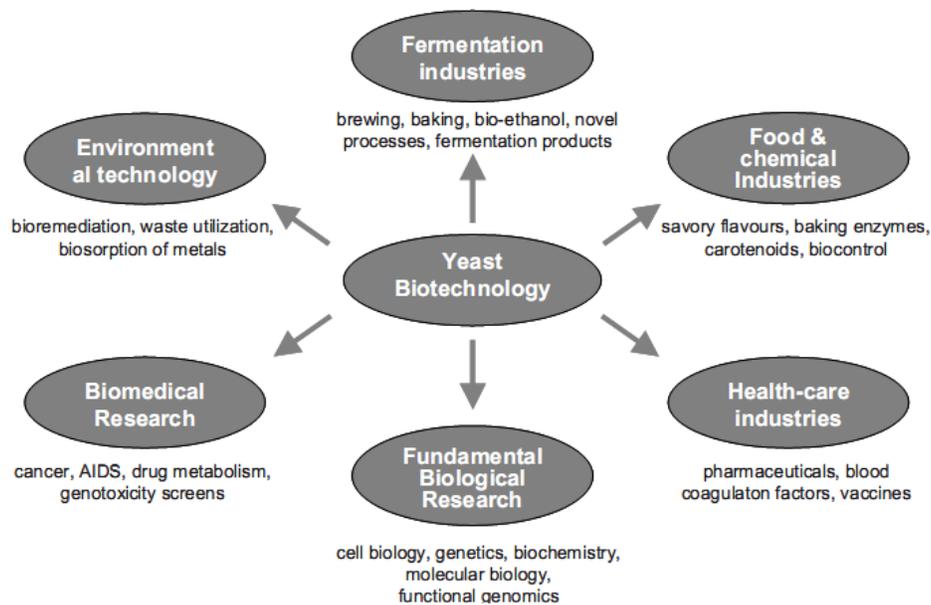


Figure I. 1. Overview of diverse application areas of yeast biotechnology (Walker, 1999).

1.1. Cell structure

1.1.1. General cellular characteristics of yeast

Yeast cells exhibit great diversity with respect to cell size, shape and color. Even individual cells from a particular yeast strain of a single species can display morphological and color heterogeneity. This is mainly due to alterations of physical and chemical conditions in the environment. Among different yeast species, cell size may vary widely (Walker, 1999).

S. cerevisiae cells are generally ellipsoidal in shape ranging from 5 to 10 μm at the large diameter and 1 to 7 μm at the small diameter. Mean cell volumes are 29 or 55 μm^3 for a haploid or a diploid cell, respectively; cell size increases with age (Fig. I. 2).



Figura I. 2. 3-D visualization of a yeast cell (Hoog *et al.*, 2007)

1.1.2. Yeast cell organelles and compartments

In an idealized yeast cell, the following ultrastructural features can be observed: cell wall; periplasm, plasma membrane, invagination, bud scar, cytosol, nucleus, mitochondrion, ER-endoplasmic reticulum, Golgi apparatus, secretory vesicles, vacuole, and peroxisome. Obviously, yeast cells share most of the structural and functional features of higher eukaryotes, which has rendered yeast an ideal model for eukaryotic cell biology. In contrast to mammalian cells, peculiarities of yeast cells are that they are surrounded by a rigid cell wall and develop birth scars during cell division; the vacuole corresponds to lysosomes in higher cells (Fig. I. 3).

The yeast **cell envelope** is a protecting capsule, consisting of three major constituents (inside out): the plasma membrane, the periplasmic space, and the cell wall. In *S. cerevisiae*, the

cell envelope takes approximately 15% of the total cell volume and has a major role in controlling the osmotic and permeability properties of the cell.

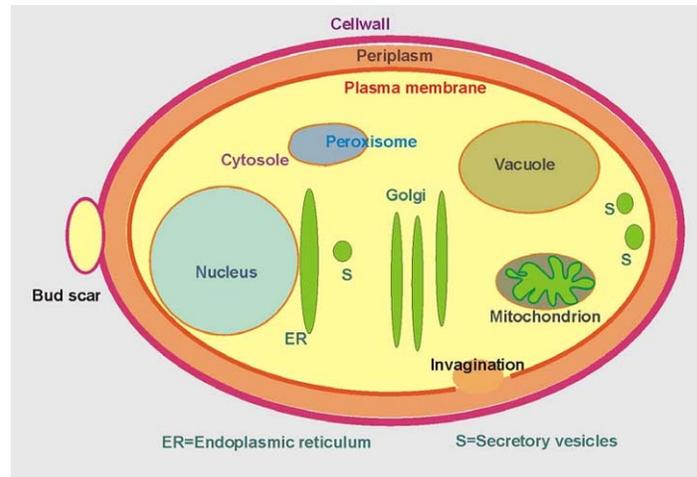


Figura I. 3. Scheme of organelles and compartments in a yeast cell (Zinser and Daum, 1995).

The **plasma membrane** is about 7 nm thick, with some invaginations into the cytosol. Like other membranes, it is a lipid bilayer with proteins inserted into this layer or traversing it as transmembrane proteins of various functions. Most important is the role of membrane proteins in regulating yeast nutrition, such as uptake of carbohydrates, nitrogenous compounds or ions, and the extrusion of molecules hazardous to the cell. Other important aspects include exo- and endocytosis of cargo molecules, stress responses, and sporulation.

The yeast **periplasm** is a thin, cell wall associated region external to the plasma membrane and internal to the cell wall.

The **wall** of a yeast cell is a remarkably thick (100 to 200 nm) envelope, which contains some 15 to 25% of the dry mass of the cell. Major structural constituents of the cell wall are polysaccharides (80-90%), mainly glucans and mannans, with a minor percentage of chitin. Glucans provide strength to the cell wall, forming a microfibrillar network (Zinser and Daum, 1995).

Bud scars are specialized, ring-shaped convex protrusions at the cell surface which remain on the mother cells (of budding yeasts) after cell division and birth of daughter cells (Fig. I. 4). The concave indentations remaining on the surface of the daughter cell after budding are called birth scars.

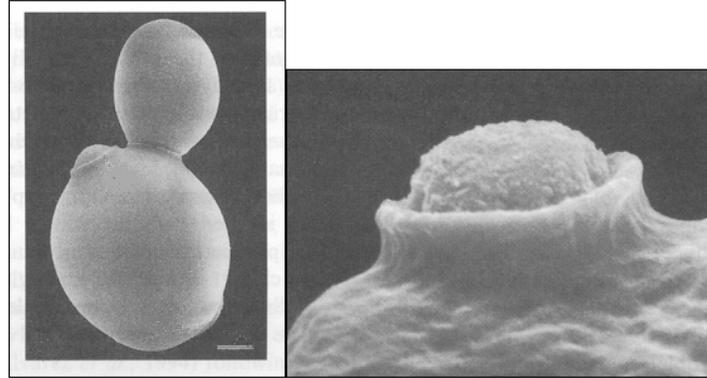


Figure I. 4. The yeast bud (Zinser and Daum, 1995).

The yeast **cytoplasm** is an acidic (pH 5.25) colloidal fluid, mainly containing ions, low or intermediate molecular weight organic compounds, and soluble macromolecules (e.g. enzyme proteins, factors, and glycogen). The cytosolic enzymes of yeast include those of: (i) the glycolytic pathway, (ii) the fatty acid synthase complex, and (iii) some enzymes for protein biosynthesis.

The **cytoskeletal network** guaranteeing internal stability to the cell and providing structural organization comprises the **microtubules** and the **microfilaments**. These are dynamic structures which fulfill their function through regulated assembly and disassembly of individual protein subunits being involved in several dynamic processes occurring during mitosis and meiosis, septation, and organelle motility.

The yeast **nucleus** is a round-lobate organelle, some 1.5 μm in diameter. The **nucleoplasm** is separated from the cytosol by a **double membrane** containing **pores** between 50 to 100 nm in diameter.

Within the nucleus there is a dense region corresponding to the **nucleolus** which disappears during mitosis and reforms during interphase. The major content of the nucleoplasm is represented by the genomic DNA which together with histones and non-histones is organized into chromatin. Yeast **chromosomes** are formed and replicated during mitosis (or meiosis) but behave virtually invisible by microscopic techniques.

In addition to the genomic material, yeast nuclei contain the machineries for DNA replication, DNA repair, transcription and RNA processing together with the necessary substrates and regulatory factors, and the resulting (precursor) products, as well as a proportion of the yeast proteasomes.

Furthermore, several non-chromosomal genetic elements may be present in the yeast nucleus (Wickner, 1995):

(i) **2 μ m DNA** is a stably maintained circular DNA plasmid, which replicates exactly once during S phase. These elements can be present in high copy number and have been useful in the construction of cloning vectors in yeast recombinant DNA technology. No functions have yet been attributed to the four genes found in 2 μ m DNA.

(ii) Double-stranded RNA and linear DNA are found in **killer strains** of yeast. They harbour genes for toxins that will be hazardous to non-killer strains.

(iii) Most interesting extrachromosomal elements are the **Ty elements**, the only class of **retrotransposons** found in yeast.

As common to all eukaryotes, yeast cells contain a system of membrane-surrounded compartments that are designed for trafficking of proteins within, into and out of the cell. In our fundamental understanding of the underlying processes and their regulation, yeast has contributed as a convenient model system (Pelham, 2001).

The **endoplasmic reticulum (ER)** is the site of biosynthesis and modification of proteins that are to be exported. After synthesis on ER-associated **polysomes** located on the surface of the ER membrane, precursor proteins are translocated into the lumen of the ER, where trimming of the precursors, chaperone-assisted folding and glycosylation of the proteins occur. From the ER, proteins are directed to the **Golgi apparatus** by **vesicles**, which fuse at the cis-side and are exported from the Golgi at the trans-side.

The key organelle in yeast involved in intracellular trafficking of proteins is the **vacuole**. It can be viewed as a form of integral component of the intramembranous system. The main role of this lysosome-like compartment is the non-specific proteolytic cleavage of proteins (Wendland *et al.*, 1998).

Peroxisomes perform a variety of metabolic functions in eukaryotic cells. In yeasts, peroxisomes contain several oxidases which serve in oxidative utilization of specific carbon and nitrogen sources.

Yeast cells contain **mitochondria** (Fig. I. 5) which structurally resemble those found in all eukaryotes. Therefore, yeast mitochondria have served as models to intensely study mitochondrial structure, function and biogenesis (Glick and Pon, 1995). Yeast mitochondria are

dynamic structures whose size, shape and number can vary greatly depending on the strain, cell cycle phase, and growth conditions (Swayne *et al.*, 2007).

General structural characteristics of mitochondria include: (i) an outer membrane – containing enzymes involved in lipid metabolism, (ii) the intermembrane space, (iii) an inner membrane – containing the components of the respiratory chain and the ATP synthase, and various membrane-integral transport proteins, and (iv) the mitochondrial matrix – containing enzymes of fatty acid oxidation, the citric acid cycle, the mitochondrial DNA together with the mitochondrial transcription and protein synthesis machineries (including mitochondrial 60S ribosomes and mitochondrial tRNAs).

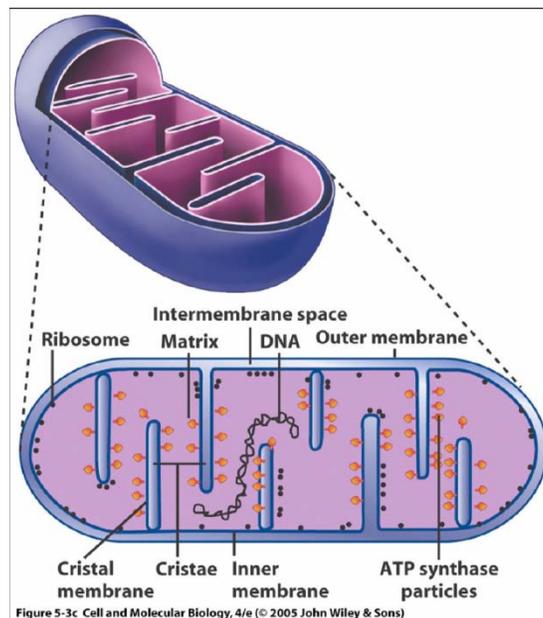


Figure I. 5. The general organization of a mitochondrion (Duchen, 2004).

Under aerobic conditions, yeast mitochondria are involved in ATP synthesis (Ackerman and Tzagoloff, 2005) coupled to oxidative phosphorylation. The activities of the citric acid cycle and the respiratory chain will largely depend on the yeast species and the expression of the Crabtree effect. This is a phenomenon related that relates glucose concentrations with the particular catabolic pathway adopted by glucose-sensitive cells, in that even in the presence of oxygen fermentation predominates over respiration. Under anaerobic conditions, mitochondria seem to be dispensable, at least for respiratory function. However, mitochondria do perform

other functions in yeast cell physiology, implicating that mitochondria are relevant to intact cell metabolism even under anaerobic conditions:

- synthesis and desaturation of fatty acids and lipids,
- biosynthesis of ergosterol,
- stress responses and adaptation to stresses,
- enzymes for the synthesis of particular amino acids and dicarboxylic acids, pyrimidines and purine bases, porphyrin, and pteridines,
- mobilization of glycogen,
- production of “flavor” components.

The importance of yeast mitochondria is best illustrated by the fact that some 8 to 10% of the nuclear yeast genes are involved in biogenesis of these organelles and maintenance of their functions. The vast majority of these proteins are synthesized by cytosolic ribosomes and become imported into yeast mitochondria, which have the potential to biosynthesize only 12 different proteins (cytochrome oxidase subunits, cytochrome b, the 6 subunits of NADH dehydrogenase, splicing factors) in addition to the mitochondrial rRNAs (15S and 26S subunits) and the complement of mitochondrial tRNAs. The biogenesis of mitochondria (Shaw and Nunnari, 2002), which involves genetic cooperativity between nuclear and mitochondrial genomes, has been widely studied in *S. cerevisiae*, since several kinds of mutations can be used in this model organism.

1.2. Molecular tools available for yeast genetic engineering

Saccharomyces cerevisiae (baker’s or budding yeast) (Fig. I. 6) is one of the most established model organisms and is widely used for studying cell physiology and molecular events (Petranovic and Nielsen, 2008). This organism has several advantages for studying cellular biology of eukaryotes. The knowledge accumulated in the last two decades demonstrates that many of the fundamental cell biological processes are highly conserved among eukaryotic cells (Mustacchi *et al.*, 2006).

Yeast is a unicellular microorganism, easy to cultivate quickly in large populations and in inexpensive media, exists both in haploid and diploid forms and provides the possibility for either sexual crossing or clonal division (budding), thereby enabling genetic manipulations and screenings (Winde, 2003). It can express heterologous genes either from an episomal plasmid or

from chromosomal integration and, furthermore, it is relatively easy to insert, delete or mutate any genomic sequence owing to the presence of a very efficient homologous recombination pathway.

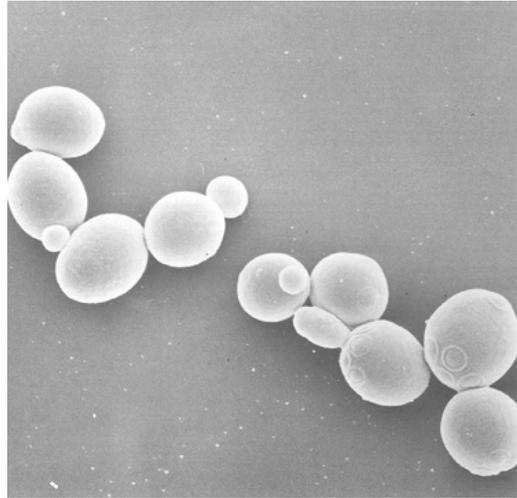


Figure I. 6. Scanning electron micrograph of budding yeast *Saccharomyces cerevisiae* (Tong *et al.*, 2006).

There are also practical benefits from the fact that this yeast has been used for many years and hence a lot of molecular biology tools and data have been amounted over time (Nielsen and Jewett, 2008). Chromosome III of *S. cerevisiae* was the first complete chromosome of any organism to be sequenced (Oliver *et al.*, 1992), and in 1996 *S. cerevisiae* became the first eukaryote organism to have its entire genome completed (Vassarotti and Goffeau, 1992). Besides the available genome sequence, a collection of single deletion mutants is available for diploid cells and for non-essential genes of haploid cells (Winzeler *et al.*, 1999). Furthermore, high-throughput data collected by functional genomic tools, such as transcriptome analysis (Lashkari *et al.*, 1997), proteome analysis (Zhu *et al.*, 2001; Usaite *et al.*, 2008), metabolome analysis (Villas-Boas *et al.*, 2005; Jewett *et al.*, 2006), flux analysis (Sauer, 2006), interactome analysis (Harbison *et al.*, 2004; Uetz and Stagljar, 2006) and locosome analysis (Huh *et al.*, 2003), are contributing to the available set of valuable information, making yeast probably the organism with the most comprehensive datasets. Furthermore, the ease with which genetic modifications can be performed on yeast means that large-scale screens of genetic interactions are possible (Sopko *et al.*, 2006), allowing the robustness of genetic interaction maps to be evaluated (Wong *et al.*, 2004)

1.2.1. Mating

Mating or breeding of laboratory strains of *S. cerevisiae* is relatively simple because these strains have straightforward genetic properties that largely follow Mendel's predictions for genetic segregation.

In most diploid laboratory strains of *S. cerevisiae* meiosis and sporulation gives rise to four-spored asci (tetrads). Each of the ascospores germinates to yield a haploid *S. cerevisiae* exhibit two mating types designated MAT \mathbf{a} and MAT α . The \mathbf{a} haploids can only mate with α haploids and vice versa. Through a process of mating type switching diploids can become $\mathbf{a/a}$ or α/α and cells homozygous for opposite mating types can mate to form tetraploids. Diploid $\mathbf{a/\alpha}$ cells are unable to mate but sporulate to segregate 2 \mathbf{a} :2 α (Sprague and Winans, 2006).

Cells of the yeast *S. cerevisiae* polarize toward an environmental signal during mating. When two yeast cells of opposite mating type (\mathbf{a} and α) come into contact, they grow toward each other in a polarized fashion (Trachtulcova *et al.*, 2004). Secretion and new cell-surface growth are concentrated in the direction of the mating partner (Tkacz and MacKay, 1979). The actin and microtubule cytoskeletons also polarize toward the mating partner (Ford and Pringle, 1991). This response facilitates efficient cell and nuclear fusion, resulting in a diploid $\mathbf{a/\alpha}$ cell (Oehlen and Cross, 1998).

Yeast mating provides an accessible genetic system for the discovery of fundamental mechanisms in eukaryotic cell fusion. Although aspects of yeast mating related to pheromone signaling and polarized growth have been intensively investigated, fusion itself is poorly understood.

1.2.2. Homologous recombination

Homologous recombination is a high fidelity and template-dependent DNA repair pathway, which serves in the non-mutagenic tolerance of DNA damage, in the repair of complex DNA damage, such as single-stranded DNA gaps, double-stranded DNA breaks and interstrand crosslinks, as well as in the recovery of stalled and collapsed replication forks (Paques and Haber, 1999). The ingenuity of this process is underscored by its conservation from bacteria to humans and its implication in a variety of unrelated nuclear processes.

Homologous recombination is implicated in the restart of stalled replication forks (McGlynn and Lloyd, 2002). It serves in the repair of DNA damage, such as single-strand gaps,

double-strand breaks and interstrand cross-links (Fig. I. 7). It is implicated in mating type switching in yeast strains and in the diversification of immunoglobulin-variable genes in vertebrates (Buerstedde *et al.*, 1990). In meiosis, the primary function of homologous recombination is to establish a physical connection between homologous chromosomes to ensure their correct disjunction at the first meiotic division. Historically prominent is the role of homologous recombination during prophase of the first meiotic division, where it contributes to high fidelity segregation of the homologs and to the generation of genetic diversity among the meiotic products (Herzberg *et al.*, 2006; Heyer *et al.*, 2006).

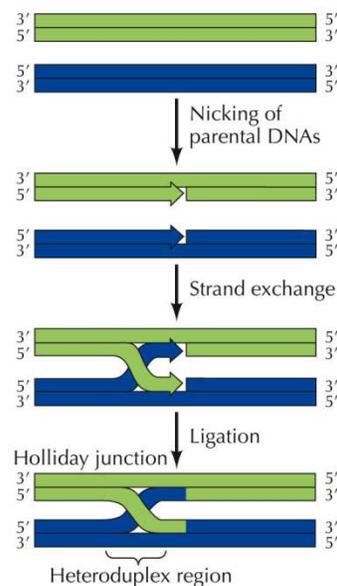


Figure I. 7. Homologous recombination scheme (Cooper, 2000).

Homologous recombination occurs due the ability of complementary sequences to align very precise and exchange fragments in a double crossover event, with no slop in the joints.

The frequency of homologous recombination is much greater in yeast than in higher eukaryotes. Therefore it has been exploited and is one of the most important tools in yeast genetics.

Homologous recombination works best with linear fragments of DNA introduced into yeast. This is true because the ends of the fragments are highly recombinogenic. But circular plasmids can also undergo homologous recombination. The precise targeted deletion of genes in yeast and their replacement with selectable markers is dependent on this phenomenon. In higher

eukaryotes, targeting of genes to known locations is one of the greatest problems for gene therapy. In yeast it is relatively easy to do. Targeted integration for creating gene-specific mutations is very efficient and requires only 50-bp fragments of target gene homology on either side of a selectable marker (Baudin *et al.*, 1993; Hua *et al.*, 1997).

Homologous or site-specific recombination in yeast are involved in several processes which are pertinent to its experimental use: including mating – type switching, gap repair mutagenesis, gene disruption or deletion.

1.2.3. Mating - type switching

The mating type of *Saccharomyces cerevisiae* is determined by a genetic locus called *MAT*. One of the two alleles, *MATa* or *MAT α* , may occur at this locus. Cells replace the allele at *MAT* with the opposite allele from one of the silent loci during an efficient event. Cells can undergo mating-type inter-conversion as often as once per generation. The HO nuclease, a site-specific double-strand endonuclease present in *S. cerevisiae*, introduces a double-strand break in the *MAT* locus of haploid yeast cells, initiating mating-type interconversion (Houston and Broach, 2006), an event which is rarely associated with crossing over. The HO gene is normally expressed in late G1 phase. After the formation of a double strand break at the *MAT* locus by the HO endonuclease, the ends are processed to form 3' single-stranded tails (Haber, 1998). The 3' single-stranded tail CEN distal to the cut site invades the donor cassette to initiate DNA synthesis by using components of both leading- and lagging-strand synthesis (Holmes and Haber, 1999). Branched DNA molecules have not been reported as intermediates during mating-type switching, suggesting that these intermediates are less stable than intermediates in meiotic recombination. The ability of mutant strains to switch mating type can also be determined by monitoring the survival of cells following HO endonuclease expression and analyzing the survivors for mating phenotype (Bakhrat *et al.*, 2004).

The HO endonuclease cut site can be inserted at other locations to create an initiation site for recombination. The advantage of this system over inducing events by irradiation is that the precise location of the initiating lesion is known. Furthermore, the high efficiency of cutting by HO endonuclease allows physical monitoring of recombination, as described for mating-type switching, and recovery of all the products of a recombination event. The Haber laboratory has used the HO system extensively to characterize the mechanisms of direct- and inverted-repeat

recombination, as well as allelic recombination in diploids (Paques and Haber, 1999; Wang *et al.*, 2004b). The high efficiency of cleavage by HO endonuclease has been exploited to identify proteins that associate with double strand breaks *in vivo* by the chromatin immunoprecipitation method (Evans *et al.*, 2000). The HO locus is not required for growth (Baganz *et al.*, 1997), and nearly all laboratory strains have a mutation at HO locus, which can be a useful target for gene integrations into yeast chromosome, offering precise information about the gene integration location.

1.2.4. Mutagenesis and Genetic screens

Study of mutant phenotypes is a fundamental method for understanding gene function (Pan *et al.*, 2004). Two common experimental goals in mutagenesis are to produce either specific or "random" mutations within a gene. DNA alterations are required for investigating, for example, structure-function relationships and essential regions of proteins, and for producing conditional mutations, such as temperature-sensitive mutation.

Specific alterations are carried out by the general procedure of oligonucleotide-directed mutagenesis (Hoffmann, 1994) that is applicable to any cloned DNA segment, including those used for yeast studies. Oligonucleotide-directed mutagenesis has been used to systemically replace amino acids within proteins, especially the replacement of charged amino acids with alanine residues. Such alanine replacements have resulted in a multitude of effects, including proteins that were unaffected, inactive and temperature sensitive.

Also, numerous general procedures for producing "random" point mutations are available, including treating plasmid DNA with hydroxylamine and mis-incorporation by PCR mutagenesis (Beukers and Ijzerman, 2005).

Gap repair mutagenesis (Ma and Ptashne, 1987) (Fig. I. 8) takes advantage of the yeast proficiency to repair a gap in a plasmid by homologous recombination if the region containing the gap and its flanking sequences are present. This region can be from the chromosome, or it can be from a linear DNA fragment transformed in with the gapped plasmid. If you can make a mutant linear fragment that overlaps a gap of the gene of insert on your plasmid, both can be transformed into the yeast and the selectable marker can be used to select for colonies that have repaired the gap (Chen and Gould, 2004). Another application, called allele rescue, is to clone mutant alleles or homologous gene regions by gap repair in the gene of interest.

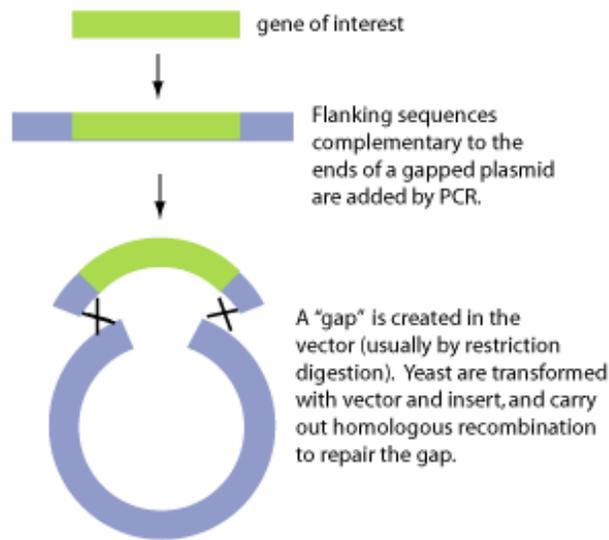


Figure I. 8. Basics of gap-repair cloning in yeast (Ma *et al.*, 1987).

By gene disruption (Wang *et al.*, 2004a), the coding sequence of the gene of interest is cut and a selectable marker is ligated in the middle of the gene. The linear DNA can be transformed into yeast cells by the Li acetate method or electroporation. Once inside, the homologous regions can line up with the chromosome copy of the gene of interest and undergo homologous recombination. In effect, this replaces the wild type copy on the chromosome with the disrupted copy that contains your selectable marker. If the disruption destroys the function of the gene of interest, a “knockout” strain of yeast is obtained, in which mutant versions of the gene of interest can be expressed on plasmid vectors. If the gene that was knockout is essential, the knockout mutations are lethal to the yeast. To avoid this, knockouts of essential genes are initially done in diploid strains. In a deletion, all or part of the coding sequence of the gene of interest may be removed.

The availability of large, well – characterized mutant strain collections and gene libraries make yeast particularly well suited to large-scale chemical-genetic screens (Armour and Lum, 2005). Completion of the yeast genome sequence in 1996 led to an effort to systematically disrupt each gene in the genome by gene replacement (Winzeler *et al.*, 1999). The resulting strain collection consists of isogenic homozygous and heterozygous mutants, each being deficient in one of the nearly 6000 genes encoded in the yeast genome. A unique DNA sequence

tag embedded in the genome of each strain can be used to quickly genotype any mutant strain by PCR. The application of innovative technologies such as DNA microarrays provides new avenues for genetic screens using this mutant strain collection (Giaever *et al.*, 2002; Butcher and Schreiber, 2004).

In forward genetics yeast mutagenesis can be effected by chemical mutagens (EMS), transposons (Hayes, 2003), and random insertion of marker genes; the last two facilitate cloning of mutants since the allele is tagged by a known sequence (inverse PCR, vectorette PCR). Reverse genetics and functional genomics involve targeted deletions in strain background that allows scoring of specific phenotype (GFP-fusion), use collection of deletion mutants covering all non-essential ORFs, and focus on genes not essential for mitotic viability (Vidan and Snyder, 2001).

1.2.5. Yeast transformation

In general, transformation is the introduction into cells of exogenously added DNA and the subsequent inheritance and expression of that DNA. The most important advances in the molecular characterization and controlled modification of yeast genes have relied on the use of shuttle vectors which can be used to transform both yeast and *E. coli*. Three main methods are available for transformation of yeast: i) those using spheroplasts; or ii) cells treated with lithium salts; and iii) the use of electroporation, the last two being the most frequently used methods.

Most investigators use cells treated with lithium salts for transformation. After treating the cells with lithium acetate, which apparently permeabilizes the cell wall, DNA is added and the cells are co-precipitated with PEG. The cells are exposed to a brief heat shock, washed free of PEG and lithium acetate, and subsequently spread on plates containing ordinary selective medium. The method using LiAc to yield competent cells (Schiestl and Gietz, 1989; Gietz and Schiestl, 2007) increased the efficiency of genetic transformation of intact cells of *Saccharomyces cerevisiae* to more than 1×10^5 transformants per microgram of vector DNA and to 1.5% transformants per viable cell. The use of single stranded, or heat denaturated double stranded nucleic acids as carrier and certain organic solvents results in about a 100-fold higher frequency of transformation with plasmids containing the 2-micron-plasmid origin of replication.

1.2.6. Shuttle vectors

Beggs (Beggs, 1978) constructed the first *S. cerevisiae*-*E. coli* shuttle vectors. These included, within a single closed-circular plasmid, the *E. coli* ColE1 plasmid replication origin, a gene for tetracycline resistance for selection in *E. coli*, the replication origin of the *S. cerevisiae* 2 µm plasmid (Broach, 1982), and the *S. cerevisiae* *LEU2* gene for selection in *S. cerevisiae* *leu2* mutant cells.

Gietz and Sugino described in 1988 the production of new alleles of the *LEU2*, *URA3* and *TRP1* genes of *Saccharomyces cerevisiae* by in vitro mutagenesis. Each new allele, which lacks restriction enzyme recognition sequences found in the pUC19 multi-cloning site, was used to construct a unique series of yeast-*E. coli* shuttle vectors derived from the plasmid pUC19. For each gene a 2 µm vector (YEplac), an ARS1 CEN4 vector (YCplac) and an integrative vector (YIplac) was constructed. The features of these vectors include (i) small size; (ii) unique recognition site for each restriction enzyme found in the pUC19 multi-cloning site; (iii) screening for plasmids containing inserts by color assay; (iv) high plasmid yield; (v) efficient transformation of *S. cerevisiae*. These vectors should allow greater flexibility with regard to DNA restriction fragment manipulation and sub-cloning (Gietz and Sugino, 1988).

To be able to move the plasmids back and forth, it is necessary to have the essential features needed by each organism on the shuttle vectors. These include the origin of replication and antibiotic resistance marker(s) for *E. coli*. In yeast all vectors will need a selectable marker that is mutated or deleted in the host. The usual strains used in the lab have about 6 selectable markers. These include *HIS3*, *URA3*, *TRP1*, *LEU2*, *ADE2* and *CAN1*, in the biosynthetic pathways for histidine, uracil, tryptophan, leucine and adenine. *CAN1* confers sensitivity to canavanine, an arginine analog that gets incorporated in proteins and is lethal to cells.

There are three types of yeast vectors for three different purposes: (i) integrative plasmids for introducing a gene into a yeast chromosome; (ii) centromeric plasmids that contain a yeast centromere and are low copy number; (iii) episomal plasmids called 2 micron plasmids from the 2 micron circles seen in some yeast strains. These are high copy number vectors in yeast (20-100 copies per cell).

YIp (yeast integrating) plasmids like YIp5 cannot survive in yeast as free plasmids because they lack an origin of replication and a centromere. They have to integrate into the host genome at some homologous sequence. These plasmids are used to carry genes into the genome so selection

does not have to be maintained, and the gene will be expressed as a single copy gene. The single copy aspect is important in some instances where gene dosage is critical. A CEN vector might provide too much of a gene if it is significantly concentration dependent. YEp (yeast episomal) plasmids such as YEp13 or YEp24 are used for overexpression of genes in yeast. They are useful for cloning high copy suppressors of temperature sensitive mutants. YCp plasmids like YCp50 are low copy (1-3 copies per cell) for expressing site-directed mutations at near-normal levels.

1.2.7. Yeast Two Hybrid Screen

The yeast two-hybrid system (Fields and Sternglanz, 1994; Colas and Brent, 1998) is a useful approach to detect novel interacting proteins, based on the fact that many eukaryotic transacting transcriptional regulators are composed of physically separable, functionally independent domains (Legrain and Selig, 2000; Ratushny and Golemis, 2008).

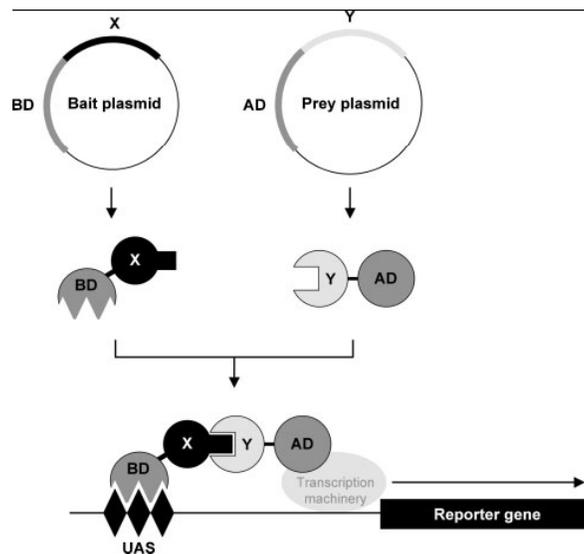


Figure I. 9. The principle of the yeast two-hybrid system. Two plasmids are constructed; the bait-encoding protein X fused to the C-terminus of a transcription factor DNA-binding domain (BD) and the prey-encoding protein Y fused to an activation domain (AD). Alternatively, the prey can consist of proteins encoded by an expression library. Each plasmid is introduced into an appropriate yeast strain either by co-transformation, sequential transformation, or by yeast mating. Only if proteins X and Y physically interact with one another are the BD and AD brought together to reconstitute a functionally active transcription factor that binds to upstream specific activation sequences (UAS) in the promoters of the reporter genes, and to activate their expression (Causier, 2003).

Such regulators often contain a DNA-binding domain (DNA-BD) that binds to a specific promoter sequence and an activation domain (AD) that directs the RNA polymerase II complex to transcribe the gene downstream of the DNA-binding site (Hope and Struhl, 1986; Keegan *et al.*, 1986; Ma and Ptashne, 1988). Both domains are required to activate a gene and, normally (as in the case of the native yeast GAL protein), the two domains are part of the same protein. The BD acts to localize the protein to specific DNA sequences within the genome, whereas the AD contacts the transcription machinery to activate gene transcription (Fig. I. 9).

If physically separated by recombinant DNA technology and expressed in the same host cell, the DNA-BD and AD peptides do not directly interact with each other and thus cannot activate the responsive genes (Brent and Ptashne, 1985; Ma and Ptashne, 1988). However, if the DNA-BD and AD can be brought into close physical proximity in the promoter region, the transcriptional activation function will be restored. In principle, any AD can be paired with any DNA-BD to activate transcription, with the DNA-BD providing the gene specificity (Brent and Ptashne, 1985).

1.3. Using yeast for the production of high added value compounds

Saccharomyces cerevisiae is a unicellular eukaryote microorganism that has traditionally been regarded either as a model system for investigating cellular physiology or as a cell factory for biotechnological use, for example for the production of fuels and commodity chemicals such as lactate or pharmaceuticals, including human insulin and HPV vaccines. Yeast has also been widely and successfully used to produce high-value pharmaceutical polypeptides, industrial enzymes and vitamins (Vandamme, 1992). In addition, this eukaryotic microorganism is emerging as a versatile host for the synthesis of lipid compounds of high commercial interest, including fatty acids, sterols, ceramides and phospholipids and other metabolites (Czabany *et al.*, 2007; Rajakumari *et al.*, 2008; Trantas *et al.* 2009).

1.3.1. Production of heterologous proteins in yeast

In addition to its application in the food and beverage industry, *S. cerevisiae* has been used for the production of protein- and small- molecule drugs. Through studies of the physiological behavior of *S. cerevisiae*, as well as the ability of this yeast to express foreign genes in conjunction with its secretory apparatus, makes *S. cerevisiae* an attractive host organism for

production of certain heterologous proteins. A number of heterologous proteins that have been used for diagnostic purposes and as human therapeutic agents and vaccines were successfully produced by *S. cerevisiae*. Human interferon was the first recombinant protein produced by *S. cerevisiae*, in 1981 (Hitzeman *et al.*, 1981) and in the following year, the hepatitis B surface antigen was produced and was the first genetically engineered vaccine that has become a safe and efficient prophylactic vaccine worldwide (Valenzuela *et al.*, 1982).

The production of a substantial part of the peptide hormone insulin by *S. cerevisiae* (Kjeldsen, 2000) covers approximately half of the insulin needed by the 154 million diabetics throughout the world. In recent years, secretion of insulin by *S. cerevisiae* has been improved by protein engineering of the leader sequence, and the improvements achieved may benefit not only insulin production but also the potential of *S. cerevisiae* as a host organism for production of other heterologous proteins.

Human insulin-like growth factor 1 (hIGF-1, or somatomedin C) is a protein of 70 amino acids purified from human serum. Native recombinant hIGF-1 (rhIGF-1) has been expressed mainly in *Escherichia coli* (Joly *et al.*, 1998), mammalian tissue cultures (Bayne *et al.*, 1987), and budding yeast (Ernst, 1986), with productions ranging between 2 and 8 mg/liter. In yeast cells, different homologous leader sequences have been tested to promote secretion of rhIGF-1, however, secretion can be achieved only using the prepro-a-factor leader sequence, which probably confers on the hIGF-1 molecule an optimal conformation, crucial for translocation (Chaudhuri *et al.*, 1992) and successfully revealed the expression and secretion of various insulin precursors.

1.3.2. Production of terpenoids in yeast

Isoprenoids (also known as terpenoids) belong to a vast group of secondary metabolites (Holstein and Hohl, 2004; Bouvier *et al.*, 2005) that include carotenoids, sterols, polyprenyl alcohols, ubiquinone (coenzyme Q), heme A and prenylated proteins. They are of valuable commercial interest as food colorants and antioxidants (carotenoids), aroma and flavor enhancers (terpenes), nutraceuticals (ubiquinone), and antiparasitic and anticarcinogenic compounds (taxol) (Lee *et al.*, 2005).

All isoprenoids are synthesized from a universal compound called isopentenyl diphosphate (IPP). In yeast, the mevalonate pathway is chiefly employed to form ergosterol

(provitamin D₂) which is an essential part of the yeast membrane and provides membrane permeability and fluidity (Veen and Lang, 2004).

Taxol is a complex substituted terpenoid that was first isolated from the bark of Pacific yew (*Taxus brevifolia*). It is a potent anticancer agent that has been approved for the treatment of refractory ovarian and metastatic breast cancer. Currently, the demand for taxol greatly exceeds the supply that can be isolated from its natural sources (Hezari and Croteau, 1997).

Although the metabolic engineering of taxol biosynthesis in yeast has only begun, the application of new protein engineering techniques to improve rate-controlling enzymes appears promising (see paragraph 2.5.1.).

1.3.3. Production of carotenoids in yeast

Carotenoids are a subfamily of isoprenoids that are the most widely distributed yellow, orange and red natural pigments synthesized in bacteria, algae, fungi. Commercially carotenoids, such as β -carotene and astaxanthin, are used as food colorants, animal feed supplements and for nutritional and cosmetic purposes. More recently, carotenoids have received attention for their significant antioxidant activities and for playing an important role in inhibiting the onset of chronic diseases (Bertram and Vine, 2005). The majority of the 600 known structures of carotenoids are C₄₀-carotenoids while a few bacterial carotenoids exist with 30, 45 or 50 carbon atoms (Fig. I. 10).

Carotenoid pathways have been successfully introduced into non-carotenogenic microbes such as *E. coli* and *S. cerevisiae*. Since *S. cerevisiae* produces ergosterol as its principal isoprenoid from farnesyl diphosphate, (FPP), it would be possible to redirect the flux of FPP away from ergosterol into geranylgeranyl diphosphate (GGPP) and subsequent carotenoids. Indeed, the insertion of a plasmid containing *Erwinia uredovora* GGPP synthase (*crtE*), phytoene synthase (*crtB*) and phytoene desaturase (*crtI*) genes under the control of various *S. cerevisiae* promoters allowed *S. cerevisiae* to produce lycopene (113 μ g/g dry weight). In addition, *S. cerevisiae* harboring a plasmid containing the additional *E. uredovora* lycopene cyclase (*crtY*) gene resulted in the production of β -carotene (103 μ g/g dry weight) (Yamano *et al.*, 1994; Verwaal *et al.*, 2007).

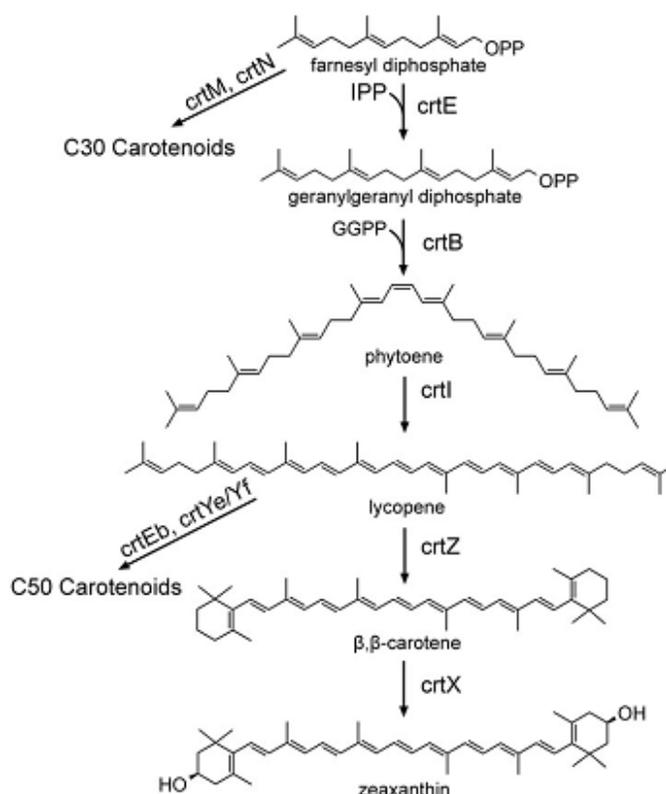


Figure I. 10. Carotenoid biosynthetic pathway (Chemler *et al.*, 2006).

1.3.4. Production of flavonoids in yeast

Flavonoids are a diverse group of plant secondary metabolites that contain a 15-carbon phenylpropanoid core, which is extensively modified by rearrangement, alkylation, oxidation and glycosylation (Turnbull *et al.*, 2004). These compounds possess extraordinary antioxidant activity and also exhibit estrogenic, antiviral, antibacterial and anti-cancer activities (Forkmann and Martens, 2001).

The health-protecting effects of flavonoids have stimulated significant research toward the elucidation of their biosynthetic networks (Fig. I.11), as well as the development of production platforms using well-characterized hosts, such as *E. coli* and *S. cerevisiae* (Yan *et al.*, 2005a). Ro and Douglas were the first to connect the two initial enzymes involved in phenylpropanoid pathway, namely phenylalanine ammonia lyase (PAL) and cinnamate 4-hydroxylase (C4H) in *S. cerevisiae* together with a cytochrome P450 reductase (Ro and Douglas, 2004). The carbon flux through the multienzyme system from phenylalanine to p-coumaric acid in yeast was evaluated in their study.

Later, two independent studies demonstrated the biosynthesis of flavanones, the common precursors of the vast majority of flavonoids, in *S. cerevisiae*. The first study by Jiang *et al.* described the production of monohydroxylated naringenin and unhydroxylated pinocembrin at levels of 7 mg/L and 0.8 mg/L respectively (Jiang *et al.*, 2005). The second study by Yan *et al.* reported the biosynthesis of flavanones in *S. cerevisiae* by constructing a gene cluster that included *C4H* (cinnamate-4-hydroxylase) from *A. thaliana*, *4cL-2* (4-coumaroyl:CoA-ligase) from parsley, and *CHI-A* (encoding for chalcone isomerase) and *chs* (chalcone synthase) from petunia (Yan *et al.*, 2005b). The recombinant *S. cerevisiae* strain was fed with phenylpropanoid acids and produced naringenin (28.3 mg/L), pinocembrin (16.3 mg/L) and the trihydroxylated flavanone eriodictyol (6.5 mg/L).

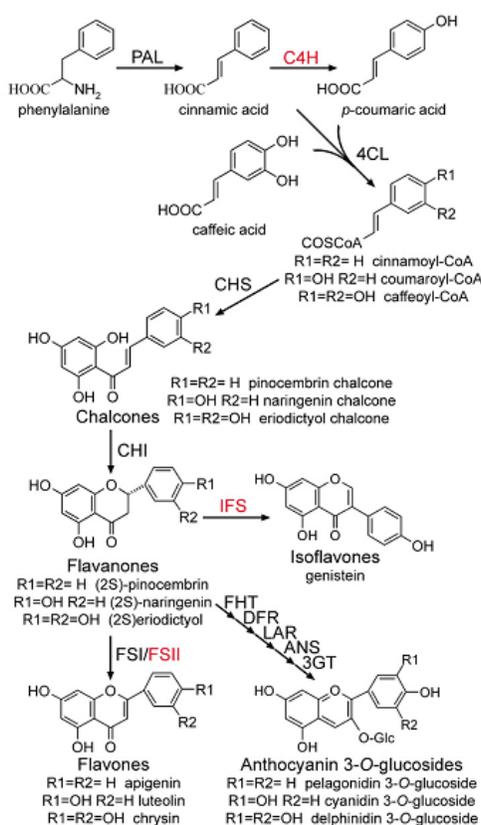


Figure I. 11. Flavonoid biosynthetic pathway (Chemler *et al.*, 2006).

In a continuation of the previous work, Leonard *et al.* demonstrated the biosynthesis of the flavone molecules apigenin, luteolin and chrysin by exploring the expression of a soluble flavone synthase I (FSI) and a membrane bound flavone synthase II (FSII) in yeast cells. The effect of the yeast P450 reductase overexpression on flavone biosynthesis was also investigated (Leonard *et al.*, 2005). Recently, Ralston *et*

al. reported the partial reconstruction of isoflavonoid biosynthesis in *S. cerevisiae* by using different types of chalcone synthase (CHI) and an isoflavone synthase (IFS) from soybean (*Glycine max*) (Ralston *et al.*, 2005). Finally, Trantas *et al.* (2009) reported the engineering of the complete pathway leading to heterologous biosynthesis of various flavonoids and stilbenoids in *Saccharomyces cerevisiae*. With the rapid unraveling of the flavonoid biosynthetic pathways a wide array of flavonoid compounds, natural and unnatural, is expected to be produced from *S. cerevisiae* in the near future. In addition, the natural coloration of some of the flavonoid molecules (anthocyanins) opens up the possibility of using directed evolution for protein activity fine tuning and thus further improvement of the overall productivity.

1.3.5. Production of biofuels using yeast

Alternative transportation fuels are in high demand due to concerns about climate change, the global petroleum supply, and energy security (Stephanopoulos, 2007; 2008). Currently, the most widely used biofuels are ethanol generated from starch (corn) or sugar cane and biodiesel produced from vegetable oil or animal fats (Fortman *et al.*, 2008). However, ethanol is not an ideal fuel molecule in that it is not compatible with the existing fuel infrastructure for distribution and storage due to its corrosivity and high hygroscopicity (Atsumi *et al.*, 2008). Also, it contains only about 70% of the energy content of gasoline. Biodiesel has similar problems (URL: <http://www.bdpedia.com/biodiesel/alt/alt.html>): it cannot be transported in pipelines because its cloud and pour points are higher than those for petroleum diesel (petrodiesel), and its energy content is approximately 11% lower than that of petrodiesel. Furthermore, both ethanol and biodiesel are currently produced from limited agricultural resources, even though there is a large, untapped resource of plant biomass (lignocellulose) that could be utilized as a renewable source of carbon-neutral, liquid fuels (Blanch *et al.*, 2008).

Microbial production of transportation fuels from renewable lignocellulose has several advantages. First, the production is not reliant on agricultural resources commonly used for food, such as corn, sugar cane, soybean, and palm oil. Second, lignocellulose is the most abundant biopolymer on earth. Third, new biosynthetic pathways can be engineered to produce fossil-fuel replacements, including short-chain, branched-chain, and cyclic alcohols, alkanes, alkenes, esters and aromatics. The development of cost-effective and energy-efficient processes to convert lignocellulose into fuels is disadvantaged by significant barriers, including the lack of genetic engineering tools for native producer organisms (non-model organisms), and difficulties in optimizing metabolic pathways and balancing the redox state in the engineered microbes

(Mukhopadhyay *et al.*, 2008). Furthermore, production potentials are limited by the low activity of pathway enzymes and the inhibitory effect of fuels and byproducts from the upstream biomass processing steps on microorganisms responsible for producing fuels (Fig. I. 12). Recent advances in synthetic biology and metabolic engineering will make it possible to overcome these hurdles and engineer microorganisms for the cost-effective production of biofuels from cellulosic biomass (Lee *et al.*, 2008b).

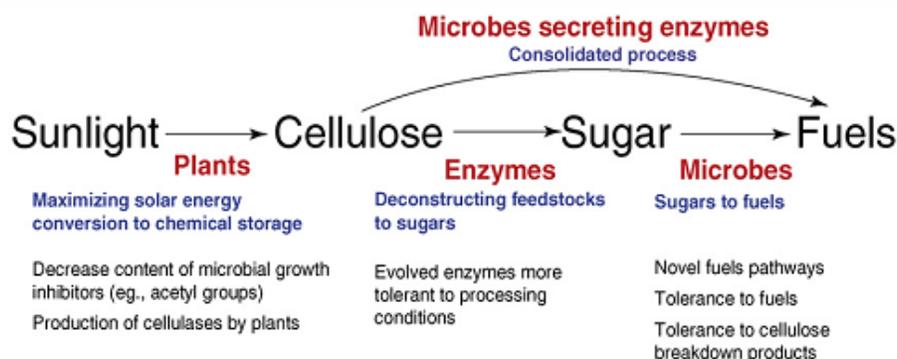


Figure I. 12. Route from sunlight to fuels. Conversion of biomass to fuels will involve the development of dedicated energy plants that maximize solar energy conversion to chemical storage and minimize the use of water and fertilizer, enzymes that depolymerize cellulose and hemicellulose into useful sugars, and microorganisms that produce advanced biofuels that are compatible with our existing transportation infrastructure. To achieve economically viable biofuel production, all aspects of these processes must be optimized. In particular, production hosts must natively have or be endowed with several important characteristics: extension of the substrate range, elimination of cellulose hydrolysates and biofuel product toxicity, and improvement of global regulatory functions. One method that has been proposed to reduce fuel production cost is to perform cellulose hydrolysis and fermentation in one step, called consolidated bioprocessing; this alternative approach avoids costs associated with cellulase production (Lee *et al.*, 2008).

To convert lignocellulosic biomass into economically viable biofuels, the production hosts must natively have or be endowed with several characteristics. Good starting points for development as production platforms are the user-friendly hosts (*E. coli* and *S. cerevisiae*) that have well-characterized genetics and the genetic tools (Fischer *et al.*, 2008) for manipulating them. Because these host organisms are also facultative anaerobes with fast growth rates, large-scale production processes can be relatively simple and economically viable (Khosla and Keasling, 2003). The successful use of *E. coli* or *S. cerevisiae* to produce alternative biofuels will

require a better understanding of their physiology under a variety of conditions and subsequent strain improvements.

Recently, *in silico* models have played an important role in engineering microorganisms to utilize new substrates to produce biofuels more efficiently. The ability to use a wider array of the biomass feedstocks would help to decrease cost by reducing the number of upstream processing steps and by turning more of the biomass into biofuel.

S. cerevisiae has been engineered with the genes encoding xylose reductase (Xyl1p) and xylitol dehydrogenase (Xyl2p) from *Pichia stipitis* to enable it to utilize xylose, the second most abundant carbohydrate in nature, as a carbon source for ethanol production (Chu and Lee, 2007). However, simply overexpressing the genes led to low growth and fermentation rates due to redox imbalance. Xylose reductase from *P. stipitis* has a higher affinity for NADPH than NADH, while xylitol dehydrogenase uses only NAD⁺. Overexpression of both genes leads to an accumulation of NADH and a shortage of NADPH. Metabolic models suggested that the cofactors could be balanced by deleting NADP⁺-dependent glutamate dehydrogenase (GDH1) and overexpressing NAD⁺-dependent GDH2 to increase the specific activity of Xyl1p for NADH. This strategy led to an increase in ethanol production and a reduction in byproduct synthesis (Jeffries, 2006). Interestingly, overexpression of the *P. stipitis* xylose reductase also increased tolerance to lignocellulosic hydrolysate (Almeida *et al.*, 2008). The *in silico* gene insertion strategy was used to improve ethanol production and decrease the production of the byproducts glycerol and xylitol. The result was a 58% reduction in glycerol, a 33% reduction in xylitol, and a 24% increase in ethanol production when glyceraldehyde-3-phosphate dehydrogenase was introduced into *S. cerevisiae* (Bro *et al.*, 2006). This technique demonstrates the importance of monitoring and balancing the levels of various important metabolites in order to achieve optimal product titers. Therefore, metabolic engineering will play an important role in engineering of microbial pathways for the production of economically sustainable biofuels.

1.4. Ergosterol biosynthetic pathway in yeast

1.4.1. Sterols in yeast

Sterols are important for the physiology of eukaryotic organisms as they form part of the cellular membrane where they modulate their fluidity and function and participate as secondary

messengers in developmental signaling (Mansour *et al.*, 2003). Ergosterol (Fig. I. 13) is an economically important metabolite, as it is the precursor for the production of vitamin D₂.

In yeast, ergosterol is the main sterol and is analogous to cholesterol in mammalian cells, an essential component of yeast plasma membranes which affects membrane fluidity, permeability, and the activity of membrane-bound enzymes (Daum *et al.*, 2007). In *Saccharomyces cerevisiae*, ergosterol is also a major component of secretory vesicles and has an important role in mitochondrial respiration (Zinser *et al.*, 1993). Ergosterol has also been predicted to play a role in oxygen sensing (Davies and Rine, 2006).

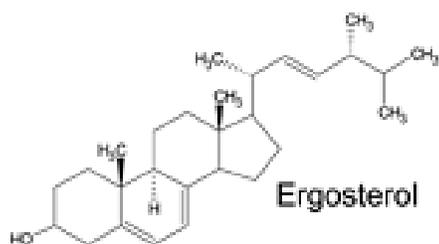


Figure I. 13. Chemical structure of ergosterol.

The production of complex sterols in yeast is tightly regulated at multiple levels. Genes in the ergosterol pathway exhibit transcriptional regulation in response to mutations in other ERG genes, resulting in sterol limitation (Kennedy *et al.*, 1999).

1.4.2. Branch point genes of sterol biosynthetic pathway

The biosynthesis of sterols in yeast is complex and is divided into two stages. This pathway is well defined and all the structural genes have been described (Lees *et al.*, 1999). The early sterol biosynthetic pathway (Fig. 14), starting with acetyl CoA which is an intermediate of the primary metabolism, provides the precursors for the synthesis of essential cellular constituents, such as the C₅ building block isopentenyl diphosphate, heme, quinones and dolichols (Meganathan, 2001; Grabinska and Palamarczyk, 2002; Lee *et al.*, 2008a). It also provides the precursor for the biosynthesis of carotenoids (Miura *et al.*, 1998; Lee and Schmidt-Dannert, 2002), sesquiterpenes (Jackson *et al.*, 2003) and squalene, the first specific intermediate of the ergosterol biosynthetic pathway. The enzymes of the later sterol pathway (Fig. I. 14) catalyze the transformation of squalene, the polyisoprene precursor, to ergosterol. This finding has opened the way to directly determine and influence critical and regulatory steps controlling the synthesis of sterols (Polakowski *et al.*, 1998; Polakowski *et al.*, 1999; Veen and Lang, 2004).

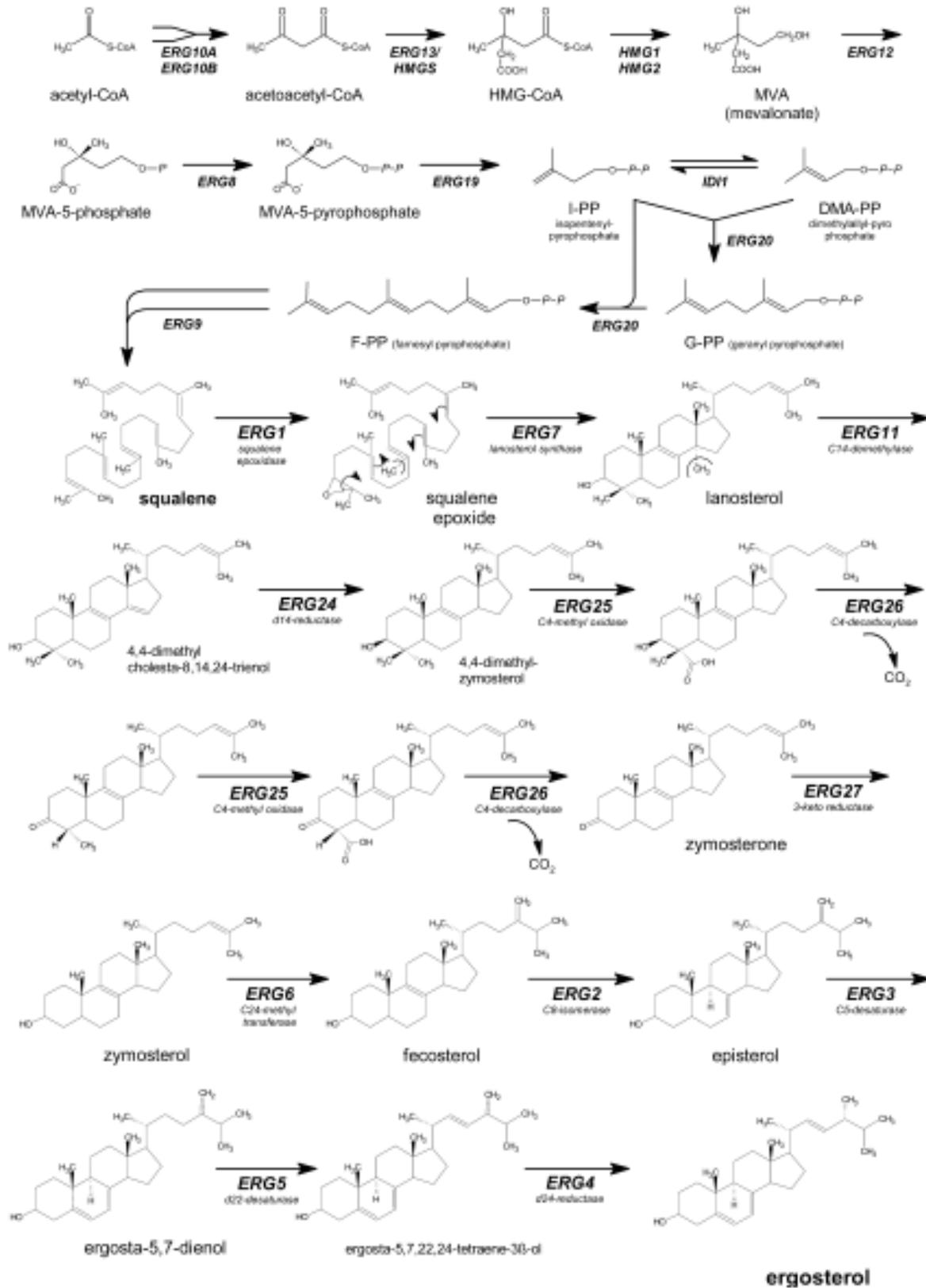


Figure I. 14. Ergosterol biosynthetic pathway in yeast (Veen and Lang, 2005).

One of the major bottlenecks in ergosterol biosynthesis is the reaction catalyzed by HMG-CoA reductase which is encoded by the isogenes HMG1 and HMG2. HMG-CoA reductase is regulated at different levels. While HMG1 is chiefly controlled at the transcriptional level by ergosterol and heme (Dimster-Denk *et al.*, 1994; Parks and Casey, 1995), HMG2 is regulated by a non-sterol isoprenoid signal at the level of premature protein degradation (Dimster-Denk *et al.*, 1994; Hampton *et al.*, 1996; Hampton, 1998; Gardner and Hampton, 1999; Hampton, 2002).

Overexpression of a truncated Hmg1p led to a 40-fold increase of HMG-CoA activity and resulted in an increased accumulation of squalene, an intermediate towards ergosterol, of up to 5.5% of dry matter (Polakowski *et al.*, 1998). The increased content of squalene has no significant effect on the content of sterols in the cell. This indicates that further regulation sites are present in the later part of the pathway and control the flux of intermediates into the end-product ergosterol.

Szkopinska *et al.* (2000) have shown that earlier literature data indicating strictly coordinated regulation of the mevalonic acid pathway enzymes, HMG-CoA reductase, farnesyl diphosphate synthase, and squalene synthase with reductase considered to be the main regulatory enzyme of cholesterol (ergosterol in yeast) synthesis, do not find full confirmation. FPP synthase (ERG20), independent of HMG-CoA reductase and to a certain degree of squalene synthase (ERG9), is the enzyme that responds to the highest degree to the changing internal and external environmental conditions, adapting the yeast cell to them.

More recently, overexpression of the post-squalene genes led to the discovery of two additional bottlenecks in the pathway, namely squalene epoxidase (Erg1p) and sterol-14 α -demethylase (Erg11p) (Veen *et al.*, 2003). Overexpression of Erg1p, Erg11p and a truncated Hmg1p led to a three-fold increase of sterol content over the wild-type. However, 90% of the sterol content was esterified sterol intermediates and not ergosterol. Obviously, further regulatory mechanisms are at play preventing the quantitative accumulation of ergosterol. Besides ergosterol, several sterol intermediates are of biological importance including methylated sterol intermediates that activate meiosis and zymosterol, another intermediate of interest as precursor of cholesterol lowering substances. By deleting sterol transmethylase (ERG6) and 8-carbon sterol isomerase (ERG2), zymosterol accumulated as the major sterol (Heiderpriem *et al.*, 1992).

1.4.2.1. 3-Hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase

Yeast sterols have the ability to regulate endogenous sterol biosynthesis, presumably by a transcriptional feedback mechanism.

HMGR is a key enzyme in the mevalonate pathway, from which sterols are synthesized, and is subject to feedback regulation as part of the cellular control of sterol synthesis. When certain sterols accumulate, the enzyme is rapidly degraded, thereby helping to terminate sterol synthesis (Goldstein *et al.*, 2006). HMGR is an integral endoplasmatic reticulum membrane protein, and its degradation occurs without exit from the endoplasmatic reticulum (Hampton and Rine, 1994). The non-catalytic, N-terminal transmembrane anchor of HMGR is both necessary and sufficient for regulated degradation (Gil *et al.*, 1985; Skalnik *et al.*, 1988). When there is abundant synthesis of pathway products, HMGR degradation is fast, and steady-state levels of the protein tend to be low. Conversely, when synthesis of pathway products is reduced, HMGR degradation is slowed and steady-state levels of the protein tend to increase.

Yeast has two isoenzymes of HMG-CoA reductase called Hmg1p and Hmg2p encoded by HMG1 and HMG2 genes, respectively. The Hmg1p isoenzyme is extremely stable while Hmg2p displays rapid, regulated degradation. In aerobic growth the proportion of Hmg1p in the cell is high and that of Hmg2p low. Hmg2p degradation is regulated by an unknown signal from the mevalonate pathway. Inhibiting early pathway enzymes, such as HMGR itself or HMG-CoA synthase decreases the rate of Hmg2p degradation. These early pathway blocks decrease the availability of the downstream signal for degradation. Conversely, inhibiting the enzyme squalene synthase, which is downstream of HMGR, stimulates degradation and ubiquitination of Hmg2p (Hampton and Bhakta, 1997). Furthermore, the degradation – enhancing effect of squalene synthase inhibition is abolished by simultaneous inhibition of HMG-CoA synthase or HMGR.

Endoplasmatic reticulum degradation of Hmg2p is tightly regulated and isozyme specific, yet occurs by the action of cellular machinery responsible for the destruction of a diverse array of misfolded proteins. Gardner and Hampton in 1999 have discovered two lysines, Lys6 and Lys 357, located at distant points along the linear sequence, which were each critical for degradation (Gardner and Hampton, 1999). Lysines 6 and 357 did not independently contribute to the regulated degradation of Hmg2p. Both are essential and their degradative function required a correct structural context. Their participation in regulated Hmg2p degradation was dependent on

information widely distributed throughout the 523-residue transmembrane domains. Thus, these lysines work together in a synergistic manner within the correct structure of Hmg2p to allow normal regulated degradation (Fig. I. 15).

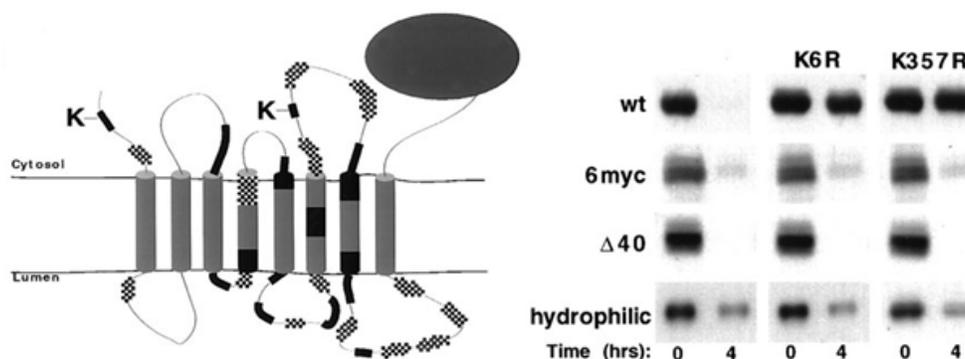


Figure I. 15. Function of Lys6 or Lys357 required the correct structure of the transmembrane region. Regions of Hmg2p required for correct regulated degradation. Hmg2p mutants were made either by replacing small regions of Hmg2p with corresponding regions from the stable, homologous (50% identical) Hmg1p, or by replacing the remaining unchanged residues of Hmg2p with semi-conservative residues. Regulated degradation of each mutant was assessed by cycloheximide–chase assay in the presence and absence of appropriate drugs. Mutation of the black regions resulted in partial or full stabilization. Alteration of the checkered regions resulted in degradation that was no longer regulated by signals from the mevalonate pathway (Gardner and Hampton, 1999).

1.4.2.2. Farnesyl pyrophosphate synthase ERG20

Farnesyl diphosphate (FPP) synthase is a key enzyme in isoprenoid biosynthesis which supplies sesquiterpene precursors for several classes of essential metabolites including sterols, dolichols, ubiquinones and carotenoids as well as substrates for farnesylation and geranylgeranylation of proteins (Fig. I. 16). The enzyme catalyzes the sequential head-to-tail condensations of isopentenyl diphosphate (IPP, C₅) with dimethylallyl diphosphate (DMAPP, C₅) and geranyl diphosphate (GPP, C₁₀) to give (*E,E*)-FPP (C₁₅). The enzyme belongs to a genetically distinct family of chain elongation enzymes that install *E*-double bonds during each addition of a five-carbon isoprene unit. Farnesyl diphosphate synthase is the prototypal representative of enzymes in the *E*-family. Metabolites derived from FPP are apparently required by all living organisms, and FPP synthase activity appears to be ubiquitous (Thulasiram and Poulter, 2006). FPP synthase is a homodimer of subunits, typically having two aspartate-rich motifs with two sets of substrate binding sites for an allylic diphosphate and isopentenyl diphosphate per homodimer. The synthase amino-acid residues at the 4th and 5th positions before

the first aspartate-rich motif mainly determine the product specificity. Hypothetically, type I (eukaryotic) and type II (eubacteria) FPPs evolved from archeal geranylgeranyl diphosphate synthase by substitution in the chain length determination region.

FPP synthase is mainly present in cytoplasm. However, recent investigations demonstrate that FPP synthase activity is associated with other subcellular compartments. Runquist *et al.* found that extensively washed rat liver microsomes contained FPP activity and that FPP produced could be used by both squalene synthase and cis-prenyltransferase present in the membranes (Runquist *et al.*, 1992). The *ERG20* gene encoding FPP synthase is an essential gene in yeast; therefore strains with deletion in this gene cannot be constructed. It appeared that the mutants deficient in the synthesis of farnesyl diphosphate required addition of exogenous FPP for the synthesis of polyprenols in vitro. Overexpression of the *ERG20* gene restored synthesis of polyprenols in vitro indicating that FPP is the allylic “starter” for prenyltransferases in yeast. Overexpression of the *ERG20* gene in the *erg9* mutant, defective in squalene synthase activity, not only restored synthesis of dehydrodolichols in vitro, but also increased the synthesis of dolichols in vivo, almost 10-fold in comparison with wild type yeast (Szkopinska *et al.*, 1997).

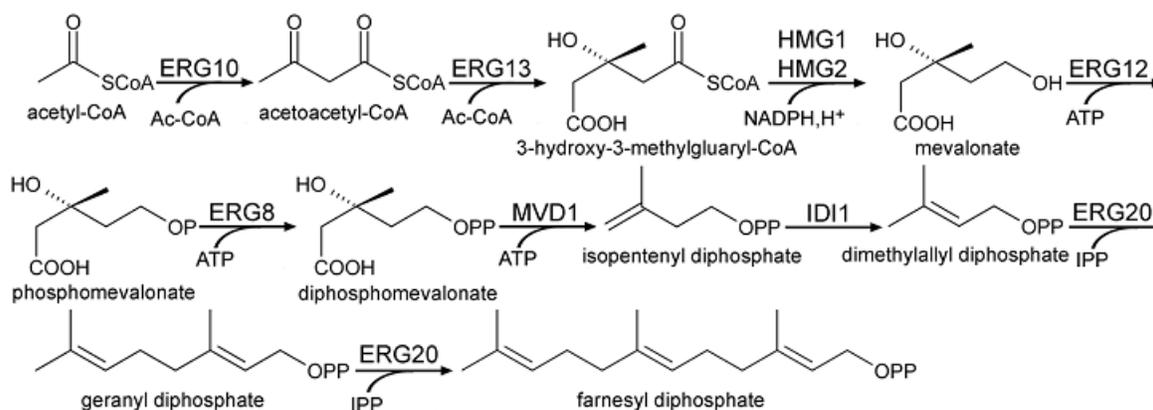


Figure I. 16. Biosynthesis of isoprenoids by mevalonate pathway (Chemler *et al.*, 2006).

1.4.2.3. Squalene synthase ERG9

Squalene synthase catalyzes the condensation of two molecules of farnesyl diphosphate to form the linear 30 carbon compound squalene, and this activity has been localized to the smooth endoplasmic reticulum (Stamellos *et al.*, 1993). Additional studies have suggested that

the carboxy-terminal portion of the squalene synthase protein anchors the protein to the ER membrane, whereas the catalytic site of the enzyme is associated with the amino-terminal portion of the protein found on the cytoplasmic face of the ER (Robinson *et al.*, 1993; Zhang *et al.*, 1993).

Sterols are essential molecules for all eukaryotic organisms, and many genetic mutations that eliminate enzymatic steps in sterol biosynthesis are lethal. In the case of ergosterol, FPP is converted to squalene by squalene synthase, which is encoded by the ERG9 gene. Severe mutations in the squalene synthase gene in yeast (*Saccharomyces cerevisiae*) are lethal, and these mutants require exogenous ergosterol to survive (Karst and Lacroute, 1977). In order to increase the level of FPP and direct flux towards terpene biosynthesis an obvious strategy is to attenuate the expression of the ERG9 gene. Since ergosterol is vital for yeast growth and yeast cells are unable to assimilate exogenous ergosterol during aerobic growth conditions, this gene cannot be completely deleted. As an alternative, the expression of the ERG9 gene can be controlled and/or diminished by different procedures.

1.4.2.4. Delta(24)-sterol C-methyltransferase ERG6

The enzyme that catalyzes sterol methylation at C-24 in *S. cerevisiae*, S-adenosylmethionine:delta(24)-sterol-C-methyltransferase, encoded by the ERG6 gene, is responsible for the alkylation of yeast sterols. This reaction, which results in the C28 methyl of ergosterol, is the most studied of the sterol synthetic enzymes in yeasts due to the uniqueness of the transformation relative to cholesterol synthesis and because of the suspected importance of C-28 to the function of ergosterol in biological membranes (Parks *et al.*, 1982). So as the sterol-C-methyltransferase imparts one of the key structural differences between ergosterol and cholesterol.

ERG6 is not required for normal vegetative growth, meiosis, or sporulation, strongly suggesting that C-24 methylation is not an essential feature of the sparking sterol in *S. cerevisiae*. However, Δ erg6 mutants have been shown to have diminished growth rates as well as limitations on utilizable energy sources (Kleinhans *et al.*, 1979), reduced mating frequency, altered membrane structural features, and low transformation rate. In addition, several lines of evidence have indicated that Δ erg6 mutations confer a variety of phenotypes consistent with alterations in membrane permeability and fluidity (Gaber *et al.*, 1989). This has been demonstrated by using

dyes, cations, and spin labels used in electron paramagnetic resonance studies. These include hypersensitivity to cycloheximide, resistance to nystatin, decreased mating frequency, decreased transformation frequency, and decreased tryptophan uptake. Although the transport of leucine, lysine, uracil, and histidine is sufficient to confer normal growth of the appropriate auxotrophic strains on typical yeast media, one might predict that the function of a number of integral membrane proteins involved in transport is affected by the altered membrane environment created by *Δerg6* mutations. Although the *Δerg6* mutations confer pleiotropic effects on the cell, *ERG6* is a unique, nonessential gene (Gaber *et al.*, 1989).

Inhibitors of the *ERG6* gene product would make the cell increasingly susceptible to antifungal agents as well as to new agents which normally would be excluded and would allow for clinical treatment at lower dosages. In addition, the availability of *ERG6* would allow for its use as a screen for new antifungals targeted specifically to the sterol methyltransferase (Jensen-Pergakes *et al.*, 1998).

1.5. Yeast metabolic engineering - perspective

Strain improvement of baker's and brewer's yeasts has traditionally relied on random mutagenesis or classical breeding and genetic crossing of two strains followed by screening of mutants exhibiting enhanced properties of interest. Recent developments of sophisticated methods in the field of recombinant DNA technology have enabled us to manipulate a given pathway of interest and hence to improve the cell by a more directed approach. Thus, it is now possible to introduce specific genetic perturbations in terms of modifying the promoter strength of a given gene, to perform gene deletions, or to introduce whole new genes or pathways into the cell. Directed improvements of the cellular properties achieved from the interplay of theoretical analysis, relying on biochemical information, and the application of genetic engineering has referred to as **metabolic engineering** (Ostergaard *et al.*, 2000).

The designing of strains of *S. cerevisiae* with new or improved properties through pathway engineering and protein engineering is nowadays a reality. The focus on *S. cerevisiae* to fulfill several biotechnological purposes is still increasing. The availability of the complete yeast genomic sequence paved the way for the development of new techniques such as the new the gene chips, which enabled genome-wide expression monitoring (Wodicka *et al.*, 1997). The gene chip generated a vast amount of biological information concerning the yeast model system and

this has also been used to further understand processes of higher eukaryotic cells like human cells. Extensive information about new “protein pathway”, protein interactions such as in signal transduction pathways, which obtained from the two-hybrid system or many other techniques also serves to identify potential targets for gene amplification or gene deletion.

Although the rigidity of *S. cerevisiae* in terms of alteration of its metabolic functions may limit certain approaches of metabolic engineering, this microorganism certainly has a great potential for pathway engineering. Undoubtedly a number of novel applications based on *S. cerevisiae* will arise in the future, which will clearly illustrate the potential of *S. cerevisiae* as a cell factory in biotechnological processes.

2. PLANT SECONDARY METABOLISM

2.1. Natural Products

Living organisms produce thousands upon thousands of different structures of low-molecular-weight organic compounds, historically referred to as natural products or secondary metabolites. The term ‘natural products’ is commonly reserved for those organic compounds of natural origin that are unique to one organism, or common to a small number of closely related organisms.

In the second half of the twentieth century, natural products formed a central pillar of the modern pharmaceutical industry. For example, 49% of the new chemical entities introduced into clinical use between 1981 and 2002 were of natural-product origin or inspiration, with the number rising to around 75% when applied to those drugs used for the treatment of severe and life-threatening conditions (Newman *et al.*, 2003). The diversity of natural-product structure and activity has also made them indispensable tools for deciphering cellular processes. This is particularly the case with the advent of chemical genetics, in which libraries of small molecules, including natural products, are screened to identify ligands that can modulate a specific function of a gene product *in vivo* (Spring, 2005).

Important recent advances in the identification and production of natural products have added new momentum in natural products research. With ever-increasing amounts of genome sequence data available, it has become clear that those compounds that have been identified and

isolated represent only the tip of the iceberg and that there is a wealth of natural product chemistry still to be mined. Moreover, the application of biosynthetic engineering has become an intensive area of research, driven by an increasing understanding of natural-product biosynthesis coupled with advances in molecular genetics. These efforts have led to the production of several natural products in heterologous systems, bypassing the limitations imposed in their application by scarce natural resources or inefficient chemical synthesis. In addition, advances in structural biology have further facilitated the development and application of protein-engineering methods for altering the specificity of biosynthetic enzymes, which can powerfully complement the combinatorial approaches.

2.1.1. Primary and secondary metabolism. Principal metabolic pathways

In the living organisms chemical compounds are synthesized and degraded by means of series of chemical reactions each mediated by an enzyme. These processes are known collectively as metabolism. All organisms possess similar metabolic pathways by which they synthesize and utilize certain essential chemical species: sugars, amino acids, common fatty acids, nucleotides, and the polymers derived from them. This is primary metabolism, and these compounds, which are essential for survival and well-being of the organism, are primary metabolites (Mann, 1987).

Most organisms also utilize other metabolic pathways, producing compounds which usually have not apparent utility: these are the so called 'natural products'. They are secondary metabolites, and the pathways of synthesis and utilization constitute secondary metabolism (Jenke-Kodama *et al.*, 2008). These pathways are perhaps only activated during particular stages of growth and development, or during periods of stress caused by nutritional limitation or microbial attack. Secondary metabolites are molecules that are not required for the growth and reproduction of a plant, but may serve some role in herbivore deterrence due to astringency or they may act as phytoalexins, killing bacteria that the plant recognizes as a threat. Secondary compounds are often involved in key interactions between plants and their abiotic and biotic environments that influence those (Facchini *et al.*, 2000). In most instances they appear to be non-essential to the plant, insect, or microorganism producing them, in marked contrast to the other organic compounds in nature, such as sugars, amino acids, nucleotides, and the polymers derived from them, which are both essential and ubiquitous. The importance of natural products

in medicine, agriculture and industry has led to numerous studies on the synthesis, biosynthesis and biological activities of these substances, making natural products a central theme of research at the interface of chemistry and biology.

There are three principal starting materials (or 'building blocks') for secondary metabolism:

i) shikimic acid, the precursor of many aromatic compounds including the aromatic amino acids, cinnamic acid, and certain polyphenols;

ii) amino acids, leading to alkaloids, and peptide antibiotics including penicillins and cephalosporins;

iii) acetate, precursor of polyacetylenes, prostaglandins, macrocyclic antibiotics, polyphenols, and the isoprenoids (terpenes, steroids, and carotenoids), via two entirely separate biosynthetic pathways (fig. 17).

Several groups of secondary metabolites have mixed biogenesis, i.e. an intermediate or metabolite from one principal pathway acts as a substrate for another metabolite from a different pathway. Thus, flavonoids are derived from a polyketide (three acetate units) and a cinnamic acid (shikimic acid). The indole alkaloids come from shikimate and a monoterpene (loganin). In each instance the precursor of these metabolites is also used for the biosynthesis of certain classes of primary metabolites, such as proteins, fatty acids etc. There are approximately 4 major classes of secondary compounds that are significant to humans. The classes are the alkaloids, phenylpropanoids, flavonoids, and the terpenoids (Fig. I. 17).

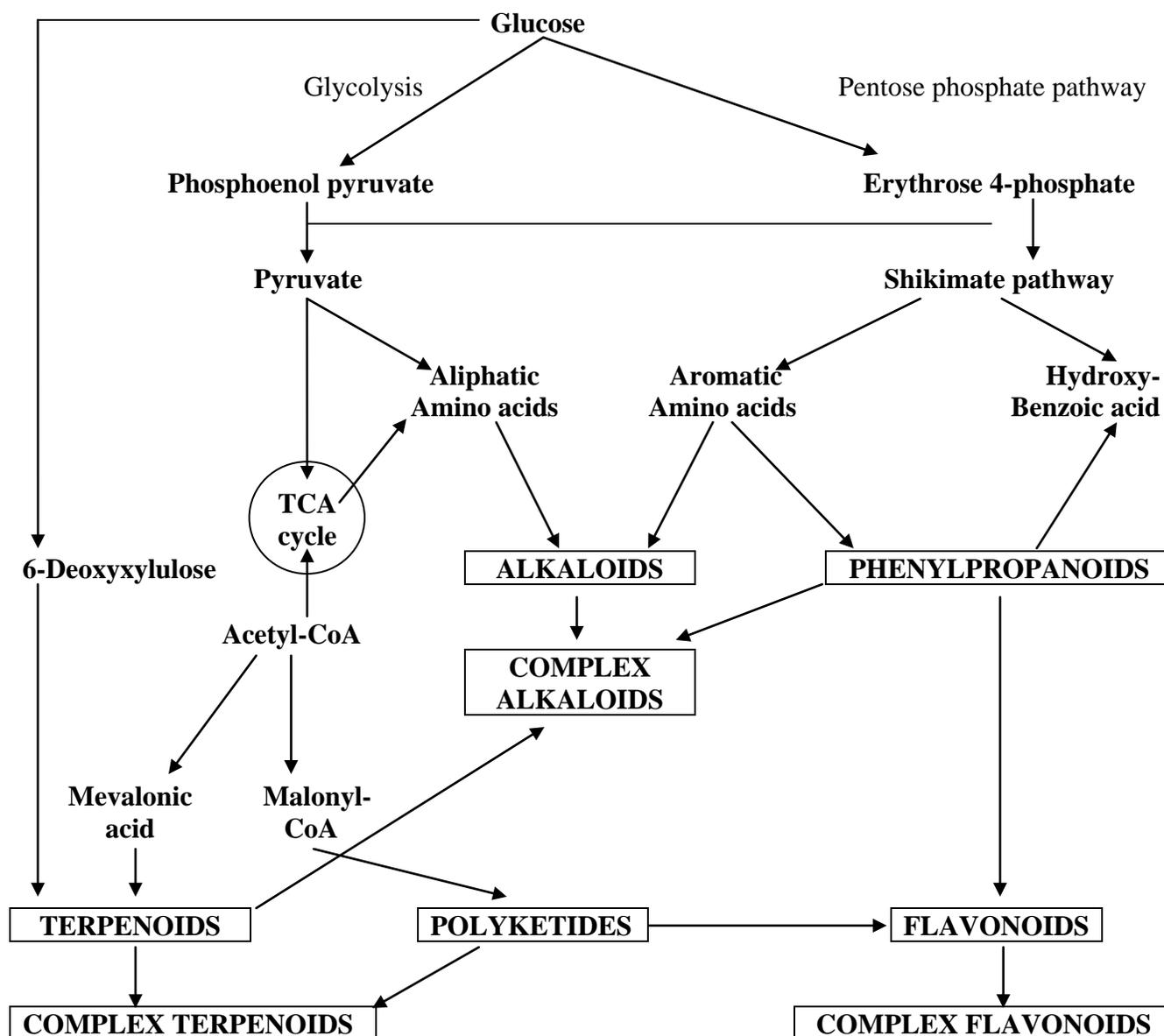


Figure I. 17. Principal biosynthetic pathways of secondary metabolism (Edwards and Gatehouse, 1999).

Secondary metabolites can be classified on the basis of chemical structure (for example, having rings, containing a sugar), composition (containing nitrogen or not), their solubility in various solvents, or the pathway by which they are synthesized (e.g., phenylpropanoid, which produces tannins). A simple classification includes three main groups: the terpenes (made from mevalonic acid, composed almost entirely of carbon and hydrogen), phenolics (made from simple sugars, containing benzene rings, hydrogen, and oxygen), and nitrogen-containing compounds (extremely diverse, may also contain sulfur) (Table I. 1).

Table I. 1. Principal classes of secondary compounds.

Class	Example compounds	Example sources	Some effects and uses
NITROGEN CONTAINING			
Alkaloids	Nicotine	Tobacco	<ul style="list-style-type: none"> • Interfere with neurotransmission; • Block enzyme action.
	Cocaine	Coca plant	
	Theobromine	Chocolate (cocoa)	
NITROGEN AND SULFUR CONTAINING			
Glucosinolates	Sinigrin	Cabbage	<ul style="list-style-type: none"> • Block initiation of tumors; • Modify steroid hormone metabolism; • Protect against oxidative damage.
TERPENES			
Monoterpenes	Menthol	Mint, relatives	<ul style="list-style-type: none"> • Interfere with neurotransmission; • Block ion transport; Anesthetic.
	Linalool	Many plants	
Sesquiterpenes	Parthenolid	Parthenium, (<i>Asteraceae</i>)	<ul style="list-style-type: none"> • Contact dermatitis.
Diterpenes	Gossypol	Cotton	<ul style="list-style-type: none"> • Block phosphorylation; toxic.
Triterpenes, cardiac glycosides	Digitogenin	Digitalis (foxglove)	<ul style="list-style-type: none"> • Stimulate heart muscle; • Alter ion transport.
Tetraterpenoids	Carotene	Many plants	<ul style="list-style-type: none"> • Antioxidant; • Orange coloring.
Terpene polymers	Rubber	Hevea (rubber) trees	<ul style="list-style-type: none"> • Gum up insects; • Airplane tires.
		Dandelion	
Sterols	Spinasterol	Spinach	<ul style="list-style-type: none"> • Interfere with hormone action
PHENOLICS			
Phenolic acids	Caffeic	All plants	<ul style="list-style-type: none"> • Cause oxidative damage; • Browning in fruits and wine
	Chlorogenic		
Coumarins	Umbelliferone	Carrots	<ul style="list-style-type: none"> • Cross-link DNA; • Block cell division.
		Parsnip	
Lignans	Podophyllin	Mayapple	<ul style="list-style-type: none"> • Cathartic; Vomiting; • Allergic dermatitis.
	Urushiol	Poison ivy	
Flavonoids	Anthocyanin	Almost all plant	<ul style="list-style-type: none"> • Flower, leaf color; • Inhibit enzymes; • Anti- and pro-oxidants; Estrogenic
	Catechin		
Tannins	Gallotannin	Oak	<ul style="list-style-type: none"> • Bind to proteins, enzymes; • Block digestion; • Antioxidants.
	Condensed tannin	Hemlock trees	
		Trefoil, legumes	
Lignin	Lignin	All land plants	<ul style="list-style-type: none"> • Structure, toughness, fiber.

2.2. Terpenes in Lamiaceae, a family of medicinal and aromatic plants

2.2.1. Lamiaceae family. *Salvia* species

The dialect name ‘sage’ is attributed to different aromatic species of the genus *Salvia* L. which are widely used as spices, as well as in the food, drug, and fragrance industry. *S. officinalis*, *S. fruticosa* Miller and *S. pomifera* L., named ‘faskomilo’ are Eastern Mediterranean species of sage and three of the most commercially utilized sage plants (Kintzios, 2000).

The properties of numerous extracts of these species have already been examined and individual terpenoids (which are the major essential oil components) have been implicated in the various health promoting roles of these plants (Cuppett and Hall, 1998; Lionis *et al.*, 1998; Wijeratne and Cuppett, 2007).

The essential oil of *S. fruticosa* (Fig. I. 18) is a folk remedy for toothaches and intestinal complaints (Rivera *et al.*, 2005). Recent studies have revealed that a leaf infusion of *S. fruticosa* has a hypoglycemic effect (Perfumi *et al.*, 1991); furthermore it has been shown that its essential oil exhibit antibacterial, cytotoxic and antiviral activities.

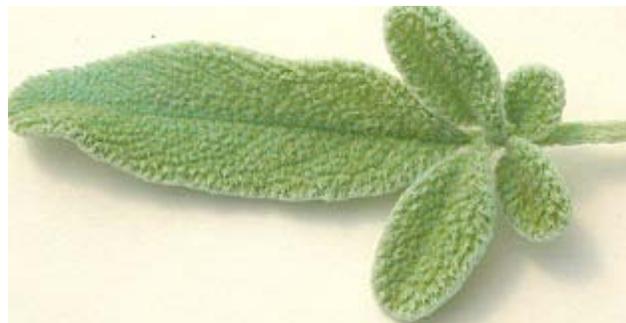


Figure I. 18. *Salvia fruticosa*

The overall distribution of the three species of sage, *S. officinalis*, *S. fruticosa* Miller and *S. pomifera*, is limited by the different climatic conditions dominating in the different areas. *S. fruticosa* Miller and *S. pomifera* exhibit a noticeable morphological variation. The leaf morphology and the essential oil content of *S. fruticosa* change gradually following the geographic-climate gradient.

The main components of *Salvia* species are flavonoids and terpenoids. The aerial parts of these plants usually contain flavonoids and terpenoids as well, while in the roots the main compounds are diterpenoids (Abietane, Clerodane, Pimarane, and Labdane). However, American *Salvia* species contain diterpenoids in the aerial parts of plant too. The majority of terpenoids (Table I. 2) are monoterpenes; both sesquiterpenes and sesterperenes are rather rare in *Salvia* species (Farhat *et al.*, 2001; Santos-Gomes and Fernandes-Ferreira, 2003; Karousou *et al.*, 2005; Raal *et al.*, 2007).

Table I. 2. Ranges of the quantity of the main compounds in *Salvia* species upper leaves (% of total oil) (Ennaifer, 2005).

	<i>Salvia fruticosa</i>	<i>Salvia officinalis</i>	<i>Salvia pomifera</i>
α-Thujene	0.30 – 1.04	0.18 – 0.48	0.24 – 0.84
α-Pinene	2.58 – 6.26	1.26 – 4.62	0.42 – 2.40
Camphene	0.17 – 1.05	0.95 – 7.96	0.51 – 4.42
Sabinene	0.31 – 0.69	0.13 – 0.43	1.00 – 2.65
β-Pinene	8.92 – 18.77	1.92 – 5.67	0.40 – 1.85
β-Myrcene	1.47 – 5.03	0.65 – 1.47	0.48 – 1.38
Limonene	0.06 – 1.28	0.60 – 1.94	0.27 – 0.59
1,8-Cineole	22.70 – 49.20	5.81 – 9.16	0.09 – 0.20
Terpinene	0.38 – 0.80	0.32 – 0.53	0.20 – 0.49
α-Thujone	0.30 – 3.53	5.42 – 43.48	9.8 – 47.35
β-Thujone	0.22 – 1.72	3.56 – 8.33	17.8 – 27.72
Camphor	0.13 – 1.72	4.72 – 27.71	0.34 – 1.18
Borneol	0.51 – 1.34	0.53 – 4.22	0.14 – 0.93
α-Cubebene	0.26 – 0.19	1.60 – 1.65	1.40 – 4.68
Caryophyllene	4.64 – 12.75	1.53 – 14.00	7.82 – 22.75
α-Humulene	2.18 – 4.13	6.00 – 11.20	0.43 – 1.49
Curcumene	0.03 – 0.18	0.03 – 0.04	0.12 – 5.16
Cadinene	0.03 – 0.70	0.04 – 0.05	2.29 – 8.00

The qualitative essential oil composition of the three *Salvia* species is similar with respect to the main components, 1,8-cineole, α - + β -thujone and camphor, which constitute the bulk of the oil (54.4 – 83.4%) (Karousou *et al.*, 2007). However, a high inter- and intraspecific variation is found in the quantitative participation of the main components in the total oil (Table I. 3). The

amount of 1,8-cineole is high in *S. fruticosa* oils, up to 66% of the total oil, whereas it is much lower, less than 16% in the other two sage species. The total thujone content is always high in *S. pomifera*, more than 58.7% of the total oil, whereas the highest amount of camphor (38.1%) has been recorded in *S. officinalis* oils.

Table I. 3. Ranges of the main essential oil components (% of the total oil) of the sage species in Greece.

	1,8-Cineole	α- + β-thujone	Camphor
<i>Salvia fruticosa</i>	22.7 – 66.2	1.4 – 37.3	0.8 – 30.3
<i>Salvia pomifera</i>	0.2 – 9.5	58.7 – 83.0	0.3 – 3.8
<i>Salvia officinalis</i>	13.4 – 15.4	15.8 – 32.7	30.3 – 38.1

2.2.2. Terpenes. Term and significance

Terpenes (also known as terpenoids or isoprenoids) comprise the largest class of natural products, encompassing tens of thousands of known compounds with an extremely diverse array of chemical structures (Connolly and Hill, 1991; Connolly and Hill, 2008).

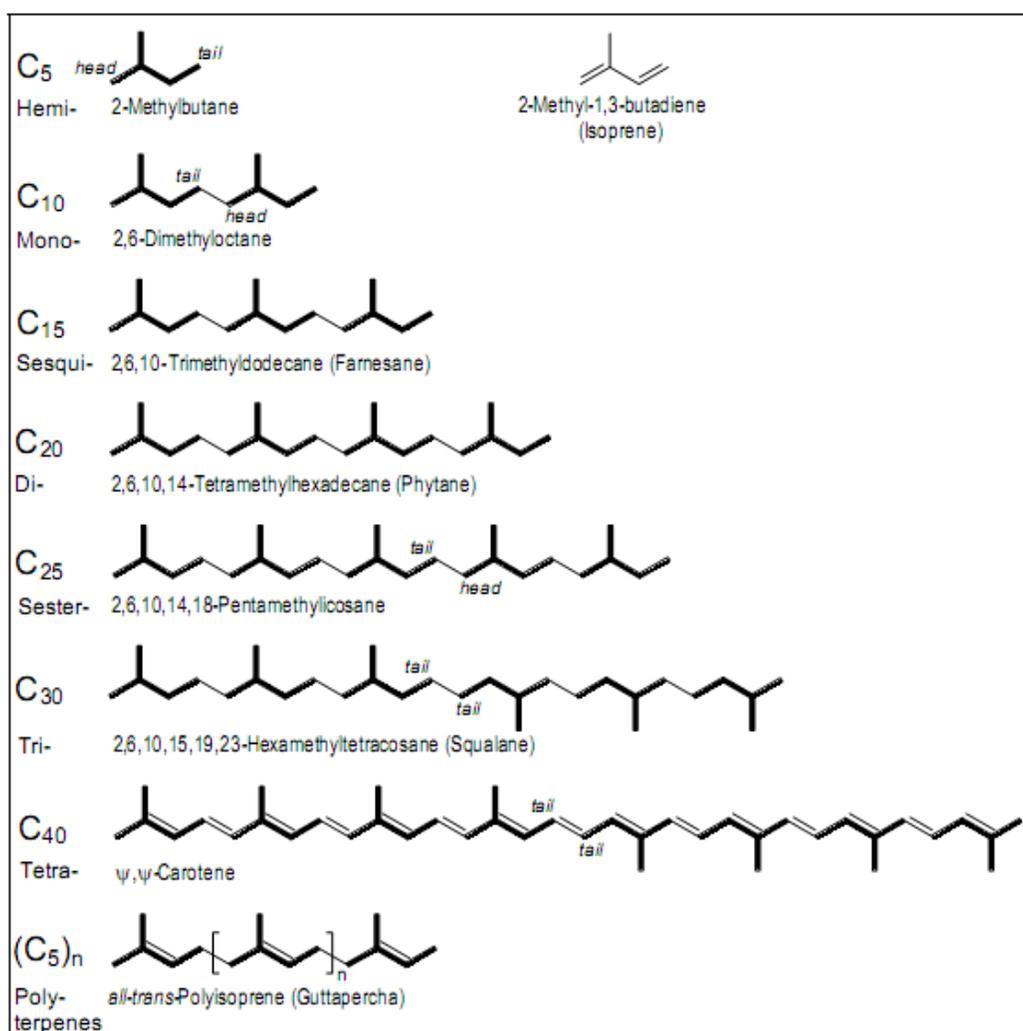
The term, ‘terpenes,’ originates from ‘turpentine’ (lat. balsamum terebinthinae). Turpentine, the so-called ‘resin of pine trees’, is the viscous pleasantly smelling balsam which flows upon cutting or carving the bark and the new wood of several pine tree species (Pinaceae).

They are traditionally valued as flavors and fragrances, as they are the primary constituents of the essential oils of plants (for example, β -damascenone (rose), limonene (citrus) and zingiberene (ginger)). However, they are also known for their many varied biological functions, such as their use as hormones (steroids, gibberellins and abscisic acid) and their roles in membrane fluidity maintenance (steroids), respiration (quinones), photosynthetic light harvesting (carotenoids), and protein targeting and regulation (prenylation and glycosylation). Although all organisms use terpenoids for these basic cellular processes, it is in the plant kingdom that their structure and function has most significantly evolved and diversified. In plants, terpenoids have many essential roles in specialized processes such as defense, communication, pollination attraction and seed dispersion in addition to their involvement in primary growth and development (Harborne, 1990). It is among these secondary metabolites that many effective and promising pharmaceuticals, such as Taxol (cancer), artemisinin (malaria) and prostratin (HIV) that have been discovered and isolated.

2.2.3. General structure, classification

The structure of terpenes is extremely variable, exhibiting hundreds of different carbon skeletons. However, this wide structural diversity has a basic structure following a general principle: 2-methylbutane residues, less precisely but usually also referred to as isoprene units, (C_5)_n, build up the carbon skeleton of terpenes. The isoprene rule found by Ruzicka, 1953 (Table I. 4) states that terpenes are multiples of C_5 units linked together head to tail fashion (Ruzicka, 1953). Several modes of cyclization are conceivable and lead to various skeletons.

Table I. 4. Parent hydrocarbons of terpenes (isoprenoids) (Ruzicka, 1953).



In nature, terpenes occur predominantly as hydrocarbons, alcohols and their glycosides, ethers, aldehydes, ketones, carboxylic acids and esters. The terpenes are classified according to the number of C_5 units: hemiterpenes, C_5 ; monoterpenes, C_{10} ; sesquiterpenes, C_{15} ; diterpenes,

C₂₀; sesterterpenes, C₂₅; triterpenes, C₃₀; and tetraterpenes, C₄₀. Hemi-, mono-, sesqui-, and diterpenes are mostly secondary metabolites, while tri-, and tetraterpenes are generally primary metabolites (Broun and Somerville, 2001). Homoterpenes (C₁₁ and C₁₆) compounds are produced by the removal of four carbons from sesquiterpenes and diterpenes. These compounds are formed in certain plants as response to insect damage (Dewick, 2002). Meroterpenes are products of partial isoprenoid origin, including cytokinins or prenylated proteins (Rodriguez-Concepcion and Boronat, 2002).

2.2.4. Monoterpenes

Monoterpenes are volatile lipophilic C₁₀ compounds that are common constituents of plant resins and essential oils. Several hundred naturally occurring monoterpenes known and all are biosynthesized from geranyl pyrophosphate, the ubiquitous acyclic C₁₀ intermediate of the isoprenoid pathway (Wheeler and Croteau, 1987). Approximately 1500 monoterpenes are documented. The monoterpenes can be acyclic (nerol, geraniol, linalool, myrcene), mono or bicycle (limonene, 1,8-cineole, thujene).

Acyclic monoterpenoid trienes such as β -myrcene and configurational isomers are found in the oils of basil (leaves of *Ocimum basilicum*, Labiatae), bay (leaves of *Pimenta acris*, Myrtaceae), hops (strobiles of *Humulus lupulus*, Cannabaceae), petitgrain (leaves of *Citrus vulgaris*, Rutaceae) and several other essential oils. Unsaturated monoterpene alcohols and aldehydes play an important role in perfumery. The cis-trans-isomers of geraniol and nerol of the oil of palma-rosa from the tropical grass *Cymbopogon martinii* var. *motia* (Poaceae) smell pleasantly flowery (Bohlmann *et al.*, 1997).

Monocyclic terpenes are derived, for the most part, from the cis-trans-isomers of p-menthane, found in abundance in species of the mint family (Lamiaceae), such as peppermint (*Mentha x piperita*) and spearmint (*Mentha spicata*). Trans-p-menthane itself occurs in the oil of turpentine. Limonene (Schwab *et al.*, 2001) is an unsaturated monocyclic terpene hydrocarbon occurring in various ethereal oils, smelling like oranges, is the dominant component of mandarin peel oil from *Citrus reticulata* and the oil of orange from *C. aurantium* (Rutaceae). Menthadienes such as α - and γ -terpinene as well as terpinolene are fragrant components of several ethereal oils originating from *Citrus*, *Mentha*-, *Juniperus*- and *Pinus* species. α -Terpineol has the very

pleasant smell of lilac blossoms, and is an important raw material in perfumery obtained from different ethereal oils (*Artemisia*, *Eucalyptus*, *Juniperus*, *Mentha*) (Fig. I. 19).

Apart from atypical terpenoid monocyclic cyclopentane derivatives, about 200 cyclopentane monoterpenes occur as iridoides and seco-iridoides, which can occur as insecticidal and antibacterial pheromones or antifeedants (Breitmaier, 2006).

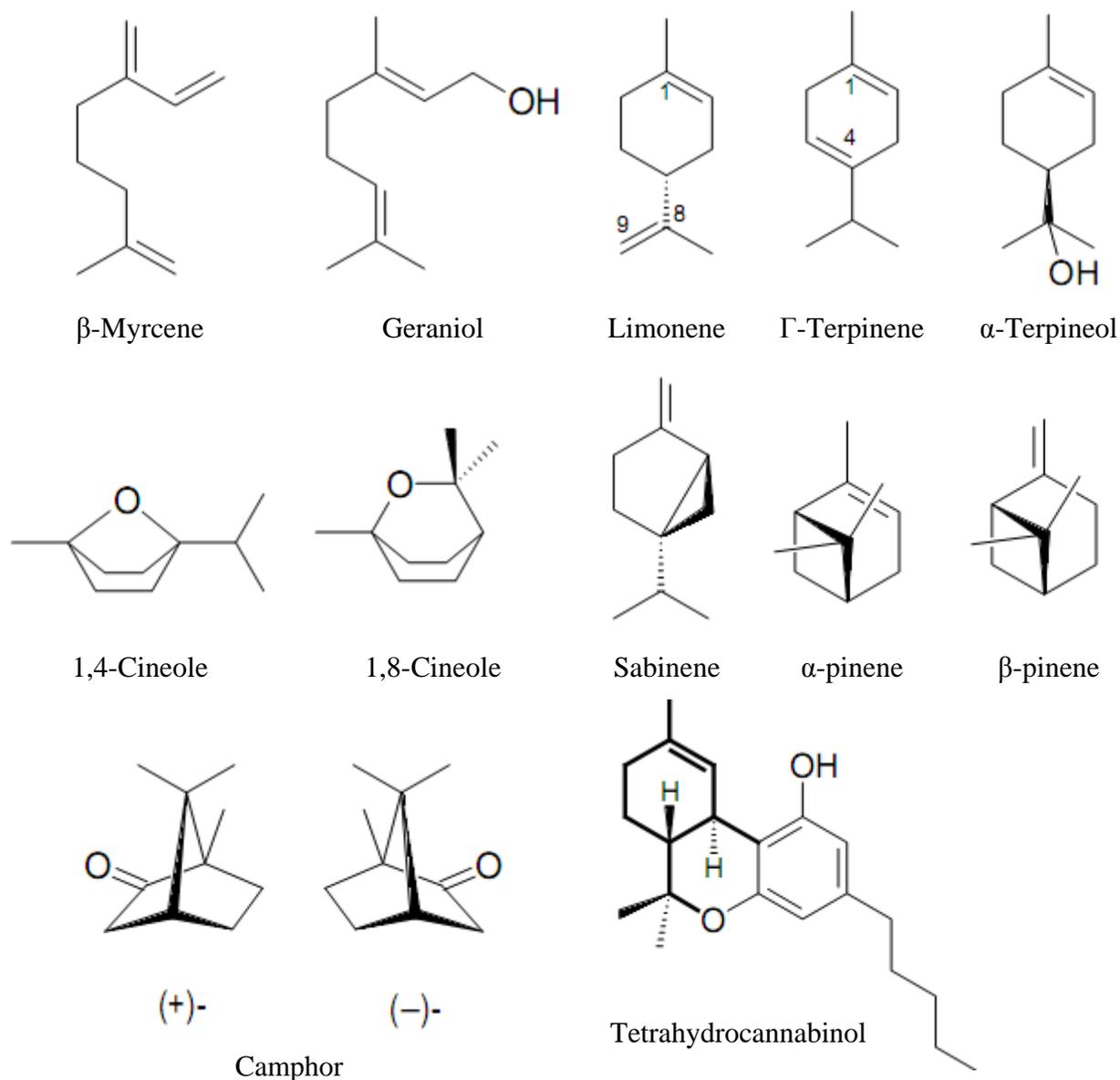


Figure I. 19. Representative members of various monoterpene subfamily structural types (Adams, 1995).

Oxygen-bridged derivatives of p-menthane such as the bicyclic ethers 1,4-cineol from *Juniperus* or *Artemisia* species and *Cannabis sativa* as well as 1,8-cineol (eucalyptol, the chief component of eucalyptus oil), stamp the spicy odor of the oils of cardamom, eucalyptus, lavender (Wise *et al.*, 1998). The oil of cardamom is used to spice food and alcoholic drinks. Oils of eucalyptus and lavender are predominantly applied as fragrances and flavors in perfumery and pharmacy (Connolly and Hill, 1991).

(+)-4(10)-Thujene, better known as (+)-sabinene, occurs in the oil of savin obtained from fresh tops of *Juniperus Sabina*. Its regioisomer 3-thujene is found in the oils of coriander-, dill-, eucalyptus-, thuja-, and juniper (Peters and Croteau, 2003). The oil of turpentine obtained on large scale from the wood of various pine trees (*Pinus caribaeae*, *P. palustris*, *P. pinaster*) and by way of cellulose production (sulfurated oil of turpentine) contains more than 70 % of α - and up to 20 % of β -pinene (Breitmaier, 2006).

(+)-Camphor (Croteau *et al.*, 1987), known as Japan camphor, is the main constituent of the camphor tree, but also occurs in other plant families, e.g. in the leaves of rosemary *Rosmarinus officinalis* and sage *Salvia officinalis* (Laminaceae). It gives off the typical camphor-like odor of spherical molecules, acts as an analeptic, a topical analgesic, a topical antipruritic, antirheumatic, antiseptic, carminative, counterirritant and, correspondingly, finds versatile application (Ohloff, 1990).

About 70 among more than 400 constituents of the Indian hemp *Cannabis sativa* belong to the cannabinoids. Illegal drugs prepared from Indian hemp include marihuana, a tobacco-like fermented mixture of dried leaves and blossoms, coming from Central Africa, Central America, U.S., as well as Southeast Asia, and hashisch, the resin secreted by the glands in the flowering tops of female plants, coming from the Middle East and South Asia, and having higher content of active substances. Regioisomers '8- and '9-tetrahydrocannabinol (THCs), differing by the position of their alkene CC double bond, are the most important addictive, analgesic, euphorizing, hallucinatory, laxative, and sedative constituents (O'Neil and Merck & Co., 2001).

2.2.5. Sesquiterpenes

Sesquiterpenes are the most diverse class of isoprenoids with more than 300 identified carbon skeletons (Cane and Bowser, 1999) and more than 7000 characterized compounds (Connolly and Hill, 1991). Highly volatile compounds, sesquiterpenes are constituents of plant

essential oils, but extraction from plants is not suitable for large-scale production of many sesquiterpenes because of slow plant growth, and variable composition and concentration based on geographical position and climate conditions (Breitmaier, 2006). From an industrial point of view, sesquiterpenes are interesting compounds because of their potent anticancer, antitumor, cytotoxic, antiviral and antibiotic properties as well as their characteristic flavors and aromas.

As a family, the sesquiterpenes encompass an bewildering array of structural types and more than fifty basic skeletons have been recognized, assumed to derive from farnesyl pyrophosphate (FPP) (Fig. I. 20).

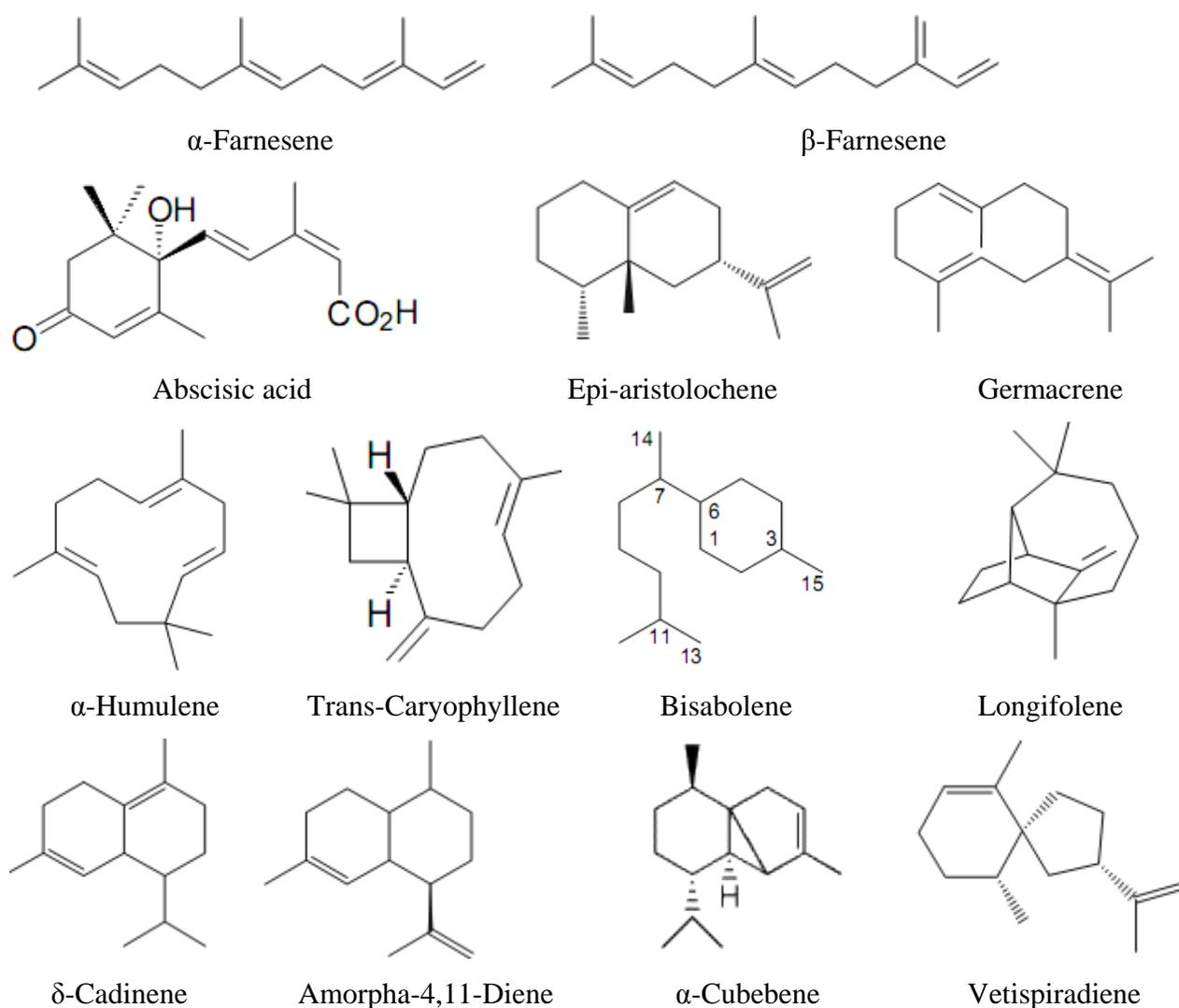


Figure I. 20. Representative sesquiterpene structure (Adams, 1995).

A few acyclic sesquiterpenes are known. Dehydration of farnesol gives farnesene, found in the oil slate, which is the parent compound of about 10000 sesquiterpenes known to date. The α -farnesene is a component of the flavors and natural coatings of apples, pears and other fruits. Associated with β -farnesene, it also occurs in several ethereal oils, for example those of camomile, citrus, and hops (Crock *et al.*, 1997).

By ring closure of carbon atoms C-6 and C-7 of farnesane, the cyclofarnesanes formally arise. One representative compound of this group is abscisic acid, which acts as an antagonist of plant growth hormones and controls flowering, falling of fruits and shedding of leaves.

Formally, C-1 and C-6 of farnesane close a cyclohexane ring in the bisabolanes, which represent a more prominent class of monocyclic sesquiterpenes. More than 300 naturally occurring, germacrane formally result from ring closure of C-1 and C-10 of farnesane. Among these are germacrene B from the peel of *Citrus junos*, germacrene D (Arimura *et al.*, 2004) from bergamot oil (*Citrus bergamia*, Rutaceae), and germacrone, derived from germacrene B, a pleasantly flowery to herby-smelling component isolated from the essential oil of myrrh (*Commiphora abyssinica*, Burseraceae).

Approximately 30 naturally abundant caryophyllanes are derived from humulanes in which C-2 and C-10 close a cyclobutane ring. Trans-caryophyllene occurs as a mixture with its *cis* isomer, isocaryophyllene in the clove oil (up to 10%) from dried flower buds of cloves (*Caryophylli flos*, Caryophyllaceae), in the oil obtained from stems and flowers of *Syzygium aromaticum* (Myrtaceae), as well as in the oils of cinnamon, citrus, eucalyptus, sage, and thyme. Clove oil, with its pleasantly sweet, spicy and fruity odor, is used not only in perfumery and for flavoring chewing gums, but also as a dental analgesic, carminative and counterirritant (Breitmaier, 2006).

Some sesquiterpenes serve as phytoalexins such as epi-aristolochene (Bohlmann *et al.*, 2002), vetispiradiene and δ -cadinene (Wang *et al.*, 2003). A-cadinene from the oil of hops (*Humulus lupulus*, Cannabaceae) as well as β -cadinene, widely spread in plants, spicy smelling and isolated from the oil of cade obtained by distillation of the wood of Mediterranean juniper *Juniperus oxycedru* (Cupressaceae), exemplify the cadinanes. Antimalarials derived from cadinane are found as constituents of the traditional Chinese medicinal herb *Artemisia annua* (Asteraceae), well-known as qinghao. Artemisinin, also referred to as qinghaosu, is a 3,6-peroxide of the acylal formed by 4,5-seco-cadinane-5-aldehyde-12-oic acid. Dihydroartemisinin

and the dehydro derivative artemisitene are the active substances which, nowadays, are applied as semisynthetic esters and ethers (e.g. artemether) to cure malaria (White, 2008).

2.2.6. Diterpenes

Widespread in the plant kingdom, approximately 5000 naturally abundant acyclic and cyclic diterpenes derived from geranylgeranyl diphosphate (Fig. I. 21) (Busch and Kirschning, 2008). Diterpenes often are encountered in the resin of conifers, woody legumes, composite, and members of the Euphorbiaceae family. Phytol, an acyclic diterpene with a simple structure, is particularly ubiquitous as it comprise one-third of the molecule chlorophyll and is found in the leaves of all green plants.

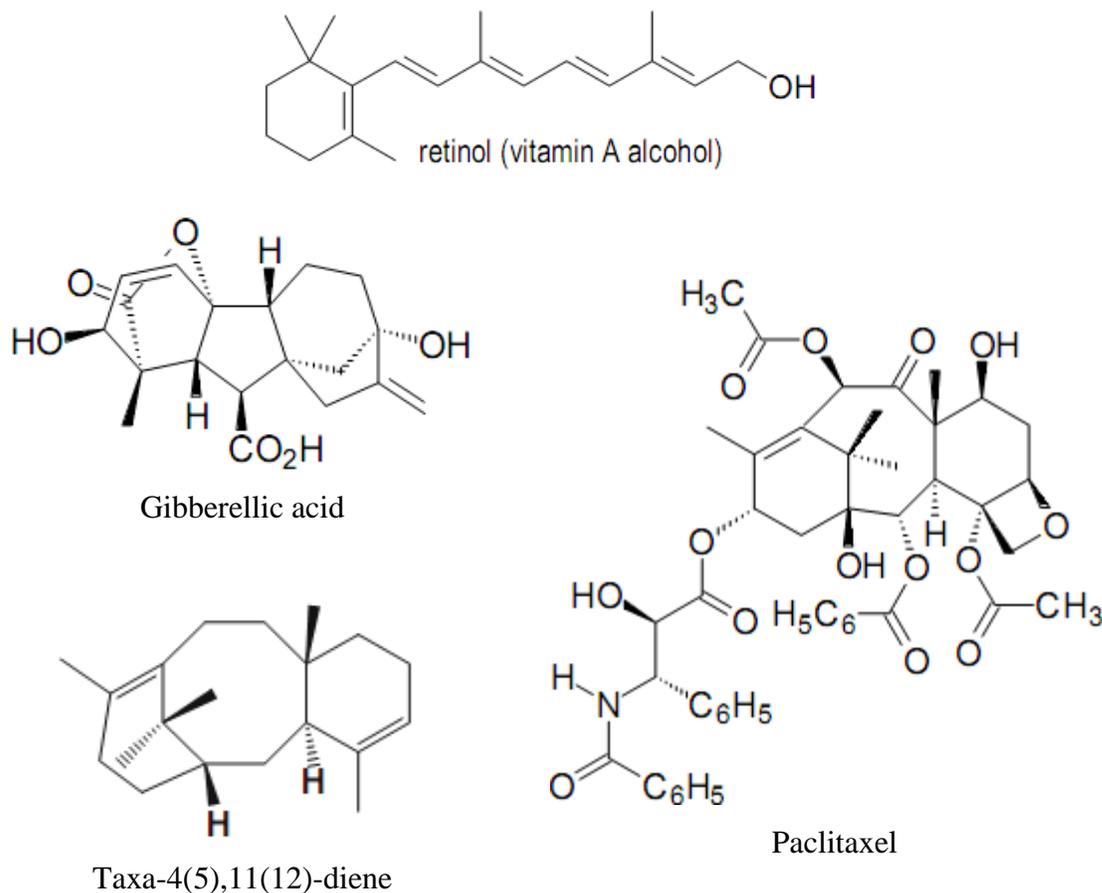


Figure I. 21. Representative diterpene structures (Adams, 1995).

Frequently occurring 10,15-cyclophytanones include very important representatives of the vitamin A series such as axerophthene, retinol, retinal and tretinoin (Duester, 2008). 11-cis-Retinal (vitamin A aldehyde) attaches as an imine to an L-lysine moiety of the apoprotein opsin within the photoreceptor protein rhodopsin (visual purple) found in the rods of the retina. The photoisomerization of 11-cis-retinal in rhodopsin induces a conformational change of the protein, resulting in a nerve pulse during the visual process in the eyes (Furr, 2004).

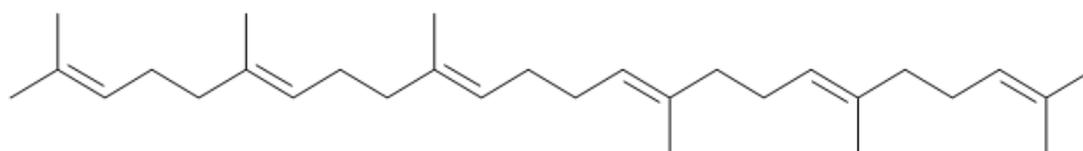
More than 60 gibberellanes isolated from higher plants and fungi to date are, for the most part, C-20-norditerpenes. They play an essential role as plant growth hormones, and also regulate the degradation of chlorophyll as well as the formation of fruits, and thus are used in agriculture (Itoh *et al.*, 2008).

The chemotherapeutic paclitaxel, from the bark of Pacific yew tree *Taxus brevifolia* and *T. cuspidate*, applied for the chemotherapy of leukemia and various types of cancer, sold under the brand name Taxol[®], is perhaps the most famous example of a commercialized terpenoid. Paclitaxel is at the terminus of a long biosynthetic pathway beginning with conversion of GGPP to produce taxa-4(5),11(12)-diene (Hezari and Croteau, 1997).

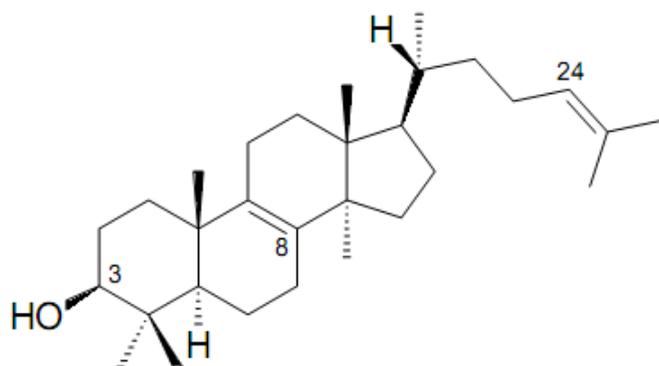
2.2.7. Other terpenes, carotenoids

Triterpenes (C₃₀) are widely distributed, especially among angiosperm plants (fig. 22). Sterols, which are biosynthesized from squalene, are triterpenes found in a number of gymnosperm and angiosperm woods (*Larix*, *Abies*, *Picea*, *Pinus*, *Gmelia*, *Fagus*, *Quercus*, and *Ulmus*). Lanosterol is a prominent constituent of lanolin, the wool fat of sheep used as ointment base, emulsifier, conditioner and lubricant in cosmetics. Lanosterol is also found in yeast and other fungi, in Euphorbiaceae such, and in various other higher plants. Moreover, it is the first isolated precursor of steroid biogenesis from 2,3-epoxysqualene in mammals. Sterols may also occur as fatty acid esters or as glycosides. Saponins contain several glycosides and produce lather in water. Cardiac glycosides have particularly strong effects on the heart muscle and can either function as a medicine or as a poison (Connolly and Hill, 2008).

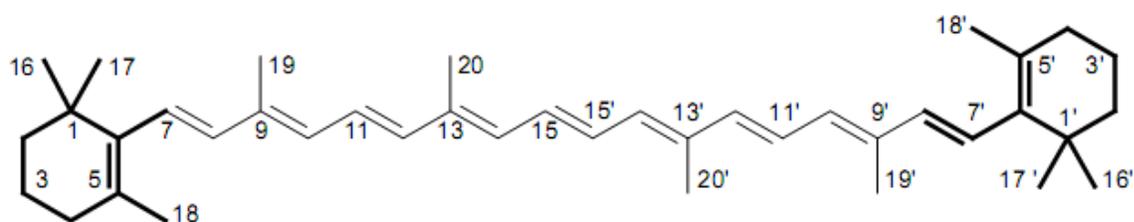
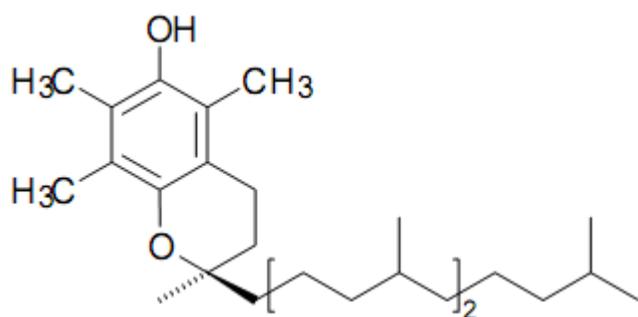
Tetraterpenes (C₄₀) are very common both in the plant and the animal kingdom. About 200 naturally abundant tetraterpenes are known to date and referred to as carotenoids; because all of them represent structural variants or degradation derivatives of β -carotene (Fig. I. 22) from the carrot *Daucus carota* (Umbelliferae) with 11 to 12 conjugated CC double bonds.



Squalene – triterpene



Lanosterol – triterpene

 β -Carotene – tetraterpene

Tocopherol – prenylquinones

Figure I. 22. Other representative terpene structures (Adams, 1995).

Carotenoids are red and yellow plant pigments, which occur in leaves, flowers, seeds and fruits, but not in wood. They are a type of medicinal and biotechnological important class of natural pigments which have many different biological functions, including species-specific

coloration, photo protection, light harvesting, pigmentation, and hormonal activity. Currently, carotenoids are produced as food colorants in animal feeds, nutritional supplements and for cosmetic and pharmaceutical purposes. Carotenoids exhibit significant anti-carcinogenic activities and play an important role in prevention of chronic diseases (Rao and Rao, 2007).

More than 1800 plant polyisoprenes have been identified. Prenylquinones contain terpenyl groups with up to ten isoprene units. Various lipid-soluble prenylbenzoquinones are coenzymes involved in electron transport during the respiration processes in mitochondria (O'Neil and Merck & Co., 2001). Vitamins of the K series ("Koagulation" vitamins) are chemically classified as prenyl-1,4-naphthoquinones. They are ingested with food originating from all green plants, are involved in oxidative phosphorylation during respiration processes and in the biosynthesis of glycoproteins in the liver, and are required as coagulation agents for blood. Vitamin E also known as tocopherol (fig. 22) occurs in fruits, vegetables and nuts, and is enriched in wheat germ and oils, particularly palm, soybean and sunflower (Breitmaier, 2006). Natural rubber is a linear polymer of 400 to more than 100000 units of 1,4-polyisoprene polymerized in *cis* configuration (caoutchouc).

2.3. Terpenes synthases, terpene modification enzymes

Terpene synthases are responsible for the diversity that makes terpenes such interesting molecules. The biosynthetic routes converge at the prenyl diphosphates for all terpenes. The cyclization product heavily influences the chemical properties of the ultimate product. Several crystal structures of sesquiterpene synthases are available, providing information on the mechanism of cyclization (Lesburg *et al.*, 1997; Starks *et al.*, 1997; Caruthers *et al.*, 2000; Rynkiewicz *et al.*, 2001; 2002). In addition, a few hundred terpene synthase genes were cloned, though many produce replicate products. A survey of all of the available crystal structures reveals remarkable structural similarities found even in cases of less than 18% primary sequence similarity. The prenyl transferases and terpene synthases appear to have evolved from a common ancestor. This is evidenced by the crystal structures of the chicken FPP synthase and tobacco epi-aristolochene synthase being remarkably similar (Tarshis *et al.*, 1994; Starks *et al.*, 1997). Both proteins are entirely formed of α -helices creating a distinct cavity referred to as the terpene fold.

In order to understand and engineer the chemical diversity of terpene synthases, it was important to decipher how these enzymes exploit the common structural scaffold to specify the

enormous number of different products. Since this chemical diversity is the product of only three substrates, it was also necessary, for further enzymatic manipulations, to determine the molecular determinants of substrate selectivity (Kampranis *et al.*, 2007).

2.3.1. Mechanism, structure

Terpene hydrocarbon scaffolds are generated by the action of mono-, sesqui-, and diterpene synthases that catalyze multistep reactions with diphosphorylated substrates of 10 [geranyl diphosphate (GPP)], 15 [farnesyl diphosphate (FPP)] or 20 [geranylgeranyl diphosphate (GGPP)] carbon atoms. The reactions catalyzed by terpene synthases are unparalleled relative to other classes of enzymes because they often consist of a series of stereochemically complex steps. These reactions include ionization of the diphosphate substituent creating an acyclic and reactive carbocation intermediate. Additional steps add increasing complexity and include regio- and stereospecific formation of single or multiple rings, proton eliminations to form double bonds, water quenching of carbocations to create terpene alcohols, and stereospecific hydride, methyl, and methylene migrations (Tholl, 2006).

All this chemical complexity is catalyzed by enzymes whose three dimensional structure is highly conserved from fungi to plants, characterized by the “terpene fold” and an active site lined mainly by inert amino acids (Lesburg *et al.*, 1997; Starks *et al.*, 1997; Rynkiewicz *et al.*, 2001; Christianson, 2006). However, apart from the presence of a short conserved motif related to metal ion binding (DDxxD), the sequence similarities between terpene synthases are dominated by relationships regardless of substrate or product specificity (Bohlmann *et al.*, 1998).

Prenyl diphosphate cyclization begins with substrate binding. Studies with aristolochene synthase show that the conformation of the bound FPP partially defines the product (Caruthers *et al.*, 2000). The later reaction steps are presumed to be very fast and there are several potential reaction cascades. The substrate must be bound in such a way as to place reactive carbons next to each other ensuring proper product stereochemistry. The stereochemical control exhibited by terpene synthases is extremely important to the production of a biologically active product. In addition, the binding of the substrate leads to a conformational change in the synthase. This change sequesters the substrate away from water in a deep hydrophobic pocket (Starks *et al.*, 1997). Either water or the released diphosphate group could prematurely terminate the reaction.

The diphosphate group is in complex with Mg^{2+} and conserved residues Arg264 and Arg441 (epi-aristolochene numbering (Starks *et al.*, 1997)).

The substrate ionization is initiated by the departure of the diphosphate group catalyzed by bound Mg^{2+} or Mn^{2+} . These divalent metal ions are bound in complex with the substrate diphosphate group and conserved aspartates in the “aspartate-rich motif” (Lesburg *et al.*, 1998). The diphosphate group is channeled out of the pocket by various hydrophilic residues leaving an allylic carbocation. Once the initial ionization is completed, the primary role of the terpene synthase is to stabilize some carbocation species while excluding water (Lesburg *et al.*, 1997; Starks *et al.*, 1997)). The series of reactions by the ionized substrate is unique to each synthase and defines the final product formed. The enzymes that maintain tight control over the carbocation species will produce a single dominant product, while others such as δ -selinine synthase and γ -humulene synthase retain poor control over the intermediate species and produce dozens of products each (Fig. I. 23) (Steele *et al.*, 1998).

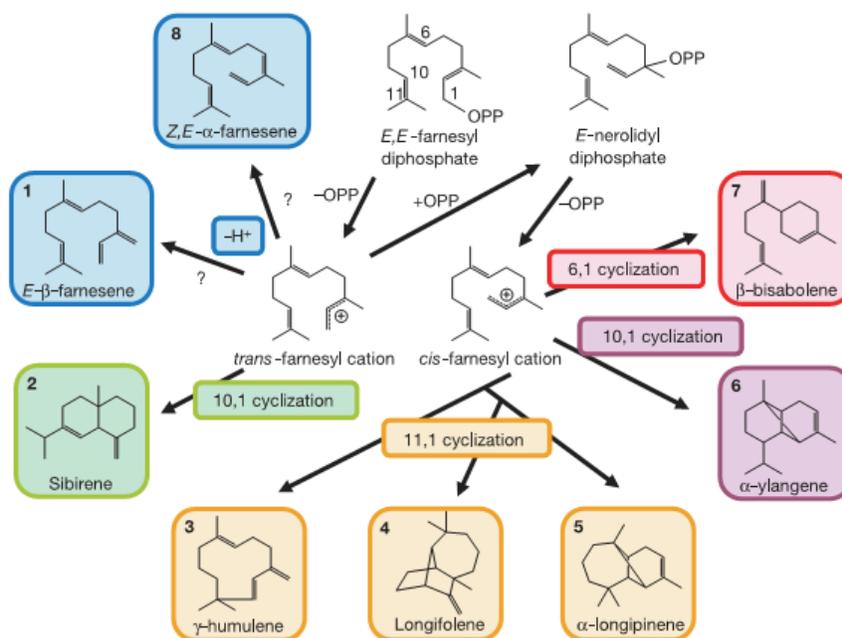


Figure I. 23. γ -Humulene synthase cyclization reaction mechanisms. When the substrate, farnesyl diphosphate, binds to the enzyme active site via divalent magnesium cations, the diphosphate group is released to yield either trans- or cis-farnesyl cation. From the trans-farnesyl cation, sibirene (2) is produced by the 10,1 cyclization reaction. From the cis-farnesyl cation, γ -humulene (3), longifolene (4) and α -longipinene (5) are produced through an 11,1 cyclization reaction; α -ylangene (6) through a 10,1 cyclization reaction; and β -bisabolene (7) through a 6,1 cyclization reaction. *E*- β -farnesene (1) and *Z,E*- α -farnesene (8) can be produced by directed deprotonation from either farnesyl cation.

An alternate cyclization mechanism has been observed in some diterpene synthases. The mechanism is very similar to the cyclization of squalene in triterpene synthesis. In this scheme, the carbocation species is created by protonation of the double bond distal to the diphosphate group (Ravn *et al.*, 2000; Williams *et al.*, 2000; Schwab *et al.*, 2001; Hamano *et al.*, 2002). This results in the formation of a cyclic diphosphate intermediate, often copalyl diphosphate in plants. The enzymes involved in this first step often lack the characteristic aspartate-rich motifs that are characteristic of the other class of terpene synthase (Schwab *et al.*, 2001). A second enzyme further cyclizes this intermediate in a mechanism similar to that of the terpene synthases discussed previously. This two-step cyclization mechanism has been seen in both plants and in *Streptomyces* (Hamano *et al.*, 2002).

2.3.2. Monoterpene synthases. Mechanism of action

Monoterpene synthases, often referred to as 'cyclases', catalyze the reactions by which geranyl pyrophosphate is cyclized to the various monoterpene skeletons. These enzymes have received considerable attention because of the basic character of their end products determined by the cyclization mechanism, which is quite complex, involving multiple alteration steps in bonding, hybridization, and configuration. Research on monoterpene cyclases has also been stimulated by the possible regulatory importance of these enzymes that function at a branch point in isoprenoid metabolism, as well as by the commercial significance of the essential oils and aromatic resins, and by the ecological roles of these terpenoid secretions, especially in plant defence (Colby *et al.*, 1993).

Monoterpene synthases are nuclear encoded, operationally soluble pre-proteins destined to be imported in the plastids, where they are proteolytically processed into their mature forms (Lucker *et al.*, 2002; El Tamer *et al.*, 2003). In all of the monoterpene synthases, the 50-60 N-terminal residues are characterized by low degree of similarity, typical of targeting sequences. These sequence, all of them, share common features of transit peptides rich in serine, threonine, and small hydrophobic residues, but with few acidic residues. Thus, all native monoterpene synthases so far examined appear to be NH₂-terminally blocked, preventing direct determination (by sequencing) of the transit peptide - mature protein cleavage junction. Significantly, a tandem pair of arginine residues is strictly conserved in the deduced sequence of all monoterpene synthases and they define the most NH₂-terminal regions of obvious homology, suggesting a

possible cleavage site as RRX₈W. This motif has been reported as required to maturation of functional protein and to diphosphate migration step, accompanying formation of the intermediate linalyl diphosphate before the final cyclization step catalysed by the monoterpene synthases (Williams *et al.*, 1998). Downstream of the tandem arginines are several regions of homology, including the highly conserved DDxxD motif found in virtually all deduced sequences for enzymes that utilize prenyl diphosphate substrate. Both these sequences RRX₈W and DDxxD are conserved among all the monoterpene synthases (McGarvey and Croteau, 1995; Bohlmann *et al.*, 1998; Lucker *et al.*, 2002; Wise *et al.*, 2002). Several other highly conserved regions are also apparent (Wise *et al.*, 2002). The previously characterized monoterpene synthases have histidyl residues at the active site based on substrate protection against inhibition by imidazole-directed reagents (Savage *et al.*, 1994).

A number of monoterpene synthases from angiosperms, gymnosperms, and bryophytes have been partially purified and characterized, and all have similar properties: native molecular mass in the range of 50 to 100 kDa, either monomers (most monoterpene synthases), or homodimers (bornyl diphosphate synthase and possibly pinene synthase) (Iijima *et al.*, 2004), requirement for a divalent metal ion (usually Mg²⁺ or Mn²⁺), pI value near 5.0, and pH optimum within a unit of neutrality.

An interesting feature of several monoterpene synthases is their ability to form multiple products from single substrate. For example the α -pinene synthase from sage and grand fir produces both α - and β -pinene. Another example is the limonene synthase that generates smaller amounts of myrcene, α -, and β -pinene in addition to the principal cyclic product limonene (Croteau *et al.*, 1987; McGarvey and Croteau, 1995; Bohlmann *et al.*, 1998; El Tamer *et al.*, 2003). This ability to synthesize multiple products from a single substrate requires the nomenclature of these enzymes to be based on the identity of the principal product synthesized by each one (Trapp and Croteau, 2001).

All monoterpene synthase are capable of catalyzing both the isomerization and cyclization reactions, and these steps occur via a series of ion pairs at the same active site. To date, at least two crystal structures for monoterpene synthases are available to understand the mechanistic details of GPP cyclization (Whittington *et al.*, 2002; Kampranis *et al.*, 2007). GPP serves as a universal precursor of regular monoterpenes and studies with monoterpene synthases from sage provided evidence for the general mechanistic of the coupled isomerisation-

cyclization (Fig. I. 24) that involves initial ionization of geranyl pyrophosphate, with *sin*-migration of the pyrophosphate moiety of the ion pair, to provide the bound tertiary allylic intermediate, linalyl pyrophosphate (Wise *et al.*, 2002). In this ‘ionization - isomerization’ step, which removes the topological barrier to cyclization, the first formal chiral center is introduced at C₃ (*i.e.* either (3R)- or (3S)-linalyl pyrophosphate). After rotation about the C₂-C₃ bond to afford the cisoid, *anti-endo* conformer, linalyl pyrophosphate is itself ionized with C₆-C₁ cyclization generating the corresponding monocyclic (R)- or (5′)- α -terpinyl cations pyrophosphate anion pair (Wheeler and Croteau, 1987).

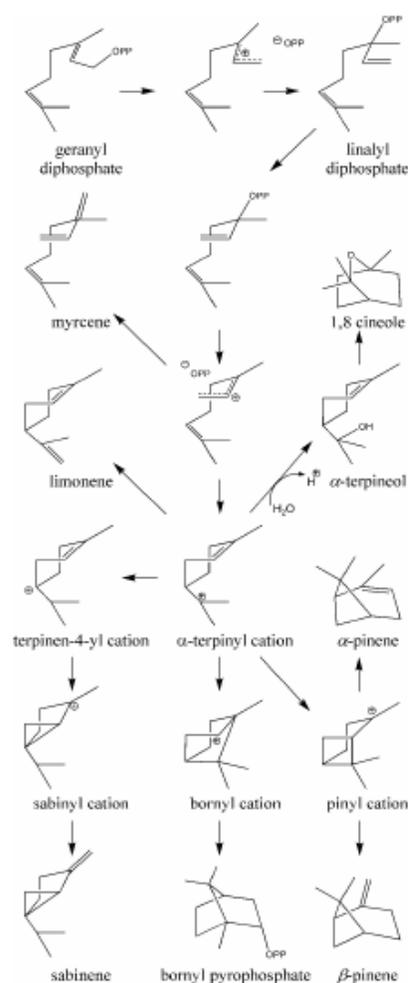


Figure I. 24. Scheme describing the reactions catalyzed by the *Salvia* monoterpene synthases and their mutants.

From this universal intermediate, the reaction may take one of several routes involving internal additions to the remaining double bond, hydride shifts, or rearrangements before the

terminal carbocation is deprotonated to an olefin or captured by water or the diphosphate anion. In the simplest of all terpenoid cyclizations, the α -terpinyl cation is deprotonated to yield limonene. Alternatively, the α -terpinyl cation may undergo further cyclization, via the remaining double bond, to afford the pinyl cation (and then α - or β -pinene after deprotonation) or the bornyl cation (to form bornyl diphosphate by capture of the diphosphate). Hydride shifts in the α -terpinyl cation yield the terpinen-3-yl or terpinen-4-yl cations, providing access to the phellandrenes and thujanes, respectively. Water capture of the α -terpinyl cation yields α -terpineol, which upon heterocyclization affords 1,8-cineole. A few monoterpene synthases produce acyclic products such as myrcene and linalool (Fig. I. 24) (Wise *et al.*, 1998).

2.3.3. Cineole synthase

The 1,8-cineole synthase from *S. fruticosa* (Sf-CinS1) was crystallized and its structure was determined at 1.95 Å resolution (Kampranis *et al.*, 2007). The secondary structure of Sf-CinS1 is that of 23 α -helices and eight 3_{10} helices. These are arranged with remarkable similarity to *Salvia officinalis* bornyl pyrophosphate synthase (So-BPPS) and tobacco 5-epi-aristolochene synthase (TEAS). The Sf-CinS1 monomer is split into two α -helical domains (Fig. I. 25). The N-terminal domain consists of eight α -helices arranged in an α - α barrel. The larger C-terminal domain consists of 15 α -helices and two 3_{10} helices arranged in an orthogonal bundle.

The metal binding DDxxD motif, common to all terpene synthases, is located on the C-terminal region. The active site, remarkably well conserved, is located within a large cavity created on the C-terminal region.

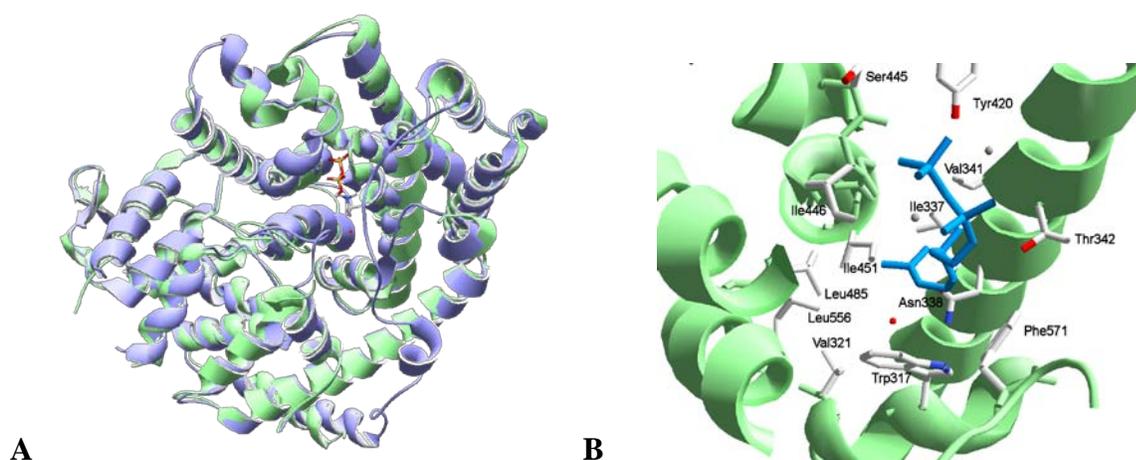


Figure I. 25. The structure of Sf-CinS1. (A) The N-terminal domain (blue) and the C-terminal domain (green). **(B)** The active-site region of Sf-Cin1 (Kampranis *et al.*, 2007).

Examination of the Sf-CinS1 active-site region compared with the So-BPPS crystal structure and the amino acid alignment of the different *Salvia* monoterpene synthases revealed a number of residues that were conserved between the different synthases and others that are varied and could play a role in product specificity. The variable residues appear to be clustered in two regions (Fig. I. 26). Region 1 is located on helix α 14 at the bottom of the active-site cavity and region 2 is part of the loop connecting helices α 18 and α 19.

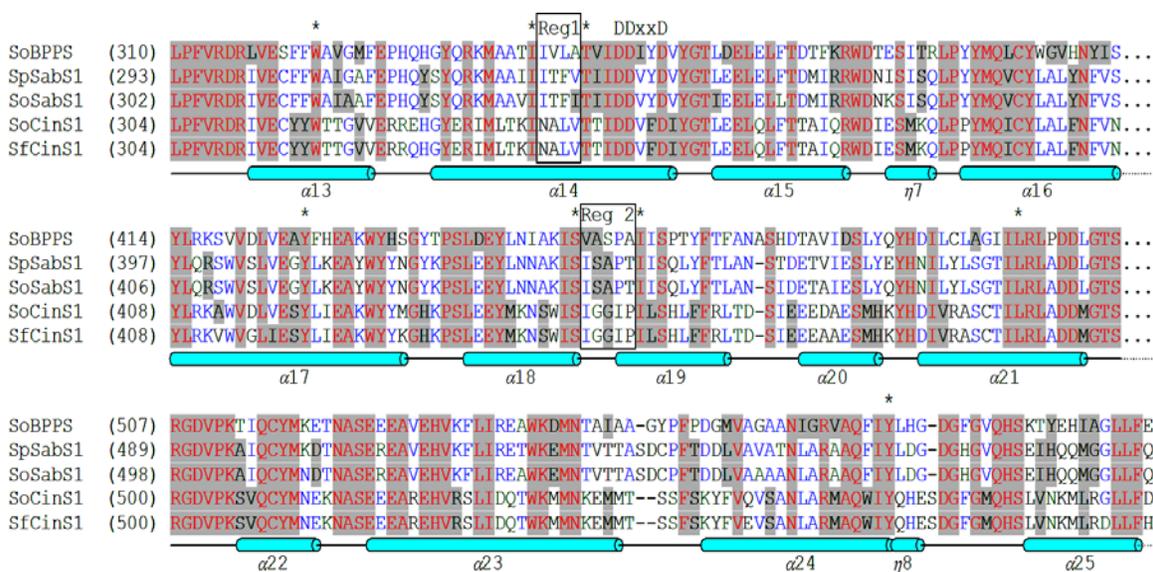


Figure I. 26. Amino acid sequence alignment of the C-terminal domain of monoterpene synthases from *Salvia* species (Kampranis *et al.*, 2007).

Central to the mechanism of 1,8-cineole synthesis is the water capture by the α -terpinyl carbocationic intermediate to yield α -terpineol, which undergoes protonation and internal addition to produce 1,8-cineole. Asparagine 338 is crucial for this step as it was found to hydrogen bond to a water molecule located at the bottom of the active site cavity (Fig. I. 27, left panel). Another major contribution to product specificity in Sf-CinS1 appears to come from a local deformation (kink) within one of the helices forming the active site that exposes the backbone carbonyls to the pocket (Fig. I. 27, right panel). This kink is observed in all other mono- or sesqui-terpene structures available pointing to a conserved mechanism. The importance of this deformation was confirmed by shifting of a proline in this region by one position, resulting in altered product specificity (Kampranis *et al.*, 2007).

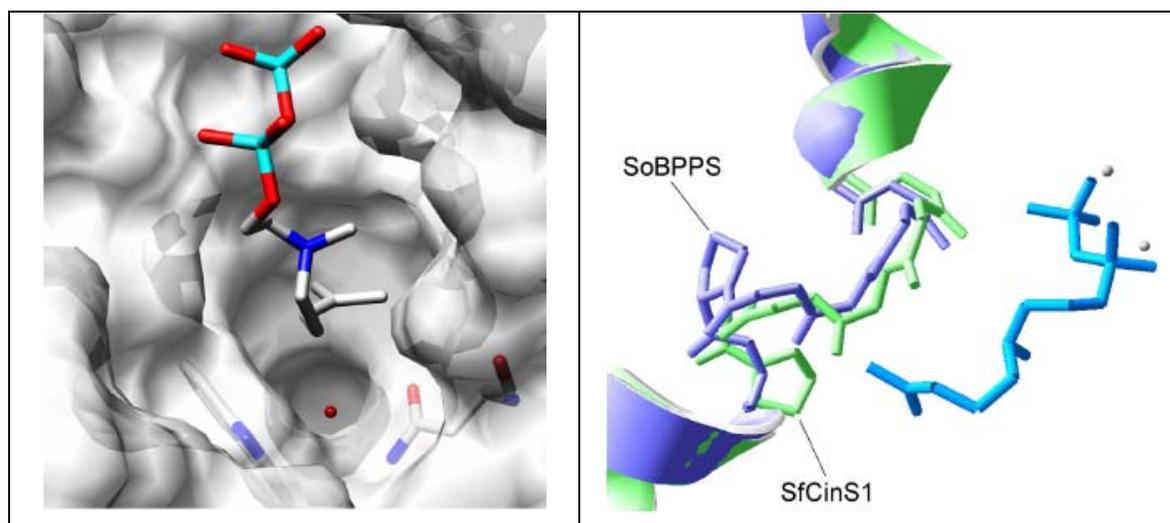


Figure I. 27. The two main structural elements responsible for product specificity in Sf-CinS1. Left, Region 1. The Sf-CinS1 active site showing Asn338, Trp317, and the water molecule (oxygen in red) likely involved in the hydroxylation of the α -terpinyl cation. **Right, Region 2.** Superimposition of Sf-CinS1 on So-BPPS to show the difference in the conformation of the kink in helix α 18. The conformation of the backbone together with the two proline residues are shown (Sf-CinS1 in green) (Kampranis *et al.*, 2007).

2.3.4. Sesquiterpene synthases

The class of sesquiterpene synthases shares common structural features that are reflected in highly conserved amino acid residues throughout the encoded protein (Bohlmann *et al.*, 1998). A very prominent feature of all terpene synthases is an Asp-rich region, the DDxxD motif, which is involved in the binding of a divalent metal cofactor (Marrero *et al.*, 1992). The substrate for sesquiterpene synthase is farnesyl diphosphate, an ubiquitous isoprenoid intermediate involved in cytoplasmic phytosterol biosynthesis. Sesquiterpene synthases lack plastidial targeting sequences and are localized to the cytoplasm. The similarities in reaction mechanisms between the plastidial monoterpene synthases and the cytosolic sesquiterpene synthases are paralleled by similarities in properties.

Interestingly, almost all terpene synthases show promiscuous function. Among those, γ -humulene (Fig. I. 28) and δ -selinene synthases are very promiscuous sesquiterpene synthases that are constitutively expressed in *Abies grandis*, each catalyzing the formation of at least 52 and 36 sesquiterpenes, respectively (Rynkiewicz *et al.*, 2001). In addition, these enzymes can use geranyl diphosphate as a substrate and catalyze the formation of monoterpenes.

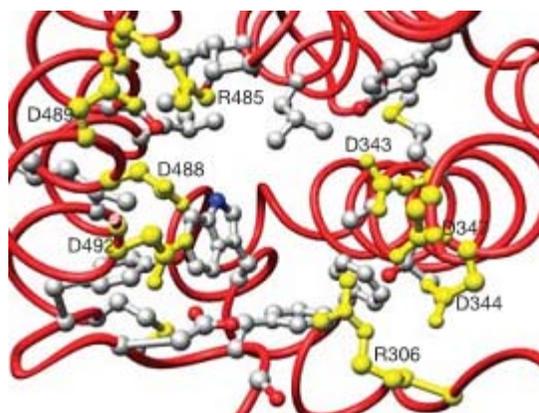


Figure I. 28. The homology structural model for the γ -humulene synthase active site. The residues that were not considered are shown in yellow. Six aspartate residues in two different aspartate-rich motifs and two arginine residues, which are generally conserved in all sesquiterpene cyclases, were not considered, because these residues are thought to be catalytically important, and mutations to these residues would have decreased enzyme activity significantly (Yoshikuni *et al.*, 2006).

Two different groups have studied the δ -selinene and γ -humulene synthases in an effort to understand and control the broad product distribution of these synthases (Little and Croteau, 2002; Yoshikuni *et al.*, 2006a). Mutation of residues in the aspartate-rich motif resulted in the narrowing of product distribution but at the expense of product complexity (Little and Croteau, 2002). Other work has identified several non-catalytic residues that influence the product distribution. By mutating these residues, Yoshikuni *et al.* (2006b) were able to narrow the product distribution but maintain the product complexity. These residues appeared to control the branch points in the cyclization of FPP. Although the specific roles of these enzymes have not been identified; it is thought that they might create chemical libraries that are important in general defense against microbial invasion.

By contrast, many other terpene synthases have highly specialized functions and are often found to have very specific roles in the formation of bioactive metabolites. For example, (+)- δ -cadinene, vetispiradiene and 5-*epi*-aristolochene synthases (Fig. I. 29) catalyze the first reaction step of phytoalexin (anti-fungal agents) production in various plant species and yield their respective sesquiterpenes with more than 98, 90 and 70% selectivity, respectively (O'Brien and Herschlag, 1999; Aubourg *et al.*, 2002). In addition, these sesquiterpene synthases are known to be expressed upon elicitation (Bohlmann *et al.*, 1998; Martin *et al.*, 2004).

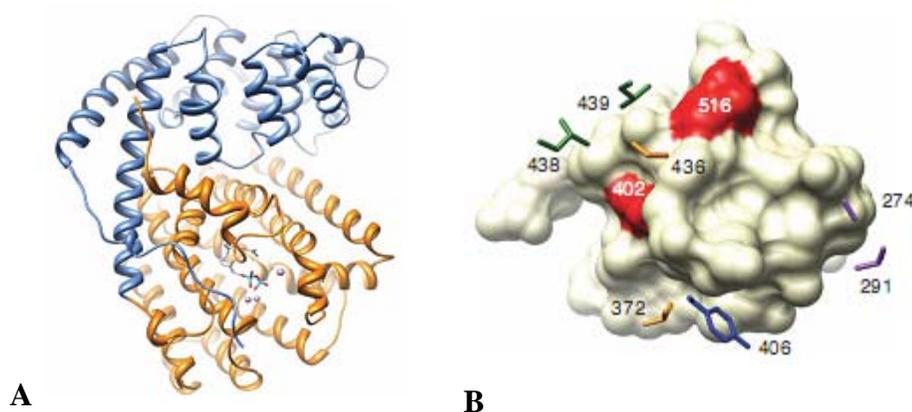


Figure I. 29. Overall structure of 5-epi-aristolochene. (A) Tertiary structure of TEAS shown as ribbons, showing the N-terminal domains (blue), C-terminal domain (yellow) and Mg²⁺ and farnesyl diphosphate modeled into the active site (O'Maille *et al.*, 2008).

Recent studies revealed the important elements in determining the product selectivity of sesquiterpenes synthases. Back and Chappell (1996) swapped domains of the epi-aristolochene and vetispiradiene synthases to identify the domains that determine the final product (Back and Chappell, 1996). Each of the synthase genes was cloned from different species, but the cyclization mechanism is identical until the final step that produces the distinct product of each synthase. They determined that the first 261 amino acids and the last 106 residues could be interchanged between the two proteins with no effect on the final product. However, a middle region between residues 261 and 442 was the region responsible for determining the final product. They were even able to engineer chimeric proteins that formed both products at various ratios.

Tobacco 5-epi-aristolochene synthase (TEAS) and henbane premnaspirodiene synthase (HPS) are closely related (75% amino acid identity), yet they cyclize ionized farnesyl diphosphate to form 5-epi-aristolochene (5-EA) and premnaspirodiene (PSD), respectively (fig. I. 30). The discovery of 4-epi-eremophllene, 4-EE biosynthetic activity supports hybridization of the final two biosynthetic steps in TEAS and HPS, involving a methyl migration shared with TEAS and a final deprotonation at C6 shared with HPS (Yoshikuni *et al.*, 2006a).

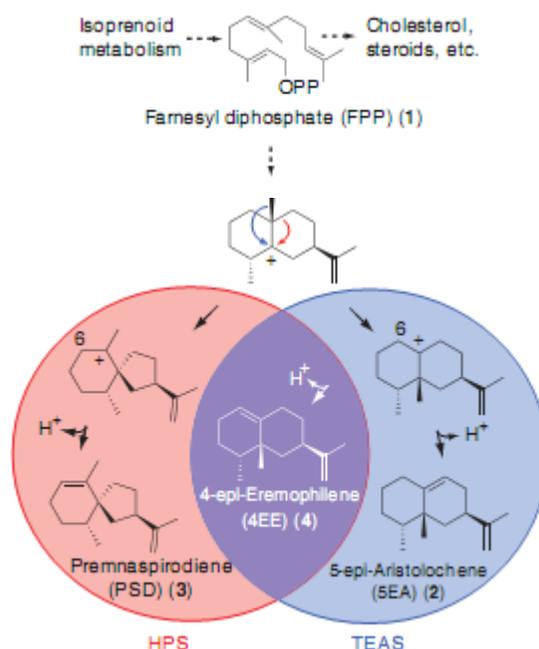


Figure I. 30. Terminal cyclization steps of TEAS and HPS terpene synthases (O'Maille *et al.*, 2008).

The sesquiterpene β -Caryophyllene has been detected in several organism and the genes for three caryophyllene synthases have been reported. Rasmann *et al.* showed that β -caryophyllene released by maize roots damaged by beetle attack attracts predatory nematodes to protect the plant (Rasmann *et al.*, 2005).

2.4. Terpenoid biosynthesis

Terpenoids are a class of compounds derived from the universal precursor isopentenyl diphosphate (IPP) and its allylic isomer dimethylallyldiphosphate (DMAPP), also called isoprene units. Terpenoid building blocks are then formed through condensation of additional IPP moieties via prenyltransferases (Fig. I. 31).

Monoterpenoids are derived from geranyl pyrophosphate (GPP, C10), sesquiterpenoids are derived from farnesyl pyrophosphate (FPP, C15), and diterpenoids are derived from geranylgeranyl pyrophosphate (GGPP, C20). Even higher order terpenoids are possible through condensation of these intermediates to larger precursor moieties. For example, sterols are derived from the triterpenoid squalene (C30), which contains six isoprene units through condensation of two molecules of FPP, and carotenoids (C40) are largely formed through condensation of two molecules of GGPP to yield eight-isoprene-unit compounds. After the formation of the acyclic

terpenoid structural building blocks (for example, GPP, FPP, GGPP), terpene synthases act to generate the main terpene carbon skeleton.

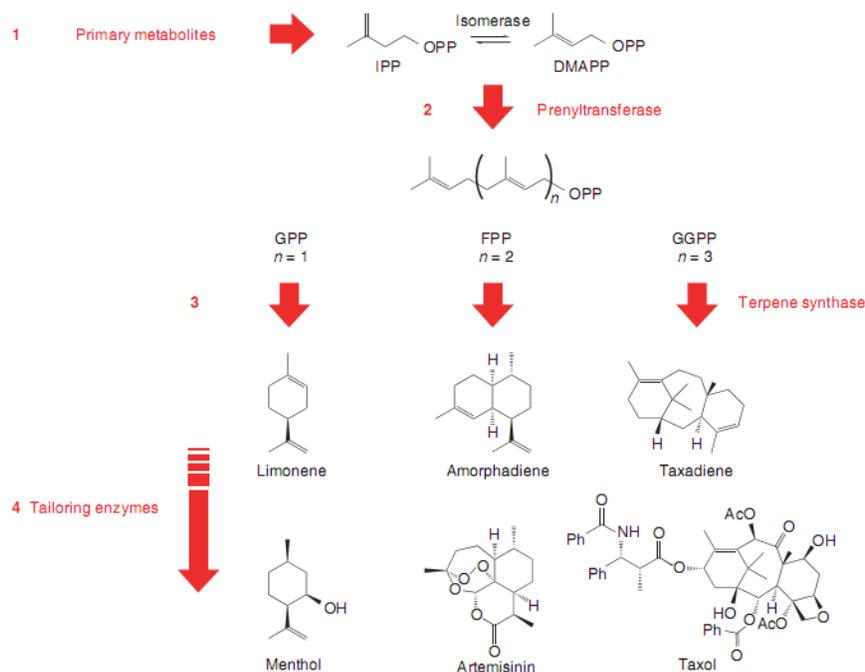


Figure I. 31. Isoprenoid biosynthesis. (1) C5 monomers IPP and DMAPP are synthesized from primary metabolites. (2) Head-to-tail condensations of IPP with a DMAPP starter unit generate the linear prenyl diphosphate precursors GPP, FPP and GGPP. (3) The terpene synthase catalyzes the cyclization or rearrangement to the parent carbon skeleton with loss of pyrophosphate. (4) Tailoring enzymes elaborate the parental backbone to generate the final product. Me, methyl; Ac, Acetyl; Ph, Phenyl. (Chang and Keasling, 2006).

Additional transformations often involving oxidation, reduction, isomerization, and conjugation enzymes decorate or alter the main skeleton with varied functional groups to yield the tremendously diverse terpenoid family of compounds.

2.4.1. Cellular structures of synthesis

Monoterpenes and sesquiterpenes, diterpenes resin acids, and a variety of other natural products such as sucrose and glucose esters of fatty acids, nitrogenous compounds (indole, nicotine), phenolic compounds, fatty acids, nectar and proteins and polysaccharides are products synthesized and secreted by glandular structures of plants. Glandular tissues are a common anatomical feature throughout the plant kingdom and many of the products formed by glandular tissues are commercially important because of their roles in attracting pollinators, protecting crop plants by deterring herbivore attack, contributing to the flavor and aroma of many plants, or

because of they are specialty chemicals used in industry for diverse purposes (Morimoto *et al.*, 2008). Example of glandular tissues include trichomes, which accumulate essential oils and resins; secretory cavities that accumulate aromatic oils in the epidermis of citrus fruits; ducts and cavities that accumulate oleoresin in conifers, and osmophores, a form of glandular epidermis common in floral tissues, which secrete volatile compounds responsible for attracting pollinators. Broadly defined, trichomes are an outgrowth of the epidermis, which can be present on any of the aerial or root surfaces of plants (Fig. I. 32).

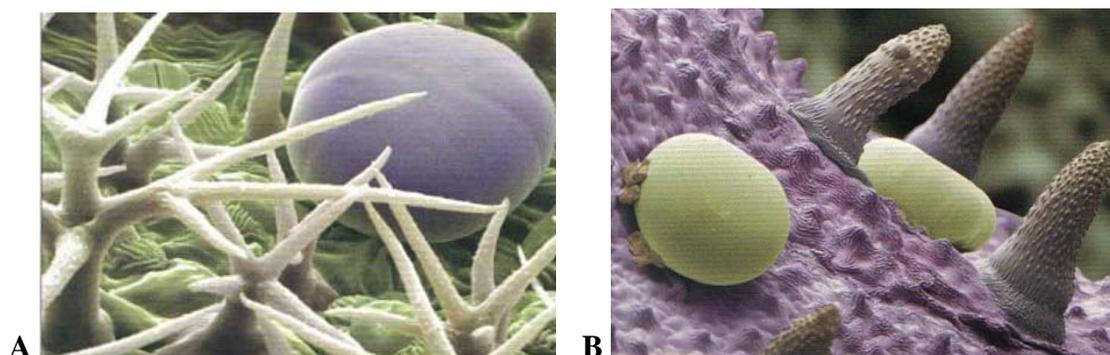


Figure I. 32. Plant secretory structures. (A) Lower leaf surface showing sessile secretory glands and dense covering of non-secretory trichomes; (B) Sessile secretory glands and non-secretory trichomes on calyx surface. (Svoboda, 2000)

Glandular trichomes are modified epidermal hairs and can be found covering leaves, stems, and even parts of flowers such as the calyx in many plants of the Labiatae family. These include basil (*Occimum basilicum*) lavender (*Lavandula* spp.) marjoram and oregano (*Origanum* spp.), mint (*Mentha* spp.) and thyme (*Thymus* sp.). The secretory cells are attached by a single stem or basal cell in the epidermis (Fig. I. 33). The outer surface of the gland is heavily cutinized. A toughened cuticle in which no pores or perforations are present, usually completely covers the trichome. They are a suitable model system for studying many aspects of carbon metabolism in secretory tissues (McCaskill and Croteau, 1999).

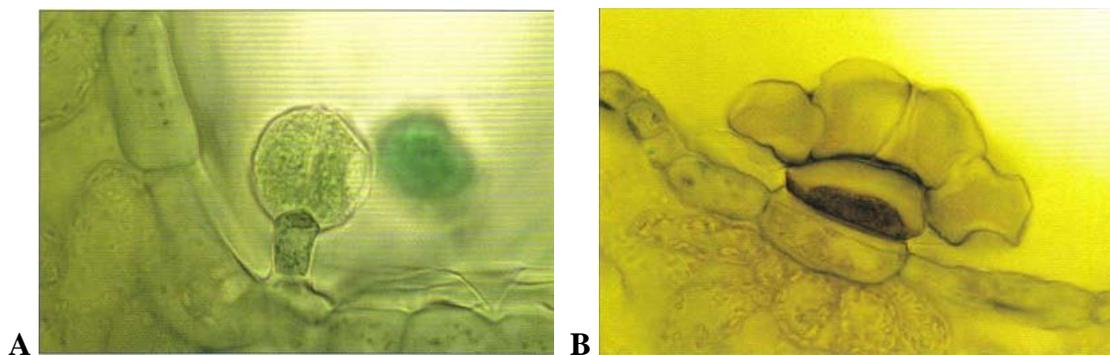


Figure I. 33. Details on secretory gland of sage (*Salvia* sp.). (A) Stalked secretory gland with unicellular head on lower leaf surface; (B) Sessile secretory gland (TISI on lower leaf surface showing basal cell and 4 secretory cells. (Svoboda, 2000)

The essential oils accumulate in subcuticular spaces and it is thought that they diffuse outwards through the cuticle (Gersbach, 2002). The glandular cells differ from normal plant cells in that they have a very large nucleus and dense protoplasm that lacks a large central vacuole. There are numerous plasmodesmata across the walls of the gland cells especially between the stalk cell and the collecting cell. In the very young gland the intracellular organisation is almost identical to that of the adjacent cells but as the secretory cells develop complex changes occur. The membrane system progressively degenerates and in the fully-developed glands only a fine granular cytoplasm remains.

Although generalization can be made about metabolism in glandular tissues, based on the biosynthetic pathway involved in producing the secreted product, many features of glandular metabolism are unlike that of other plant tissues. Glandular tissues are capable of devoting prodigious amount of carbon and energy to the production of specialized chemicals that are often toxic and not essential for normal growth or development. Terpenoids in particular are among the most expensive compounds to produce in terms of the carbon and energy requirements (Opitz *et al.*, 2008). Terpenoid products formed by glandular tissues typically occur from trace levels up to approximately 1% of tissue dry weight, although there are exceptions. Examples include the accumulation of up to 20% dry weight of monoterpenoid essential oils in eucalyptus leaves (Fox and Morrow, 1981), 16% dry weight of papyriferic acid, a triterpene, in immature buds of Alaska paper birch (*Betula resinifera*), and over 8% dry weight of cannabinoids in *Cannabis* (Lehmann *et al.*, 1997).

Numerous reports have described the isolation of the secretory cell from glandular trichomes and other secretory tissues from various plants. Most studies have employed isolated glandular trichomes as enriched sources for the isolation of individual enzymes involved in biosynthesis of the specific products formed by the glands (Guo *et al.*, 1994). Isolated glandular trichomes have been used as a source of mRNA for construction of enriched cDNA libraries.

2.4.2. Biochemical pathways

Two distinct pathways generate the universal C5 precursors IPP and DMAPP (Fig. I. 34). The classic mevalonate (MVA) pathway was discovered in the 1950s and was assumed to be the sole source of the terpenoid precursors IPP and DMAPP. The MVA pathway is active in bacteria, plants, animals and fungi and functions in the cytosol to generally supply the precursors for production of sesquiterpenes and triterpenes. Recently, labeling experiments in bacteria and plants revealed the presence of an alternate pathway to IPP and DMAPP supply (Eisenreich *et al.*, 2001; Rodriguez-Concepcion and Boronat, 2002).

A nonstandard mevalonate pathway involving the phosphorylation of isopentenyl phosphate was discovered in *M. jannaschii* (Grochowski *et al.*, 2006). Before 1993, the mevalonate pathway was the only known source of isoprenoids. After isotope-labeling studies by Rohmer *et al.*, it was shown that there was an alternate pathway to isoprenoids that did not originate from acetyl-CoA (Rohmer *et al.*, 1993). The complete pathway (Fig. I. 34) was finally elucidated in 2002 (Rohdich *et al.*, 2002).

This pathway, which is named after the first committed precursor, 2-C-methyl-d-erythritol-4-phosphate (MEP; the pathway is also sometimes referred to as the DXP pathway), is plastidial in nature and is generally used to supply precursors for the production of monoterpenoids, diterpenoids and tetraterpenoids. Although interaction between these pathways is still largely undiscovered, recent evidence has revealed an exchange of intermediates between the cytosol and the plastid (Bick and Lange, 2003; Hemmerlin *et al.*, 2003; Laule *et al.*, 2003). The deoxyxylulose-5-phosphate (DXP) pathway is generally found in prokaryotes. The DXP pathway also produces the isomer to IPP and DMAPP, whereas in the mevalonate pathway IPP is the sole product (Rohdich *et al.*, 2002). IPP is isomerized to DMAPP by the action of isopentenyl diphosphate isomerase (Idi).

The mevalonate pathway is the more studied of the two pathways to IPP. The biosynthesis of IPP begins with the conversion of three molecules of acetyl-CoA to mevalonate through acetoacetyl-CoA and β -hydroxy-beta-methylglutaryl coenzyme A. The last step in this sequence resulting in mevalonate has been the focus of much research (Szkopinska *et al.*, 2000; Friesen and Rodwell, 2004). The statins, potent cholesterol-lowering medications, are inhibitors of HMG-CoA reductase. Sequential phosphorylation of mevalonate to diphosphomevalonate followed by decarboxylation produces IPP. It is worth noting that this makes the activity of an IPP–DMAPP isomerase essential.

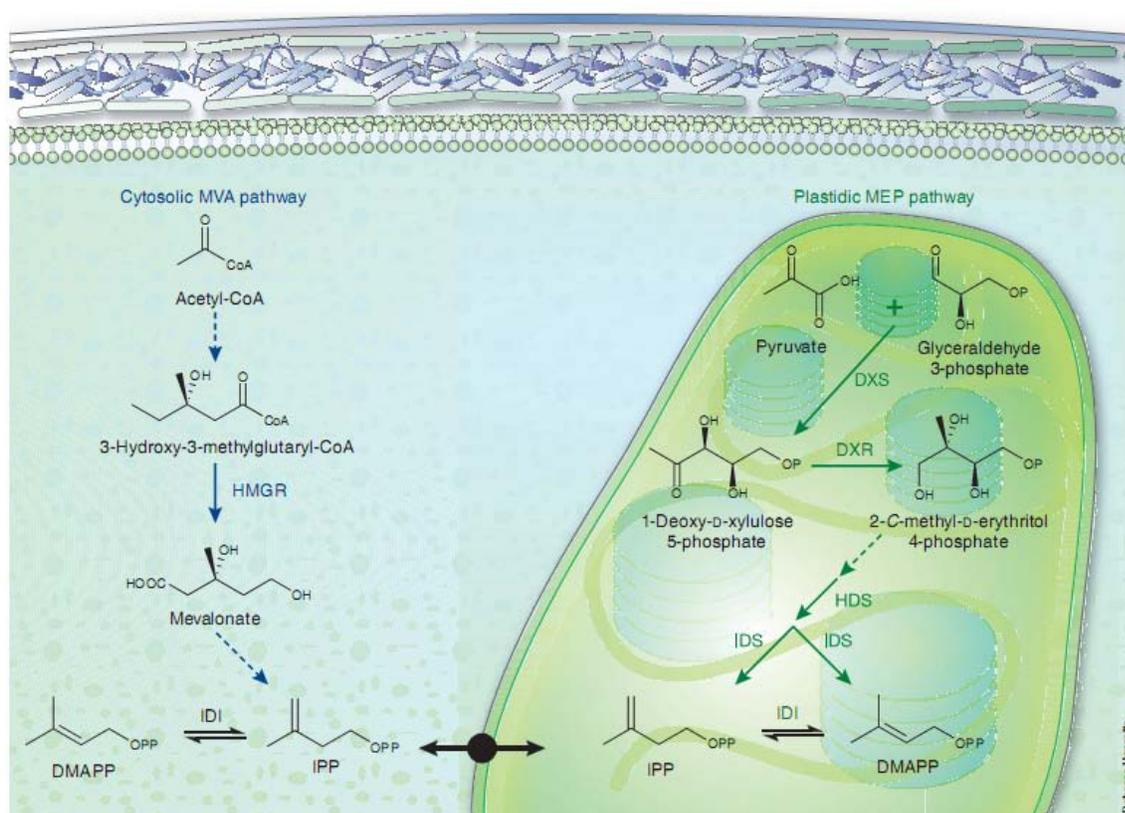


Figure I. 34. Compartmentalized biosynthesis of IPP and DMAPP. (Left) Via the cytosolic mevalonate (MVA) pathway. HMGR, 3-hydroxy-3-methylglutaryl coenzyme A reductase; IDI, isopentenyl diphosphate isomerase. (Right) Via the plastidic MEP pathway. DXS, 1-deoxy-d-xylulose-5-phosphate synthase; DXR, 1-deoxy-d-xylulose-5-phosphate reductoisomerase; HDS, hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase; IDS, isopentenyl diphosphate: dimethylallyl diphosphate synthase; IDI, isopentenyl diphosphate isomerase. Dashed arrows indicate more than one step. (Roberts, 2007).

The DXP pathway begins with the production of DXP from one molecule each of pyruvate and glyceraldehyde-3-phosphate catalyzed by 1-deoxy-d-xylulose-5-phosphate synthase (Dxs). DXP is a known precursor for non-isoprenoids such as vitamins. The first committed step is the intramolecular rearrangement of DXP into 2-C-methyl-D-erythritol 4-phosphate (MEP). Deoxyxylulose reductase (IspC) catalyzes this reaction in a NADPH- and Mn²⁺-dependent manner (Takahashi *et al.*, 1998). IspD (4-diphosphocytidyl-2-C-methyl-D-erythritol synthase) catalyzes the reaction of MEP with cytidine 5'-triphosphate to form 4-diphosphocytidyl-2-C-methylerythritol, which is then phosphorylated by IspE in an ATP-dependent manner (Rohdich *et al.*, 1999; Luttgen *et al.*, 2000). The product of this reaction (4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate) is cyclized by IspF to form 2C-methyl-D-erythritol 2,4-cyclodiphosphate with the last two steps to form IPP and DMAPP being catalyzed by IspG and IspH (Wungsintaweekul *et al.*, 2001; Rohdich *et al.*, 2002).

In contrast to the mevalonate pathway, the DXP pathway produces IPP and DMAPP in a 5:1 ratio. This partially explains why early studies on the *E. coli* *idi* determined that the gene was not essential for growth (Hahn *et al.*, 2001). Subsequent efforts have shown that increased expression of *Idi* results in improved terpene production in *E. coli*. Work by Martin *et al.* demonstrated that there is insufficient natural *Idi* activity to complement an IspC mutant *E. coli* by providing mevalonate-derived IPP (Martin *et al.*, 2003). It was suggested that the DXP pathway would be an excellent target for new antibiotic development, as the pathway genes have no known homologs in humans (Rodriguez-Concepcion, 2004).

The prenyl transferases are responsible for the formation of the higher molecular weight polyprenyl diphosphate precursors. They catalyze the head-to-tail addition of IPP units to allylic diphosphates, typically DMAPP. These reactions produce the 10-, 15-, and 20-carbon precursors' geranyl diphosphate (GPP), farnesyl diphosphate (FPP), and geranylgeranyl diphosphate (GGPP), respectively (Fig. I. 35).

Reactions catalyzed by enzymes such as squalene synthase generate larger terpene precursors by combining the polyprenyl diphosphates. The terpene synthases, also called terpene cyclases, catalyze the reaction of the polyprenyl diphosphates to produce the carbon skeletons of the monoterpenes (C₁₀), sesquiterpenes (C₁₅), and diterpenes (C₂₀). The creation of the carbon skeleton is only the first committed step in the biosynthesis of many commercially relevant

isoprenoids. The carbon skeletons undergo several more enzymatic modifications, especially by cytochrome P450s, before being completed.

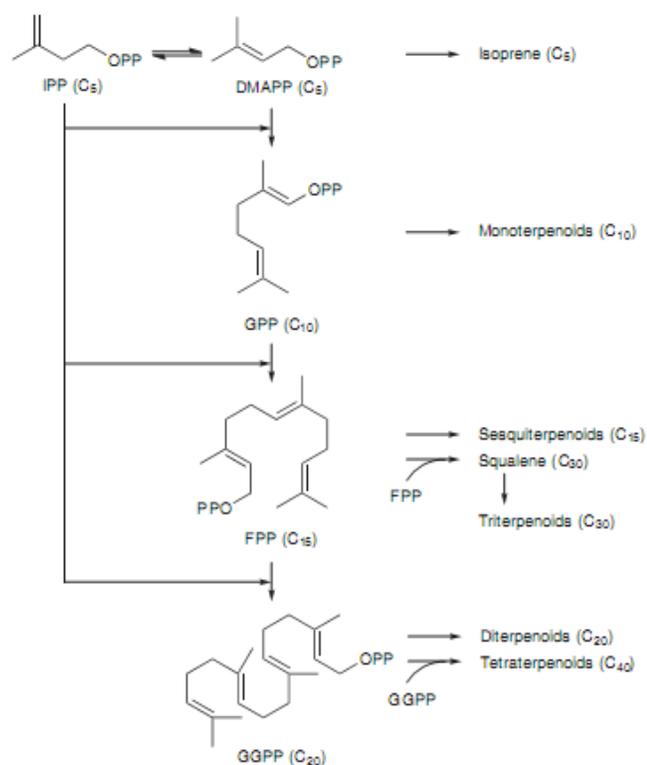


Figure I. 35. Overview of terpenoid biosynthesis and the generation of terpenoid starting blocks. Isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) condense to form higher order terpenoid precursors, including the monoterpene precursor geranyl diphosphate (GPP), the sesquiterpene precursor farnesyl diphosphate (FPP) and the diterpene precursor geranylgeranyl diphosphate (GGPP). Two molecules of FPP condense to form the triterpene precursor squalene, and two molecules of GGPP condense to form higher order tetraterpenoids. (Roberts, 2007)

The prenyl transferases catalyze the head-to-tail condensation of the 5-carbon prenyl diphosphates producing the longer prenyl diphosphates that are the immediate precursors to terpenes. The proteins catalyze the addition of IPP units to prenyl diphosphates with allylic double bonds to the diphosphate moiety. Most of the prenyl transferases accept DMAPP as the initial substrate, but they will also bind GPP or FPP depending on the particular synthase (Tarshis *et al.*, 1994). The physical dimensions of the protein catalytic cavity regulate product chain-length. Bulky amino acids opposite the substrate-binding domain define the length of the catalytic fold and thus the maximum length of any product formed (Ohnuma *et al.*, 1996b). From

these observations, Ohnuma et al. were able to change the product chain length by mutating a tyrosine from the active site of an FPP synthase, resulting in a GGPP synthase (Ohnuma *et al.*, 1996a). The ability to bind and extend any allylic substrate has produced a challenge for engineering monoterpene metabolism, as the native FPP synthase will consume the GPP precursor to produce FPP (Reiling *et al.*, 2004).

2.5. Engineering of terpenoid production

The application of modern biotechnology techniques to isoprenoid production can be divided into the engineering of native hosts and the engineering of heterologous hosts. Native host engineering has focused on elevating levels of terpenoids normally found in a species, typically plants. A specific example of native host engineering is the effort to increase menthol production in peppermint, *Mentha x piperita*, by the Croteau lab (Lange and Croteau, 1999; Lange *et al.*, 2000; Mahmoud and Croteau, 2003). Native host engineering also includes efforts to create plant tissue cultures that produce the terpenoids of interest. This approach is contrasted by the efforts to engineer isoprenoid production in a heterologous host. These hosts typically have little or no natural terpenoid production beyond the essential central metabolite isoprenoids such as sterols and quinones. Terpenoid metabolic engineering of downstream secondary metabolite pathway steps has been more limited, in part owing to the lack of both pathway definition and understanding of regulation, but some successes have been reported for taxanes, alkaloids and artemisinin.

2.5.1. Taxol

Taxol (paclitaxel), produced in *Taxus brevifolia* (Pacific yew), is a classical example of the plant kingdoms' ability to produce chemotherapeutics. Taxol is a complex and highly functionalized terpenoid that has been approved for the treatment of refractory ovarian and metastatic breast cancer. Its production is based on a semisynthetic route that chemically converts a paclitaxel intermediate isolated from the needles or cell cultures of various *Taxus* species (Kingston, 2000). However, the desire for a robust biosynthetic route using microorganisms still exists in order to meet growing demands.

Very elegant total chemical syntheses of taxol have been developed but, because of low yields and high costs, none of these approaches is suitable for its commercial production.

Similarly, *Taxus* cell cultures have yet to reach commercialization due to low and unstable productivity (Dejong *et al.*, 2006).

With the partial elucidation of the taxol biosynthetic pathway the possibility of taxadiene (a taxol biosynthetic intermediate) biosynthesis in *E. coli* was successfully demonstrated (Yang *et al.*, 2001). One major limitation with that approach however is that the prokaryotic *E. coli* can not functionally express the P450 enzymes that widely participate in the taxol biosynthetic pathway. Hence, *S. cerevisiae* was chosen as an alternative, since it has been reported that yeast was successfully employed as host for the engineered, multi-step production of other (mevalonate-derived) terpenoids, including steroids and carotenoids (Dejong *et al.*, 2006).

Recent progress has been made in the heterologous expression of the preliminary paclitaxel biosynthesis genes and optimization of paclitaxel intermediates in microbial hosts.

The biosynthesis of paclitaxel involves approximately 20 biochemical steps (Ketchum *et al.*, 2007) that do not necessarily operate in a linear fashion. As some steps of the metabolic pathway are missing, identification and reclassification of the involved steps continues (Jennewein *et al.*, 2004; Jennewein *et al.*, 2005; Rontein *et al.*, 2008).

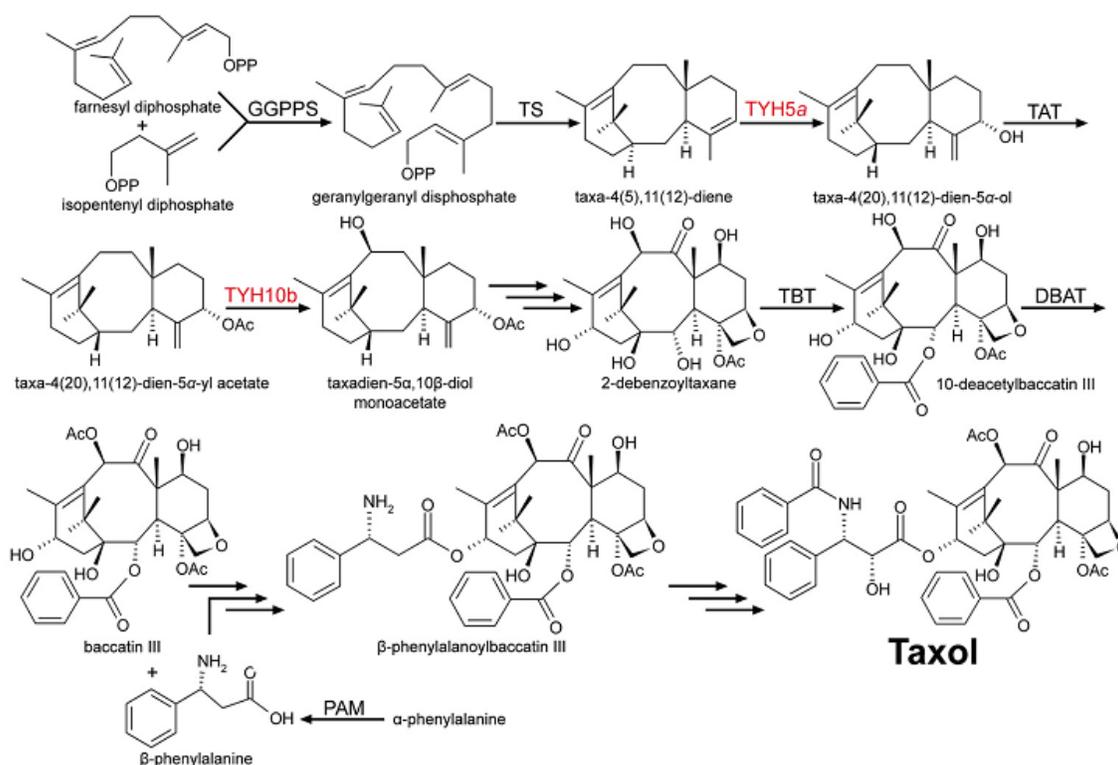


Figure I. 36. Taxol biosynthetic pathway. (Chemler *et al.*, 2006)

Recently, partial taxol biosynthetic pathway was constructed in *S. cerevisiae* by expressing five sequential pathway steps leading from primary isoprenoid metabolism to the intermediate taxadien-5 α -acetoxy-10 β -ol. The *Taxus cuspidate* genes the yeast host expressed included geranylgeranyl disphosphate synthase (GGPPS), taxadiene synthase (TS), cytochrome P450 taxadiene 5 α -hydroxylase (TYH5a), taxadienol 5 α -O-acetyl transferase (TAT) and taxoid 10 β -hydroxylase (THY10b). The recombinant strain produced taxadiene at 1.0 mg/L while taxadien-5 α -ol was produced in very small amounts (~25 μ g/L). These results suggest that the first two enzymes (GGPPS and TS) cooperated well with each other and that the metabolic flux was reduced at the 5 α -hydroxylation step, which is catalyzed by a cytochrome P450 hydroxylase. It is anticipated that overexpressing *Taxus* P450 oxygenases with their corresponding P450 reductases in the yeast host would improve the overall production amounts (Fig. I. 36) (Jennewein *et al.*, 2005).

More recent work successfully boosted the production of taxadiene in yeast 40-fold through metabolic engineering by using codon optimization, combinatorial biosynthesis, and introducing regulatory elements to inhibit competitive pathways (Engels *et al.*, 2008), so as *S. cerevisiae* recombinant strains not only provide a new approach for taxol production, but also create platforms that allow the synthesis of taxol analogues and other rare taxoids for clinical evaluation (Jennewein and Croteau, 2001).

2.5.2. Lycopene

Carotenoids are a subfamily of isoprenoids that are the most widely distributed yellow, orange, and red natural pigments synthesized in bacteria, algae, and fungi. Commercially available carotenoids such as lycopene, β -carotene, and astaxanthin are used as food colorants, animal feed supplements, and for nutritional and cosmetic purposes. More recently, carotenoids have received attention for their significant antioxidant activities and for playing important roles in inhibiting the onset of chronic diseases (Fig. I. 37) (Rao and Rao, 2007).

Lycopene is the pigment principally responsible for the deep-red color of ripe tomato (*Lycopersicon esculentum*) fruits and tomato products. Many have attributed the health benefits of lycopene to its antioxidant properties (lycopene quenches singlet oxygen almost twice as well as β -carotene does), although other mechanisms of lycopene action are possible: the modulation

of intercellular communication, hormonal and immune system changes, and alterations of metabolic pathways may also be involved (Agarwal and Rao, 2000).

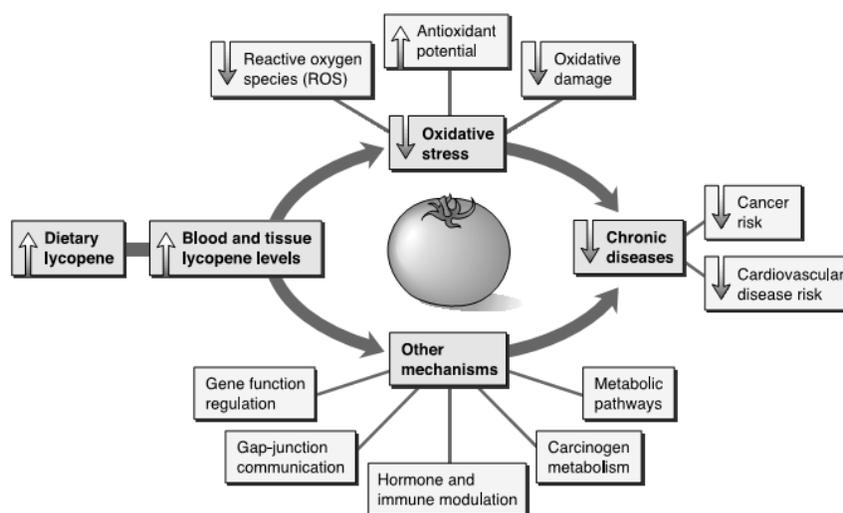


Figure I. 37. Proposed mechanisms for the role of lycopene in preventing chronic diseases. Dietary lycopene may increase the lycopene status in the body and, acting as an antioxidant, may trap reactive oxygen species, increase the overall antioxidant potential or reduce the oxidative damage to lipid (lipoproteins, membrane lipids), proteins (important enzymes) and DNA (genetic material), thereby lowering oxidative stress. This reduced oxidative stress may lead to reduced risk for cancer and cardiovascular disease. Alternatively, the increased lycopene status in the body may regulate gene functions, improve intercell communication, modulate hormone and immune response, or regulate metabolism, thus lowering the risk for chronic disease. These mechanisms may also be interrelated and may operate simultaneously to provide health benefits (Agarwal and Rao, 2000).

Overexpression of three exogenous genes including geranylgeranyl diphosphate synthase (GGPPS), phytoene synthase (PS), and phytoene desaturase (PD) was sufficient to produce the red-color lycopene in *E. coli* (Lotan and Hirschberg, 1995). To identify additional genes to enhance production, a shotgun genomic library was screened in *E. coli* for enhanced coloration. Three regulatory elements and one directly involved in the MEP pathway were found to increase lycopene accumulation (Kang *et al.*, 2005). An alternative strategy was employed to increase lycopene concentrations by performing gene knockouts. Using both systematic (model-based) and combinatorial (transposon-based) approaches, a number of targets were identified leading to the construction of two triple knockout mutants (Alper *et al.*, 2005). Most recently, enhanced lycopene accumulation in *E. coli* was reported using a hybrid approach to identify the best combination of gene expressions and knockouts (Jin and Stephanopoulos, 2007).

2.5.3. Artemisinin

Artemisinin is an antimalarial drug isolated from *Artemisia annua* L (family Asteraceae; commonly known as sweet wormwood), which has been used for more than 2000 years in traditional Chinese medicine, and is highly effective against multidrug resistant strains of *Plasmodium falciparum*. Thus this compound has retained its place as a terpenoid of wide-ranging pharmaceutical and socio-economic value. Supply from natural sources is limiting and too expensive for most malaria sufferers while the cost of its total synthesis is prohibitive. Exciting new studies are exploring the biosynthesis and metabolic engineering of artemisinin with the goal of developing cost-effective methods for stable production at large scale and with consistent quality. The biosynthesis of artemisinin (Fig. I. 38) occurs in glandular trichomes on the surface of *A. annua* leaves. A semisynthesis route (Roth and Acton, 1989) could potentially be cost effective by starting with artemisinic acid that can be produced from a genetically engineered microbe.

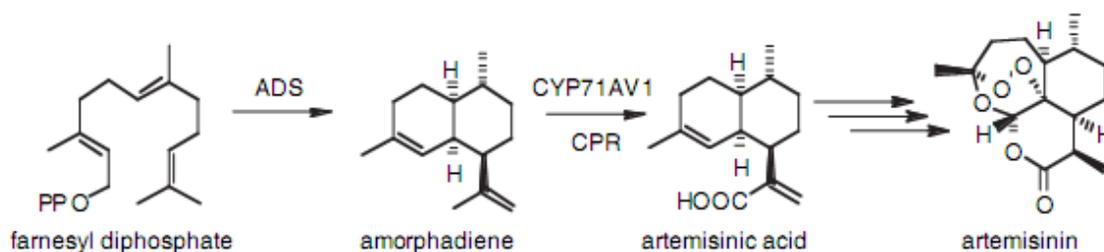


Figure I. 38. Pathway of artemisinin biosynthesis in *Artemisia annua*. ADS, amorpha-4,11-diene synthase; CPR, cytochrome P450 reductase (Bohlmann and Keeling, 2008).

The first committed step of artemisinin biosynthesis (Fig. 38) begins with the cyclization of FDP, catalysed by amorpho-4,11-diene synthase, a sesquiterpene synthase (Mercke *et al.*, 2000; Wallaart *et al.*, 2001; Berteau *et al.*, 2005; Lindahl *et al.*, 2006). Subsequent three-step oxidation of amorpha-4,11-diene to artemisinic acid is catalysed by a multi-functional cytochrome P450 (CYP71AV1) (Ro *et al.*, 2006; Teoh *et al.*, 2006). The remaining reactions from artemisinic acid to artemisinin remain to be characterized, but are thought to include non-enzymatic photo-oxidation reactions (Wallaart *et al.*, 2001).

Recent work by Keasling and co-workers at the University of California (Berkeley, CA) on microbial production of artemisinin provides an impressive example of successful synergy between biochemistry, genomics and biochemical engineering of plant terpenoids (Ro *et al.*,

2006; Shiba *et al.*, 2007). In brief, these authors used amorphaadiene synthase, CYP71AV1 and P450 reductase from *A. annua* in combination with introduction of a MEV pathway into *E. coli*, or optimization of flux through the MEV pathway in *S. cerevisiae*, for substantial production of artemisinic acid in these microbial hosts. Using a semi-synthetic route from artemisinic acid to artemisinin, their approach resulted in complete synthesis of artemisinin (Ro *et al.*, 2006; Shiba *et al.*, 2007). In parallel with exploring microbial systems for the production of artemisinin, new and elegant approaches for plant metabolic engineering of amorphaadiene and other plant terpenoids have been developed. Wu *et al.* employed transgenic co-expression of FDP synthase and amorphaadiene synthase in plastids for successful, high-level synthesis of amorphaadiene in transgenic tobacco (*Nicotiana tabacum*) plants (Wu *et al.*, 2006). This approach was designed to avoid competition for a cytosolic pool of FDP by amorphaadiene synthase and endogenous FDP-utilizing enzymes. The redirection of additional FDP biosynthesis into plastids, together with targeting of amorphaadiene synthase to the same compartment, apparently provided a substantial substrate pool for this engineered pathway without compromising plant growth.

The example of artemisinin highlights recent progress in utilizing plant metabolic engineering as well as microbial biochemical engineering for production of a plant terpenoid at high yield. These studies may result in novel agricultural crops or efficient microbial fermentations for plant terpenoid production. It is also important to note that the modern socio-economic value of artemisinin as a medicinal compound, and research towards cost-effective and large-scale biotechnological production, are largely founded on traditional knowledge regarding a terpenoid-producing medicinal plant (Bohlmann and Keeling, 2008).

2.6. Future directions in metabolic engineering of terpenoids

The increasing cost of energy and raw materials for building complex chemical structures, combined with the environmental concerns associated with conventional manufacturing, mean that biosynthesis using engineered microbial cells will probably become a preferred route for obtaining valuable chemicals. Despite the recent advancements and promising results in terpenoid metabolic engineering, there are still numerous areas in which significant progress can be made. They are found in both plants and microorganisms, and include important therapeutic, antimicrobial, flavor and fragrance compounds. These compounds are built by the condensation of a series of isoprene units to geranylgeranyl and geranylgeranyl diphosphates,

followed by enzymatic cyclization by a terpene cyclase, and subsequent chemical functionalization (Dewick, 2002). Large-scale extraction of terpenoids is not sustainable owing to the low yield of these products, especially therapeutic molecules such as artemisinin and taxol.

In plants, pathway engineering and combinatorial biosynthesis have been favored as strategies for overcoming such limitations (Julsing *et al.*, 2006). Alternatively, the plant terpenoid biosynthesis pathway can be transferred into recombinant microorganisms for production of the molecule of interest or its immediate precursors. Direct synthesis of plant terpenes in recombinant microorganisms has been limited by host-borne regulatory networks and the inability to control the expression of multisubunit heterologous enzymes (Martin *et al.*, 2003). The solution will require combining principles of pathway evolution with the synergistic regulation of cellular networks for precise and timely rerouting of metabolic fluxes to the product of interest.

The recent successful production of artemisinin offers hope that existing problems can be circumvented and microbial cell engineering extended to other potential therapeutic molecules such as taxol. Ultimately, the goal is to construct microbial cell factories for biosynthesis of terpenoids starting from inexpensive materials, and to deliver a mature platform for large-scale commercial fermentations.

Accessing new chemical space through heterologous expression and the *de novo* assembly of new biosynthetic systems is provocative, and improved access to synthetic biology methods has provided a key first step along this road. The key practical measure of success for biosynthetic engineering and related approaches is more prosaic: to make molecules that bind to biological targets of interest and, above all, that answer fundamental scientific questions, rather than to make only those molecules that the pharmaceutical industry may desire. Thus, the use of molecular genetic and sequencing approaches for uncovering new chemistries is most effectively applied through the use of imaginative screening and engineering methods to identify and diversify new chemotypes acting against important biological targets. As the example of platensimycin illustrates, even the culturable microbial diversity, so thoroughly investigated already, has many surprises yet in store.

The output of biosynthetic engineering and related efforts to generate new compounds in quantities sufficient for screening and subsequent validation has remained predominantly limited to the rational construction of modestly sized focused libraries or specific compounds (Wilkinson

and Bachmann, 2006). This is in contrast with early aspirations for the technology to produce large compound libraries. A parallel can be drawn with combinatorial chemistry and medicinal chemistry, where the production of large libraries for screening purposes has given way to the recognition that this technology is best applied as a means of optimizing lead structures through the production of smaller, high-quality libraries (Koehn and Carter, 2005).

Without doubt, through mining, heterologous expression, biosynthetic engineering and the other methods, we are significantly increasing our portfolio of natural products and analogs, making these molecules an even more valuable asset not only in drug discovery but also in fundamental research, including chemical genomics.

2.7. Natural Products Isolation

Isolation of natural products (Pauli *et al.*, 2008) differs from that of the more prevalent biological macromolecules because natural products are smaller and chemically more diverse than the relatively homogeneous proteins, nucleic acids and carbohydrates, and isolation methods must take this into account.

How to begin the isolation of a natural product? First, something about the nature of the compound needs to be known so that the approach to take can be determined. How much needs to be discovered depends on how much is already known and what our aim is. The general features of a molecule that are useful to ascertain at this early stage might include: solubility (hydrophobicity / hydrophilicity), acid/base properties, charge, stability, and size. Selecting general separation conditions, there is no correct purification method for each natural product and no compulsion to follow such a method. Solid phase extraction, gradient HPLC, TLC, and different other forms of chromatography are some of many separation techniques which can be used to isolate natural products (Cannell, 1998; Stashenko and Martinez, 2008).

2.7.1. Solid Phase Extraction

This method involves sorption of solutes from a liquid medium onto a solid adsorbent by the same mechanism by which molecules are retained on chromatography stationary phase.

Solid-phase extraction (SPE) has become one of the most important sample preparation techniques to extract analytes from various fluids. It is widely used in metabolic target and profiling analysis when a sufficient separation of analytes from interfering matrix is needed. SPE

describes the non-equilibrium, exhaustive removal of analytes from a flowing liquid sample by retention on a solid sorbent. The retained analytes are subsequently eluted from the sorbent using a solvent or solvent mixtures with sufficient elution strength (Guo and Mitra, 2000). Based on the analytes and the sorbent selected, compounds are retained by van der Waals interactions, dipole–dipole interactions, hydrogen bonding, or electrostatic forces. A variety of SPE sorbent materials are commercially available for the broad extraction of metabolites, such as silica, alkylated silica (e.g., C-18), carbon-based sorbents, ion exchange materials, polymer materials, and restricted access materials (RAM) used to remove macromolecules.

Volatile metabolites can include alcohols, aldehydes, furans, ketones, pyrroles, terpenes, and others (Mills and Walker, 2001). The analysis of volatile compounds is challenging by itself. Conventional sample preparation techniques, such as liquid extraction or SPE, are often not feasible, because the analytes might be incompletely extracted and losses will occur if the extract is concentrated. Moreover, the solvent can interfere with the analytes during gas chromatographic separation due to incomplete separation of the solvent peak from the analyte peaks. Specifically, in GC-MS the analytes will be cut off with the solvent delay. Therefore, solvent-free sample preparation techniques are often the method of choice for the analysis of volatile metabolites. Wahl *et al.* used headspace sampling with a gastight syringe in combination with cold trapping in a temperature-controlled cold injection system and GC-MS for the analysis of volatile metabolites in urine samples (Wahl *et al.*, 1999). Thirty-four compounds were identified using MS libraries and reference compounds. Another rapid and solvent-less sample preparation technique for the determination of volatile and semi-volatile compounds is headspace-solid phase microextraction (HS-SPME) (Bicchi *et al.*, 2008). SPME uses a silica fiber that is coated with a stationary phase for sampling. The fiber is housed in a syringe-type assembly. A number of different stationary phases comprising non-polar to polar materials are commercially available as fiber coatings. In the case of HS-SPME the sampling takes place in the headspace compartment above the sample, enabling an easy fractionation of volatile analytes and complex matrix. However, sample extraction is not exhaustive, but the amount of analytes enriched on the fiber depends on two equilibrium distribution steps, sample/gaseous phase and gaseous phase/fiber coating. Both distribution coefficients are temperature dependent, and extraction time and temperature need to be optimized to maximize sample enrichment. Moreover, the distribution of the analytes between sample and gaseous phase is influenced by

analyte-matrix interactions and has to be considered for quantitative analysis. Desorption of the analytes from the fiber can be achieved by thermal energy in the hot injector of aGC.

SPME-GC-MS has been used to study volatile organic compounds in plants (Verdonk *et al.*, 2003; Bino *et al.*, 2005). Stashenko *et al.* used different extraction methods, such as hydrodistillation, simultaneous distillation–solvent extraction, microwave-assisted hydrodistillation, supercritical fluid extraction (SFE) with CO₂, static headspace, simultaneous purge and trap in CH₂Cl₂, and HS-SPME for sampling volatile plant metabolites (Stashenko *et al.*, 2004b; a).

2.7.2. Gas Chromatography – Mass Spectrometry

Chromatography involves the distribution of a compound between two phases—a moving, mobile phase that is passed over an immobile stationary phase. Separation is based on the characteristic way in which compounds distribute themselves between these two phases.

The combination of gas chromatography (GC) with mass spectrometry (MS) provides high-chromatographic metabolite resolution, analyte-specific detection, and quantification of metabolites, as well as the capability to identify unknowns. However, a major prerequisite for GC-MS analysis is a sufficient vapor pressure and thermal stability of the analytes. Standard capillary GC columns provide a high-peak capacity.

While MS provides individual mass spectra that can differentiate between co-eluting metabolites that are chemically diverse, the main advantages of GC over other separation techniques are a high separation efficiency (as highlighted by its ability to distinguish isomeric compounds), ease of use, robustness, and low cost (Lommen *et al.*, 2007; Mas *et al.*, 2007). GC–MS is a technique that usually yields extensive and highly reproducible fragmentation because of the standardized use of electron ionization (EI), which has enabled metabolites to be identified by matching their relative retention times, retention or Kovats indices, and mass spectral fragmentation patterns, to known and predicted information available from extensive databases (e.g. the NIST and Wiley database) (Kind and Fiehn, 2007).

MS spectra have been studied for several decades and, therefore, many MS spectra libraries exist containing thousands of spectra. The creation of MS libraries was facilitated by high-instrument reproducibility and the standardized electron energy of 70 eV commonly used for ionization. The largest spectra libraries are available from Wiley (400,000 spectra) and

the National Institute of Standards and Technology (NIST) (200,000 spectra) as well as a combination of both with some additional spectra from Palisade (600,000 spectra).

Large commercial libraries contain many spectra of synthetic compounds, which are unlikely to appear in biological samples. Although their spectra may provide leads for the structure identification of unknown metabolites, searching through thousands of spectra can be time consuming if hundreds of unknown metabolites need to be identified. In addition, less common metabolites are often not represented in large commercial libraries. However, there are metabolite libraries available, which contain mass spectra of metabolites found in biological samples. The existing MS libraries also contain retention indices (MSRI libraries), which tremendously improve the identification of spectrally similar compounds.

3. OXIDATIVE STRESS IN PLANTS

3.1. Oxidative Stress

Oxidative stress is defined as a disturbance in the prooxidant-antioxidant balance in favor of the former, leading to potential damage (Sies, 1997). This damage can affect a specific cell or the entire organism. Oxidative stress is imposed on cell as a result of one of three factors: (i) an increase in oxidant generation, (ii) a decrease in antioxidant protection, or (iii) a failure to repair oxidative damage.

Prooxidants are reactive oxygen species (ROS), such as free radicals and peroxides, a class of molecules derived from the metabolism of oxygen and exist inherently in all aerobic organisms. The main source of ROS *in vivo* is aerobic respiration, although ROS are produced by peroxisomal oxidation of fatty acids, stimulation of phagocytosis by pathogens or lipopolysaccharides, arginine metabolism and tissue specific enzymes.

A complicated multifactorial antioxidant network composed of low-molecular mass antioxidants and enzymes, such as catalase, ascorbate peroxidase, and glutathione radicals. Peroxidases decompose ROS and lipids peroxides and quench radicals. The main damage to the cell results from the ROS-induced alteration of macromolecules, such as polyunsaturated fatty acids in membrane lipids, essential proteins, and DNA.

The antioxidant defense system of plants has been attracting considerable interest in the characterization of mutants and transgenic plants with altered expression of antioxidant enzymes, a potentially powerful approach to understand the function of the antioxidant system and its role in plants against stress. Ultimately, the function determination of all genes that participate in stress adaptation or tolerance reaction is expected to provide an integrated understanding of the biochemical and physiological basis of stress responses in plants (Cakar *et al.*, 2005; Rouhier and Jacquot, 2005).

3.1.1. Peroxidoredoxins (Prx)

Peroxioredoxins are part of the antioxidant defense. They decompose ROS and lipids peroxides and tune ROS and peroxides levels in signaling events.

Thiol peroxidases are ubiquitous, recently characterized by catalytic cysteine residues. The identification of TPx as a specific inducer of thioredoxin and thioredoxin reductase gene expression in response to oxidative stress in yeast indicates that this protein may be part of an important conserved sensing mechanism for redox conditions in plants as it is in yeast (Verdoucq *et al.*, 1999; Ross *et al.*, 2000). TPx is one of the main cellular enzymes for detoxification of the peroxides through the thioredoxin system.

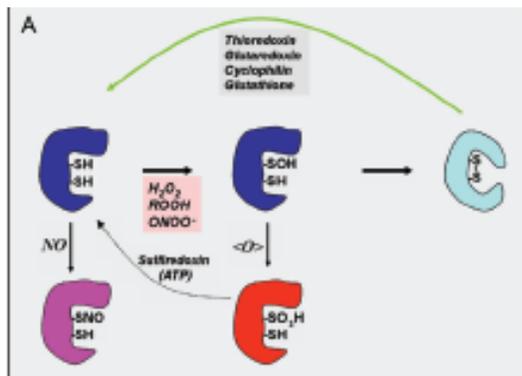


Figure I. 39. Unified reaction cycle of peroxidoredoxins. The catalytic Cys is oxidized to sulphenic acid. A disulphide bridge is formed under a conformational switch of the protein. The disulphite-oxidized Prx is regenerated by electron donors such as Trx, Grx, Cyp, and glutathione. Nitrosylation and overoxidation to sulphinic or sulphonic acid derivatives of the catalytic Cys withdraw Prx from the cycle and are considered as regulatory and potentially signaling mechanisms. (Dietz *et al.*, 2006)

Thioredoxin enzymes are small oxidoreductases with two conserved cysteine residues in their active sites (Holmgren, 2000), that are induced by different types of stress including oxidative one, and operate in defense against ROS. The active site sequence for thioredoxin Cys-Gly-Pro-Cys is found to be conserved among species. The catalytic Cys residues undergo oxidation during the peroxide reduction reaction and need to be reduced by electron donors such

as glutaredoxins, thioredoxins, or cyclophilins before the next catalytic cycle (Konig *et al.*, 2002; Gelhaye *et al.*, 2005; Rouhier and Jacquot, 2005) (Fig. I. 39). By work with transgenic plants, their activity was shown to (i) affect metabolic integrity, (ii) protect DNA from damage *in vitro* and *in vivo*, and (iii) modulate intracellular signaling related to reactive oxygen species and reactive nitrogen species (Dietz *et al.*, 2006).

3.1.2. Acrolein as a toxicant

Acrolein is a highly electrophilic α , β -unsaturated aldehyde, a widespread environmental pollutant. It is also a metabolic product of cyclophosphamide and allylamide and is believed to be the main cause of their toxicity (Fraiser *et al.*, 1991). Acrolein will rapidly bind to or deplete cellular nucleophiles such as glutathione. It can also react with cysteine, histidine, and lysine residues of proteins and with nucleophilic sites in DNA. This may affect gene expression or acrolein may directly interact with various genes and transcription factors. In addition, it is capable of inducing the expression of the glutathione *S*-transferases (Uchida, 2000). Acrolein has been identified as both a product and initiator of lipid peroxidation (Adams and Klaidman, 1993; Uchida, 2000).

High-dose acute acrolein toxicity has been suggested as involving oxidative stress subsequent to the loss of glutathione (GST), or to be due to alkylation reactions at various nucleophilic sites in a cell – particularly in the nucleus. This dichotomy is somewhat confusing, since the loss of GSH is itself due to an alkylation reaction with acrolein. While oxidative stress can result from both an increase in the production of oxidants and a decrease in levels of antioxidants such as GST, this latter occurrence is more likely to increase a cell's susceptibility to oxidants and may not lead to any adverse effects in the absence of an oxidant. Nevertheless, thiol redox balance is critical for numerous cell functions and changes induced by acrolein seem certain to affect various signaling pathways (Kehrer and Biswal, 2000).

3.2. *Arabidopsis* sp. as a model system to study oxidative stress

Arabidopsis thaliana, a member of the Cruciferae family, is used as a model system for molecular and physiological studies because of its short life cycle, small size, small genome, ease of mutagenesis analysis and functional genomic analysis. Genes can be easily introduced in these plants. The most common methods for introduction of DNA into plant cells use

Agrobacterium tumefaciens bacteria or rapidly propelled tungsten micro projectiles that have been coated with DNA (Gelvin, 2006b).

A. tumefaciens is a plant pathogen which causes the natural transformation of plant cells leading to the formation of crown galls in most dicotyledonous plants. The *Agrobacterium*-based transformation method has several advantages over other forms of transformation, including: (i) the ability to transfer large segments of DNA with minimal rearrangement; (ii) the precise insertion of transgenes resulting in fewer copies of inserted genes; and (iii) the simplicity of the technology showed up in a lower cost (Amoah *et al.*, 2001; Lorence and Verpoorte, 2004).

3.2.1. Molecular basis of *Agrobacterium*-mediated transformation

The molecular basis of natural genetic transformation of plant cells by *Agrobacterium* is transfer from the bacterium and integration into the plant nuclear genome of a region of a large tumor-inducing (Ti) or rhizogenic (Ri) plasmid resident in *Agrobacterium* (Fig. I. 40A). Ti plasmids are on the order of 200 to 800 kbp in size (Goodner *et al.*, 2001; Wood *et al.*, 2001; Tanaka *et al.*, 2006). The transferred DNA (T-DNA) (Fig. I. 40B) is referred to as the T-region when located on the Ti or Ri plasmid. T-regions on native Ti and Ri plasmids are approximately 10 to 30 kbp in size (Suzuki *et al.*, 2000).

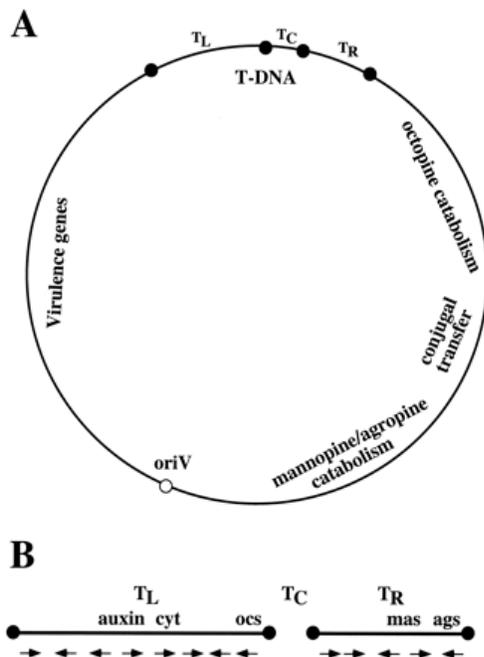


Figure I. 40. Schematic representation of a typical octopine-type Ti plasmid (A) and the T-DNA region of a typical octopine-type Ti plasmid (B). (A) The T-DNA is divided into three regions. T_L (T-DNA left), T_C (T-DNA center), and T_R (T-DNA right). The black circles indicate T-DNA border repeat sequences. *oriV*, the vegetative origin of replication of the Ti plasmid, is indicated by a white circle. (B) The various T-DNA-encoded transcripts, and their direction of transcription, are indicated by arrows. Genes encoding functions involved in auxin synthesis (auxin), cytokinin synthesis (cyt), and the synthesis of the opines octopine (ocs), mannopine (mas), and agropine (ags) are indicated. (Gelvin, 2003)

Thus, T-regions generally represent less than 10% of the Ti plasmid. Some Ti plasmids contain one T-region, whereas others contain multiple T-regions. The processing of the T-DNA from the Ti plasmid and its subsequent export from the bacterium to the plant cell result in large part from the activity of virulence (*vir*) genes carried by the Ti plasmid. The export of T-DNA from bacteria to the plant cells is performed by the activity of the virulence (*vir*) genes carried by the Ti-plasmid (Gelvin, 2006a). T-regions are defined by T-DNA border sequences. These borders are 25 bp in length and highly homologous in sequence. They flank the T-region in a directly repeated orientation and serve as targets of the VirD1/VirD2 border specific endonuclease that processes the T-DNA from the Ti plasmid. Both VirD1/VirD2 proteins cleave the double strand border sequence *in vivo* and *in vitro*.

Because of the complexity of introducing foreign genes directly into the T-region of a Ti plasmid, several laboratories developed an alternative strategy to use *Agrobacterium* to deliver foreign genes to plants. This strategy was based on seminal findings of Hoekema et al. These authors determined that the T-region and the *vir* genes could be separated into two different replicons. When these replicons were within the same *Agrobacterium* cell, products of the *vir* genes could act in *trans* on the T-region to effect T-DNA processing and transfer to plant cells. This was called binary-vector system; the replicon harboring the T-region constituted the binary vector, whereas the replicon containing the *vir* genes became known as the *vir* helper. The *vir* helper plasmid generally contained a complete or partial deletion of the T-region, rendering strains containing this plasmid unable to incite tumors (“disarmed” Ti plasmid) (Hoekema *et al.*, 1984).

4. AIM OF STUDY

This study describes progress in using the yeast *S. cerevisiae* as a model organism for the fast and efficient analysis of genes and enzyme activities involved in the ergosterol biosynthetic pathways. We further assess the impact of baker's yeast on the production of natural product compounds. We focus on the *in vivo* production of several terpenoids by expressing three novel plant terpene synthases isolated by the EST-based approach: one monoterpene synthases and two sesquiterpene synthases. Research progress using yeast as a host organism has been considerable in recent years, so as we chose this organism as a production platform for terpenes as well. In

particular, we investigate for particular mechanisms to increase the endogenous pool of FPP and the means for diverting this intermediate to terpene biosynthesis. To enhance the endogenous FPP pool we first introduced a knockout mutation of sterol methyltransferase (*erg6*) to block the ergosterol biosynthesis to zymosterol, and further on, we introduced chromosomally an extra copy of a stable mutant (K6R) form of Hmg2 reductase and upregulated the FPP synthase (ERG20) by inserting a Galactose inducible promoter upstream of the one allele. Limiting the use of FPP pool for sterol biosynthesis by introducing a haploinsufficient mutation of squalene synthase (*erg9*) should provide an additional enhancement in FPP pool. Finally, we use current knowledge to highlight perspectives for future biotechnological applications in the field of terpenoid compounds.

CHAPTER II

MATERIALS AND METHODS

1. STRAINS

1.1. *Escherichia coli* strains

Mach1TM-T1R - *E. coli* strain is modified from the wild-type W strain (ATCC #9637, S. A. Waksman) and has a faster doubling time compared to other standard cloning strains. It confers resistance to phage T1. Genotype is F- ϕ 80 (lacZ) Δ M15 Δ lacX74 hsdR (rK-mK+) Δ recA1398 endA1 tonA, (Invitrogen).

1.2. *Saccharomyces cerevisiae* strains

BY4741 – Mat **a**, his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0. This strain is derived from strain S288C in which commonly used selectable marker genes are deleted by design, based on the yeast genome sequence which has been constructed and analyzed. These strains minimize or eliminate the homology to the corresponding marker genes in commonly used vectors without significantly affecting adjacent gene expression.

EGY48 – MAT α , trp1, his3, ura3, lexAops-LEU2. This strain is derivative of U457 in which the endogenous LEU2 gene has been replaced by homologous recombination with LEU2 reporters carrying 6 LexA operators. Transcription of the LEU2 gene is under control of lexA operators. EGY48 is a basic strain used to select for interacting clones from a cDNA library.

Other yeast strains used in the present project are presented in Table II. 1.

Table II. 1. Yeast strains used in the project:

BY4741	Mat a, <i>his3Δ1</i> , <i>leu2Δ0</i> , <i>met15Δ0</i> , <i>ura3Δ0</i> .
EG60	Mat α, <i>ura3</i> , <i>trp1</i> , <i>his3</i> .
EG61	Mat a, <i>ura3</i> , <i>trp1</i> , <i>his3</i> .
KSY10	Mat a, GALp-(K6R)HMG2::URA3, <i>trp1</i> , <i>his3</i> .
AM62	Mat α, Δ <i>erg6</i> :: HIS5, <i>ura3</i> , <i>trp1</i> , <i>his3</i> . Derivative of EG60.
AM63	Mat α, GALp-(K6R)HMG2::URA3, <i>trp1</i> , <i>his3</i> . Derivative of EG60.
AM64	Mat α, URA3- <i>hisG</i> -GALp-(K6R)HMG2::HO, <i>trp1</i> , <i>his3</i> , Δ <i>erg6</i> :: HIS5. Derivative of AM63.
AM65	Mat α, GALp-(K6R)HMG2, <i>ura3</i> , <i>trp1</i> , <i>his3</i> . Derivative of AM64.
AM66	Mat α/a, GALp-(K6R)HMG2x2::HO, <i>ura3</i> , <i>trp1</i> , <i>his3</i> . Derivative of AM63 and KSY10.
AM67	Mat α/a, GALp-(K6R)HMG2x2, GALp-HA-ERG20-URA3, <i>trp1</i> , <i>his3</i> . Derivative of AM66.
AM68	Mat α/a, GALp-(K6R)HMG2x2, GALp-HA-ERG20, <i>ura3</i> , <i>trp1</i> , <i>his3</i> . Derivative of AM67.
AM69	Mat α/a, GALp-(K6R)HMG2x2, GALp-ERG20::URA3, <i>trp1</i> , <i>his3</i> , Δ <i>erg9</i> :: HIS5. Derivative of AM67.
AM70	Mat α/a, GALp-(K6R)HMG2x2, GALp-HA-ERG20, <i>ura3</i> , <i>trp1</i> , <i>his3</i> , Δ <i>erg9</i> :: HIS5. Derivative of AM68.
AM71	Mat α/a, GALp-(K6R)HMG2x2, GALp-ERG20, ADH1-ERG19-URA3, <i>trp1</i> , <i>his3</i> .
AM74	Mat α/a, GALp-(K6R)HMG2x2::HO, GALp-HA-ERG20, ADH1- <i>myc</i> ::tHMG1, URA3, <i>trp1</i> , <i>his3</i> . Derivative of AM68.
AM75	Mat α/a, GALp-(K6R)HMG2x2::HO, GALp-HA-ERG20, ADH1- <i>myc</i> ::tHMG1, URA3, <i>trp1</i> , <i>his3</i> , Δ <i>erg9</i> ::HIS5. Derivative of AM70.

1.3. *Agrobacterium tumefaciens* strain

AGL-1 – This genotype is AGL-0 (C58 pTiBo542) recA::bla, T-region deleted Mop(+) Cb(R). AGL0 is an EHA101 with the T-region deleted, which also deletes the aph gene (Lazo et al., 1991). This strain contain the helper plasmid pAL154

1.4. *Arabidopsis thaliana* strain

Columbia Ecotype – Col 0, direct descendant of Col-1 donated via AIS.

2. NUTRITIONAL MEDIA

2.1. Bacterial media

2.1.1. Luria-bertani (LB) medium

0.5 % (w/v) sodium chloride,
1% (w/v) bacto-tryptone,
0.5% (w/v) bacto-yeast extract.

For solid LB medium, the above mixture was supplemented with 2% (w/v) agar after adjusting the pH to 5.8-6.2. The medium was autoclaved under standard conditions (20min, 151 lb/sq.in. on liquid cycle). Antibiotics were added to the medium after being cooled down to around 55° C at a final concentration according to the kind of antibiotic being used. Solidified dried plates were stored at 4° C for up to two weeks.

2.1.2. YEP medium

0.5 % (w/v) sodium chloride,
1% (w/v) peptone,
1% (w/v) bacto-yeast extract.

The pH was adjusting to 7 with 5M NaOH.

2.1.3. SOC medium

2.0 % (w/v) bacto-tryptone

0.5 % (w/v) bacto-yeast extract

10 mM MgCl₂

10 mM MgSO₄

20 mM glucose

This is a rich medium used to recover the transformed cells after the heat shock.

2.1.4. Bacterial glycerol stocks

For the preparation of bacterial glycerol stocks, 1 volume of saturated bacterial culture in selective LB medium was added to 1 volume of a 50% sterile glycerol solution. Samples were stored at -80° C.

2.1.5. Antibiotics

Kanamycin sulfate (100 mg/ml) (SIGMA)

Stock solution of 100 mg/ml was prepared with distilled water, sterilized by filtration through a 0.22 µm filter membrane and then stored at – 20° C as 1 ml aliquots until required.

Chloramphenicol (34 mg/ml) (SIGMA, C-1919)

Stock solution of 34 mg/ml was prepared in 100% ethanol. This does not require filter-sterilization. 500 µl aliquots were stored at – 20° C until required.

Ampicillin, sodium salt (100 mg/ml) (SIGMA, A9518)

Stock solution of 100 mg/ml was prepared with distilled water; sterilized by filtration through a 0.22 µm filter membrane, and then stored at – 20° C as 1 ml aliquots until required.

Tetracycline (100 mg/ml) (SIGMA)

Stock solution of 100 mg/ml was prepared with distilled water; sterilized by filtration through a 0.22 µm filter membrane, and then stored at – 20° C as 1 ml aliquots until required.

2.2. Yeast media

2.2.1. YPD medium

1% (w/v) yeast extract,

2% (w/v) peptone,

2% (w/v) glucose.

For the solid media, the pH is adjusted to be 5.8-6.2. Bacteriological agar was used at a concentration of 2% (w/v). Standard autoclaving conditions are applied (20 min, 151 lb/sq.in. on liquid cycle).

2.2.2. Complete minimal (CM) dropout medium

Complete dropout powder is a defined medium which contains all the amino acids necessary for growth and serves as an important selective medium. Leaving out one or more nutrients (amino acids) from the CM dropout powder (which can be complemented by plasmid acquirement) selects for yeast able to grow, only cells containing the plasmid of interest will be able to grow on such selective dropout media.

0.13% (w/v) dropout powder,

0.67% (w/v) yeast nitrogen base without amino acids (YNB),

2% (w/v) glucose.

For the solid media, the pH is adjusted to be 5.8-6.2. Bacteriological agar was used at a concentration of 2% (w/v). Standard autoclaving conditions are applied (20 min, 151 lb/sq.in. on liquid cycle).

2.2.3. Alternative carbon source

To induce transcription of a gene sequence fused on the GAL promoter, Galactose and Raffinose are used as an alternative carbon source as the following:

0.13 % (w/v) dropout powder,

0.67 % (w/v) YNB (yeast nitrogen base without amino acids).

2% (w/v) galactose,

1% (w/v) raffinose,

For the solid media, the pH is adjusted to be 5.8-6.2. Bacteriological agar was used at a concentration of 2% (w/v). Standard autoclaving conditions are applied (20 min, 151 lb/sq.in. on liquid cycle).

2.2.4. FOA medium

Nontoxic, 5-FOA is converted to the toxic form (i.e., 5-fluorouracil) in yeast strains expressing the functional URA3 gene coding for orotidine-5'-monophosphate decarboxylase that is

involved in the synthesis of uracil. Yeast strains that are phenotypically Ura⁺ become Ura⁻ and 5-FOAR (resistant), after selection.

The FOA medium was prepared as following:

Mixture A

0,25 mg of 5- fluorooroacetic acid (10 ml of 25 mg/ml 5-FOA solution

0,63 ml of 2,4 mg/ml uracil solution

51,87 ml of sterile water were mixed well and filter sterilized.

Mixture B (Glu/CM media for 125 ml of total volume):

0,17 g of dropout completely minimal media powder

0,84 g of yeast nitrogen base

2,5 g of glucose

2,5 g of agar

Standard autoclaving conditions are applied (20 min, 151 lb/sq.in. on liquid cycle).

After sterilizing, mixture B was cooled down to ~50°C and combined with mixture A. The complete medium was swirled and poured on the plates. Plates containing 5-FOA are stable for 2 to 3 months when stored at 4°C. The recipe is fully scalable and can be adjusted to any desired amount.

2.2.5. Yeast glycerol stocks

For the preparation of yeast glycerol stocks, 1 volume of saturated yeast culture was added to 1 volume of a 50% sterile glycerol solution. Glycerol stocks are stored at -80°C.

2.3. Plant media

2.3.1. Murashige and Skoog (MS) medium

Growth plant medium contains 0.5X or 1X Murashige and Skoog (MS) mineral salts, agar at 0.6 – 0.8 %, and sucrose at 0 – 3 %. The recipe for 1 liter of standard MS medium is:

4,33 g MS Basal Salt Mixture (Sigma-Aldrich, M5524)

1% (w/v) sucrose

For the solid media, the MS salts and sucrose were mixed with water and stirred to dissolve. The pH was adjusted to 5.7 with KOH, and then the Phyto or Bacto agar was added at a

concentration of 0,7 % (w/v). Standard autoclaving conditions are applied (20 min, 151 lb/sq.in. on liquid cycle).

After cooling down 0,1 % MS Vitamin Solution 1000X (Sigma-Aldrich, M3900) containing 2.0 glycine, 100.0 myo-inositol, 0.50 nicotinic acid, 0.50 pyridoxine hydrochloride, 0.10 thiamine hydrochloride.

2.3.2. Acrolein

1M Acrolein MW 56.06 was prepared using the 90% stock solution.

3. DNA VECTORS

3.1. pCRII-TOPO

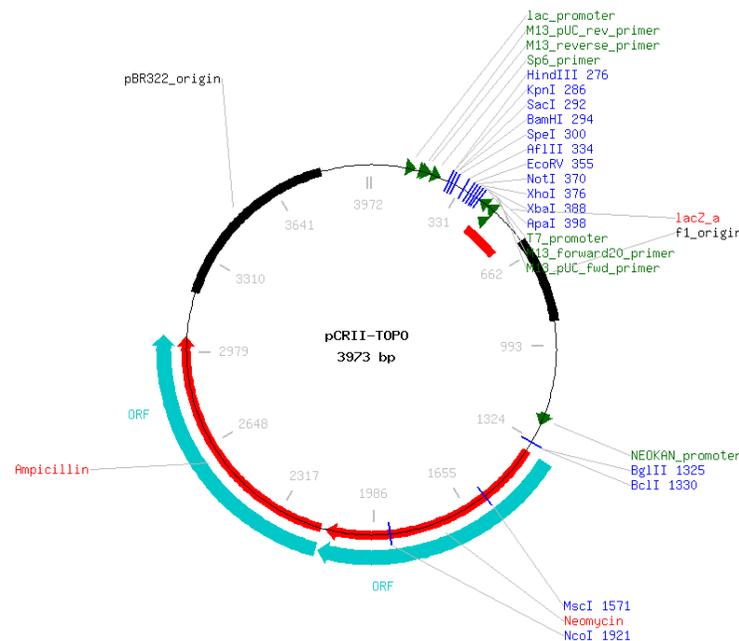


Figure II. 1. The pCRII-TOPO plasmid map

Size: 3973 nucleotides. LacZ α fragment: bases 1-589. M13 reverse priming site: bases 205-221. Multiple cloning sites: bases 269-383. T7 promoter/priming site: bases 406-425. M13 Forward (20) priming site: bases 433-448. F1 origin: bases 590-1027. Kanamycin resistance

ORF: bases 1361-2155. Ampicillin resistance ORF: bases 2173-3033. pUC origin: bases 3178-3851 (Fig. II. 1).

3.2. pBluescript II SK

All pBluescript II phagemids are 2961 bp in length, designed for DNA cloning, dideoxy DNA sequencing, *in vitro* mutagenesis and *in vitro* transcription in a single system. pBluescript II phagemids contain: f1 (IG) - the intergenic region of phage f1; *rep* (pMB1) - the pMB1 replicon responsible for the replication of phagemid; *bla* (ApR) - gene, coding for beta-lactamase that confers resistance to ampicillin; *lacZ* - 5'-terminal part of *lacZ* gene encoding the N-terminal fragment of beta-galactosidase. This fragment, corresponding to wt beta-galactosidase, allows blue/white screening of recombinant phagemids (Fig. II. 2).

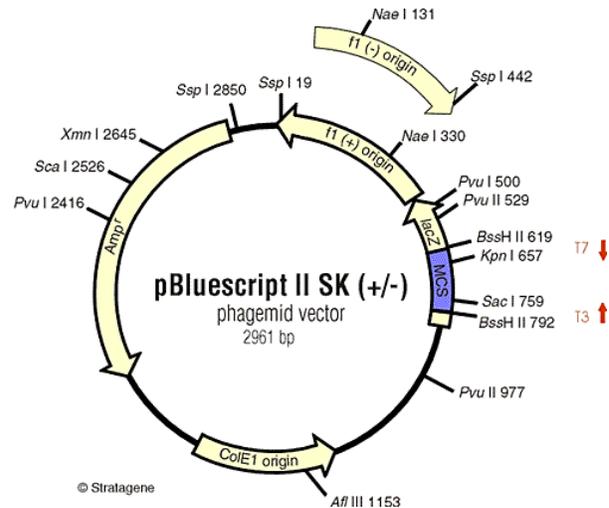


Figure II. 2. The pBluescript II SK (+/-) plasmid map

3.3. pDNR-LIB

Size: 4200 nucleotides. *LoxP* recombination sites: 9–42; 1218–1251. MCS A: 45–99. Stuffer fragment: 100–291. MCS B: 292–351. Chloramphenicol (Cm^r) open reading frame (ORF): 1189–530. *SacB* negative selection marker: 3168–1266. T7 RNA polymerase promoter: 4091–4109. pUC origin of replication: 3337–3980. M13-reverse sequencing primer site: 382–362. M13-forward sequencing primer site: 4065–4081.

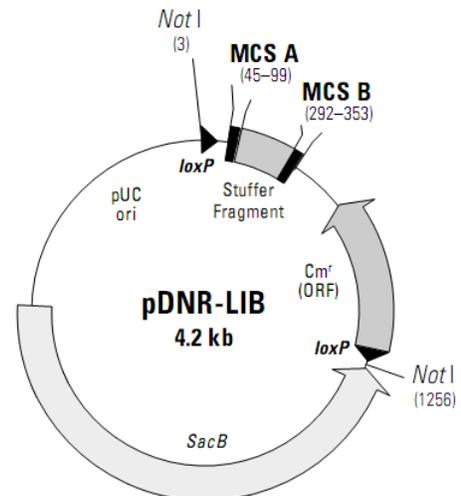


Figure II. 3. The pDNR-LIB plasmid map

The pDNR-LIB vector contains two loxP sites, which flank the 5' end of MCS A and the 5' end of the open reading frame encoding the chloramphenicol resistance gene (Cmr). PDNR-LIB also contains the sucrose gene from *B. subtilis* (SacB), which provides negative selection against both incorrect recombinants and the parental donor vector following recombination with an acceptor vector (Fig. II. 3).

3.4. pJG4-6

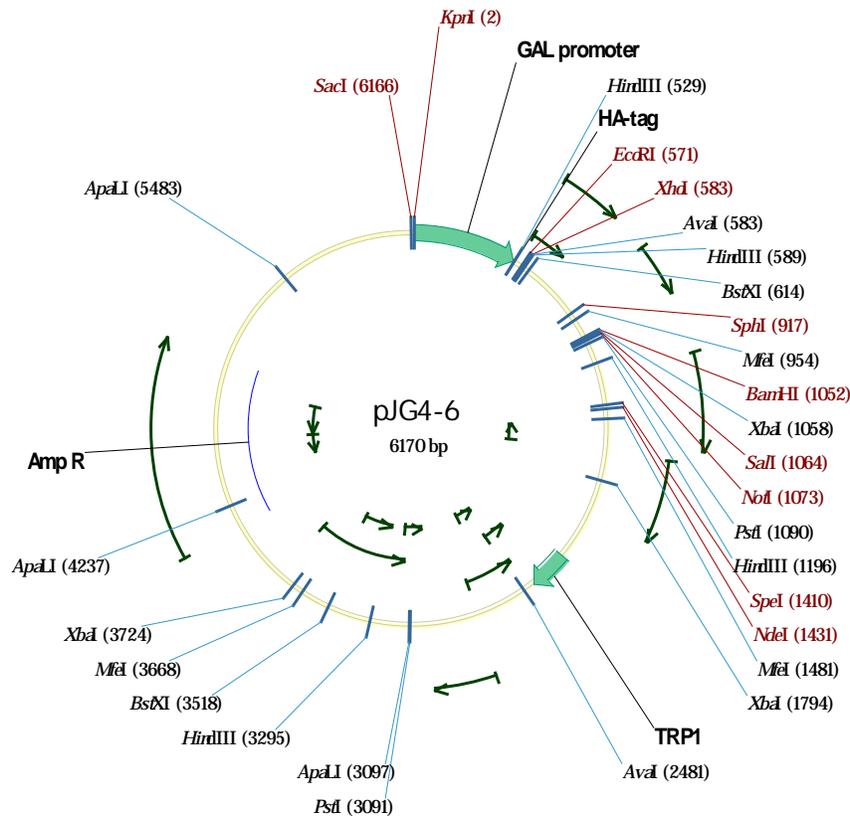


Figure II. 4. The pJG4-6 plasmid map

Size plasmid: 6170 bp. The vector contains the *TRP1* auxotrophy marker and 2μ m origin for propagation in yeast, and the ampicillin resistance (Amp^R) gene and the pUC origin to allow propagation in *E. Coli*. The GAL1 promoter allows expression of the protein fused to a hemagglutinin tag only when the yeast cells are grown in galactose medium (with no glucose present) (Fig. II.4).

3.6. pYX143 and variants

A low copy number vector (ARS/CEN, LEU2) expressing genes under the control of the constitutive TPI promoter (Fig. II. 7).

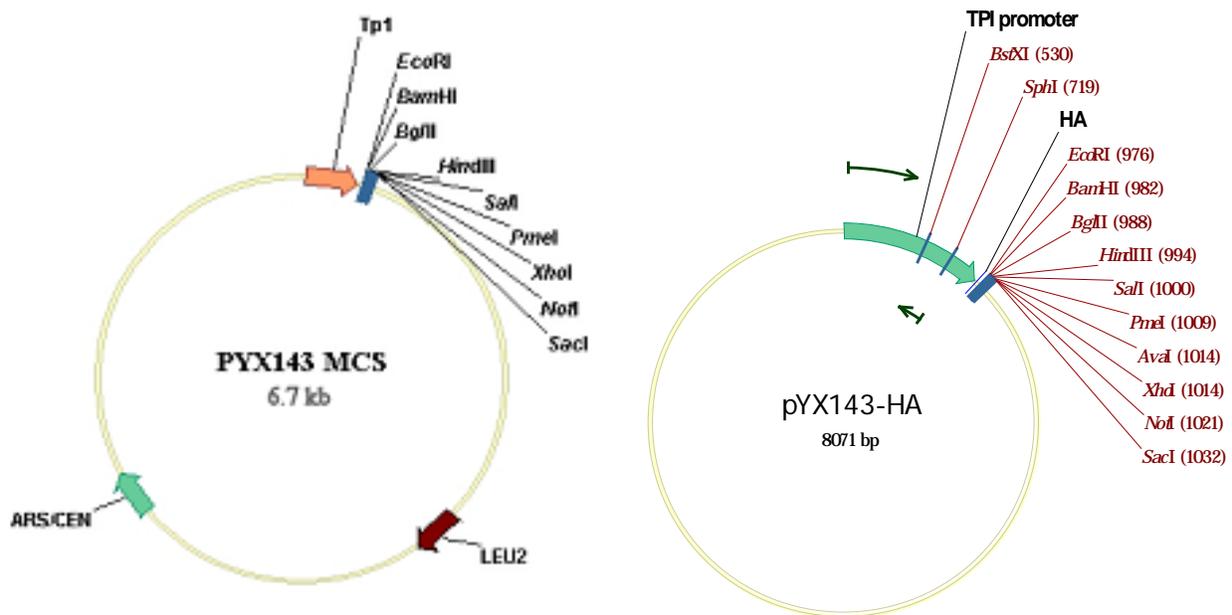


Figure II. 7. The pYX143 (left) and the pYX143-HA (right) plasmid maps.

3.7. pUG27

The plasmid contains the *his5⁺* gene from *Schizosaccharomyces pombe* that complements the *Saccharomyces cerevisiae his3* mutation. Due to the flanking *loxP* sites the deletion marker can be removed by expressing *cre* recombinase. Transformed *E.coli* cells are resistant to ampicillin. The plasmid is useful as template for PCR to create gene deletions. Integrative transformation of the respective PCR fragment and selection of the mutants (Fig. II. 8).

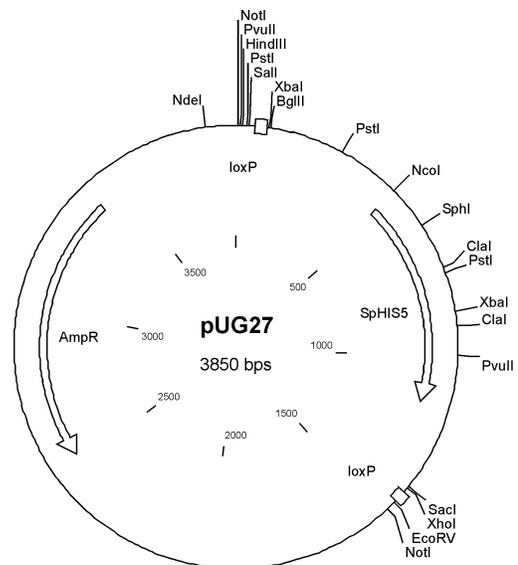


Figure II. 8. The pUG27 plasmid map.

3.8. pUG72

The plasmid contains the *URA3* gene from *Klyveromyces lactis* that complements the *ura3* mutation of *Saccharomyces cerevisiae*. Due to the flanking *loxP* sites, the deletion marker can be removed by expressing *cre* recombinase. The plasmid is useful as a template for PCR to create gene deletions. Integration of the transformation of the respective PCR fragment can give a selection of mutants (Fig. II. 9).

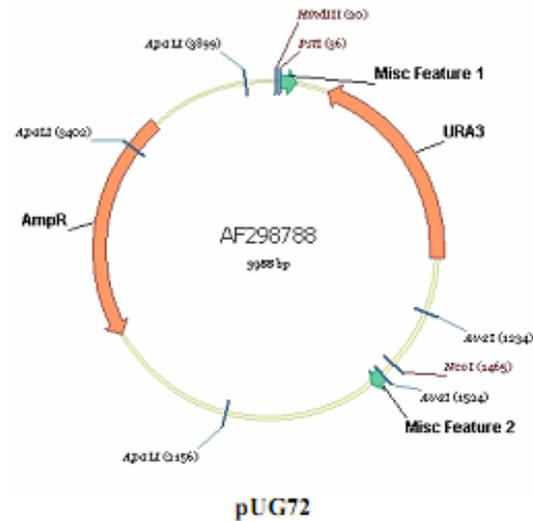


Figure II. 9. The pUG72 plasmid map.

3.9. pB227/GAL-cre

pGAL-*cre* was constructed in Dr F. Heffron's laboratory (Hoekstra *et al.*, 1991). It consists of a LEU2/CEN/amp vector and carries the phage P1 *cre* gene under the control of the GAL1/10 promoter.

3.10. M4366 (HO-hisG-URA3-poly-HO)

Size plasmid: 8487 bp. The parent vector M4366 (HO-hisG-URA3-hisG-poly-HO) was designed for genomic integration into the *HO* locus. Plasmid *HO-hisG-URA3-hisG-poly-HO* contains the ampicillin resistance gene for selection in *E. coli* and *URA3* yeast selectable marker. Selection for uracil prototrophy can be used to identify integration of the *HO-hisG-URA3-hisG-poly-HO* plasmid at the *HO* locus. Moreover, one can select for cells that have returned to the *ura3* state by selection on 5-FOA, identifying yeast in which recombination has occurred between the *hisG* repeats flanking the *URA3* gene (Voth, 2001). Thus, these cells which have the integrated sequences that are stably maintained in the absence of selection can be transformed with a *URA3* plasmid (Fig. II. 10).

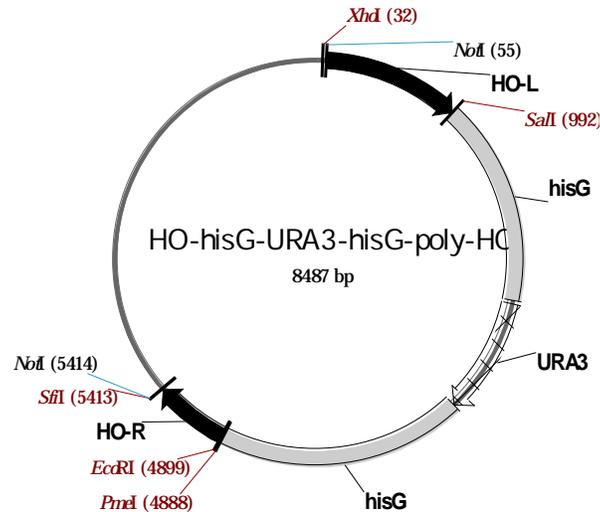


Figure II. 10. The M4366 plasmid map.

3.11. *Agrobacterium* dual binary vector system

3.11.1. pGREEN300:35S-nos

The pGREEN is a small binary vector able to replicate in *Agrobacterium* only when another vector, called helper plasmid is present. Several major components of the pGREEN vector have been optimized for the improvement of *Arabidopsis* plant transformation via *Agrobacterium*. These are the T-DNA and cloning site, plasmid selection, and binary replication origin.

A pGREEN300:35nos plasmid containing a modified T-DNA cassette, flanked by *Bgl*III restriction sites was kindly received from Dr. Makris. The *Bgl*III restriction sites are useful in post-transformation analysis. The T-DNA cassette was fused with the 35S promoter and nos-bar herbicide selective gene.

The plasmid selection region *Npt*I gene of pACYC177 expresses resistance to kanamycin, and is used to select for plasmid transformation in both *Agrobacterium* and *E. coli*. While some Ti vectors are able to replicate autonomously in *Agrobacterium* pGREEN requires the presence of another plasmid for such replication. It carries two replication origins (ori), one for use in *E. coli* *Agrobacterium* and one for *Agrobacterium* (e.g. pSa-Ori). This replication origin is recognized in *trans* by the product of the *repA* gene which resides on a separate plasmid

pSoup. This plasmid or a modified version can be introduced in *Agrobacterium* by co-electroporation with pGREEN vector (Hellens et al., 2000) (Fig. II. 11).

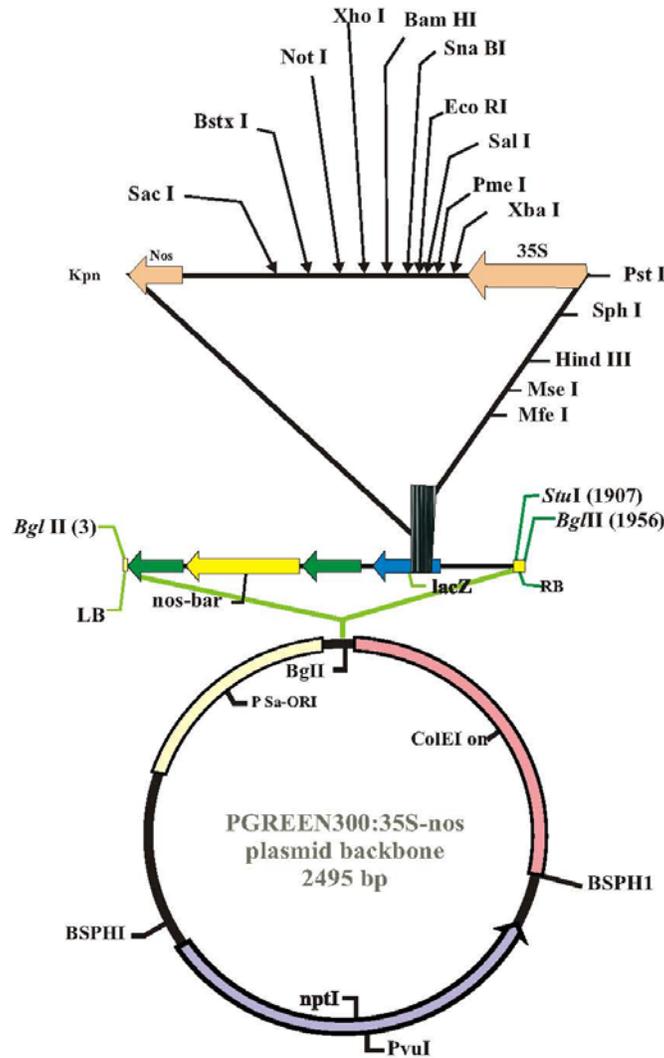


Figure II. 11. The pGREEN300:35S-nos plasmid map.

3.11.2. pAL154

pAL154 is a pSoup-derived plasmid that contains the 15-kb Komari fragment (Komari et al., 1996) and functions as a helper plasmid providing replication in trans for pGREEN plasmid (Fig. II. 12).

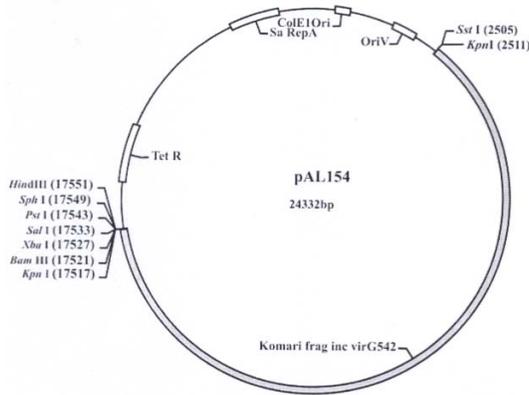


Figure II. 12. The pAL154 plasmid map.

3.11.3. pGREEN300:35S-nos variants

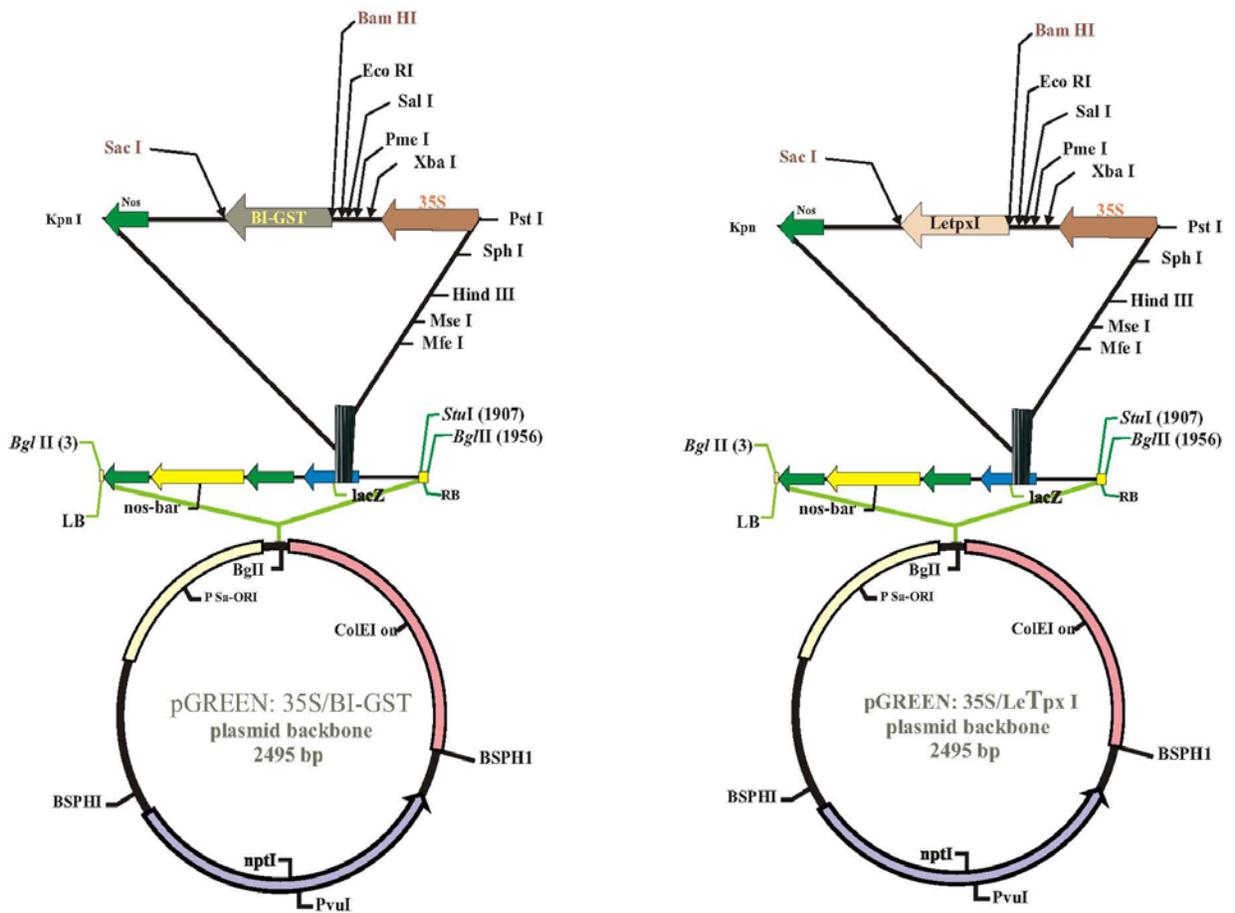


Figure II. 13. The pGREEN300:35S-nos/BI-GST (left) and pGREEN300:35S-nos/LeTpx1 (right) plasmid maps.

The plasmids pGREEN:35S/BI-GST and pGREEN300:35S/LeTpx1 was kindly received from Dr. Antonios Makris. The BI-GST and LeTpx1 fragments were fused in the C-terminus to V5 epitope and were subcloned into *Bam*HI and *Sac*I sites of pGREEN300/35S agrobacterial expression vector. The two constructs have kanamycin selection in *E.coli* and Basta selection in plants (Fig. II. 13).

4. DNA EXTRACTION

4.1. Plasmid extraction

4.1.1. Bacterial plasmid extraction

Plasmid extraction using the QIAGEN miniprep kit

Plasmid extraction was performed according to the manufacturer's instructions QIAprep Miniprep Handbook, 2006. Bacterial cultures were collected by centrifugation, and then resuspended by resuspension buffer (P1) solution, alkaline lysed by lysis buffer (P2) and subsequently neutralized by the neutralization buffer (N3), lysates were cleared by centrifugation. The cleared lysates were then applied to the QIAprep columns where plasmids DNA adsorb to the silica-gel membrane. Impurities are washed away and pure DNA is eluted in a small volume of elution buffer or hot water.

QIAprep spin columns regeneration

- Recycling solution:

3 M NaCl,

0.15 % Triton-X100,

The columns were washed with recycling solution (750 µl) at least three times and the regenerated with 500 µl binding buffer (PB).

4.1.2. Yeast plasmid extraction

Materials and solutions

- Sterile ddH₂O
- Centrifuge (Megafuge 1.0 R, HERAEUS Instruments)

- Microfuge (Centrifuge 5417C, TECHNOLAB)
- Ultra low freezer -75 °C
- Freezer -20 °C
- Vortex (UniEquip)
- 1 M Tris-HCl solution

The solution was prepared as 1M w/v Tris-HCl in ddH₂O and the pH was adjusted to 8.0. The solution was filter-sterilized and stored at room temperature.

- 0.5 M EDTA solution

The solution was prepared as 0.5 M w/v EDTA in ddH₂O, filter-sterilized and stored at room temperature.

- β -mercaptoethanol (BDH Limited Poole England)

It was provided as ready to use solution.

- β -glucuronidase (Sigma) 100000 Units/ml
- 10 % SDS solution

The solution was prepared by dissolving 10 % w/v SDS in ddH₂O. It was filter-sterilized and stored at room temperature.

- 7.5 M Ammonium Acetate
- Isopropanol (Fluka)

It was provided ready to use solution.

- Cold 70 % ethanol

100 % ethanol was diluted in ddH₂O to final concentration at 70 % and stored at 20 °C

- Rescue buffer (prepared fresh)

2.5 ml 1 M Tris-HCl

1 ml 0.5 M EDTA

150 μ l β -mercaptoethanol

ddH₂O up to 50 ml

- Lysis solution (prepared fresh)

β -glucuronidase was dissolved in rescue buffer in ratio 1:50

- TE buffer

10 mM Tris-HCl, pH 8.0

1 mM EDTA

The pH was adjusted to 7.5 and the solution was sterilized by autoclaving.

- Elution buffer (QIAprep Spin Miniprep Kit)

Method

3 ml of yeast culture were collected in eppendorf tubes. The cells were harvested by centrifugation in minifuge at 13,000 rpm for 2 minutes, and washed two times with sterile ddH₂O, under the same conditions.

The supernatant was discarded and cells were resuspended in 1 ml of Rescue Buffer, pelleted, dried and resuspended again in 25 µl of lysis solution. The tubes were vortex and incubated for 1 hour at 37 °C. As a result of the lysis, yeast coagulates into a white precipitate. After lysis, 25 µl of 10 % SDS was added and mixed gently to completely disperse the precipitates. The samples were kept 20 minutes at room temperature. 100 µl of 7.5 M ammonium acetate was added, and the mixture was gently swirled. The samples were frozen for 15 minutes at -75 °C. After completed the freezing, the samples were taken out, gradually warmed up for 2-3 minutes, and spun 3,000 g for 15 minutes at 4 °C. The resulted clear supernatant of each sample was transferred to new tubes and 0.7 volume of isopropanol was added. The content of each sample was gently mixed and subsequently spun for 10 minutes at 13,000 rpm, and the resulted supernatant was decanted. The pellets were washed with 500 µl 70 % cold ethanol without pipeting or mixing, spun for 2 minutes 2 times at 13,000 rpm, completely dried (30 minutes in over), and resuspended in 20 µl TE or EB.

4.2. Genomic DNA extraction

4.2.1. Yeast genomic DNA extraction

Materials:

- YPD growth medium (1 % yeast extract, 2 % peptone, 2 % glucose),
- Lysis buffer: 2 % Triton-X100, 1 % SDS, 100 mM NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0.
- Absolute Ethanol
- 70 % Ethanol
- Chloroform
- TE buffer: (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0)

- RNase A

Method:

1.5 ml of liquid culture grown for 20-24 h at 30 °C was transferred into a microcentrifuge tube, and cells were pelleted by centrifugation at 13,000 rpm for 5 minutes. The pellet was resuspended in 200 µl of lysis buffer. The tube was frozen at -80 °C by immersing in ice cold ethanol bath for 2 minutes and then transferred to 95 °C water bath for 1 minute. This process was repeated twice. The cells were subsequently vortexed for 30 seconds, and 200 µl of chloroform was added to the tube and vortexed for 2 minutes. The tubes were centrifuged at 13,000 rpm for 3 minutes at room temperature and the upper aqueous phase was transferred to a new centrifuge tube containing 400 µl ice cold 100 % ethanol (which favors DNA precipitation), mixed by inversion or gentle vortexing, then incubated at -20 °C for 30 min for precipitation, and then centrifuged again at 13,000 rpm for 5 minutes at room temperature. The pellets were washed with 0.5 ml ice cold 70 % ethanol (salts dissolve in the aqueous phase) and were dried in vacuum dryer (5 minutes at 60 °C). Finally, the pellets were resuspended in 40 µl TE buffer and 3 µl RNase was added to the tube to destroy any remaining RNA during the purification procedure.

4.2.2. Plant genomic DNA extraction

Materials:

- Extraction buffer: 200 mM Tris-HCl, pH 8
 250 mM NaCl
 25 mM EDTA
 0,5 % SDS
- Phenol
- Chlorophorm
- Isoamylalcohol
- 70 % Ethanol
- RNAase A

Method:

Frozen young leaves tissue ground with liquid nitrogen was used to extract genomic DNA. Approximately 200 mg of tissue was mixed with corresponding amount of extraction buffer (400 μ l), vortexed briefly and spun at 13,000 rpm for 10 minutes to pellet the debris. The resulting supernatant was treated with 20 μ l RNAase and incubated for 1 hour at 37 °C, then extracted with phenol : chlorophorm : isoamylalcohol (25:24:1) and centrifuged 10 minutes at 13,000 rpm. The supernatant was collected and treated with equal amount of chlorophorm : isoamylalcohol (24:1), and then centrifuged for 30 minutes at 13,000 rpm. The pellet was washed with 600 μ l of 70 % cold ethanol, centrifuged 20 minutes at 13,000 rpm. The ethanol was discarded and the pellet let to dry for 30 minutes at 37 °C. The amount of extracted genomic DNA was quantified by spectrophotometer at 260 nm and its integrity was checked in a 1.2 % agarose gel. The sample was stored at -20 °C.

5. CLONING PROTOCOLS

5.1. Agarose gel electrophoresis

Materials

- Agarose (Invitrogen)
- Running Gel Buffer, TAE Buffer 50X stock

24.2 % (w/v) Tris base,

3.72 % (w/v) EDTA, the pH was adjusted to 8.0 with ~57 ml Glacial acetic acid and brought to 1 liter with distilled water. The working solution was 1x TAE prepared by dilution of the 50X buffer (40 ml 50X TAE and the volume completed to 2 liters with distilled water).

- STEB (6X Loading buffer) for DNA

100 mM Tris HCl, pH 7.5,

20 % Glycerol (w/v),

1 mM EDTA,

Few grains of Bromophenol blue.

- Staining solution:

50 μ l of 1 mg/ml Ethidium bromide stock added to 500 ml of distilled deionized water.

- Destaining solution: clean deionized distilled H₂O.

Method

Agarose gel electrophoresis approach was used to separate, identify and purify DNA fragments. 1 % agarose gel was prepared by adding 100 ml of 1X TAE to 1g of agarose, mixed and then heated for 2-3 min in a microwave oven until the solution became clear. The tray was assembled with the comb into the gel apparatus. After being cooled down to 55 °C, the gel was poured into the tray and left until solidified. The tank was filled with 1X TAE solution covering the top of the gel. The DNA samples were mixed with the corresponding volume of 6X loading buffer (to a final concentration of 1X) and the samples were applied to the wells. The gel was run at 150 V for at least half an hour to ensure an optimal separation of the fragments, and then stained with Ethidium Bromide for 15-20 min. The gel was transferred to the destaining solution for 5 minutes to remove the excess ethidium bromide before visualizing it under UV light.

5.2. Restriction digestion of plasmid DNA

Restriction enzymes or restriction endonucleases are essential tools in recombinant DNA methodology. Several hundred have been isolated from a variety of prokaryotic organisms and named according to a specific system of nomenclature. Restriction endonucleases recognize and digest specific sequences in double stranded DNA (usually a four to six base pair sequence of nucleotides). The result is the production of fragments of various lengths. The cleavage may be at adjacent sites leaving a blunt ends, or might be offset by 1 to 4 bases, leaving either a 3' overhang or a 5' overhang of a single strand yielding "sticky end" cuts.

Materials:

- 10X BSA stock (10 mg/ml) New England Biolabs
- 10X restriction enzyme buffer, NEBuffer, New England Biolabs
- Restriction enzyme
- Plasmid DNA (at a concentration no more than 200 ng/μl)

Method:

Restriction digestion of the plasmid was performed for excision of certain inserts to be cloned in another plasmid, also used for the identification of an insert being cloned after cloning. The reaction mixture usually contained:

5 μl plasmid DNA,

3 μl of 10X BSA,

3 μl of 10X restriction enzyme buffer,
1 μl of each restriction enzyme,
Sterile water was added to a final volume of 30 μl .

The reaction mixture was incubated at 37 °C for 2 hours and then the fragments were resolved by agarose gel electrophoresis.

5.3. Vector-insert ligation

DNA ligases have become an indispensable tool in modern molecular biology research for generating recombinant DNA sequences. The mechanism of DNA ligase in connecting broken DNA strands is to form covalent phosphodiester bonds between 3' hydroxyl ends of one nucleotide with the 5' phosphate end of another. DNA ligations were performed by incubating DNA fragments with appropriately linearized cloning vector in the presence of buffer, ATP, and T4 DNA ligase.

The DNA concentration in the ligation should be of 0.2 μM to support intramolecular interaction, thus the formation of monomeric circles. The digested samples were ligated with 1 μl of T4 ligase in a 40 μl reaction volume overnight at 4 °C. For a successful ligation, after digestion of plasmid and excision of the insert, the concentrations of linear vector and insert were estimated by running on agarose gel. Approximately 1:1 or 1:2 vector: insert DNA was used.

The reaction mixture usually contained:

5 μl vector,
10 μl insert,
2 μl 10X T4 ligase buffer,
1 μl of T4 ligase
sterile water was added to make up the final volume 20 μl .

A control reaction which contained only

5 μl vector,
2 μl T4 ligase buffer,
1 μl ligase enzyme
sterile water was added to make up the final volume 20 μl .

The reaction mixture were incubated overnight at 4 °C and μl of the small aliquots (usually 10 μl) were used to transform Mach-1 competent cells. Colonies grown on the selective

medium were first amplified overnight and then DNA plasmid was extracted. Confirmation of the new clone was justified by digestion with restriction enzymes other than those used for the cloning purpose and by PCR.

6. PCR AMPLIFICATION

Polymerase chain reaction (PCR) is a very sensitive, inexpensive, fast and versatile *in vitro* molecular biological technique for amplifying a short (usually up to 10 kb), well-defined part of a DNA strand within a heterogeneous collection of DNA sequences without using a living organism, such as *E. coli* or yeast. The technique allows a small sample of DNA to be copied multiple times so it can be used for analysis, using thermo stable DNA polymerases. Before the PCR product is used in further applications, it has to be checked if it has the right size and sequence by agarose gel electrophoresis where a single, discrete band should be seen.

6.1. PCR reaction

Materials and Solutions:

- DNA Template (Usually the amount of template DNA is in the range of 0.01-1ng for plasmid DNA and 0.1-1 μ g for genomic DNA, for a total reaction mixture of 50 μ l).
- Primers 100 pmol/ μ l stock solution, used to the final concentration of 10 pmol/ μ l,
- Taq DNA polymerase (5 U/ μ l stock)
- 10X ThermoPol Taq buffer with MgCl₂,
- 10 mM dNTPs,
- Sterile PCR water.

Method:

A master mix was prepared, consisting of 1 μ l of each primer to the final concentration of 10 pmol/ μ l, 5 μ l 10X Thermo Pol Taq buffer, 1 μ l dNTPs (10mM), and then water was added to reach 47,5 μ l volume per reaction, and multiplied by the number of reactions. After adding the master mix to the PCR tubes, 2 μ l DNA template (diluted 1:10) were added. Reactions were incubated allowed to cycle in Gene Amp PCR 2400 system (Perkim-Elmer) under the conditions needed depending on the reaction, 0,5 μ l of Taq polymerase was added to the reaction mixture

when the temperature reached 94 °C, to have final volume of 50 μ l. The standard PCR reaction is described by the below parameters (Fig. II. 14):

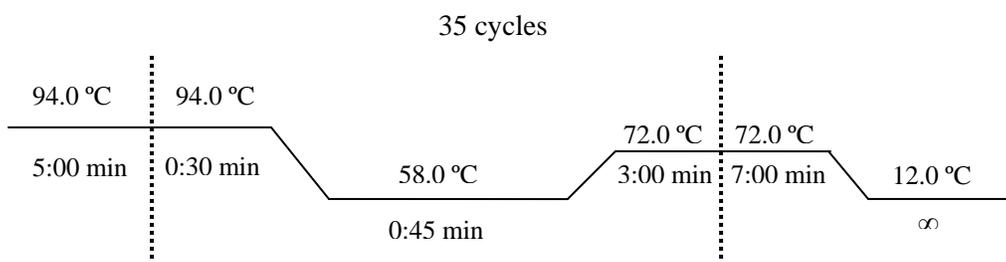


Figure II. 14. Schematic representation of the PCR standard reaction conditions.

When HotStarTaq Polymerase was used for higher PCR specificity, the Qiagen HotStarTaq® PCR Handbook protocol was followed.

6.2. Colony PCR

Colony PCR is a typical PCR reaction in which the DNA template is replaced by a small amount of bacterial or *Agrobacterium* cell culture. In colony PCR the samples are first heated at least 5 minutes at 94 °C in order to lyse the cells and expose the endogenous DNA. The rest of the reaction is carried out as a regular PCR reaction, however this must be done under restricted sterilized conditions.

6.3. Generation of double stranded oligos

5' and 3' HA oligos were designed to generate the double stranded HA oligo:

HA tag (HindIII):



The HA oligos were dissolved to a final concentration of 1 μ g/ μ l and 1 μ l of each primer and 98 μ l of H₂O were heat in a PCR machine for 3 min at 94°C and anneal at 58 °C for 5 min, then cool at 12 °C (Fig. II. 15).

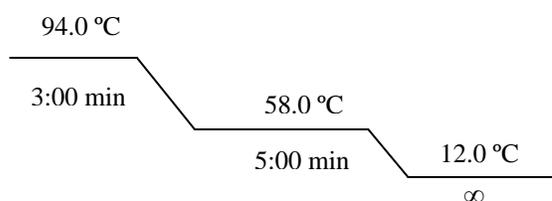


Figure II. 15. Schematic representation of the PCR reaction conditions used to amplify HA oligo.

6.4. Purification of the PCR product

6.4.1. Purification by gel extraction

Several techniques can be used to purify DNA (PCR product or linearised plasmid after enzymatic digestion) from agarose gels, most of them start out by excising the desired "band" from an ethidium-stained gel viewed with a UV transilluminator, with minimum exposure. Then, the desired piece of agarose is cut out using a clean and sharp scalpel, avoiding the agarose excess. The scalpel may be rinsed with autoclaved water or TE prior cutting the next band. The protocol of gel extraction kit provided by (QIAGEN, gel extraction kit, cat.28704) was followed for the purification of the nucleic acids. The gel slice was weighed in a tube and 3 volumes of QG buffer were added to 1 volume of excised gel. The mixture was incubated at 50 °C for 10 minutes, or until the gel dissolved completely. The appearance of a vivid yellow color indicates the complete melting of the gel slice. One gel volume of isopropanol was added to the sample and mixed thoroughly. This step was found to increase the yield of DNA extracted, especially when the size of the fragment is between 500 bp and 4000 bp. The sample was applied to the QIAquick spin column accompanied with a provided 2 ml collection tub, and centrifuged for 1 minute. The eluant was discarded and the column was washed with 750 µl of washing buffer (PE) and centrifuged for 1 minute. The eluant was discarded again, and the column was spun for an additional 1 minute at 13,000 rpm. The column was placed into a clean 1.5 ml tube and 30 µl of prewarmed elution buffer (EB) was added to the centre of the matrix of the column to elute the DNA. The column was spun at 13,000 rpm for 1 minute and the eluant was collected. The presence of the DNA fragment was checked on a 1 % agarose gel and the extracted DNA was stored at -20 °C.

6.4.2. Purification by ethanol:chloroform precipitation

Phenol is slightly water-soluble, used as water-saturated solution buffered with Tris and gives a 'fuzzy' interface that is sharpened by the presence of chloroform. Ethanol precipitation is a method used to concentrate DNA. DNA is polar, soluble in water, which is also polar. Based on the principle of "like dissolves like", it is insoluble in the relatively less polar ethanol. The DNA precipitation in this step is due to the ethanol interacting with the water so that fewer water molecules are available to dissolve the DNA molecules.

Materials:

- Tris-buffered 50 % phenol
- Tris-buffered 50 % chloroform
- 100 % ethanol
- 70 % ethanol

Method:

To 200 μ l of DNA sample into a 1.5 ml microcentrifuge tube, one volume of phenol:chloroform is added which will give a sharp interface, vortexed vigorously one minute to mix the phases. Centrifugation was carried out at top speed for 3 min at 13000 rpm to separate the phases, and then the aqueous phase is removed to a new tube. The aqueous phase is usually the upper one. The loss of DNA into the organic phase can be reduced by adding a second volume of water, mixing, centrifuging, and removing again. The extraction step with (1:1) V of phenol:chloroform which is repeated once, and then followed by extraction with 1 volume of chloroform, which will ensure that no phenol whatsoever remains. DNA is initially ethanol-precipitated by adding 2 V of ice cold 100 % ethanol and 1/10 volume of 3 M Sodium Acetate pH 5.8. The samples were incubated 30 min at -80 °C or 2 hours at -20 °C. This will precipitate DNA, as well as the salts that form ionic bonds with it. The suspension was centrifuged for 20 min in a microcentrifuge tube at 4 °C and the supernatant was removed, leaving a pellet of the crude DNA. In the next step, 70 % ethanol was added to the precipitated pellet, without vortexing or disturbing the pellets (fast), to resuspend the DNA. This allows the 20-30 % water to access the salts present in the pellet. This suspension was centrifuged again for 20 min at 13000 rpm, and the supernatant was removed by aspiration. The pellet was air-dried giving the purified DNA and were resuspended with the 10-20 μ l of 1X TE or elution buffer.

6.5. Cloning of the amplified PCR products in pCRII-TOPO vector

Materials and Solutions:

- Salt solution .
- Sterile water
- PCR TOPO Vectors: pCRII TOPOR or pCR1.2 TOPO (Invitrogen,USA) .
- Insert
- SOC medium
- 40 mg/ml X-gal in dimethylformamide (DMF)
- Commercial chemically competent *E. coli* Mach one cells.

Method:

Once the desired PCR product has been produced and extracted from the agarose gel, it was cloned into the pCR[®]2.1-TOPO[®] or pCR[®]II-TOPO[®] vector and used to transform competent *E. coli* cells following the protocol provided by (Stratagene).

To confirm the presence and correct orientation of the insert, plasmids-insert were extracted and analyzed by restriction enzymes using one or a combination of two enzymes that cut once in the vector and once in the insert. Once the correct clone identified, a glycerol stock was prepared for long term storage and the stock of plasmid DNA was stored at – 20 °C.

7. NOVEL DNA RECOMBINANT VECTORS

7.1. Generation of loxP-URA3-loxP-GALp-HA integration cassette

The cassette will be used for the integration of the galactose promoter-hemagglutinin tag upstream of native yeast genes.

7.1.1. pYES-HA

Two complementary oligos were designed which contained the DNA sequence coding for the hemagglutinin epitope M(YPYDVPDYA), an internal *Bam*HI restriction site and 5'*Hind*III and 3'*Eco*RI overhang to ligate to pYES2 digested with the corresponding enzymes. The primer sequences were 5'-HA(*Hind*III) 5'-AGCTCATGTATGATGTGCCAGATTATGCC TCTGGATCCG-3' and 3'-HA(*Eco*RI) 5'-AATTCGGATCCAGAGGCATAATCTGGCACATC ATACATG-3'. The two oligonucleotides were mixed in equal amounts to a final concentration

of 1 $\mu\text{g}/\mu\text{l}$ and added to a PCR tube containing 98 μl H_2O . The complementary oligos were heated at 94 $^\circ\text{C}$ to unfold and were then allowed to anneal by incubation at 58 $^\circ\text{C}$ for 5 min. The pYES2 vector was digested with *Hind*III and *Eco*RI, resolved in an agarose gel and purified by gel extraction using the Qiagen gel extraction kit. A series of ligations were prepared using increasing concentrations of the double stranded oligo (10 ngs, 20 ngs, 30 ngs) to 250 ngs of plasmid. Mach-1 *E.coli* cells were transformed with 5 μl of the ligation reaction. Several colonies growing on the selection plates were inoculated in LB ampicillin 5 ml cultures and plasmid DNA was extracted. To verify proper integration of the HA double oligo plasmid DNA was digested with *Hind*III, *Sac*I or *Bam*HI. The plasmids that successfully incorporated the tag are not digestible by *Hind*III or *Sac*I in contrast to the wild type plasmid vector. pYES-HA was designed to have *Bam*HI and *Eco*RI cloning fusion sites at 0 frame (Fig. II. 16).

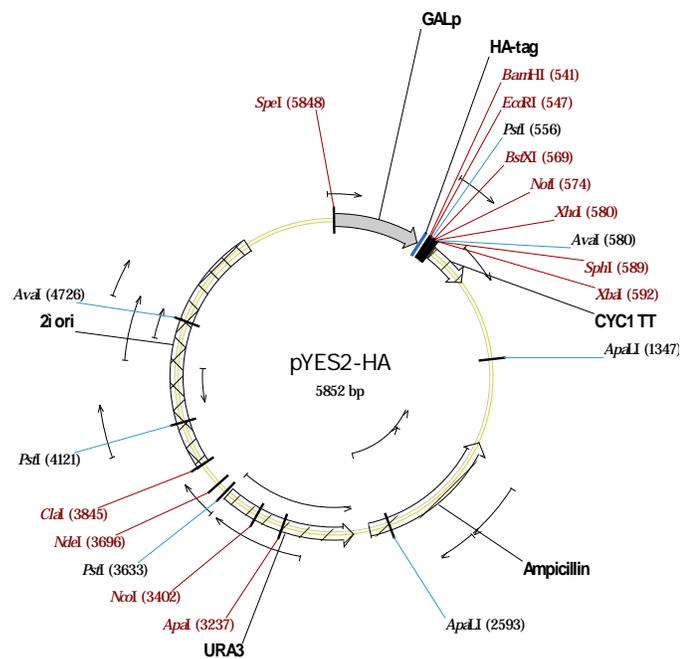


Figure II. 16. The pYES-HA plasmid map.

7.1.2. pCOD2

The DNA fragment encompassing the loxP-URA3-loxP sequence was excised from the pUG72 cassette vector using *Hind*III and *Sfi*I. The 1.7 kb insert was gel purified by gel extraction. Using the plasmid pJG4-6 as template, the galactose promoter and hemagglutinin tag (HA) sequence were amplified by PCR using the primers 5'-GAL(*Sfi*I) 5'-AGT GGC CTA TGC

GGC CAC GGA TTA GAA GCC GCC GA-3' and 3'-HA (EcoRI) 5'-AAT TCG GAT CCA GAG GCA TAA TCT GGC ACA TCA TAC ATG-3' which introduce an *SfiI* site at 5' of the GAL promoter and an *EcoRI* site at the 3' of the HA tag. The amplified product was cloned into the pCRII vector in a TOPO TA reaction. Plasmid DNA was extracted from several grown colonies and digested with *EcoRI* to verify the presence of the insert.

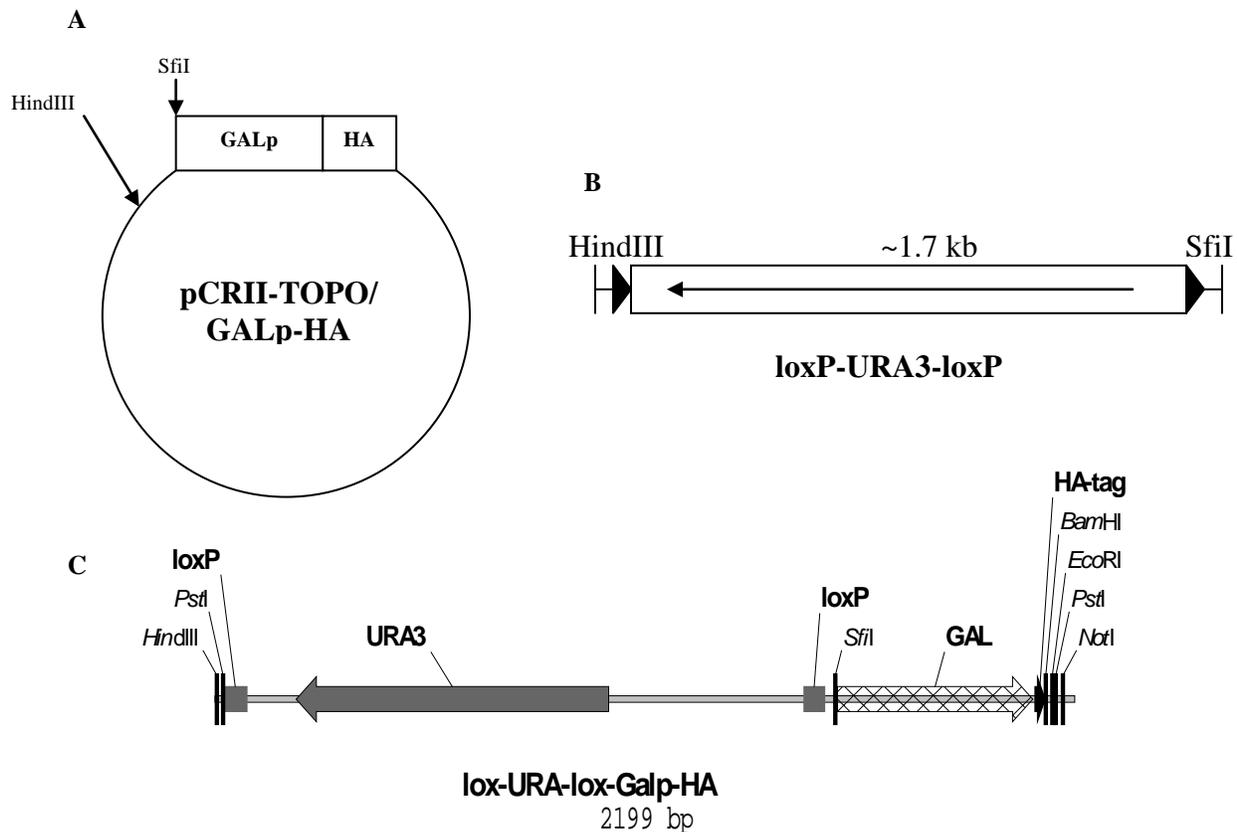


Figure II. 17. Cloning of loxP-URA3-loxP fragment into pCRII-TOPO/GALp-HA construct to generate pCOD2 integration vector: (A) pCRII-TOPO/GALp-HA vector; (B) loxP-URA3-loxP fragment; (C) loxP-URA-loxP-GALp-HA pCOD2 integration cassette.

The plasmids confirmed to carry the GALp-HA insert were further analyzed for their orientation relative to the pCRII vector by double digestion with *HindIII* which cuts once on the plasmid MCS and *SfiI* which cuts at the 5' of the galactose promoter. The linearized plasmids carrying the insert in the desired orientation were used in a ligation reaction with the loxP-URA3-loxP insert. The ligation was transformed into chemically competent Mach1 cells. Plasmid DNA was extracted from grown colonies, digested with *EcoRI* and *HindIII* to verify the

presence of the second insert. The identified full construct was confirmed by DNA sequencing using m13forward and m13reverse primers annealing to the plasmid proximal sites. The complete loxP-URA3-loxP-GALp-HA integration cassette was named pCOD2 (Fig. II. 17) (see sequence in Appendix 1).

7.2. Generation of a loxP-URA3-loxP-ADH-myc integration cassette

The following cassette was developed to enable the insertion of a strong constitutive promoter and a myc epitope tag upstream of yeast promoters in a recyclable manner.

Using the vector pEG202 as DNA template the ADH promoter was amplified by PCR using the primers *SpeI-SfiI* 5'ADH 5'-ACTAGTGGCCTATGCGGCCTATTTTCGGATATCCTT TTGTTG-3' and 3'-ADH(*HindIII*) 5'-AAGCTTGGAGTTGATTGTAT GCTTG-3'. The amplified PCR product was purified by gel extraction and cloned into the pCRII vector in a TOPO TA reaction. Plasmid DNA was extracted from few colonies and the presence of the ADH-myc insert was verified by restriction digest by *SpeI* and *HindIII*.

7.2.1. pCADH-myc

In parallel to the cassette construction a yeast plasmid vector pADHmyc (URA3, 2 μ , amp) was developed by digesting the pYESmyc and the pCRII/ADHmyc plasmids with *SpeI-HindIII*. The reaction products were resolved by agarose gel electrophoresis and the ADHmyc insert and the pYES backbone minus the GAL promoter were gel extracted and purified. The two fragments were ligated and 5 μ l from the ligation product was used to transform Mach-1 cells. Four colonies were selected, grown in LB ampicillin liquid cultures and plasmid DNA was extracted. Integration of the ADH-myc insert was verified by restriction digest with *SpeI* and *HindIII* and separation on an agarose gel to assess the size of the insert which coincided with the size of the ADH-myc fragment (Fig. II. 18).

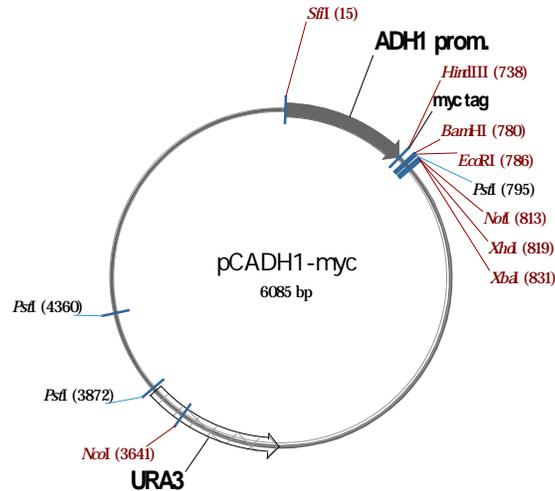


Figure II. 18. The pCADH1-myc plasmid map.

7.2.2. pCOD3

The cassette vector pCOD2 and the plasmid pADH-myc were digested with *SfiI*-*BamHI*, which in the case of pCOD2 it excised the GALp-HA insert and left the loxP-URA3-loxP plasmid backbone. The DNA fragments plasmid-loxP-URA3-loxP and ADH-myc were resolved by gel electrophoresis and the fragments were gel extracted and purified. They were subsequently ligated and a small aliquot was used to transform Mach-1 bacteria. Colonies growing were inoculated in LB ampicillin media and plasmid DNA was extracted and digested with *SfiI* and *BamHI* to verify the presence of the ADH-myc. The new integration cassette construct was pCOD3 (Fig. II. 19). The construct was additionally validated by DNA sequencing using primers from the plasmid backbone (see sequence in Appendix 1).

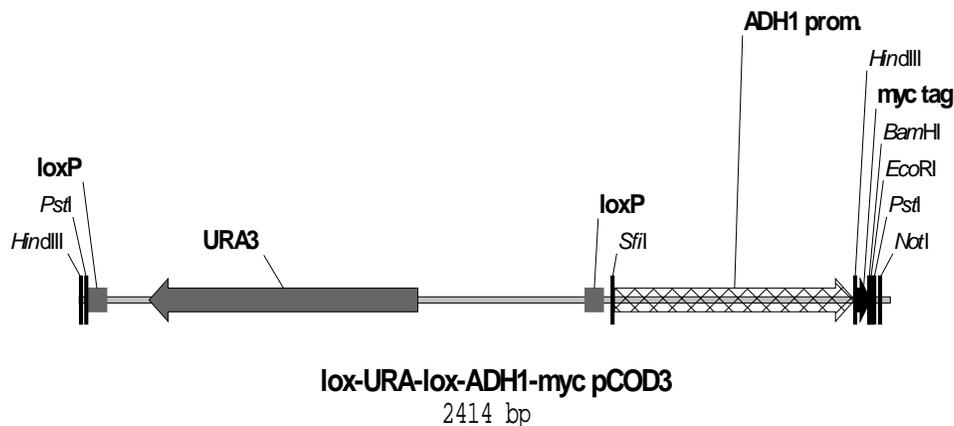


Figure II. 19. The loxP-URA-loxP-ADH1-myc pCOD3 integration cassette.

7.3. pDNR-GAL

The GALp-HA-cyc1 fragment was PCR amplified from pYES-HA recombinant vector in a PCR reaction using the following primers: 5'-GAL (HindIII) 5'-AAGCTTACGGATTAGAA GCCGCCGAG-3' and 3'-CYC1 (SalI) 5'-GTTCGACGGCCGCAAATTAAGCCTTCG-3'. The amplified PCR product was purified by gel extraction and cloned into the pCRII vector in a TOPO TA reaction. Plasmid DNA was extracted from few colonies and the presence of the GALp-HA-cyc1 insert was verified by restriction digest by *HindIII* – *SalI*. The resolved insert was gel extracted and purified.

The pDNR-LIB vector was digested with *HindIII* and *SalI*. The reaction was resolved by agarose gel electrophoresis and the linearized vectors were gel extracted and purified. A ligation reaction was performed between the pDNR-LIB linearized vector and the GALp-HA-cyc1 fragment respectively. Small aliquot was used to transform Mach-1 bacterial cells. Several growing colonies from each transformation were used to extract plasmid DNA. To verify the presence of GALp-HA-cyc1 the construct was digested with *HindIII* and *SalI*. The novel construct was named pDNR-GAL-HA (Fig. II. 20).

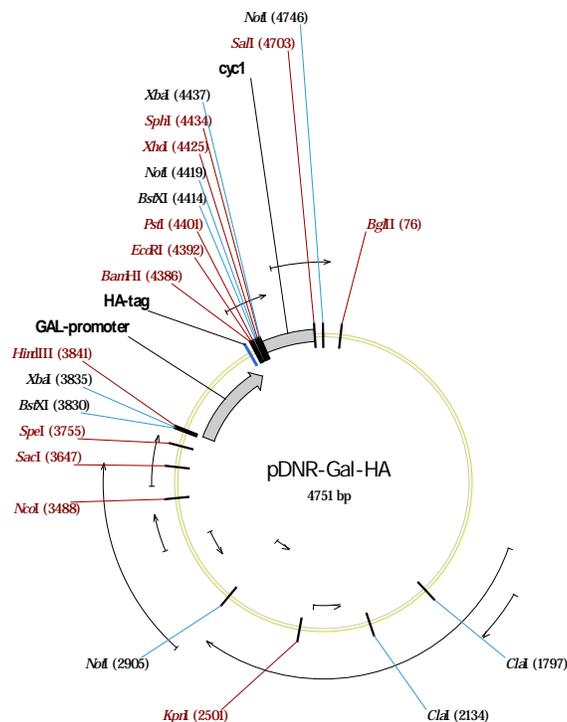


Figure II. 20. The pDNR-GAL-HA plasmid map.

8. Subcloning

8.1. Cloning of GPP synthase into pJG4-6 and pYES-myc plasmids

The primers 5'GPPS 5'-GAATTCATGTGCTCAAACACAAATGCCAG-3' and 3'-GPPS 5'-CTCGAGTCAGTTCTGTCTTTGTGCAATGTA-3' were used to amplify in a PCR reaction the core enzyme sequence of spruce GPPS with the plastidial sequence removed. The PCR product was cloned in pCRII in a TOPO TA reaction. Plasmid DNA was extracted from several grown colonies and was digested using *EcoRI* and *XhoI* and resolved by agarose gel electrophoresis. The resolved insert was gel extracted and purified.

The vector plasmids pYESmyc and pJG4-6 were digested with *EcoRI* and *XhoI*. The reactions were resolved by agarose gel electrophoresis and the linearized vectors were gel extracted and purified.

Two sets of ligations were performed between the pJG4-6 and the pYESmyc and the GPPS fragment respectively. Small aliquots were used to transform Mach-1 bacterial cells. Several growing colonies from each transformation were used to extract plasmid DNA. To verify the presence of GPPS insert, the constructs were digested with *EcoRI* and *XhoI* restriction enzymes (Fig. II. 21).

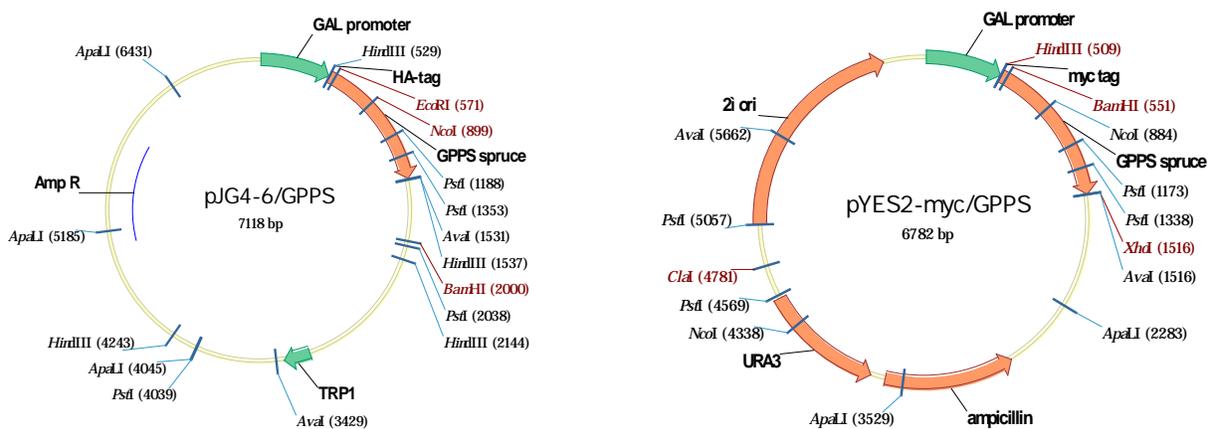


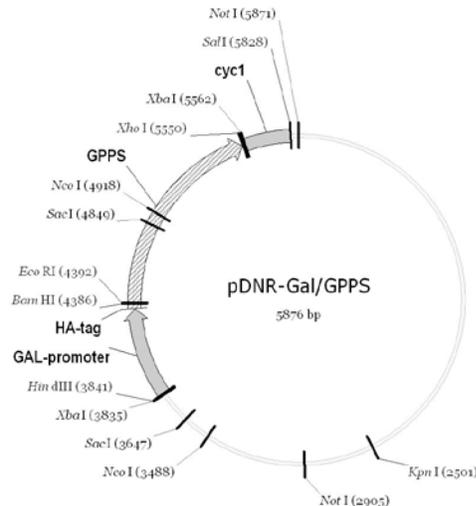
Figure II. 21. The pJG4-6/GPPS and pYES2-myc/GPPS plasmid maps.

8.2. Cloning of GPP synthase into pDNR-GAL

The GPP synthase fragment of excised from pYESmyc with *EcoRI* and *XhoI*, resolved by agarose gel electrophoresis, and then gel extracted and purified. The vector plasmid pDNR-GAL was digested with the same enzymes *EcoRI* and *XhoI*, and identically processed.

A reaction ligation was performed between the pDNR-GAL vector and the GPP synthase fragment. The reaction tubes were incubated overnight at 4 °C, and 10µl from each tube were used to transform Mach-1 *E. coli* competent cells. The plasmids from 5 randomly picked colonies grew on LB-Chloramphenicol selective plates were isolated and digested with *Eco*RI and *Xho*I and after running in 1 % agarose gel, the plasmid pDNR-GAL with the correct inset GPPS was picked and stored at -20 °C (Fig. II. 22).

Figure II. 22. The pDNR-GAL/GPPS plasmid map.



8.3. Cloning of GALp-GPPS-CYC1 cassette into pUG27 vector

The pDNR-Gal/GPPS recombinant construct was restriction digested with *Hind*III and *Sal*I to liberate the cassette GALp-GPPS-CYC1 of ~2kb size. Also the disruption vector pUG27 was digested with *Hind*III and *Sal*I. The digestion products were run on 1% agarose gel and the linear pUG27 vector (~4kb) and the GALp-GPPS-CYC1 fragment were gel extracted and purified. A ligation reaction was performed between the two fragments, and overnight incubated at 4°C. Small aliquot was transformed into Mach-1 *E. coli* competent cells. Several growing colonies from each transformation were used to extract plasmid DNA. To verify the presence of GPPS cassette the constructs were digested with *Hind*III and *Sal*I (Fig. II. 23).

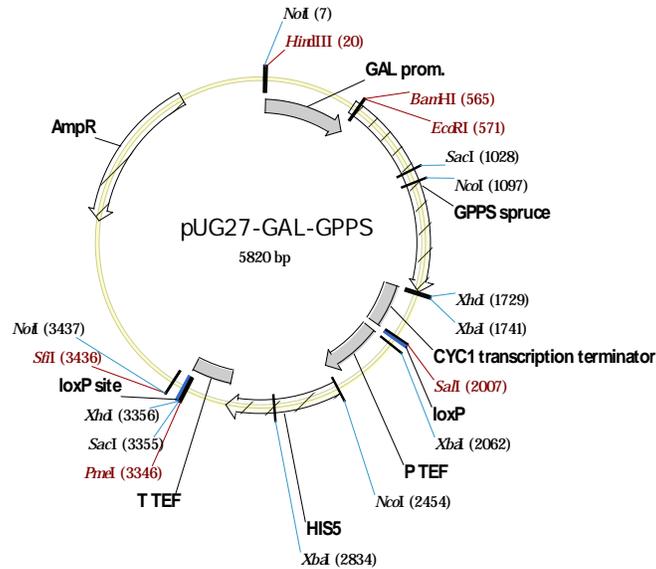


Figure II. 23. The pUG27-GAL-GPPS integration vector map.

8.4. Cloning of *SfCinS*(RC) and *SfCinS*(RR) into pJG4-6 and pYES plasmids

The pTOPO/*SfCinS*(RC) and pTOPO/*SfCinS*(RR) constructs, received from Walid Mahrez, were restriction digested with *EcoRI* and *XhoI*. In parallel, the pJG4-6 and pYES2 yeast vector plasmids were digested with *EcoRI* and *XhoI*. The digestion products were resolved by agarose gel electrophoresis. The *SfCinS*(RC), and the *SfCinS*(RR) fragments as well as the linearized vectors, pJG4-6 and pYES2 were gel extracted and purified.

Four sets of ligations were performed between the pJG4-6 and the pYES2 and the *SfCinS*(RC), and the *SfCinS*(RR) fragments respectively. Small aliquots were used to transform Mach-1 bacterial cells. Several growing colonies from each transformation were used to extract plasmid DNA. To verify the presence of the inserts, the constructs were digested with *EcoRI* and *XhoI*.

8.5. Cloning of P330 into pYES2 plasmid

The pDNR/P330 construct, kindly provided by Dr. Antonios Makris, was restriction digested with *EcoRI* and *XbaI*. In parallel, the pYES2 yeast vector plasmid was digested with *EcoRI* and *XbaI*. The digestion products were resolved by agarose gel electrophoresis. The P330 fragment, as well as the linearized vector pYES2 were gel extracted and purified.

A ligation reaction was performed between the pYES2 and the P330 fragment respectively. Small aliquots were used to transform Mach-1 bacterial cells. Several growing

colonies from each transformation were used to extract plasmid DNA. To verify the presence of the inserts, the constructs were digested with *EcoRI* and *XbaI*.

8.6. Cloning of HSP90 into pYX143HA and pYES2 plasmids

The HSP90 fragment was PCR amplified from modified pBluescript1/*Ath*HSP90 (RALF06-10-K09, RIKEN Arabidopsis) recombinant vector in a PCR reaction using the following primers: 5'HSP90(*Bam*HI) 5'-GGATCCATGGCGGACGCAGAAACCTT-3' and 3'HSP90(*Xho*I) 5'-CTCGAGGTCAACTTCCTCCATCTTGCT-3'. The amplified PCR product was purified by gel extraction and cloned into the pCRII vector in a TOPO TA reaction. Plasmid DNA was extracted from few colonies and the presence of the GALp-HA-cyc1 insert was verified by restriction digest by *Bam*HI and *Xho*I. The resolved insert was gel extracted and purified.

The pYES2 and pYX143HA vectors were restriction digested with *Bam*HI and *Xho*I. The reaction was resolved by agarose gel electrophoresis and the linearized vectors were gel extracted and purified. Two sets of ligation reactions were performed between the pYES2 and pYX143HA linearized vectors and the HSP90 fragment respectively. Small aliquots were used to transform Mach-1 bacterial cells. Several growing colonies from each transformation were used to extract plasmid DNA. To verify the presence of HSP90 the construct was digested with *Bam*HI and *Xho*I.

8.7. Cloning of 215R clone into pYESmyc plasmid

The pCRII-TOPO/215R construct, kindly received from Dr. Antonios Makris, was restriction digested with *Bam*HI and *Xho*I. In parallel, the pYESmyc yeast vector plasmid was digested with *Bam*HI and *Xho*I. The digestion products were resolved by agarose gel electrophoresis. The 215R fragment, as well as the linearized vector pYESmyc were gel extracted and purified.

A ligation reaction was performed between the pYESmyc and the 215R fragment respectively. Small aliquots were used to transform Mach-1 bacterial cells. Several growing colonies from each transformation were used to extract plasmid DNA. To verify the presence of the inserts, the constructs were digested with *Bam*HI and *Xho*I.

9. TRANSFORMATION PROTOCOLS

9.1. Quantification of nucleic acid concentration

The absorbance of UV light at 260 nm by nucleic acids gives an estimate of their concentration. Both RNA and DNA contain aromatic purine and pyrimidine groups that absorb UV light thus making possible to detect and quantify either at concentrations as low as 2.5 ng/μl. Based on the extinction coefficient of nucleotides at 260 nm that is equal to 20, the absorbance at 260 nm in a 1-cm quartz cuvette of a 50 μg/ml solution of double stranded DNA or a 40 μg/ml solution of single stranded DNA or RNA is equal to 1. An absorbance ratio of 260 nm and 280 nm gives an estimate of the purity of the solution. Pure DNA and RNA solutions have OD₂₆₀/OD₂₈₀ values of 1.8 and 2.0, respectively. This method is not useful for small quantities of DNA or RNA (<1 μg/ml).

The concentration of the DNA or RNA in our sample can be calculated as below:

DNA concentration (μg/ml) = (OD₂₆₀) x (dilution factor) x (50 μg DNA/ml)

RNA concentration (μg/ml) = (OD₂₆₀) x (dilution factor) x (40 μg RNA/ml)

Materials:

Quartz cuvettes, DNA to be quantified, positive controls of known concentration (pUC19), Spectrophotometer (Hewlett Packard 8452 A).

Method:

To measure the concentration of DNA present in my samples, 10 μl of DNA were mixed with 790 μl of distilled water, put into a clean quartz cuvette and the absorbance measured at 260 nm using water as blank and water with pUC19 as positive control.

9.2. Measuring the optical density of cell cultures

The measurement of the optical density of a bacterial culture by spectrophotometer is used by microbiologists to determine the amount of bacterial cells present in a liquid culture. As visible light passes through a cell suspension the light is scattered. Greater scatter indicates that more bacteria or other material is present. Typically, when working with a particular type of cell, the optical density at a particular wavelength that correlates with the different phases of bacterial growth would be determined. Generally cells at their mid-log phase of growth are required for protein expression.

Materials:

- Plastic cuvettes
- Bacterial culture
- Spectrophotometer

Method:

The flask containing LB broth with the bacterial culture was mixed well, then 1 ml of the culture was taken into the plastic cuvettes and the OD₆₀₀ was measured using 1 ml of LB medium mixed with the appropriate antibiotics as a negative control.

9.3. Bacterial transformation

9.3.1. Preparation of chemical competent cells

Materials:

- Transformation Buffer –X
- 50 mM CaCl₂,
- 5 % (v/v) Glycerol,
- 10 mM MOPS pH 6.6 (adjusted with NaOH only, MOPS = 3 [N- morpholino]

propanesulfonic acids pH 6.2-7.6, FW: 231.2).

Autoclaving conditions are applied (15 min, 151 lb/sq.in. on liquid cycle).

Method:

A 5 ml overnight culture of cells was grown in LB broth. 1ml of the overnight culture was inoculated into 400 ml of pre-warmed LB broth and grown to ~ 0.2 OD₆₀₀ (2-3 hours). The cells were moved to 50 ml sterile falcon tubes and chilled on ice for 10 min. Later on cells were pelleted at 5500 rpm at 4 °C for 10 min and resuspended in 50 ml of ice-cold transformation buffer-X. The resuspended cells were incubated on ice for 20 min and spun again for 10 min (5000-6000 rpm, 4 °C). The pellet was resuspended with 3 ml of ice-cold transformation buffer-X (for each falcon). 50µl aliquots of the cells were prepared into 1.5 ml Eppendorf tubes, fast frozen in liquid nitrogen and directly kept at -80 °C.

9.3.2. Bacterial transformation

A 50 µl aliquot of competent cells was chilled on ice for 10 min to thaw. 1-3 µl of DNA was subsequently added and the cells were incubated on ice for 30 min. The cells were heat shocked for 45 seconds at 42 °C in a water bath and then incubated on ice for 2 minutes. After the end of heat-shock, 300 µl of recovery LB or SOC medium was added and the tubes were further incubated shaking for 1 hour at 37 °C. Aliquots of 100µl of cells were plated on selective LB-agar plates and the plates were incubated overnight at 37°C till distinct colonies appeared.

9.4. Yeast transformation

Reagents for LiOAc yeast transformation

- 10X Lithium acetate solution

The lithium acetate (Lithium acetate FW: 102.02 g/mol) solution is prepared as a 1 M stock in distilled water and filter sterilized. The final pH should be between 8.4-8.9 (pH 7.5).

- 10X TE buffer (pH 7.5)

Tris HCL

0.5 M EDTA stock

The buffer is a mixture of 100 mM Tris HCL (pH 8.0), and 10 mM EDTA, pH is then adjusted to 7.5 and sterilized by standard autoclaving for 20 minutes.

- Polyethylene glycol (PEG 50 % w/v)

The polyethylene glycol (PEG), MW 3350 (Sigma) is made up to 50 % (w/v) with distilled water and then filter sterilized. For optimal transformation efficiencies, care must be taken to ensure that the PEG solution is at the proper concentration. In addition, it is important to store the PEG in a tightly capped container to prevent evaporation of water and a subsequent increase in PEG concentration. 50 g of polyethylene glycol is dissolved in 35 ml of water in a 150 ml glass beaker by stirring for at least 30 minutes. The liquid is then transferred to a 100 ml graduated cylinder and the volume is then adjusted to 100 ml by adding distilled water. The solution is then filter sterilized and stored in a securely capped bottle.

- Sonicated single –stranded carrier DNA (sssDNA)

sssDNA was used as a carrier which enhances the transformation efficiency, 2 mg of high-molecular weight DNA (deoxyribonucleic acid sodium salt, type 3 from salmon testes, Sigma)

per 1ml of sterile TE buffer was sonicated, aliquoted and stored at -20 °C. Prior to use, it should be boiled for 10 minutes and quick-cooled on ice.

- Solution 1: 1 M Lithium Acetate, 10X TE Buffer, dd H₂O, 1:1:8 ratios respectively.
- Solution 2: 1 M Lithium Acetate, 10X TE Buffer, 50 % PEG, 1:1:8 ratios respectively.

Method

5ml of an overnight culture were diluted at OD₆₀₀ to ~ 0.1 in 50 ml fresh medium and grown till OD₆₀₀ ~ 0.5. Cells were pelleted by centrifugation at 4000 rpm for 5 minutes, and then washed twice with ddH₂O water and resuspended with 300 µl of solution 1. 50 µl of these cells were added to new eppendorfs containing 3 µl of plasmid DNA mixed with 3 µl of the carrier sssDNA. 350 µl of Solution 2 was added to each tube, and then incubated at 30 °C for 45 minutes. Cells were then heat shocked at 42 °C for 15 minutes. Cells were then plated on the desired plate at a volume depending on the expected transformation efficiency.

9.5. *Arabidopsis* plant transformation

9.5.1. Preparation of *Agrobacterium* electrocompetent cells

Materials:

- Ice-cold sterile water
- Ice-cold sterile 10 % glycerol

Method:

A single colony of *Agrobacterium tumefaciens* AGL-1 was picked from the plate using a sterile toothpick to inoculate 5 ml of YEP medium supplemented with appropriated antibiotic (tetracycline 100 mg/ml). The culture was grown overnight at 28 °C with vigorous shaking.

The following day 200 ml of YEP/Tc were inoculated with 2 ml of the fresh overnight culture and incubated at 28 °C on the shaker until the optical density reached to OD₆₀₀ ~ 0.5-0.8 (~ 4 hours). The flask was chilled on ice for 20 minutes and the cells were pelleted at 6,000 x g and 4 °C for 15 minutes. The supernatant was removed and the pellet was gently resuspended in 200 ml of ice-cold sterile water. The cells were centrifuged again under the same conditions and

and the pellet was resuspended in 100 ml of ice-cold water. After one more centrifugation as described above, the pellet was resuspended in 2 ml of ice-cold sterile 10 % glycerol. The tube was inverted several times and the cell suspension aliquoted as 50 μ l cells, frozed in liquid nitrogen, and then stored at -80 °C.

9.5.2. *Agrobacterium* transformation by electroporation

The electroporation was performed using BIORAD gene pulser (Capacitance Extender II, Pulse controller II, Gene pulser II, BIORAD) set to the parameters 2,5 kv/cm, 25 μ F, 200 Ohms.

1 μ l of the recombinant plasmid was added to the vial containing 50 μ l agrobacterial electro-competent cells and mixed gently. After incubation on ice for 3 minutes, the mixture was transferred to the ice-chilled electroporation cuvette. The cuvette was subsequently placed in a chamber slide and the electric pulse was applied. The electroporated cells were recovered with 500 μ l YEP media and shaken for 1 hour at 28° C. 100 μ l transformed cells were plated on YEP agar plates supplemented with kanamycin to a final concentration of 50 μ l/ml and incubated at 28 °C for 48 hours.

9.5.3. *In planta* transformation

Materials

- YEP media
- 5 % Sucrose solution
- Silwet L-77

Transformation protocol

Healthy *Arabidopsis* plants were grown until they began to flowering. They were grown in soil in pots covered with cellophane sheet under short day period (hours). The *Agrobacterium* strain carrying the gene of interest on the binary vector was prepared in large scale culture at 30 °C in YEP medium containing kanamycin antibiotic to select for the binary plasmid. The culture was spun down, resuspended to OD₆₀₀ = 0.9 in 5 % sucrose solution. Silvet L-77 was added to a concentration of 0.05 % (500 μ l/L) and mixed well. Finally, the flowering parts of plant were dipped in an *Agrobacterium* solution for 4 to 5 minutes. The dipped plants were placed under

cover for 24 hours to maintain high humidity. The plants were watered normally until seeds became mature. The dry seeds were then harvested.

Herbicide stock solutions

- Basta (glufosinate ammonium)

The concentration of glufosinate ammonium ($C_5H_{15}N_2O_4P$; FW = 198.16) in commercial preparations varies from product to product. Final concentration of 300 μ M was used. The product is stored at room temperature.

Selection of transformed plants

The transgenic *Arabidopsis* plants carry the herbicide resistance marker transferred to the plant in the T-DNA fragment (fig.). Glufosinate ammonium (Basta) was used for selection of transformed plants to a final concentration of 200-300 μ M. After the 5th day of germination the plants were sprayed with Basta solution for three times with 2-3 days interval between spraying.

10. MOLECULAR MANIPULATION OF YEAST STRAINS

10.1. Diploid strain generation by mating

The mating was done by streaking both haploid strains Mat α and Mat **a** over each other on YPD media, and then the plates were placed in 30 °C incubator for two days, selection for mated cells was made using the suitable selective media.

10.2. URA3 selection marker recycling by Cre-lox recombination

The selection marker flanked by the *loxP* sites of the in frame promoter-integration cassette incorporated upstream of native yeast genes can be excised by using the plasmid pGAL-*cre*, thus leaving behind only the promoter-tag DNA. The same cassette can be reused to sequentially alter the whole biosynthetic pathway. Following induction on galactose, cells having undergone *Cre*-mediated recombination and loss of the URA3 marker are selected on medium containing 5-fluoroorotic acid (5-FOA).

Strains of interest are transformed with pGAL-*cre* and the transformants are selected on CM-U,L dropout medium. Transformants are then grown until saturation in 2 ml of CM-U,L dropout medium with 2% raffinose as carbon source and further on diluted 100-fold into the same dropout medium with 2 % galactose as carbon source, while as control, an aliquot is diluted in the same medium with 2 % glucose as carbon source and incubated over night at 30 °C. 10 µl of culture is then diluted and plated on to a 5-FOA plate, and single colonies are isolated after incubation for 2 days at 30 °C.

10.3. Chromosomal gene upregulation by promoter integrations

10.3.1. Genome insertion of a stabilized extra copy of HMG2 under the control of a Galactose promoter

The open reading frame of yeast HMG2 (YLR450) was amplified from yeast genomic DNA using the following primers: 5HMG2(K6R) 5'-ATGTCACTTCCCTTAAGAACGATAGTACATTTG-3' and 3HMG2(*Xho*I) 5'-CTCGAGTTATAATAATGCTGAGGTTTTACAGGGGGG-3'. The 5' primer was designed to contain a mismatch so as to convert in the amplified product the AAA codon into AGA altering K6 into an R. The amplified product was cloned by TOPO TA cloning into the pCRII vector and the insert was sequenced.

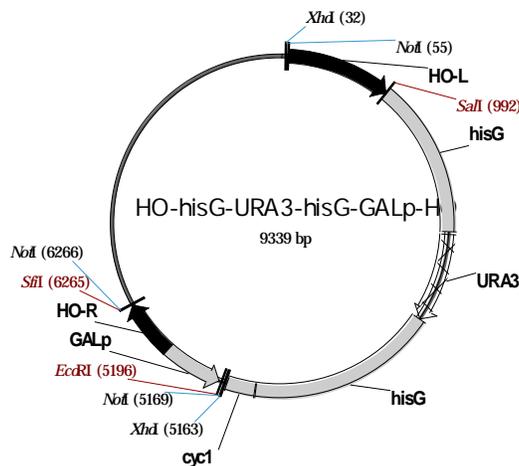


Figure II. 24. The HO-hisG-URA3-hisG-GALp-HA plasmid map.

The DNA region of the yeast pYES2 vector encompassing the GAL promoter-multicloning sites and the *cyc1* termination sequence was amplified using primers 5GAL(*Mfe*I)

5'-CAATTGTTTAAACGGATTAGAAGCCGCCGAGCG-3' and *CYC1*(*Bg*/III) 5'-AGATCTGCGCGCAAATTAAGCCTTCG and subsequently inserted into the M4366 vector (HO-hisG-URA3-hisG-HO) (see paragraph 3.11.) into the *Eco*RI-*Bg*/III sites creating a new plasmid named M4366-GAL (Fig. II. 24).

The plasmid M4366-GAL (HO-hisG-URA3-hisG-GALp-HO) was digested with *Eco*RI which linearized the vector at the multicloning site, was dephosphorylated using Shrimp Alkaline Phosphatase and was purified by gel extraction. The vector was subsequently ligated to an *Eco*RI insert containing the open reading frame of K6R HMG2. Integration of the gene in the orientation proximal to the GALp was verified using PCR primers from the promoter and the gene (Fig. II. 25).

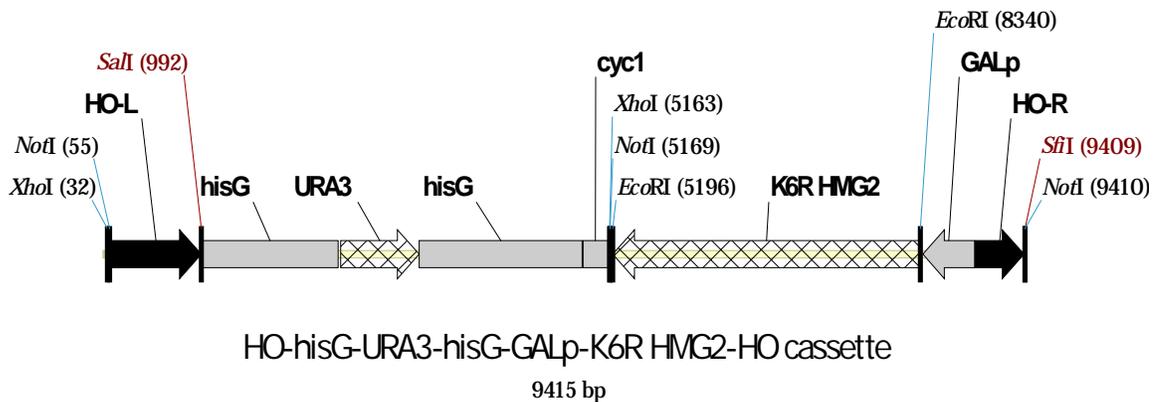


Figure II. 25. The HO-hisG-URA3-hisG-GALp-K6R HMG2-HO integration cassette.

The full construct was digested with *Not*I to release the cassette from the backbone plasmid and was used to transform yeast cells. Growing colonies on media lacking uracil were selected in media lacking uracil and the integration was confirmed.

AM63 yeast strain

The wild type EG60 strain (Mat α , *ura3*, *trp1*, *his3*) was used the basal strain for the development of a strain harboring an additional copy of a mutant (K6R) form of HMG2 under the control of the inducible galactose promoter {GALp-(K6R) HMG2}.

The integration cassette construct M4366-GAL/(K6R) HMG2 was digested with *Not*I and the fragments were used to transform EG60 cells. Transformed cells contain a stably integrated hisG-URA3-hisG-GALp-(K6R) HMG2-*cyc1* insert into the HO locus. The verified strain was named AM63 (Mat α , GALp-(K6R)HMG2::*URA3*, *trp1*, *his3*).

AM65 strain

Excision of the hisG-URA3-hisG fragment was accomplished by plating cells on FOA plates which counterselect the presence of URA3. The flanking hisG identical sequences enable the endogenous recombination and excision of the selection marker under pressure, generating the *ura3* sensitive version of AM 63. The verified strain was named AM65 (Mat α , GALp-(K6R) HMG2, *ura3*, *trp1*, *his3*).

AM66 strain

By successive backcrosses of AM63 (Mat alpha, Galp-(K6R) HMG2, URA3, *trp1*, *his3*) strain with EG61 (Mat **a**, *ura3*, *trp1*, *his3*) the KSY10 (Mat a, GALp-(K6R)HMG2, URA3, *trp1*, *his3*) strain has been generated by Kristine Stepanian in the frame of a Master Project. AM65 (Mat alpha, Galp-(K6R) HMG2, *ura3*, *trp1*, *his3*) strain containing pJG4-6/*SfCinS*(RC) was mated with KSY10 (Mat a, GALp-(K6R)HMG2, URA3, *trp1*,*his3*) strain. The positive colonies on Glu/CM-U-W were grown on Glu/CM-U three consecutive rounds for pJG4-6 plasmid curing and then serial dilutions were plated on 5-FOA media for URA3 removal. The verified strain was named AM66 (Mat α/a , GALp-(K6R)HMG2x2::*HO*, *ura3*, *trp1*, *his3*).

10.3.2. Upregulation of ERG20 gene with a GALpromoter

Using as a template the pCOD2 plasmid the loxP-URA-loxP-GALp-HA cassette was PCR amplified by using the following primers: ERG20-COD2-for 5'-CTCAACCAACAGGTA TTGGACTGACATAGGCACAATAAACTCAAAAATAAAGCTTCGTACGCTGCAGG-3' and ERG20-COD2-rev 5'-GGGAAAACGTTCAAGAATCTCTCTCTCCTAATTTCTTTTTCT GAAGCCATGGATCCAGAGGCATAATCT-3'. The PCR reaction was performed six times. The size of the PCR products of each reaction was test by resolving small aliquots (5 μ l) on 1 % agarose gel. The PCR products were collected and purified by ethanol:chlorophorm precipitation.

AM67

The purified integration cassette loxP-URA-loxP-GALp-HA was used to transform AM66 cells. Transformed cells contain a stably integrated loxP-URA-loxP-GALp-HA fragment upstream of ORF of ERG20 gene.

The expression of HA-ERG20 stable recombinant protein was verified by Western blott. The stable integration of loxP-URA-loxP-GALp-HA cassette was confirmed by PCR amplification of a 500 bp fragment using the primers 5'GALprom 5'-GACTACTAGCAGCTG

TAATACGACTCACTATAGGGAATATTAAGCTCATGT and 3' ERG20confREV 5'-TTGG AAGGTGACCTCATGGAACAATTCG-3'. The PCR product was resolved by electrophoresis on 1 % agarose gel. The verified strain was named AM67 (Mat α/a , GALp-(K6R)HMG2x2, GALp-HA-ERG20-URA3, *trp1*, *his3*).

AM68

Excision of the loxP-URA3-loxP fragment was accomplished by transforming cells with pGAL-cre plasmid. Transformants were endured the galactose induction to allow the lox-cre recombination and then were plated on FOA plates which counterselect the presence of URA3. The verified strain was named AM68 (Mat α/a , GALp-(K6R)HMG2x2, GALp-HA-ERG20, *ura3*, *trp1*, *his3*).

10.3.3. Truncation and upregulation of HMG1 gene with a ADH1 promoter

Using as a template the pCOD3 plasmid the loxP-URA-loxP-ADH1p-myc cassette was PCR amplified by using the following primers: HMG1t-COD3-for 5'- ACATAGTGTATCATTGTC TAATTGTTGATACAAAGTAGATAAATACATAAAAGCTTCGTACGCTGCAGG - 3' and HMG1t-COD3-reverse 5'- GGAGCAGTAAAAGACTTCTTGGTGACTTCAGTTT TCACCA ATTGGTCCATAAGCTTGGAGTTGATTGTATGC - 3'. The PCR reaction was performed six times. The size of the PCR products of each reaction was test by resolving small aliquots (5 μ l) on 1 % agarose gel. The PCR products were collected and purified by phenol:chloroform extraction and ethanol precipitation.

AM74 and AM75

The purified integration cassette loxP-URA-loxP-ADH1p-myc was used to transform AM68 and AM70 cells. Transformed cells contain a stably integrated loxP-URA-loxP-ADH1p-myc fragment in the HMG1 promoter, truncating the N-terminal region of the HMG1 protein. The stable integration of loxP-URA-loxP-ADH1p-myc cassette was confirmed by PCR amplification of a 500 bp fragment using the primers ADH1conf56 5' – AGGGGTATCTTCGA ACACACGAAA – 3' and HMG1conf-rev 5' – GCAACCGCTCTCGTAGTATCACCT – 3'.

The PCR product was resolved by electrophoresis on 1 % agarose gel. The verified strain was named AM74 (Mat α/a , GALp-(K6R)HMG2x2::HO, GALp-HA-ERG20, ADH1-

myc::HMG1, URA3, *trp1*, *his3*) and AM75 (Mat α/a , GALp-(K6R)HMG2x2::HO, GALp-HA-ERG20, ADH1-myc::HMG1, URA3, *trp1*, *his3*, $\Delta erg9::HIS5$) strains, respectively.

10.4. Gene disruption

10.4.1. Deletion of *erg6* gene

The disruption cassette of the pUG27 plasmid containing HIS5 gene as a selection marker was PCR amplified using the following primers: ERG6-pUG F: 5'-AAAAACAAGAA TAAAATAATAATATAGTAGGCAGCATAAGATGAGTGCAGCTGAAGCTTCGTACGC-3' and ERG6-pUG R: 5'-GGTATATATCGTGCGCTTTATTTGAATCTTATTGATCTAGTGAA TTTAGCATAGGCCACTAGTGGATCTG-3'.

The PCR reaction was performed six by two times. The size of the PCR products of each reaction was test by resolving small aliquots (5 μ l) by agarose gel electrophoresis. The PCR products were collected in two batches and purified by phenol:chloroform extraction and ethanol precipitation.

AM62

The purified disruption cassette ERG6-HIS5-ERG6 was used to transform EG60 cells. Transformants were grown on Glu/CM-H. Transformed cells contain a HIS5 disruption cassette stably integrated into the *erg6* locus, by this way inactivating the gene. The verified strain was named AM62 (Mat α , $\Delta erg6::HIS5$, *ura3*, *trp1*, *his3*).

AM64

By the same way as above the HIS5 disruption cassette was integrated into the AM63 cells generated the strain AM64 (Mat α , URA3-*hisG*-GALp-(K6R)HMG2::HO, *trp1*, *his3*, $\Delta erg6::HIS5$).

10.4.2. Deletion of one allele of *erg9* gene

The disruption cassette of the same plasmid, pUG27, containing HIS5 gene as a selection marker was PCR amplified using the following primers: EGR9-pUG F 5'-AGA GAAAAGACGAAGAGCAGAAGCGGAAAACGTATACACGTCACATATCACAGCTGACT

TCGTACGC-3' and ERG9-pUG R 5'-GTACTTAGTTATTGTTTCGGAGTTGTTTGTATTATGT
TATTTGGCGCAGACTGCATAGGCCACTAGTGGATCTG-3'

The PCR reaction was performed six times. The size of the PCR products of each reaction was test by resolving small aliquots (5 µl) on 1% agarose gel. The PCR products were collected and purified by phenol:chlorophorm extraction and ethanol precipitation.

AM69 and AM70

The purified disruption cassette ERG9-HIS5-ERG9 was used to transform AM67 and AM68 diploid cells. Transformants were grown on Glu/CM-H. Transformed cells contain a HIS5 disruption cassette stably integrated into the *erg9* locus on one allele of the two, by this way partial inactivating the gene.

The stable integration of distruption cassette was confirmed by PCR amplification of a 500 bp fragment using the primers ERG9prom 5'-CTAAACGAGCAGCGAGAA CACGACCAC-3' and pUG27confR 5'-GGATGTGATGTGAGAACTGTATCC-3'. The PCR product was resolved by electrophoresis on 1% agarose gel. The verified haploinsufficient strains was named AM69 (Mat α/a , GALp-(K6R)HMG2x2, GALp-ERG20::URA3, *trp1*, *his3*, $\Delta erg9$:: HIS5), derivative of AM67, and AM70 (Mat α/a , GALp-(K6R)HMG2x2, GALp-HA-ERG20, *ura3*, *trp1*, *his3*, $\Delta erg9$:: HIS5), derivative of AM68.

11. YEAST PROTEIN PROTOCOLS

11.1. Small scale of protein expression in yeast

11.1.1. Promoter induction

To initiate protein expression under inducible promoters, 5 ml of Glu/CM-selective amino acids medium were inoculated with a mixture of transformed selected yeast cells and were incubated overnight at 30 °C. Next morning 50 ml of Glu/CM-selective amino acids medium were inoculated with ~400 µl of the above mentioned overnight cultures to an optical density $OD_{600} = 0.1$, and incubated at 30 °C until the optical density OD_{600} reach 0.5-0.7 (5-6 hours). The cultures were spun down at 4000 rpm and room temperature for 5 minute and the pellets were washed twice with sterile ddH₂O. The washed pellets were resuspended in 50 ml of Gal-Raff

based selective growth medium and incubated at 30 °C ~12 hours. The induced cultures were spun down and the pellets washed twice with sterile ddH₂O.

11.1.2. Terpene synthase induction

To initiate terpene synthase activity, the induced pellets were resuspended in 20 ml of enzymatic reaction buffer: 10 mM MOPS, 2 mM MgCl₂, 0.2 mM MnCl₂, 2% Glycerol, 1 mM DTT and incubated at 30 °C from 12 hours to several days. If the incubation on buffer was prolonged more than 12 hours the accumulated volatiles were released every 24 hours.

11.2. Protein extraction

11.2.1. Protein extraction using glass beads

Pellets of cells overexpressing the desired protein already washed with ddH₂O after induction. Pellets were resuspended in 0.5 ml of yeast lysis buffer (protein extraction buffer) (120 mM NaCl; 0.5 % NP-40; 1 mM PMSF) and then transferred to 2 ml eppendorf tubes containing equal volume of glass beads. Cells were then vortexed (alternative 1 minute vortexing followed by 1 minute incubation on ice) 20 times. After being centrifuged (twice for better purity), supernatants were transferred to new tubes and stored at -20 °C.

11.2.2. Protein extraction by sonication

The cells were disrupted by sonication with a microprobe for minimal sonication time, using short bursts to prevent heating of the sample. The sonication was repeated until the solution changed color from very milky to slightly more translucent; it is important to make sure that the cell suspension is kept chilled packed with crushed ice during sonication. The chilled sonicate was then cleared by centrifugation at 13 000 x g, 20 min, 4 °C to precipitate insoluble proteins and cell debris, and soluble proteins were collected in the supernatant which is subsequently analysed for protein concentration, enzymatic activity and purification or stored with glycerol 100 % (v/v) at -20 °C.

11.3. Determination of protein concentration

Protein quantification is often necessary prior to handle protein samples for isolation and

characterization. It is a required step before submitting protein samples for chromatographic, electrophoretic and immunochemical separation or analyses.

11.3.1. Spectrophotometric determination of protein concentration

This method describes the measurement of the concentration of a protein in solution using absorbance spectroscopy. The absorbance, A , is a linear function of the molar concentration, c , according to the Beer-Lambert law: $A = \epsilon \times l \times c$, where ϵ is the molar absorption coefficient ($M^{-1} \text{ cm}^{-1}$) and l is the cell path length (cm). The absorbance of a protein is measured at 280 nm.

Materials

- Solution of pure protein in either distilled water or buffer
- Spectrophotometer with UV and visible lamps
- Quartz cuvette

The spectrophotometer and UV and visible lamps were allowed to a 30-min warm up period. The measurement was done for 260nm, 280 nm, and 330 nm. The reading at 260 nm and 330 nm serves to detect nucleic acid interference and light scattering, respectively.

The spectrophotometer was calibrated with a 200 μl of solvent blank (distilled water or buffer). The solvent was discarded and the cuvette was washed once and thoroughly dried. 10 μl of protein was mixed with 190 μl distilled water. 200 μl of protein solution was added in the cuvette and was replaced in the spectrophotometer. The absorbance of the samples was recorded at the appropriate wavelengths. The protein concentration (c) was calculated in mg/ml according to the following equation: $c = (A_{280} \times V) / (\epsilon_{280} \times v)$, where $\epsilon_{280} = 1,51$; $V =$ final volum, 200 μl ; $v =$ volum of protein, 10 μl .

11.3.2. Bradford assay

Several colorimetric methods have been described for quantifying proteins in solution, including the widely used Bradford. Bradford assay is a popular, simple, rapid, inexpensive, sensitive and reproducible assay for the quantification of proteins. It works by the action of coomassie brilliant blue G-250 dye, which binds specific basic amino acids (primarily arginine, lysine and histidine) exposed on the surface of the protein undergoing quantification. This results in a spectral shift from the reddish/brown form of the dye (absorbance maximum at 465 nm) to the blue form of the dye (absorbance maximum at 610 nm). The difference between the two

forms of the dye is greatest at 595 nm, so that is the optimal wavelength to measure the blue color from the coomassie dye-protein complex.

Materials and Solutions:

Bradford solution (BIO RAD),

Plastic cuvettes,

Spectrophotometer (Hewlett Packard 8452 A)

Method:

In a 1.5 ml Eppendorff tube, 790 μ l of distilled water were mixed with 200 μ l of Bradford solution, then 10 μ l of the protein solution were added, and the mixture was subsequently analyzed by spectrophotometry at OD₅₉₅ nm using the distilled water mixed with Bradford solution as a blank. Known BSA concentrations were used to generate a linear curve in order to calculate the unknown concentrations of our samples.

11.4. SDS-POLY ACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE)

SDS-PAGE (Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis) is the most widely used electrophoresis technique in biochemistry and molecular biology to characterize individual proteins in a complex sample or to examine multiple proteins within a single sample. PAGE can be used as a preparative tool to obtain a pure protein sample, or as an analytical tool to provide information on the mass, charge, size, shape, purity or presence of a protein.

Acrylamide is the material of choice for preparing electrophoretic gels to separate proteins by size. Acrylamide mixed with bisacrylamide forms a cross-linked polymer network when the polymerizing agent, ammonium persulfate is added. Ammonium persulfate produces free radicals faster in the presence of TEMED (N,N,N,N'-tetramethylethylenediamine). The size of the pores created in the gel is inversely related to the amount of acrylamide used. To ensure a flat surface and to exclude air the gel is overlaid with saturated isopropanol, which is washed off with water after the gel has set.

11.4.1. Materials and solutions

Devices:

- Mini slab gel apparatus (Bio-Rad or Hoefer)

- Power pack (Bio-Rad or Hoefer)

Protein samples solubilized in 2X SAB (Sample Application Buffer)

Tris HCl pH 8.8 : 1 M Tris HCl autoclaved under standard conditions and stored at room temperature.

Tris HCl pH 6.8 : 1 M Tris HCl autoclaved under standard conditions and stored at room temperature.

Acrylamide/bisacrylamide monomer solution, ready to use, solution of 40 % acrylamide
10 % Sodium dodecyl sulphate stock solution (SDS) in deionized water.

Tetra Methyl Ethylene Diamine (TEMED), ready to use solution

10 % (w/v) Ammonium persulfate stock solution (APS), stored at 4 °C for a maximum of 1 week.

Water saturated Butanol, the top layer of equal volumes of water and n-butanol, shaken vigorously and separated.

1X Tris –Glycine electrophoresis running buffer (pH 8.3):

250 mM Glycine (14.4 g/L)

0.1 w/v SDS (1 g/L)

25 mM Tris HCl (3 g/L)

The pH was adjusted to 8.3 and the buffer was stored at room temperature.

2X SAB (Sample Application Buffer):

2.5 ml 0.5 M Tris HCl, pH 6.8

2 ml Glycerol

0.4 g SDS

0.4 ml β -Mercaptoethanol

Few grains of bromophenol blue were added and the volume was brought to 10 ml with ddH₂O.

Coomassie Brilliant Bleu Staining Solution

0.25 g Coomassie brilliant bleu R 250

10 % Acetic acid

45 % Methanol

45 % water

The volume was brought to 100 ml with ddH₂O, then filter sterilized and stored at room

temperature.

Destaining Solution

45 % Methanol

45 % water

10 % acetic acid

Prestained Protein Molecular Weight Marker (Fermentas)

Obtained as a ready to use solution molecular markers β -galactosidase (116 kDa), BSA (66.2 kDa), Ovalbumin (45 kDa), lactate deshydrogenase (35 kDa), β -lactoglobulin (18.4 kDa), Lysozyme (14.4 kDa).

Resolving and Stacking gel mixture composition

Components	Resolving gel (12 %)	Resolving gel (15 %)	Stacking gel (4 %)
ddH ₂ O	3.04.ml	2.29 ml	7.54 ml
1 M Tris HCl, pH 8.8	3.75 ml	3.75 ml	-
1 M Tris HCl, pH 6.8	-	-	1.25 ml
40% acrylamide	3.00 ml	3.75 ml	1.00 ml
10% SDS	100.00 μ l	100.00 μ l	100.00 μ l
10% APS	100.00 μ l	100.00 μ l	100.00 μ l
TEMED	10.00 μ l	10.00 μ l	10.00 μ l
Total volume	10.00 ml	10.00 ml	10.00 ml

11.4.2. Method

The mini protein slab gel apparatus was assembled according to the manufacturer's instructions. The resolving gel mixture was prepared as described above and transferred to the glass plate sandwich using a Pasteur pipette up to approximately 2/3 of the glass plate sandwich total volume. Water saturated butanol was added immediately to remove oxygen from the surface and to make the surface of resolving gel horizontal. After polymerisation, the water saturated butanol was drained off and the gel surface rinsed with distilled water. Any remaining traces of water were removed with filter paper, taking care not to damage the gel surface. The stacking gel mixture was prepared in the same way as described in the material section and overlaid on the resolving gel up to the brim of the glass plate sandwich. Taking care not to trap any air bubbles, a 10-slot teflon comb was inserted in such a way that approximately 1-1.5 cm

were left between the top of the resolving gel and the bottom of the comb. The monomer was then left to polymerize. After the stacking gel had polymerized, the comb was removed carefully and the wells that had formed were washed with ddH₂O to remove any unpolymerized acrylamide solution. The polymerized gel slab-glass plate sandwich is first removed from the casting stand and transferred to the upper buffer chamber, and then the entire unit was introduced into the larger lower buffer chamber. 1xTris-glycine electrophoresis buffer was added to the lower (approximately 2/3) and then to the upper chambers to the brim of the sandwich, so that all wells are properly covered in buffer before the combs were carefully pulled out. Before loading the protein samples for SDS-PAGE analysis on the electrophoresis gel, they were first denatured by the use of a sample buffer containing 0.5% SDS with or without a reducing agent such as 3 mM DTT (or β -Mercaptoethanol) then boiled for 3-5 minutes, to denature the proteins by disrupting non-covalent intra- and intermolecular associations and by reducing disulfide linkages, thus overcoming some forms of tertiary protein folding, and breaking up quaternary protein structure (oligomeric subunits). The protein samples were then cooled to room temperature and spun for 2 min at 13000 rpm to avoid protein streaking during electrophoresis due to cell debris. To assess the relative molecular weight of a protein on a gel, protein molecular weight markers were run in the outer lanes of the gel for comparison. Using a Prot/Elec loading tip, the desired amounts of the protein samples or prestained SDS-PAGE molecular weight standard markers were loaded into each lane and the gel was run at a constant current (30 mA, 200 V) until the dye front reaches the end of the gel or the lowest band of the prestained marker is approximately 0.5 cm from the gel bottom. As a protein sample passes through a gel, the buffer front can be visualized using low molecular weight dyes that migrate with the buffer front. The most commonly-used tracking dye is bromophenol blue.

Once protein bands have been separated on a gel they can be visualized by staining with Coomassie Brilliant Blue Dye that turns the entire gel blue. To see the protein bands the gel must be destained with a methanol/acetic acid mixture to remove background stain for about 4 hours to overnight with 3 - 4 times changes of the destaining solution.

11.5. WESTERN BLOTTING OF RECOMBINANT PROTEINS

In a Western blot, proteins that are separated on polyacrylamide gels on the basis of size, are transferred to a membrane and then specifically detected in the immunoassay step using specific antibodies to the protein. The technique involves first the transfer of the proteins from the SDS-PAGE gel onto a membrane made of nitrocellulose or PVDF, then the blocking, where the transfer membranes are blocked with a concentrated protein solution (e.g. ovalbumin, BSA, hemoglobin, non-fat milk powder) to prevent further non-specific binding of proteins. The blocking step is followed by the incubation of the membrane in a diluted antibody solution, the washing of the membrane, the incubation in a diluted conjugated probe antibody or other detecting reagent, further washing, and finally the detection.

11.5.1. Materials

- Tank Electrophoretic Transfer system
- Power Pack 300 Power Supply (BIO-RAD, cat165-5051)
- PVDF Protein Transfer (NEN TM, Life Science Products, cat NEF1002)
- Cellophane Nylon Wrap
- Hypercassette™18x24cm (Ameshram LIFESCIENCES, Batch240195, RPN1642)
- Fluorescence Sticker
- Scientific Imaging Film (Kodak, cat1651454)

11.5.2. Solutions

Methanol

100 % Methanol (Absolute Methanol)

Transfer Buffer

48 mM (5.82 g) Tris-HCl

39 mM (2.93 g) Glycine

The pH was adjusted to 8.0, then 200 ml methanol was added up to final volume 1liter.

1x Tris Buffered Saline (1 X TBS)

10 mM Tris-HCl

150 mM NaCl

The pH was adjusted to 7.5

Tris Buffered Saline/Tween 20 (TTBS)

0.1 % (v/v) Tween 20 in 1X TBS.

Blocking solution

5 % powder milk in 1X TBS

Primary antibody stock

The primary antibodies were prepared in 1X TBS solution containing 3 % milk according to the dilution factor. The antibody was then stored at -20 °C.

Anti HA epitope was prepared to a concentration of 1:1000 dilution.

Secondary antibody stock

The secondary antibodies were prepared in 1X TBS solution containing 3 % milk according to the dilution factor. The antibody was then stored at -20 °C.

HRP-anti mouse antibody was prepared to a concentration of 1:8000 dilution.

Western blotting detection reagent (SuperSignal Western Pico Chemiluminescent substrate, PIERCE 34080)

The two solution ready to use, were mixed in equal volumes (1 ml each).

Developer and Replenisher (Kodak GBX, cat1901859)

To 103 ml of the stock solution, ddH₂O was added to a final volume of 473 ml and kept in a dark place.

Fixer and Replenisher (Kodak GBX, cat1901859)

To 103 ml of the stock solution, ddH₂O was added to a final volume of 473 ml and kept in a dark place.

11.5.3. Method

Western blotting involves transferring of electrophoretically separated components from a gel to a solid support, such as a nitrocellulose or nylon filter, and probing with reagents that are specific for particular sequences of amino acids. The probes are usually antibodies that bind to antigenic epitopes displayed by the target protein attached on the solid support. Western blotting is extremely useful for the identification and quantitation of specific proteins in a complex mixture of proteins. A 1-5 ng of an average sized protein can be detected by western blotting.

The gel cassette was opened and lifted gently and the gel was cut, stacking gel was thrown away, resolving gel was immersed in transfer buffer for 15 minutes, thus allowing it to equilibrate. A piece of the PVDF membrane was cut to the dimension of the gel. One corner of the membrane was notched for later correspondence with the corner of the gel. The membrane was then wetted in 100 % methanol for 15 seconds to prevent gel swelling with heating, and to keep proteins adsorbed to the membrane, and then soaked in distilled water for 2 minutes and finally equilibrated into transfer buffer for at least 5 minutes. The 3 MM paper, the PVDF membrane and the gel were assembled in the cassette in the following order starting from the black part: foam pad, paper, gel, PVDF membrane, paper, foam pad. Care should be taken to exclude bubbles between gel and nitrocellulose, and between nitrocellulose and paper. The cassette was closed and placed in the tank blotting apparatus so that the side of the cassette holder with the gel was facing the cathode. Transfer buffer was then added to the blotting apparatus until the cassette was totally covered. The tank was then connected to the power output and the system was turned on for 3 hours at 24 volts. The tank was soaked in ice to prevent overheating of the buffer during the transfer process. After removing the foam pad and filter papers, the membrane was blocked in 10 % milk to allow the saturation of all non-specific protein binding sites on the blots, for 2 hours till overnight incubation and then washed three times for 10 minutes each in 1X TTBS. The membrane was then transferred to a plastic bag containing the primary antibody, and it was incubated for 3 hours slowly shaking at room temperature. The membrane was washed with 1X TBS once for 15 minutes and then two times for 15 minutes each in 1X TTBS, then once with 1X TBS for 5 minutes. The membrane was transferred to a new plastic bag, and the HRP-labelled second antibody was added and incubated with the membrane for 1-2 hours. The membrane was washed again with 1 X TBS once for 15 minutes and then two times for 15 minutes each in 1X TTBS then once with 1X TBS for 5 minutes. Chemiluminescence reagent was prepared by mixing an equal volume of solution from bottle 1 and bottle 2. The membrane was then shaken with at least 0.125 ml of chemiluminescence reagent per cm² for 5 minutes. Excess reagent was removed by draining and the blot was enclosed in a plastic wrap, placed protein side up in the film cassette and exposed to an autoradiography film (Fuji) for 1 minute till overnight exposure. The film was then processed for 1 minute in developing solution, 2 minutes in fixing solution after a brief wash with water. The size of protein of interest was identified by comparison with the molecular markers.

12. TERPENE SYNTHASES *IN VITRO* ASSAY

12.1. Materials and Solutions

Stock Solution for mixture buffer:

100 mM MOPS (pH 7.0 adjusted with KOH),

200 mM MgCl₂,

2 mM MnCl₂,

Glycerol 10% (w/v),

DTT 10 mM,

Distilled water,

Farnesyl diphosphate (FPP) substrate (Sigma, F6892),

Geranyl diphosphate (GPP) substrate (Sigma, G6772).

Reaction Substrate preparation:

Na₂CO₃ (50 mM) and NaHCO₃ (50 mM)

These two solutions were mixed in equal volume and their pH adjusted to 10-12, then 1 ml of this mixture was added to 1 vial containing 1 mg of GPP or FPP. The mixture should be mixed well and aliquoted into eppendorffs and kept at -20 °C.

12.2. Method

For functional expression analysis, the resulting soluble enzyme preparations (for GC or SPME analysis) or whole cells (for SPME) were assayed for activity either by using geranyl diphosphate (GPP) (Sigma, G6772) at 27 μM final concentration to check the monoterpenes production or by using farnesyl diphosphate (FPP) (Sigma, F6892) at 46 μM final concentration to check the sesquiterpenes production, as substrate. The assay mixture (1 ml) in screw septed glass vials was overlaid with 1 ml of organic solvent, typically pentane or ethyl acetate, according to the polarity of the volatiles, to trap volatile products. After incubation at room temperature for 2-3 h to over night to obtain sufficient product, vigorous vortexing was then applied to partition hydrocarbon product into the organic phase. This mixing step also quenched the reaction by denaturing the enzyme, which was often visible as a white precipitate at the organic/aqueous layer interface. With ethyl acetate as solvent, the aqueous and organic phases separate within minutes of vortexing, whereas hexane and pentane extractions produce a thick precipitate that requires brief centrifugation to separate; when the reaction mixture was extracted,

the combined extract was passed through a 1.5 ml column of anhydrous Na₂SO₄ to remove any traces of water and silica gel (or cotton) to trap and fractionate hydroxylated products. The column was then washed with a further 0.5 ml of pentane to a final volume of 1 ml. The combined pentane eluant contains the terpenoid hydrocarbons, which do not bind to silicic acid under these conditions. These non hydroxylated species were then concentrated slowly in a Gyro Vap apparatus prior being put into the insert in the screw septed glass vials to be analyzed by GC or stored at 4 °C if the analysis was not immediate.

Capillary GC-FID analysis or Head Space – Solid Phase Microextraction (HS-SPME) were performed as described below. Products were quantitated by integration of the detector signal from 4.5 to 15 minutes for monoterpene which, under these parameters, spans the range of retention times of α -pinene (4.98 min) through geraniol (>15min), and from 15 minutes to 30 minutes for sesquiterpene. All other monoterpene products reported have retention times within 4.5 and 15 minutes range, while sesquiterpene products reported have retention times within 15 and 30 minutes range. Products were tentatively identified by comparing retention times with those of authentic standards, and identities were confirmed by GC–MS analysis.

13. HEAD SPACE – SOLID PHASE MICROEXTRACTION (HS-SPME)

SPME was first developed in 1989 at the University of Waterloo (Ontario, Canada) by Pawliszyn and his group and has been marketed since 1993 by Supelco.

Solid Phase Micro Extraction is a simple, effective adsorption/ desorption technique, solventless alternative to conventional sample extraction techniques. It allows fast, high-throughput, and semi quantitative headspace measurements, and it eliminates the need for solvents or complicated apparatus for concentrating volatile or nonvolatile compounds in liquid samples or headspace. SPME is compatible with analyte separation/detection by gas chromatography or HPLC, and provides linear results for wide concentrations of analytes. By controlling the polarity and thickness of the coating on the fiber, maintaining consistent sampling time, and adjusting several other extraction parameters.

13.1. Devices

The SPME unit (Fig. II. 26) consists of two elements: a length of coated fused silica fiber bonded to a stainless steel plunger, and a holder. The choice of the fiber is based on the selectivity for certain target analytes and their volatility range e.g. non polar coating like (Poly Dimethyl Siloxane) retains hydrocarbons, while for the polar compounds, a polar fiber coating like Poly Achrylate and Carbowax is used. The SPME apparatus and fibers used (7 mm bonded polydimethylsiloxane) were obtained from Supelco (Bellefonte, PA). An SPME kit includes the sample holder and a package of fiber assemblies (one fiber each of 85 mm polyacrylate coating, and 100 mm and 7 mm bonded polydimethylsiloxane coating).

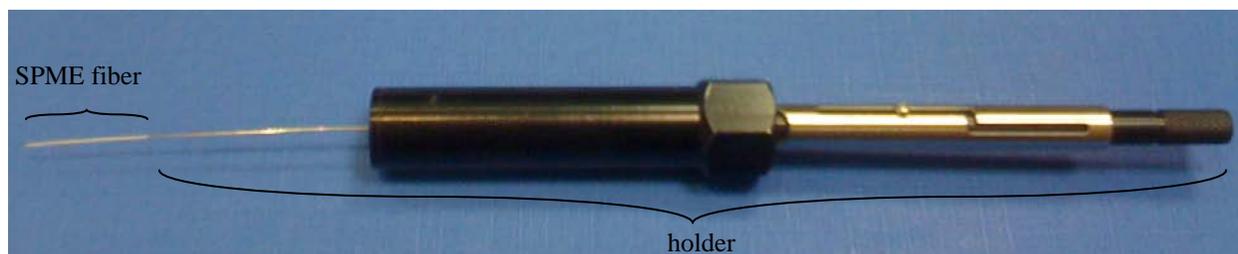


Figure II. 26. The SPME fiber-holder.

13.2. Method

The fiber is introduced into the sample or headspace (Fig. II. 27), and organic analytes adsorb in the phase and establish equilibrium usually after 2 to 30 minutes. Sampling was performed by inserting the syringe needle of the SPME assembly through the septum cap into the headspace above the sample. Yeast cultures were sampled by exposing the SPME fiber for 30 minute to 2 hours above the whole 20 ml of cultures induced for promoter and terpene synthase activation. Volatiles were sorbed by extending the fiber into the headspace. After adequate sorption time, the fiber was withdrawn into the outer septum-piercing needle, removed from the vial, and subsequently the analyte were desorbed from the fiber to a capillary Gas Chromatograph column in the heated chromatograph injection port, where they are focused on the inlet of the capillary column. The products were identified by comparing retention times and mass spectra with authentic reference compounds.

Figure II. 27. The introduction of the SPME fiber into the head-space of: (A) SPME vial; (B) 250 ml flask.



14. GAS CHROMATOGRAPHY – MASS SPECTROMETRY (GC-MS)

GC analysis is a common confirmation test; it separates all of the volatile components in a sample and provides a representative spectral output. Once the sample is injected into the injection port of the GC device, the GC instrument vaporizes the sample and then separates and analyzes the various components. Each component ideally produces a specific spectral peak that may be recorded on a paper chart or electronically. The time elapsed between injection and elution is called the "retention time". This time can help to differentiate between some compounds. The size of the peaks is proportional to the quantity of the corresponding substances in the specimen analyzed.

Given the volatility of the terpene hydrocarbon products, gas chromatography has proven to be particularly useful. The principal application of this technique has been product detection and identification, which is a fundamental aspect of the characterization of newly cloned or mutagenized cyclases. The GC apparatus used is a Hewlett Packard 5890 II gas chromatograph equipped with flame ionization detector. For the analysis 2 μ l of sample was injected split-type (splitless) with a HP 7673 auto sampler to a silica capillary column HB5 (30 m long, with 0.25

mm in diameter) and 0.25 μm film thickness packed with beads composed of silica gel.

This type of very long column tends to have a higher resolution, with better separation of gases than does a conventional column. Temperature of injector and detector were 230 $^{\circ}\text{C}$ and 270 $^{\circ}\text{C}$, respectively; oven temperature was programmed initially at 60 $^{\circ}\text{C}$ for 3 min and then increased with a rate of 3 $^{\circ}\text{C}/\text{min}$ with a final isotherm at 230 $^{\circ}\text{C}$ for 20 min. The carrier gas used for the analysis was Helium (He) at constant pressure of 127 kPa, at 37.1 cm/s velocity. As the carrier gas passes through the column, compounds with more polarity tend to bind with greater avidity to the column coating, thereby slowing them down more than less polar compounds.

For further assessment the analytes were thermally desorbed at 230 $^{\circ}\text{C}$ from SPME fiber into the injector of a QP2010 Shimadzu gas chromatograph equipped with QP2010 mass selective detector, and a ZB5 (0,25 mm x 30 m, 0,25 μm film thickness) column, in the splitless mode. A further refinement present in the GC/MS is that the coated capillary column is placed in an oven whose temperature is slowly raised with a rate of 3 $^{\circ}\text{C}/\text{min}$ from 60 $^{\circ}\text{C}$ up to 240 $^{\circ}\text{C}$ and maintained at this temperature for 5 min to equilibrate. This slow rise in temperature helps to elute compounds with higher boiling points that may not be in gaseous phase in the beginning of the run. Once out of the column the FID will detect the compound. The products were identified by comparing retention times and mass spectra with authentic reference compounds.

15. QUANTIFICATION OF TERPENES BY SPME IN GAS- CHROMATOGRAPHY

15.1. Cineole quantification

The amount of cineole produced in selected yeast strain was calculated based on the graph resulted from the amount of cineole existing in dilutions in pentane phase of 99 % 1,8-cineol standard (Aldrich, C8,060-1) (density 0,921) and the peak areas corresponding to each dilution when exposed for 30 min to the SPME and measured by GC or GC-MS. The protocol of correlation between the mass/volume of cineole and the peak areas of the cineol standard dilutions is presented in this subtitle.

1 %	10 µl of 99 % cineole-----990 µl pentane
10 ⁻¹ %	100 µl of 1 % cineole-----900 µl pentane
10 ⁻² %	100 µl of 10 ⁻¹ % cineole-----900 µl pentane
10 ⁻³ %	100 µl of 10 ⁻² % cineole-----900 µl pentane
10 ⁻⁴ %	100 µl of 10 ⁻³ % cineole-----900 µl pentane
10 ⁻⁵ %	100 µl of 10 ⁻⁴ % cineole-----900 µl pentane

The final volume of each dilution was 1 ml.

100 ml-----0,921 g

99 ml-----x

$$x = 99 * 0,921 / 100 = 0,91179 \text{ g cineole } 99 \%$$

1 % dilution -----10/1000 ml x 0,91179 g = 9,1179 mg/ml

10⁻¹ % dilution-----100/1000 ml x 9,1179 mg = 0,91179 mg/ml = 911,79 µg/ml

10⁻² % dilution-----100/1000 ml x 911,79 µg = 91,179 µg/ml

10⁻³ % dilution-----100/1000 ml x 91,179 µg = 9,1179 µg/ml

10⁻⁴ % dilution-----100/1000 ml x 9,1179 µg = 0,91179 µg/ml

10⁻⁵ % dilution-----100/1000 ml x 0,91179 µg = 0,091179 µg/ml

The cineole quantification by SPME was determined for the above dilutions with different fibers used during this work (Fig. II. 28):

% of cineole	µg/ml cineole	Area
0.0001	0.91179	2495
0.001	9.1179	10888
0.01	91.179	91096
0.1	911.79	942105

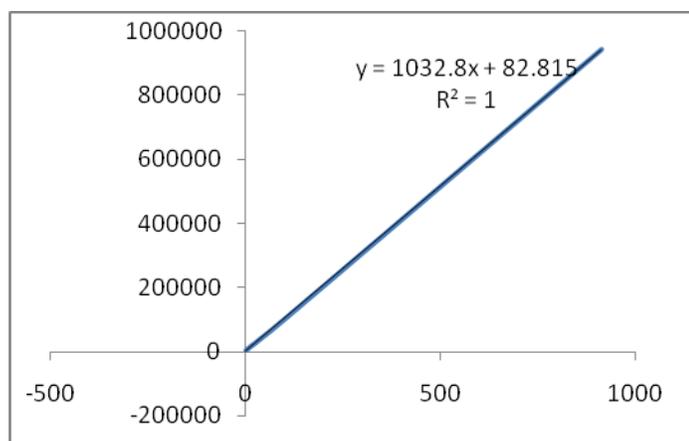


Figure II. 28. The graph of cineole quantification.

The equation of the chart is: $y = 1032.8x + 82815$

We considered that equation as $y = 1032.8x$.

We calculate the yield of cineole accumulate into 50 ml of yeast culture in 24 h by trapping the released volatiles on the SPME fiber and quantitative detecting them by GC based of the above equation and the area peak value obtained.

15.2. Trans-caryophyllene quantification

The amount of (-)-trans-caryophyllene produced in selected yeast strain was calculated based on the graph resulted from the amount of (-)-trans-caryophyllene existing in dilutions in pentane phase of (-)-trans-caryophyllene standard (Sigma, C9653-5) (density 0,902) and the peak areas corresponding to each dilution when exposed for 30 min to the SPME and measured by GC or GC-MS. The protocol of correlation between the mass/volume of cineole and the peak areas of the cineol standard dilutions is presented in this subtitle.

1 %	10 µl of 100 % (-)-trans-caryophyllene -----	990 µl pentane
10 ⁻¹ %	100 µl of 1 % (-)-trans-caryophyllene -----	900 µl pentane
10 ⁻² %	100 µl of 10 ⁻¹ % (-)-trans-caryophyllene -----	900 µl pentane
10 ⁻³ %	100 µl of 10 ⁻² % (-)-trans-caryophyllene -----	900 µl pentane
10 ⁻⁴ %	100 µl of 10 ⁻³ % (-)-trans-caryophyllene -----	900 µl pentane
10 ⁻⁵ %	100 µl of 10 ⁻⁴ % (-)-trans-caryophyllene -----	900 µl pentane

The final volume of each dilution was 1 ml.

100 ml-----0,902 g

1 % dilution -----10/1000 ml x 0,902 g = 9,02 mg/ml

10⁻¹ % dilution-----100/1000 ml x 9,02 mg = 0,902 mg/ml = 902 µg/ml

10⁻² % dilution-----100/1000 ml x 902 µg = 90,2 µg/ml

10⁻³ % dilution-----100/1000 ml x 90,2 µg = 9,02 µg/ml

10⁻⁴ % dilution-----100/1000 ml x 9,02 µg = 0,902 µg/ml

10⁻⁵ % dilution-----100/1000 ml x 0,902 µg = 0,0902 µg/ml

The (-)-trans-caryophyllene quantification by SPME was determined for the above dilutions with different fibers used during this work (Fig. II. 29):

% of caryophyllene	mg/L caryophyllene	area
0.000001	0.00902	2140
0.00001	0.0902	21578
0.0001	0.902	45428
0.001	9.02	286601

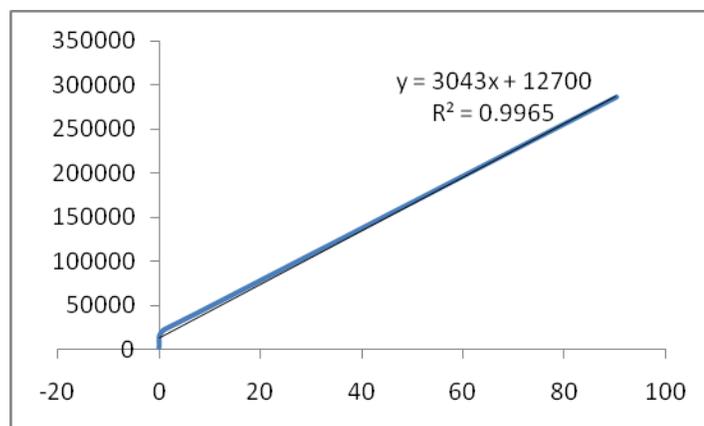


Figure II. 29. The graph of trans-caryophyllene quantification.

The equation of the chart is: $y=30430x + 12700$

We considered that equation as $y=30430x$.

We calculate the yield of trans-caryophyllene accumulate into 50 ml of yeast culture in 24 h by trapping the released volatiles on the SPME fiber and quantitative detecting them by GC based of the above equation and the area peak value obtained.

16. YEAST TWO HYBRID SYSTEM

The yeast two-hybrid system is a useful approach to detect novel interacting proteins (Fields & Sternglanz, 1994; Colas & Brent, 1998). The original system described by Fields & Song (1989) was based on the yeast GAL4 transcription factor, and is now known as the GAL4 system. A similar system was developed that utilized the BD of the bacterial repressor protein LexA in combination with the Escherichia coli B42 AD (the LexA or interaction trap system; Gyuris et al., 1993). The LexA system relies on a transcriptional readout for the detection of protein–protein interactions, through the reconstitution of a functional transcription factor. The repression assay, in the LexA system has been developed to determine whether the bait localizes to the nucleus and binds the LexA-operator sequences. In contrast to in vitro studies of protein – protein interactions, the yeast two-hybrid system offers a sensitive in vivo assay to study protein interactions, and can be used routinely to test for interactions between known proteins or to identify novel interaction partners through prey library screening approaches.

16.1. Plasmids used during yeast two hybrid approaches

16.1.1. LexA fusion plasmids

The plasmid pEG202 and/or pGilda are used to express the probe or “bait” protein as a fusion to the heterologous DNA-binding protein LexA on the N-terminal end of the bait protein. The major requirements for the bait protein are that it should not be actively excluded from the yeast nucleus, and it should not possess an intrinsic ability to strongly activate transcription. The plasmids expressing the LexA-fused bait protein are used to transform yeast possessing a dual reporter system responsive to transcriptional activation through the *LexA* operator.

The pYES-LexA fusion plasmid was specially engineered in the context of this project to express the LexA on the C-terminal end of the bait protein.

16.1.1.1. pEG202

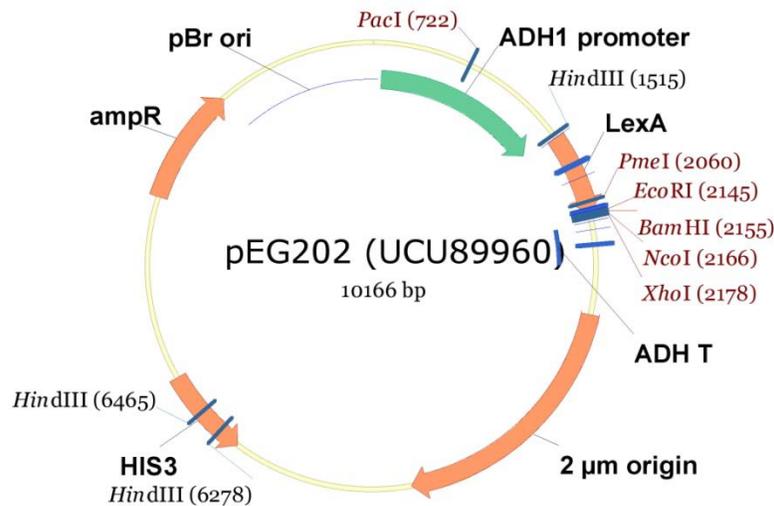


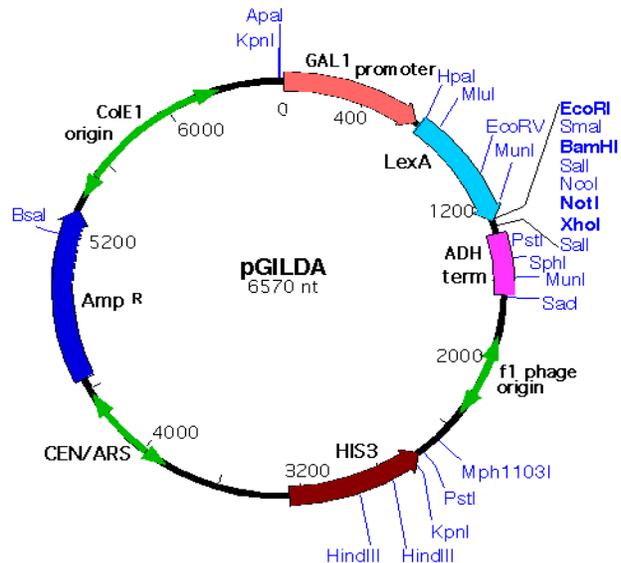
Figure II. 30. The pEG202 plasmid map.

The original parent plasmid for generating LexA fusion, pEG202 (Fig. II. 30) is a derivative of 202+PL that contains an expanded polylinker region and the reading frame relative to LexA (5'-CGT CAG CAG AGC TTC ACC ATT G-3').

16.1.1.2. pGILDA

pGILDA created by David Shaywitz, is an expression vector designed for use in the LexA-based yeast two-hybrid system. It is primarily used when a particular bait is potentially toxic to yeast and can be used in place of the normal bait expression vector pEG202. The GAL1 promoter allows expression of the LexA-bait fusion protein only when the yeast cells are grown in galactose medium (with no glucose present). Also, the plasmid contains a CEN (centromere) sequence that maintains the plasmid at only one copy per cell. The LexA-MCS (multiple cloning site)-ADH terminator is exactly the same as that found in pEG202 (see MCS sequence above). The ARS sequence allows for replication of the plasmid in yeast, while the HIS3 gene is used for selection in yeast. The Amp^R gene is used for selection in *E.coli*, while colE1 ori allows the plasmid to replicate in *E.coli* (Fig. II. 31).

Figure II. 31. The pGILDA vector map.



pGILDA/SfCinS(1094)

Plasmid DNA pTOPO-SfCinS(1094) and the vector plasmid pGILDA were digested using *EcoRI* and *XhoI* and resolved by agarose gel electrophoresis. The resolved insert SfCin(1094) and linearized vector pGILDA were gel extracted and purified.

A ligation reaction and a control reaction were performed between the pGILDA linearized plasmid and the SfCinS(1094) fragment respectively. 10 μ l of ligation reactions was used to transform Mach-1 bacterial cells. Several growing colonies from each transformation

were used to extract plasmid DNA. To verify the presence of GPPS the constructs were digested with *EcoRI* and *XhoI*.

16.1.1.3. pYES-LexA

Using the plasmid pEG202, the DNA fragment encompassing the LexA sequence was PCR amplified using the primers 5'LexA(*XhoI*) 5'-CTCGAGATGAAAGCGTTAACGGCC-3' and 3'LexA(*XbaI*) 5'-TCTAGATTACAGCCAGTCGCCGTTGCG-3'. The amplified product was cloned into the pCRII vector in a TOPO TA reaction. Plasmid DNA was extracted from several grown colonies and digested with *XhoI* and *XbaI* and resolved by agarose gel electrophoresis. The resolved insert was gel extracted and purified.

The pYES2 vector was digested with *XhoI* and *XbaI*, resolved in an agarose gel and purified by gel extraction using the Qiagen gel extraction kit. A ligation reaction was performed between the pYES2 and the LexA fragment. Small aliquots were used to transform Mach-1 bacterial cells. Several growing colonies from each transformation were used to extract plasmid DNA. To verify the presence of LexA fragment the construct was digested with *XhoI* and *XbaI* (Fig. II.32.).

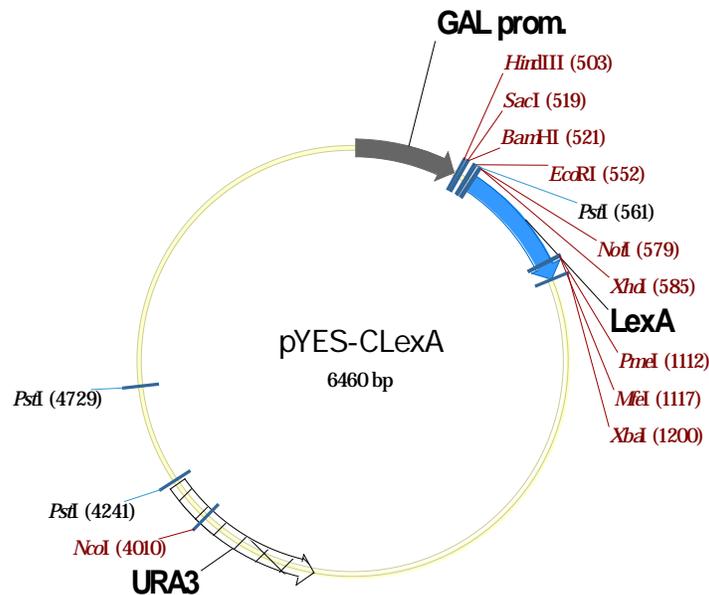


Figure II. 32. The pYES-CLexA plasmid map.

pYES-SfCinS(RC)-LexA

Using the plasmid pTOPO- *SfCinS*(RC) as a template, the *SfCinS*(RC) gene encoding for RC truncated cineole synthase was PCR amplified without stop codon in order to express the LexA protein on the C-terminal end of the bait protein, using the primers 5'CS-C(*EcoRI*-*NdeI*) 5'-GAATTCCATATGTCACTACAAACGGGTAATGAGATC-3' and 3'CS(*XhoI*) no stop 5'-CTCGAGCTCATAGCGGTGGAACAGCAAGTC-3'. The amplified PCR product was purified by gel extraction and cloned into the pCRII vector in a TOPO TA reaction. Plasmid DNA was extracted from few colonies and the presence of the *SfCinS*(RC) insert was verified by restriction digest by *EcoRI* and *XhoI*. The resolved insert was gel extracted and purified.

The pYES-LexA vector was digested with *EcoRI* and *XhoI*, resolved in an agarose gel and purified by gel extraction using the Qiagen gel extraction kit. A ligation reaction was performed between the pYES-LexA plasmid and the *SfCinS*(RC) fragment. Small aliquots were used to transform Mach-1 bacterial cells. Several growing colonies from each transformation were used to extract plasmid DNA. To verify the presence of *SfCinS*(RC) fragment the construct was digested with *EcoRI* and *XhoI* and resolved by agarose gel electrophoresis (Fig. II. 33).

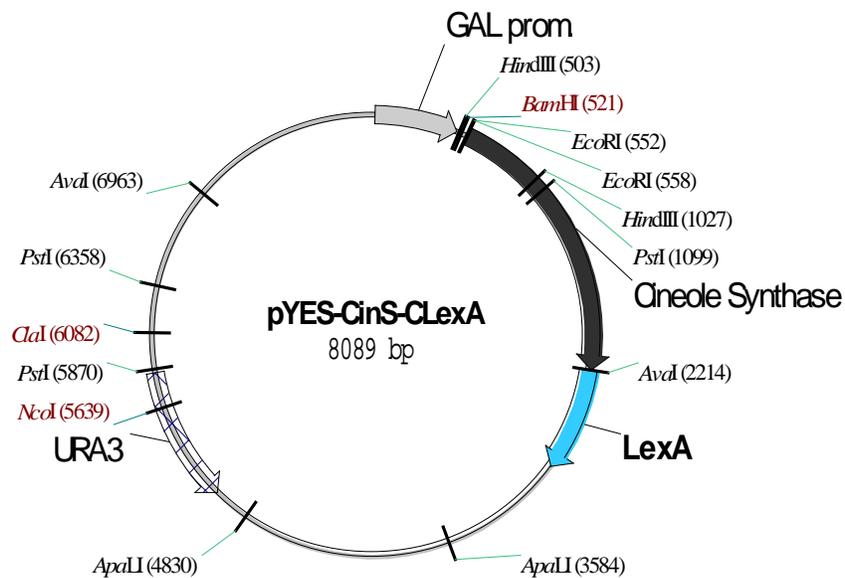


Figure II. 33. The pYES-CinS-CLexA plasmid map.

16.1.2. Activation domain fusion plasmid

16.1.2.1. pJG4-5

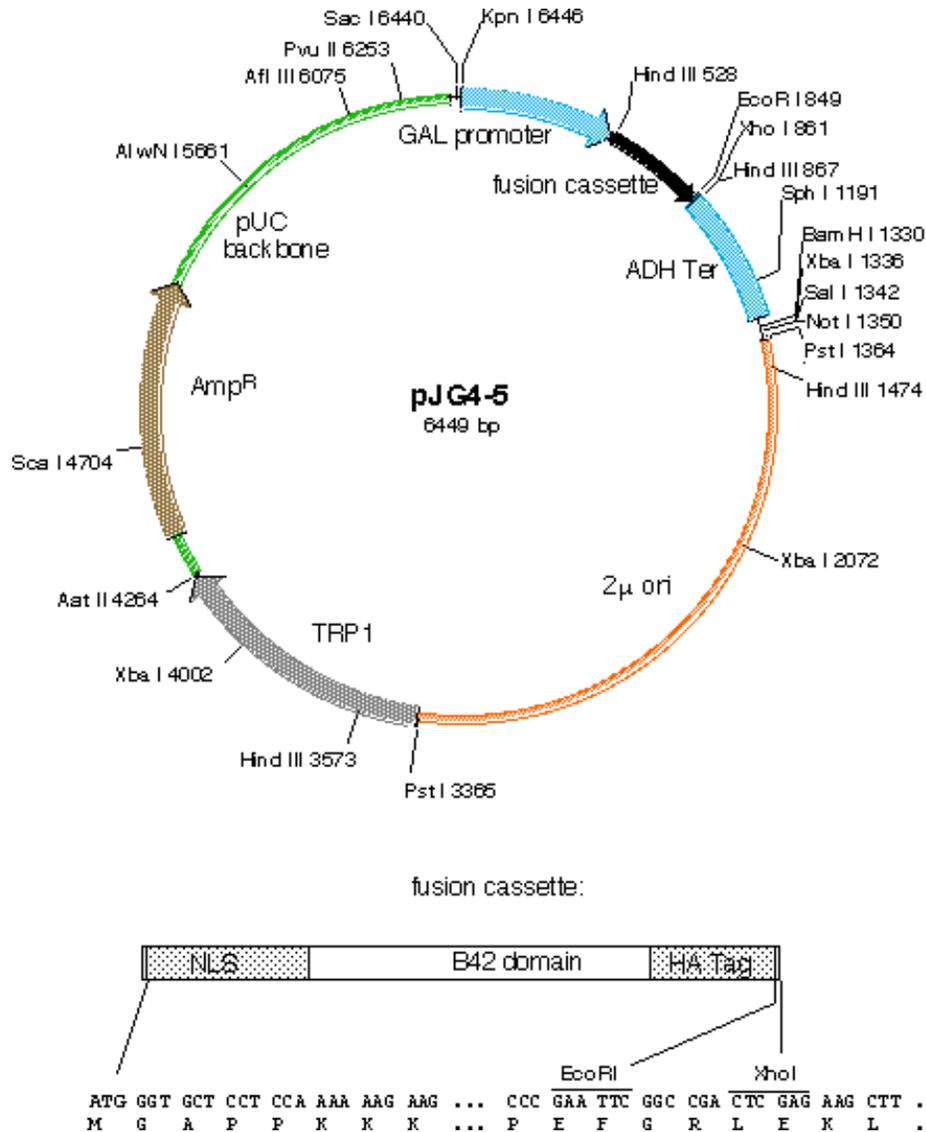


Figure II. 34. The pJG4-5 vector map.

The plasmid is generally used in yeast two-hybrid screens to identify protein-protein interactions. Expression is driven from the GAL1 galactose-inducible promoter. The expression cassette consists of an ATG to start translation, a SV40 nuclear localization signal to extend the interaction trap's range to include proteins that are normally predominantly localized in the

cytoplasm, an acid blob B42 activation domain, the hemagglutinin epitope tag to permit rapid assessment of the size of encoded proteins, *EcoRI-XhoI* sites designed to receive directionally synthesized cDNAs, and the alcohol dehydrogenase (ADH) termination sequence to enhance the production of high level of library protein. The plasmid also contains the TRP1 auxotrophy marker and 2 μ m origin for propagation in yeast, and the ampicillin resistance (Amp^R) gene and the pUC origin to allow propagation in *E. Coli*. The pJG4-5 cDNA library expression cassette is under control of the *GAL1* promoter, so library proteins are expressed in the presence of galactose but not glucose. This conditional expression has a number of advantages, the most important of which is that many false-positives obtained in screens can be easily eliminated because they do not demonstrate a Gal-dependence phenotype (Fig. II. 34).

16.1.2.2. Yeast strain

EGY48 is a basic strain used to select for interacting clones from a cDNA library.

16.1.2.3. Screening libraries

cDNA Tomato library cloned into the pJG4-5 vector

cDNA Salvia library cloned into the pJG4-5 vector

16.2. Bait protein characterization

16.2.1. Leu requirement test

Actual selection in the interactor hunt is based on the ability of the bait protein and acid-fusion pair, but not the bait protein alone, to activate transcription of the *LexA* operator-LEU2 gene and allow growth on medium lacking Leu. The EGY48 yeast cells expressing the LexA-bait fusion under the GAL promoter were plated on glucose and galactose medium lacking not only the selective amino acids but also Leu. The absence of colonies on these media shows that the bait protein is free of self activation of LEU gene. This test reveals whether the bait protein is likely to have an unworkably high background.

16.2.2. Immunoblot analysis

The expression of the bait protein was analysed by SDA-PAGE and Western Blott as described at subchapter 15 and 16 respectively.

16.3. Interactor hunt

16.3.1. Library transformation

EGY48 cells containing pGILDA-*SfCinS*(1094) and pYES-*SfCinS*(RC)-LexA transformed with the Tomato library clones and *Salvia* library clones, respectively, into the pJG4-5 plasmid. The transformation protocol is a version of the lithium acetate transformation protocol described by Schiestl and Gietz (1989) and Gietz et al., (1992) that maximizes the efficiency in *Saccharomyces cerevisiae* and produce up to 10^5 colonies/ μ g plasmid DNA. The yeast transformation was set at large scale. Over night culture of EGY48 transformants carrying the gene of interest on a bait plasmid pGILDA or pYES was diluted into 200 ml Glu/CM –Ura, –His liquid dropout medium to 2×10^6 cells/ml ($OD_{600} \sim 0.10$) and incubate at 30 °C until the culture contain $\sim 1 \times 10^7$ cells/ml ($OD_{600} = \sim 0.50$). 20 independent library transformations were performed by aliquot into each 1 μ g of library DNA and 50 μ g high quality salmon sperm carrier DNA (a total volume ~ 10 μ l), 50 μ l of yeast cells resuspended in lithium acetate solution, 300 μ l of 40 % PEG solution. The transformation mixture was incubated 1 hour at 30 °C, exposed at heat shock 15 min at 42 °C. Just before heat shock 10 % DMSO was added. In the first plating, yeast cells were plated on appropriate complete minimal medium dropout plates with glucose as a sugar source to select for the library plasmid. For an accurate estimation of the transformation efficiency (the number of transformants obtained) a dilution series was done.

16.3.2. Primary transformants cell collection

Conventional replica plating does not work well in the selection process because so many cells are transferred to new plates that very high background levels inevitably occur. Instead, primary transformant cells were collected by scraping of yeast colonies off the plates that creates slurry in which cells derived from $>10^6$ primary transformants are homogeneously dispersed.

16.3.3. Replating efficiency determination

An aliquot of transformed yeast was diluted 1/10 with Gal/Raff/CM dropout medium and incubate with shaking 4 hour at 30 °C to induce the *GAL1* promoter on the library. Serial dilutions of the yeast cells using the Gal/Raff/CM dropout medium were plated on Gal/Raff/CM dropout plates. The colonies were counted and the number of colony-forming units (cfu) per aliquot of transformed yeast was determined as $\sim 3 \times 10^6$.

16.3.4. Interacting proteins screening

In the second plating, which selects for cells that contain interacting proteins, aliquots of primary transformants was plated on appropriate complete medium dropout plates lacking leucine too, with galactose and raffinose as the sugar source.

An appropriate quantity of transformed yeast based on the replating efficiency was diluted 1/10 with Gal/Raff/CM dropout medium and incubate with shaking 4 hour at 30 °C to induce the *GAL1* promoter on the library.

The culture was further diluted in Gal/Raff/CM medium lacking the appropriate selection markers and leucine, as necessary to obtain a concentration of 10^7 cells/ml ($OD_{600} \sim 0.50$), and 100 μ l of yeast dilution were plated on 20 plates for full representation of transformants, and incubated 2 to 3 days at 30 °C until the colonies appear.

Because not all cells that contain interacting proteins plate at 100 % efficiency on –Leu medium (Estojak et al., 1995), it was desirable that for actual selection, each primary colony obtained from the transformation be represented on the selection plate by three to ten individual yeast cells. This will in some cases lead to multiple isolations of the same cDNA.

Appropriate colonies were picked carefully to a new master dropout plates lacking bait-plasmid selection marker, tryptophan and leucine with galactose and raffinose as carbon source and incubated 2 to 7 days at 30 °C until the colonies appear. Colonies obtained were picked on two dropout master plates, one of them containing glucose as carbon source and the other with galactose and raffinose as carbon source. Colonies obtained on different days were picked up on different master plates. Only the colonies from day 2 and day 3 were further analyzed.

16.3.5. Plasmid isolation

Library plasmids from colonies identified in the second plating are purified by bacterial transformation and used to transform yeast cells for the final specificity screen.

The selected colonies were then cured of the bait plasmid by being grown on Glu/CM-TRP for 3 rounds and subjected to library plasmid DNA isolation from yeast cells (see protocol at paragraph 4.1.2.).

The plasmids recovered from these transformants were then introduced into Mach-1 *E. coli* bacteria cells by bacterial transformation. The library-containing plasmids were then purified following the bacterial miniprep procedure. Two individual bacterial transformants were chosen for yeast positive and restriction digested with restriction enzyme *EcoRI* and *XhoI* to release cDNA inserts. The size of inserts was determined on an 1% agarose gel to confirm that all three plasmids contain the same insert and whether repeated isolation of the same cDNA is occurring.

16.3.6. Positive colonies assessment

The purified library-containing plasmids were used to transform yeast that already contain the baits plasmids carrying the bait gene and the bait plasmid empty as a control.

Transformation mix was plated on Glu/CM dropout plates lacking tryptophan and the selection marker of the bait plasmid. A glucose master dropout plate was created for each library plasmid being tested and incubated 2 to 3 days at 30 °C until the colonies appear. Colonies from this master plate were restreaked to the same series of test plates previously used: complete minimal dropout lacking bait-plasmid selection marker, tryptophan and leucine, one of them with glucose as sugar source and other with galactose and raffinose as carbon source.

Two more control transformants were also assessed, one of them carrying the bait plasmid empty and the library plasmid empty and the other carrying the bait gene on suitable plasmid and the library plasmid empty.

The positive colonies were further analyzed by sequencing.

17. ACROLEIN ASSAY OF TRANSGENIC PLANTS

MS agar standard plant medium supplemented with acrolein of 0.5 to 2.5 mM concentrations was prepared in plates. *Arabidopsis* seeds were surface sterilized by, first soaking

the seeds for 1 min in 50 % ethanol and then washing them for 10 min in 50 % bleach with 0.001-0.2 % SDS. The surface sterilized seeds were incubated at 4 °C for 4 days to break the dormancy. The conditioned seeds were plated on MS medium without and with acrolein of 0.5 mM, 1 mM, 1.5 mM, 2 mM, and 2.5 mM, approximately 100 seeds per plate. The morphological aspects of seedlings were examined by Leica MZ7₅ Stereomicroscope and captured by ProgRes[®] C12^{plus} software program at the stage of 12 days with a magnification of 6.3X.

18. PLANT SPRAY INOCULATION WITH *Pseudomonas syringae* *pv.*tomato DC3000

Overnight bacterial culture of *Pseudomonas syringae pv. tomato* DC3000 were grown in LB media + rifampicin. Next day, 5 ml of the overnight bacteria culture were centrifuged at 3000 rpm for 10 minutes at 4 °C. The supernatant was discarded and the bacterial pellet was resuspended in 5 ml of MgCl₂ 10 mM. The bacterial resuspension was centrifuged again at 3000 rpm for 10 minutes at 4 °C, and then the supernatant was discarded. The washing was repeated two more times and then the bacterial pellet was resuspended in MgCl₂ 10 mM, so that the final concentration of bacterial cells to be 10⁸ cfu (optical density measured on the spectrophotometer equal OD₆₀₀ = 0.35-0.4). LS-77 solution at final concentration of 0.02% was added in the bacterial suspension and then *Arabidopsis thaliana* fully expanded leaves stage were sprayed until became wet. The treated plants were then transferred in the growing chamber at high humidity and required light. The symptoms of the *Pseudomonas* infection were occurred and observed in 8-12 days after inoculation.

CHAPTER III

PRODUCTION OF PLANT TERPENOIDS IN YEAST

1. Expression of Cineole Synthase 1 in yeast cells

Chemical synthesis and modification of natural products can be extremely difficult because of their structural complexity that is highly significant for selectivity against biomolecular targets. Moreover, extraction of bioactive compounds from their natural sources is often inefficient and requires considerable expenditure of the natural resource. Metabolic engineering of biologically amenable microorganisms, such as *Escherichia coli*, *Streptomyces coelicolor* and *Saccharomyces cerevisiae*, aiming to divert their cellular resources and biosynthetic machinery towards the production of desirable heterologous natural products is an environmentally friendlier and more efficient alternative which can lead to less expensive sources of drugs and industrial chemicals of high value. Recent efforts in *S. cerevisiae* engineering have been successful in the incorporation of genes from two important biosynthetic pathways, that of taxol and artemisinin, and in the manipulation of the yeast biosynthetic machinery so as to efficiently produce the artemisinin precursor artemisinic acid.

Previous work pointed to the high antioxidant value of aromatic herbs of the *Salvia* family growing in the island of Crete, such as *Salvia fruticosa* and *Salvia pomifera*. The properties of numerous extracts from these species have been thoroughly examined, and individual terpenoids have been implicated in the antioxidant, antimicrobial and other health promoting properties of these plants. The terpene synthases are the enzymes responsible for the production of terpenoids.

During the course of the current work several groups reported important progress in developing yeast systems for terpene production by engineering metabolic pathways and enzymes (Jackson et al., 2003; Carter et al., 2003; DeJong et al., 2006; Ro et al., 2006; Takahashi et al., 2006) and the field has become one of the hottest areas of biotechnology research.

To establish a system of terpene expression in yeast and engineer further cells to produce high product quantities of terpenes, the Cineole Synthase 1 (Sf-Cin1) gene isolated from a *Salvia fruticosa* cDNA library by EST approach was chosen as a typical monoterpene synthase producing cineole (Kampranis *et al.*, 2007). The Sf-CinS1(RC) N-terminal variant of cineole synthase, previously characterized by bacterial expression and *in vitro* enzymatic assay (Kampranis *et al.*, 2007) was subcloned into an appropriate yeast expression vector (pJG4-6) and expressed under the control of a galactose-inducible promoter.

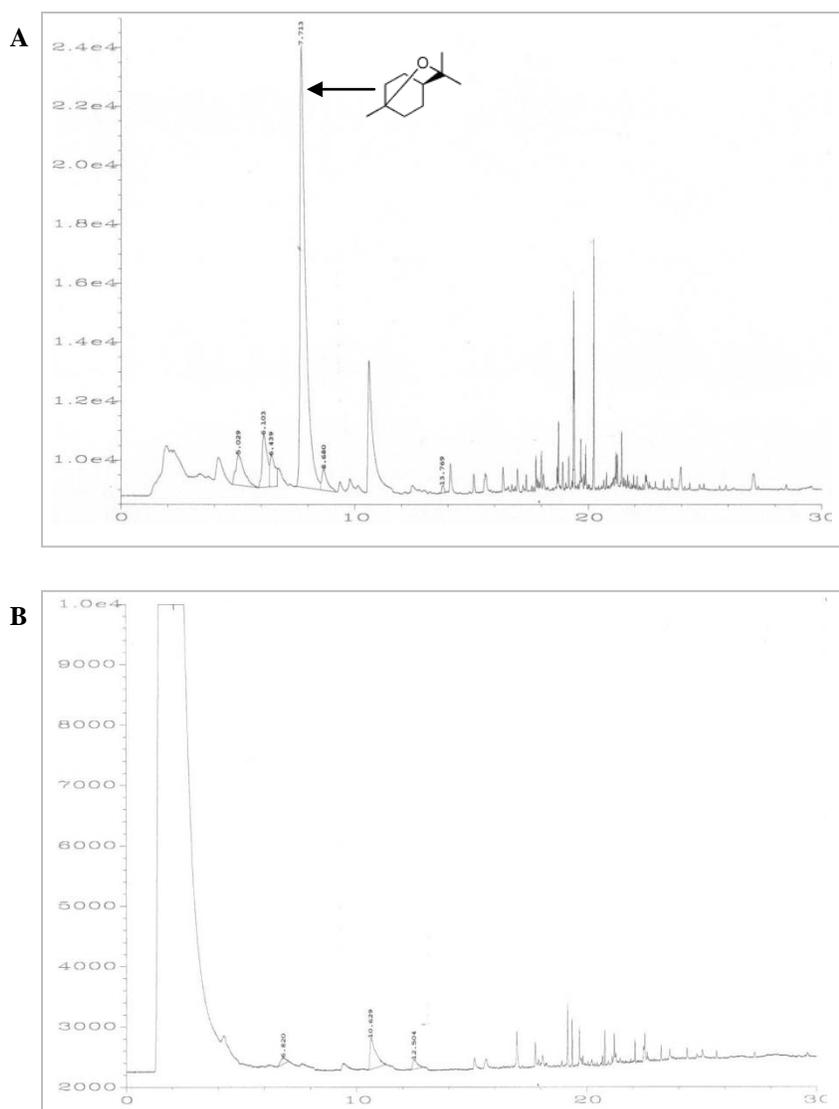


Figure III. 1. Volatiles produced by yeast cells were sampled by solid-phase micro-extraction (SPME) and analyzed by Gas Chromatography. A) Yeast cells expressing Sf-CinS1(RC); B) Yeast cells carrying empty vector.

The ambient volatiles produced were harvested by Solid Phase Microextraction (SPME) and examined by Gas Chromatography. Preliminary results showed that Sf-CinS1(RC) could be successfully expressed in *S. cerevisiae* enabling the production of 1,8-cineole in living cells using the endogenous GPP pool (Fig. III. 1 A). No detectable activity was attained from yeast cells carrying empty vector (Fig. III. 1 B), despite of few putative non-specific yeast compounds. The enzyme catalyzed the formation of multiple monoterpenes using the endogenous GPP pool as substrate, but clearly the 1,8-cineole is the major product. A significant peak of 1,8-cineole (75.09 %) was obtained at a retention time of 7.71 minutes accompanied with several additional peaks: α -pinene (7.78 %) at 5.02 minutes, β -pinene (8.10 %) at 6.10 minutes, myrcene (4.9 %) at 6.43 minutes, γ -terpinene (3.57 %) at 8.68 minutes and α -terpineol (0.54 %) at 13.76 minutes.

In vivo production of monoterpenes by expression of Sf-CinS1(RC) in yeast cells fits well with the profile of products found *in planta* as well as in the bacterial expression enzymatic assay. The higher cineole percentage observed in the bacterially purified enzyme of Sf-CinS1(RC) coincides to that produced in yeast and is a bit higher compared with the yield of cineole in *Salvia fruticosa* (Table III. 1).

Table III. 1. Range of variation of the quantity of the main monoterpene compounds in *Salvia fruticosa* comparing with bacterial and yeast expression product profile of Sf-CinS(RC).

Monoterpene	<i>Salvia fruticosa</i> (%)	Bacterial expression (%)	Yeast expression (%)
α -Pinene	2.58 – 6.26	4.6	7.78
Sabinene	0.31 – 0.69	3.6	-
β -Pinene	8.92 – 18.77	9.1	8.10
Myrcene	1.47 – 5.03	2.2	4.9
Limonene	0.06 – 1.28	<1.0	-
1,8-Cineole	22.70 – 66.20	72.4	75.09
α -terpinene	0.38 – 0.80	-	3.57
α -terpineol	-	7.1	0.54

The monoterpene profile of Sf-CinS1 was further analyzed and confirmed by Gas Chromatography/Mass Spectrometry (GC/MS) (fig. III. 2).

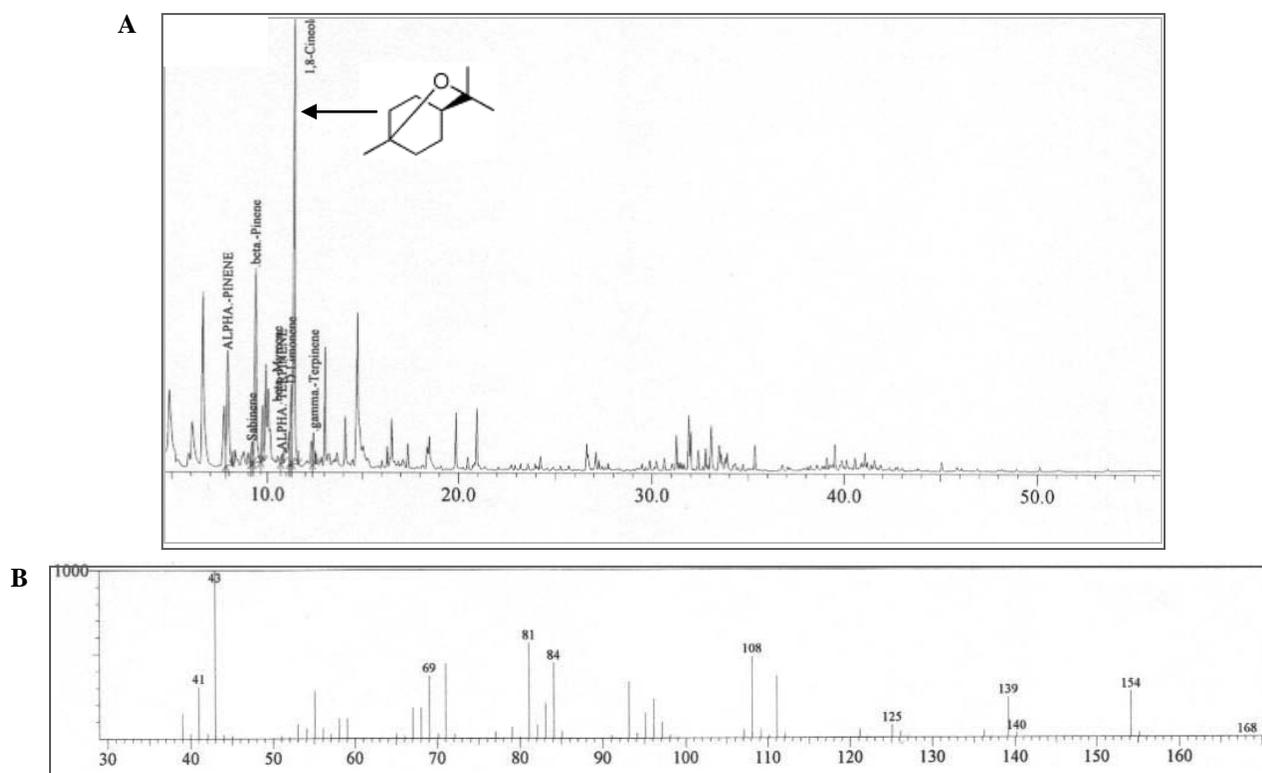


Figure III. 2. Volatiles produced by yeast cells expressing Sf-CinS1(RC) were sampled by solid-phase micro-extraction (SPME) and analyzed by Gas Chromatography/ Mass Spectrometry. A) Chromatogram of the Sf-CinS1(RC) synthase products; B) The mass spectrum of 1,8-cineole detected in yeast cells.

No detectable activity was attained from yeast cells carrying empty vector, despite of traces of farnesol possible resulted from internal sterol biosynthesis of yeast cell, and which do not correspond to the major products assessed in this study (Fig. III. 3).

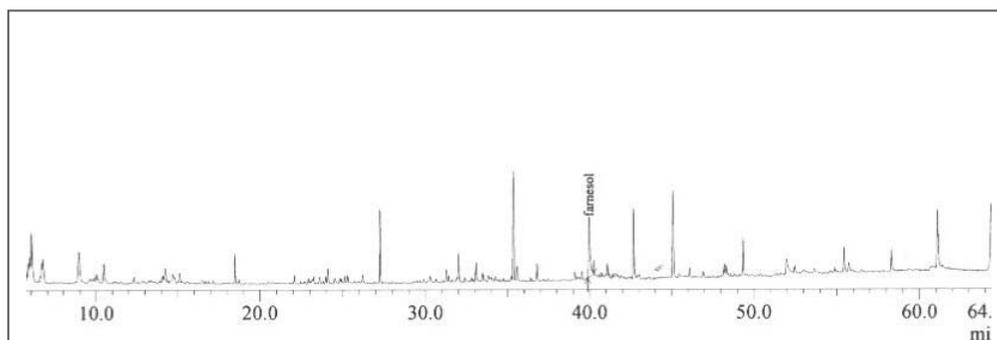


Figure III. 3. The volatiles produced by yeast cells expressing empty vector were sampled by solid-phase micro-extraction (SPME) and analyzed by Gas Chromatography/ Mass Spectrometry. Traces of farnesol were detected.

1.1 N-terminus truncated monoterpenes are active and stable in yeast cells.

In plants, monoterpene synthases are active in the chloroplast but synthesized in the cytoplasm. Monoterpene synthases are encoded as pre-proteins bearing an amino-terminal transit peptide consisting of a conserved tandem RRX₈W motif for import of these gene products into plastids where they are proteolytically processed to the mature forms.

All native monoterpene synthases thus far examined appear to be NH₂-terminally blocked, preventing direct determination (by protein sequencing) of the transit peptide-mature protein cleavage junction (Williams *et al*, 1998). Unfortunately attempts to identify the N-terminal amino acid of the active protein have so far not been successful, truncation of monoterpene synthases exactly before the conserved tandem arginine motif (RRX₈W) has been demonstrated to result in an active and stable enzyme (Fig. III. 4).

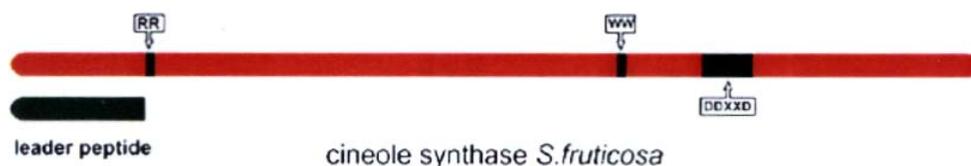


Figure III. 4. 1, 8- cineole synthase predicted protein diagramme (Anssour, 2005).

Previously work on the effect of RR motif at the N-terminal side (Anssour, 2005) had shown that the highly conserved RR motif at the N-terminus are not essential for the enzymatic activity and product specificity of *Salvia fruticosa* 1,8-cineole synthase, as suggested by Williams *et al* (1998), but that the absence of the RR motif affects the substrate binding to the enzyme by lowering the enzyme affinity and also significantly affects enzyme stability. The expression of the 1,8-cineole synthase without the RR motif in *E. coli* produces a highly active enzyme with a product profile similar to that of RR motif containing enzyme. Thus, it remained likely that the cleavage of the transit peptide is upstream of this motif. A few amino acids further upstream of the RRX₈W motif a conserved CS motif next to the proposed cleavage site has been predicted (Mahrez, M.Sc thesis) (Fig. III. 5). Both CS and RR motifs were tested to assess the effect on the enzymatic activity of the two versions.

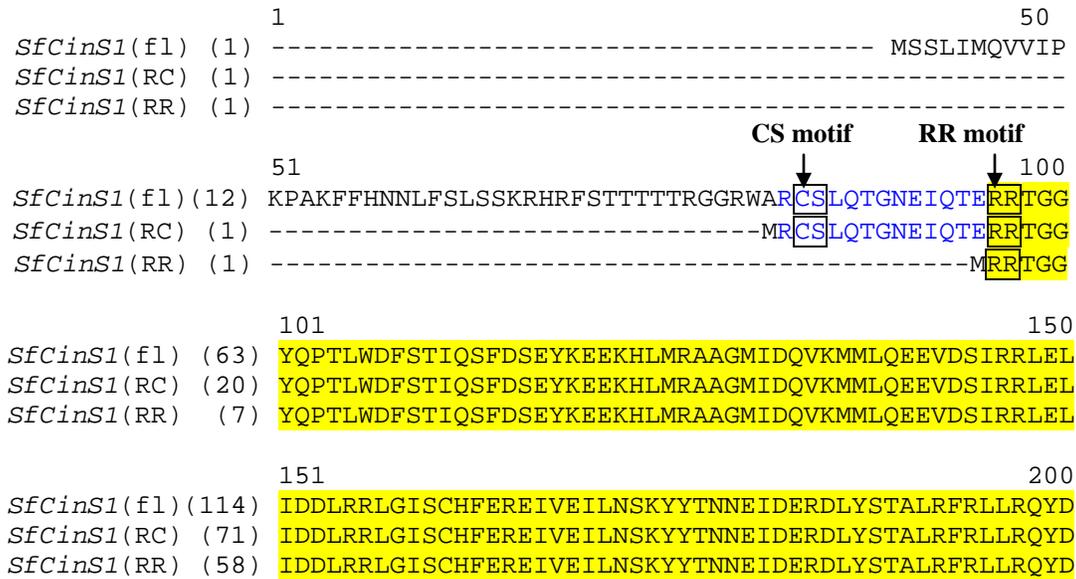


Figure III. 5. Alignment of the deduced amino acid sequences of the transit peptide of the truncated *Salvia fruticosa* 1,8-cineole synthase full length *SfCinS1*(fl) and the two truncated versions *SfCinS1*(RC) and *SfCinS1*(RR).

To functionally characterized the expression and stability of the two truncated cineole synthases, *Sf-CinS1*(RR) and *Sf-CinS1*(RC) in yeast cells, were cloned in the pJG4-6 yeast vector under the control of the *GAL* promoter (Fig. III. 6).

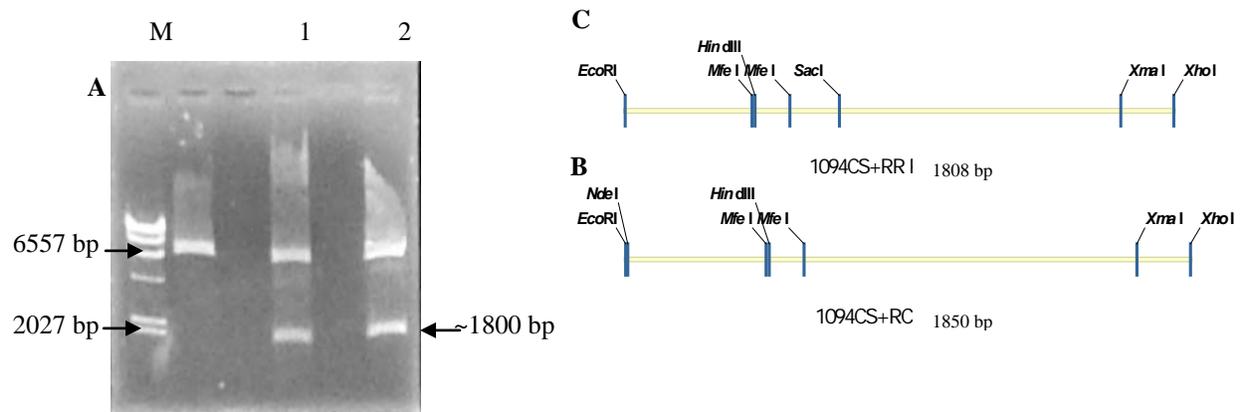


Figure III. 6. Cloning of *SfCinS1*(RR) and *Sf CinS1*(RC) fragments into yeast vector pJG4-6. **A)** Agarose gel electrophoresis of pJG4-6/ *Sf-CinS1*(RR) (1) and pJG4-6/ *Sf-CinS1*(RC) (2) digested with EcoRI-XhoI, where M - Gene ruler, λ DNA/*Hind*III marker; **B)** Linear map of *Sf-CinS1*(RR), 1808 bp; **C)** Linear map of *Sf-CinS1*(RC), 1850 bp.

The expression of Sf-CinS1(RR) and Sf-CinS1(RC) proteins in yeast cells was detected by immunoblotting yeast extracts and probing with the anti-hemagglutinin antibody which recognizes the N-terminal HA fusion tag for detection (Fig. III. 7).

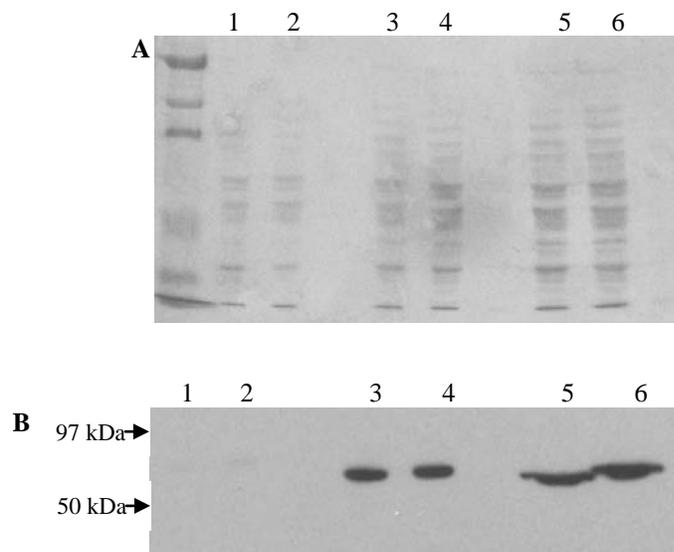


Figure III. 7. Immunodetection of Sf-CinS1(RR) and Sf-CinS1(RC) proteins by HA tag. Total protein lysates from yeast cells overexpressing the Sf-CinS1(RR) (1, 3, 5) and Sf-CinS1(RC) (2, 4, 6) subjected to: **A**) SDS-PAGE electrophoresis followed by **B**) western blot (utilizing the HA tag) to detect and visualize the expression of the proteins of interest. Different concentrations of protein extract used: 10 $\mu\text{g}/\mu\text{l}$ (1, 2), 20 $\mu\text{g}/\mu\text{l}$ (3, 4), 30 $\mu\text{g}/\mu\text{l}$ (5, 6) attested similar level of expression for both proteins.

Table III. 2 summarizes the expected protein sizes and molecular weight of the different *SfCinS1* forms.

Table III. 2: The size of the truncated 1, 8-cineole synthase variants

Analysis	SfCinS1(RR)	SfCinS1(RC)
Length (aa)	574	588
Molecular Weight (Da)	67765	69363

To functionally compare the two truncated proteins, the modified strain AM63 was used to express the Sf-CinS1(RR) and Sf-CinS1(RC) for a period of 13 days. The accumulated volatiles were liberated every 24 hours after each measurement. The activity of the two truncated

enzymes was detected by SPME daily and assessed by GC-MS (Fig. III. 8). Both versions are stable proteins expressed at similar level, but Sf-CinS1(RC) showed a consistent modest improvement in cineole accumulation in comparison to Sf-CinS1(RR).

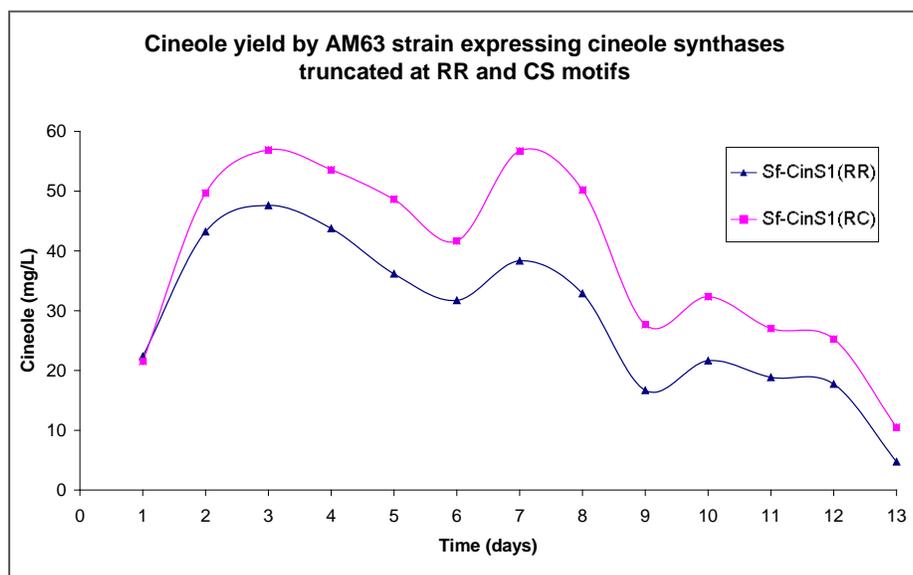


Figure III. 8. Cineole production by yeast strain AM63 transformed with two truncated cineole synthases Sf-CinS1(RR) (▲) and Sf-CinS1(RC) (■) for a period of 13 days.

The enzyme activity was also assessed by conventional enzymatic assay (see Chapter II, subchapter 12). The induced yeast cultures were disrupted by sonication. Yeast lysates were subsequently prepared, incubated with exogenous GPP and the reaction products further evaluated GC-MS (Fig. III. 9 and III. 10). Both truncated proteins, Sf-CinS1(RR) and Sf-CinS1(RC), from broken yeast cells were active when exogenous GPP was added.

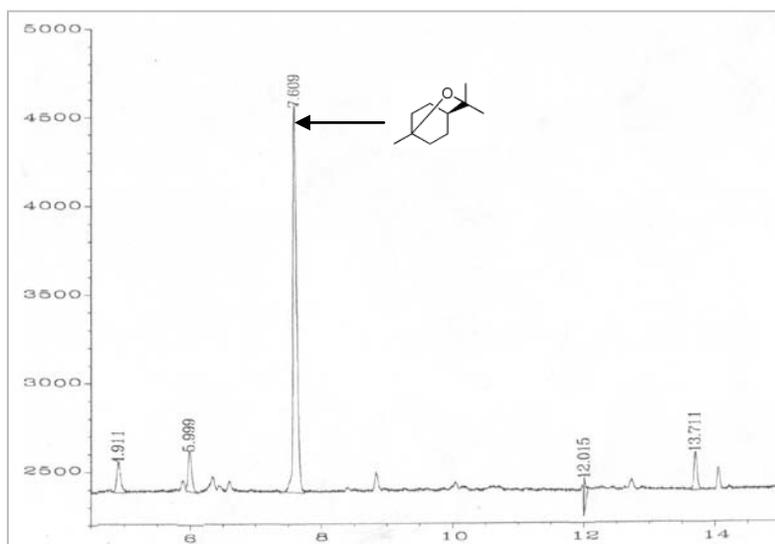


Figure III. 9. The GC-MS analysis of yeast cells expressing cineole synthase Sf-CinS1(RR) and evaluated by conventional enzymatic assay. The enzyme extracted yeast cells is functional in the presence of exogenous GPP substrate.

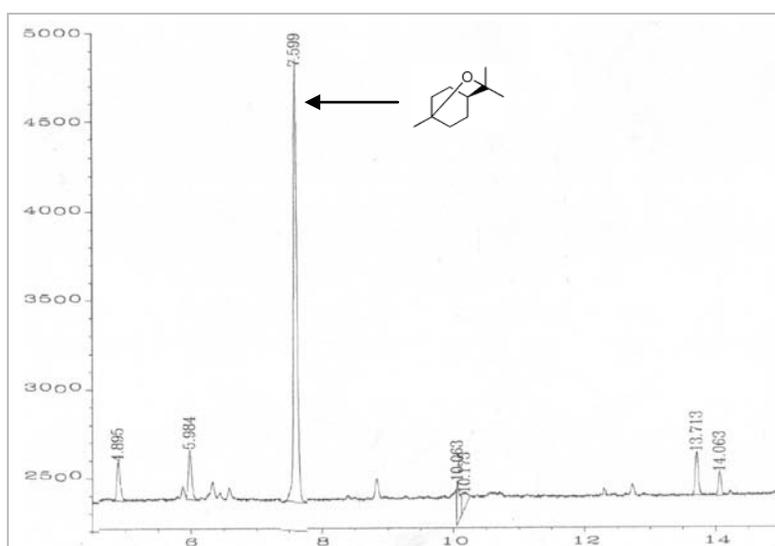


Figure III. 10. The GC-MS analysis of yeast cells expressing cineole synthase Sf-CinS1(RC) and evaluated by conventional enzymatic assay. The enzyme extracted yeast cells is functional in the presence of exogenous GPP substrate.

Additionally, the generated constructs of Sf-CinS1(RR) and Sf-CinS1(RC) in the pJG4-6 vector, described above, the two truncated versions of Sf-CinS1 were also subcloned into EcoRI and XhoI sites of pYES2 vector (see ChapterII, paragraph 8.4) (Fig. III. 11) taking advantage of URA3 selection marker.

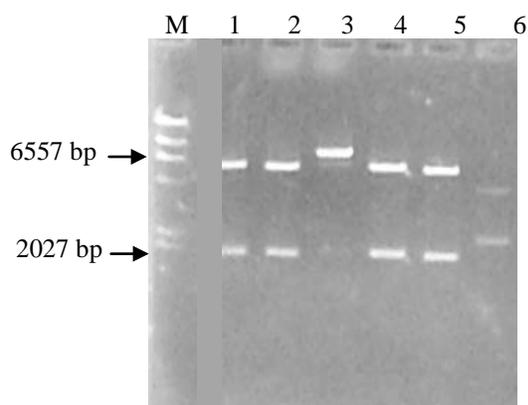


Figure III. 11. Agarose gel electrophoresis of pYES/Sf-CinS1(RC) (1, 2, 3) and pYES/Sf-CinS1(RR) (4,5,6) digested with EcoRI-XhoI, where M - Gene ruler, λ DNA/*Hind*III marker.

2. Establishment and optimization of an SPME-based assay for terpene detection

The terpene synthases expressed in yeast cells used the endogenous pool of substrate (GPP or FPP) to produce volatile compounds which were released into the head space of the vial. The Solid Phase Microextraction (SPME) method was applied to measure *in vivo* the production of terpenes by harvesting the ambient volatile compounds, which were further analyzed qualitatively and quantitatively by Gas Chromatography / Mass Spectrometry (GC/MS). The products were identified by comparing retention times and mass spectra with authentic reference compounds.

Optimal conditions for enzyme induction (described in Chapter II, paragraphs 11.1.1 and 11.1.2) were established after a set of trials testing the volume of yeast culture and the volume of the head space, the second induction on enzymatic buffer, and the glycerol concentration.

Initially, the yeast cells were grown on glucose-based medium until they reached mid-log phase at optical density $OD_{600}=0,5$ and the cultures were switched into galactose-containing medium and incubated for 4-8 hours for protein induction. 500 μ l of culture were removed into a small vial and sampled with an SPME fiber. However, assessing the terpene production in 500 μ l yeast cells resulted in high experimental errors which were probably due to variation in cellular growth of the different samples. Besides the cineole profile, various non-specific compounds were interfered with the GC analysis.

To overcome this, the incubation protocol was modified by extending the induction period until yeast cultures reached saturation and stopped growing ($\sim OD_{600}=2$). The SPME assay

was carried out on the whole 50 ml of culture taking advantage of the head space of the 250 ml Erlenmeyer flask. Highly reproducible results were achieved this way, thus diminishing the experimental errors.

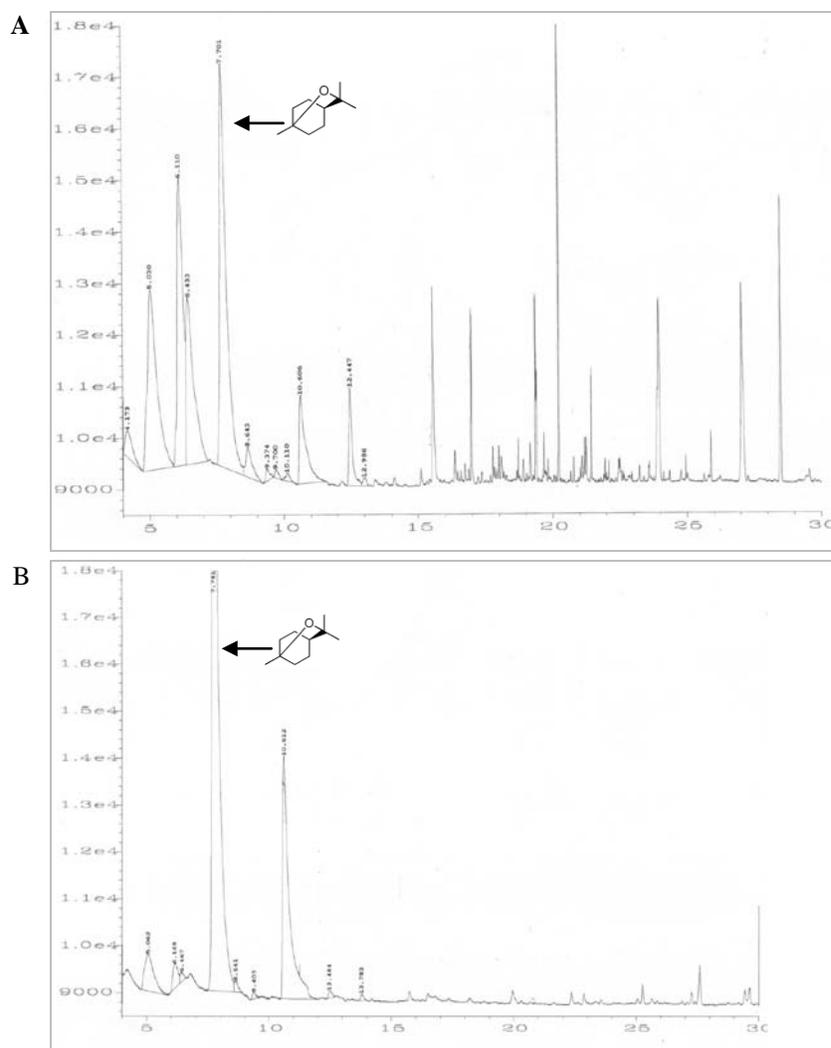


Figure III. 12. Suppression of non-specific volatile compounds from terpene profile by resuspending in enzymatic reaction buffer. A) The GC chromatogram of ambient volatile after galactose induction register several non-specific compounds; **B)** The GC chromatogram of ambient volatile after galactose induction and enzymatic reaction buffer induction.

Conditions were further optimized when a second incubation of transformants was carried in buffer containing terpene synthase cofactors (Fig. III. 12). The viability of the cell, assessed by Evans blue staining, was not affected for a period of time of 10-12 days and the risk of contamination was minimized allowing daily SPME measurements. The compared results

revealed a better activity of terpene synthases after incubation on enzymatic reaction buffer and the suppression of the production of the non-specific compounds (fig. III. 12).

2.1. Influence of glycerol concentration on terpene production

The enzymatic reaction buffer besides the required enzymes cofactors, bivalent metal ions Mn^{2+} and Mg^{2+} , contain glycerol, which confer the source of energy necessary to the yeast cells to survive for a longer period of time allowing the production of terpenes to be carried on for several days. After 10 days, the cineole production almost ceased. The yeast cells were further challenged with 2 % fresh glycerol solution which resulted in restoration of terpene release at low level for a period of four more days (Fig. III. 13).

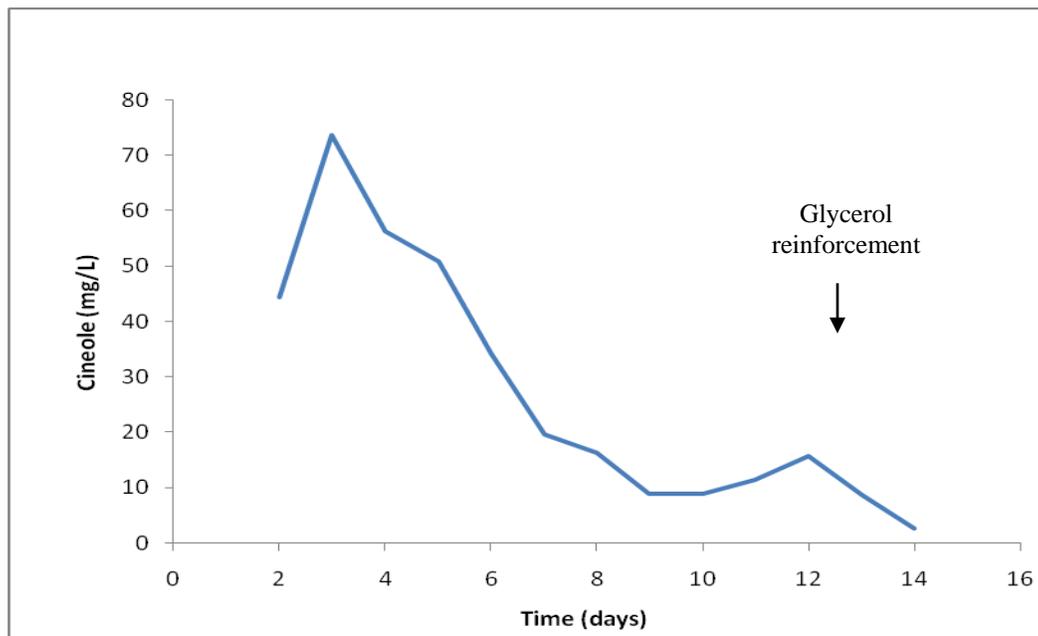


Figure III. 13. The average amount of cineole produced per hour by cineole synthase in EG60 yeast cell for a period of time of 14 days. After 9 days cineole accumulation into the head-space of the flask was ended. 2 % of glycerol was added on day 10 to revive the yeast cells. The cineole production showed a small increase, which culminate on day 12, and decreasing next two day almost zero.

To further investigate if higher concentrations of glycerol could give an extra boost to cineole production, EG60 yeast cells expressing Sf-CinS1(RC) were further evaluated after incubation in buffer containing different concentrations of glycerol. Similar yield of cineole was achieved in samples incubated in no glycerol containing buffer and when 2 % of glycerol was

added into the buffer. The cineole production decreased by ~50 % when the concentration of glycerol was increased to 5 %. In conclusion, high concentration of glycerol may inhibit cineole production, so that the optimal concentration of glycerol used for further determinations was 2% (Fig. III. 14).

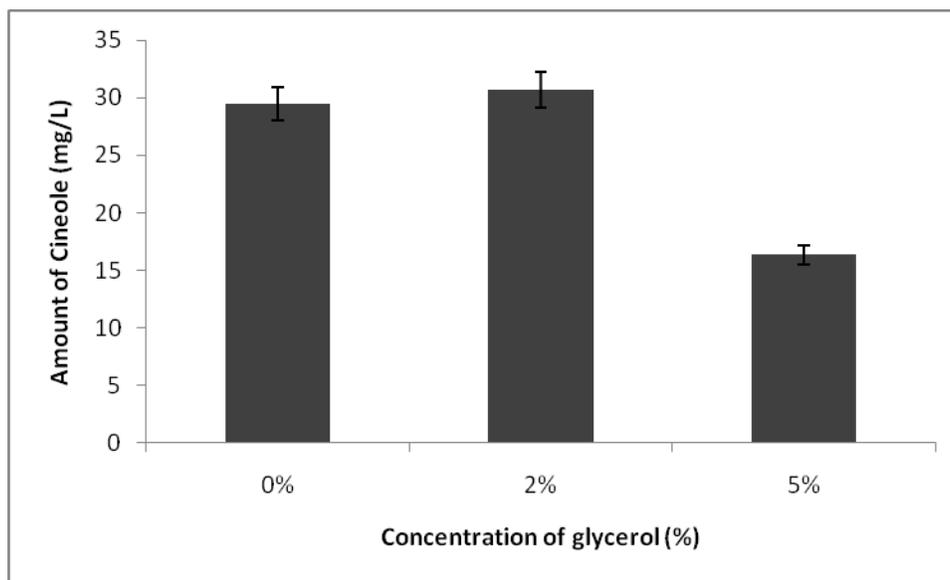


Figure III. 14. The influence of glycerol concentration of cineole production in EG60 yeast cells. When increasing the concentration of glycerol from 2 % to 5 % in enzymatic reaction buffer, the amount of cineole produced in 30 minute was inhibited with around 50 %. Bars indicate the average cineole amount; this experiment was repeated three times with similar results.

3. Assessing different yeast genetic backgrounds for highest production of terpenes

Yeast laboratory and industrial strains have been shown to possess a great deal of phenotypic variability in the response to adverse stress conditions, the capacity to produce high levels of heterologous proteins with limited proteolytic degradation, the tolerance to the presence of toxic compounds and growth properties.

To achieve efficient production of terpenes in yeast it was important to employ and target for further modifications a yeast strain characterized by the production of high levels of the isoprenoid precursors. Production of yeast pheromones by haploid MAT alpha cells has already been shown to drain the available intracellular isoprene units. This variability makes imperative the evaluation of different yeast strains for their capacity to produce high levels of a heterologous

terpene synthase and to release increasing quantities of the terpene compound. To achieve efficient production of terpenes in yeast, the yeast strain EG60 (Mat α , *ura3*, *trp1*, *his3*) for which we had evidence that higher levels of ergosterol and other sterols are produced was tested against the standard Research Genetics strain BY4741 (Fig. III. 15).

The two different genetic backgrounds were transformed with individual plasmids bearing two truncated forms of cineole synthase Sf-CinS1(RR), and Sf-CinS1(RC), respectively under control of *GAL* promoter. Transformed strains were grown on 50 ml minimal media and 20 mg/L glucose as carbon source, and then induced on corresponding galactose-raffinose media for 12-15 h until cells reached the stationary phase. Further incubation in enzymatic reaction buffer at 30°C aided the removal of non-specific compounds. Volatile compounds were absorbed on SPME fibers and analyzed by GS-MS (Fig. III. 15).

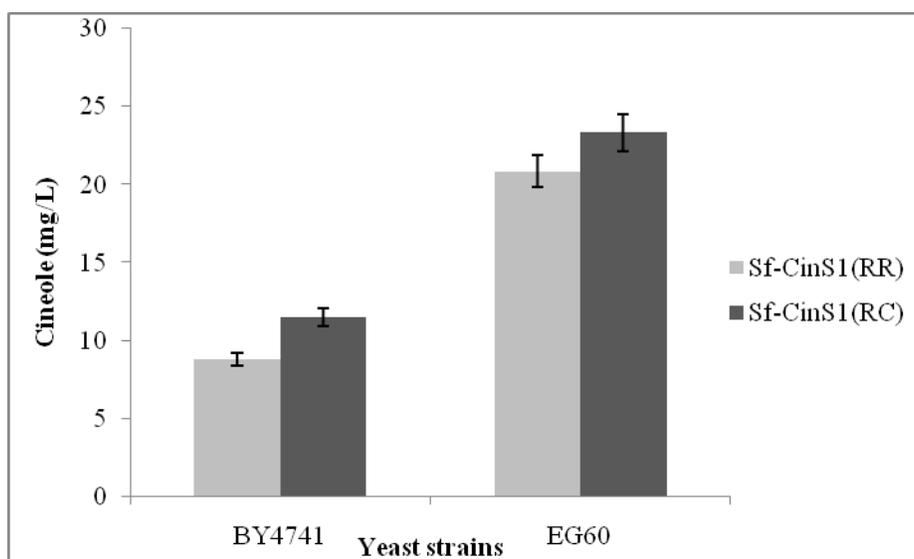


Figure III. 15. Cineole production in two different yeast strain, BY4741 and EG60. Two different truncated forms of cineole synthases RR (blue) and RC (red) were assessed for cineole production. EG60 strain produced almost 2 fold more cineole than BY4741. Bars indicate the average cineole amount; this experiment was repeated three times with similar results.

As shown in Figure III. 15, EG60 cells reproducibly produced 2.5-fold higher levels of cineole than the ResGen BY4741 strain which confirmed the initial hypothesis of selection. EG60 carrying Sf-CinS1(RC) construct exhibited higher cineole production among the proteins expressed and was chosen for further work in the engineering of yeast strains.

4. Integration of the GALp-(K6R)HMG2 construct into the HO locus of the EG60 strain

HMG2 is an important enzyme in the mevalonate pathway, for sterols' synthesis, and is subject to feedback regulation as part of the cellular control of sterol synthesis. When certain sterols accumulate, the enzyme is rapidly degraded, thereby helping to terminate sterol synthesis (Goldstein et al., 2006). Eight biosynthetic steps are involved in the formation of FPP and GPP (Fig. III. 16).

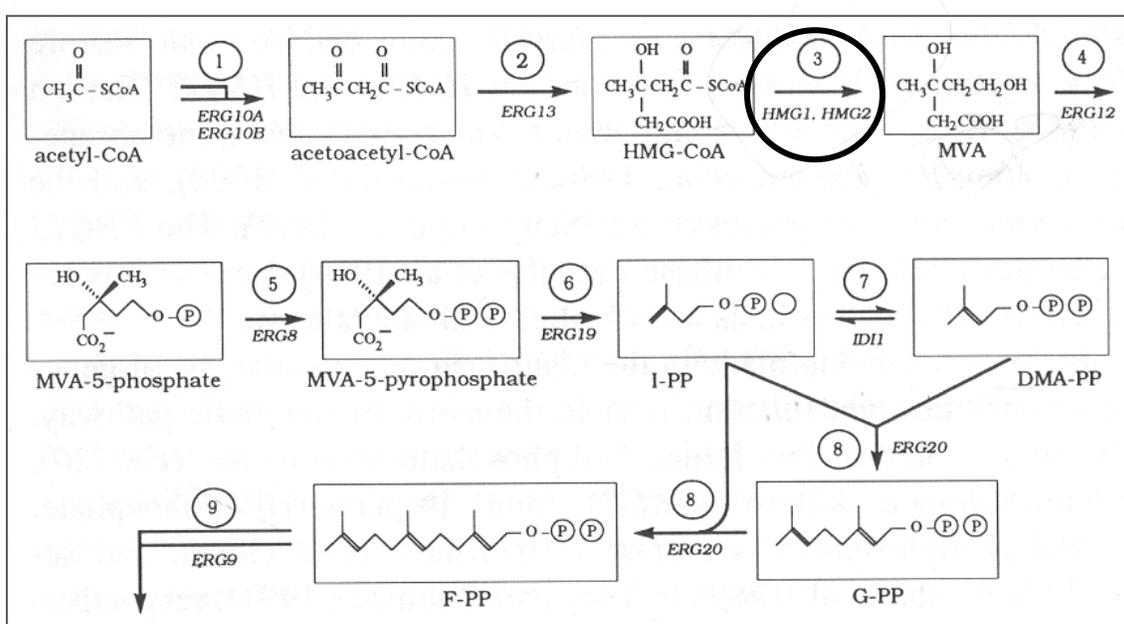


Figure III. 16. Sterol biosynthetic pathway. The first eight biosynthetic steps involved in the formation of isoprene units, FPP and GPP.

Yeast cells contain two HMG genes HMG1 and HMG2. Among them, HMG2, the enzyme converting 3-hydroxy-3methylglutaryl (HMG) coA into mevalonic acid, has been identified as a rate limiting step, and the stability of the protein is regulated by the proteasomal system. Gardner and Hampton uncovered in 1999 two lysines 6 and 357 of HMG2 protein which are critical for degradation. Mutation of Lys 6 results in alteration of HMG2 degradation that is no longer regulated by signals from mevalonate pathway, acquiring available additional endogenous GPP and FPP precursors.

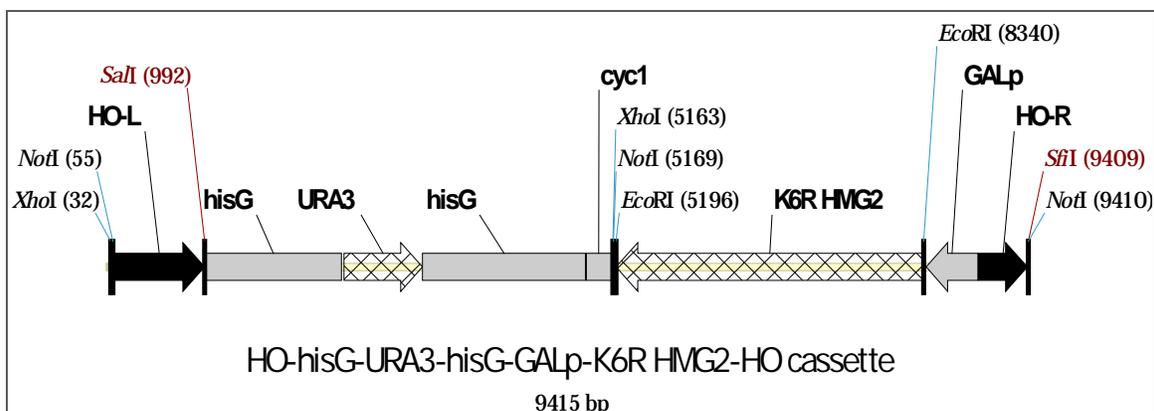


Figure III. 17. The HO-hisG-URA3-hisG-GALp-K6R HMG2-HO integration cassette.

A mutant stabilized version (K6R) of HMG2 stabilized from ubiquitin-mediated degradation under the control of the inducible Gal promoter (Fig. III. 17) was stably integrated into the HO locus of EG60 (Mat α , *ura3*, *trp1*, *his3*) strain, to generate strain AM63 (Mat α , URA3, *trp1*, *his3*, GALp-(K6R)HMG2). By plating AM63 yeast cells on FOA- media, the URA3 gene which was flanked by *hisG* repeats recombined and was eliminated to generate **AM65 strain** (Mat α , GALp-(K6R)HMG2, *ura3*, *trp1*, *his3*).

AM63 and the parental strain EG60 were transformed with individual plasmids bearing the two variant forms of cineole synthase Sf-CinS1(RR), and Sf-CinS1(RC) respectively under control of *GAL* promoter. Transformed strains were grown on 50 ml minimal media and 20 mg/L glucose as carbon source, and then induced on corresponding galactose-raffinose media for 12-15 h until cells reached the stationary phase. After 24 h of incubation in enzymatic reaction buffer at 30°C, volatile compounds were absorbed on SPME fibers for a period of 30 minutes and analyzed by GS-MS. The modified strain expressing Sf-CinS1(RC) produces on average 1.5 fold higher yield than the parental strain and exhibited a reduced background of non-specific volatile metabolites (Fig. III. 18 A). Cineole production was monitored for a period of 13 day to assess the ability of the modified AM63 strain to sustain a lifelong increased production compared to parental EG60 strain. The transformants of truncated Sf-CinS1(RC) exhibiting better activity during the first days were assessed for persistent activity.

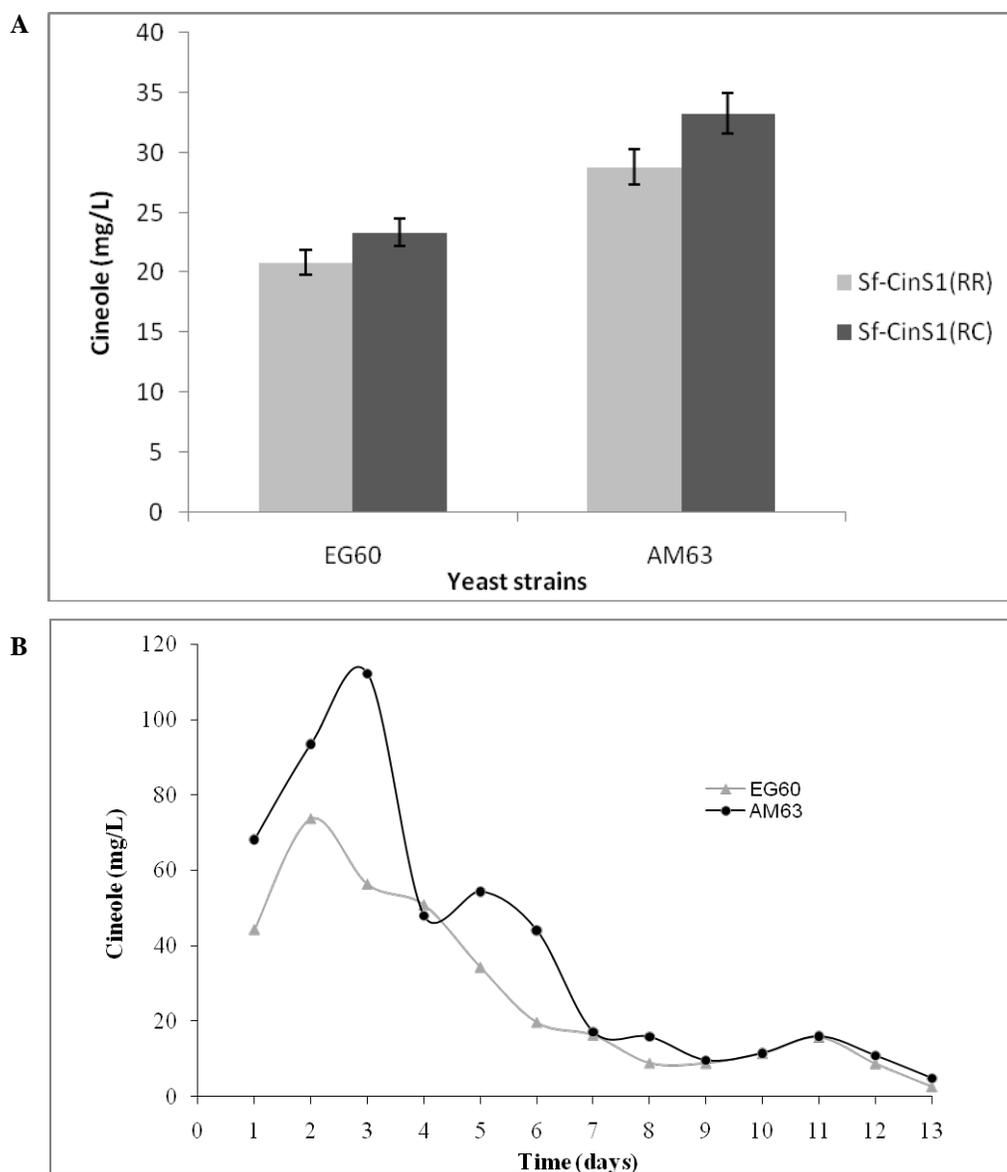


Figure III. 18. Cineole production in modified yeast strain. **A)** AM63 yeast strain engineered with specific modifications to the MVA pathway (GALp-(K6R)HMG2) and parental strain EG60, expressing two truncated forms of cineole synthase (Sf-CinS1(RR) and Sf-CinS1(RC)) were monitored for cineole accumulation at 24 hours. Bars indicate the average cineole amount; this experiment was repeated three times with similar results. **B)** The parental EG60 strain and the modified AM63 strain were used to express Sf-CinS1(RC) and were compared for accumulation of cineole for a period of 13 days. Values of cineole production per h are presented from three repeated experiments.

The SPME sampling occurred for a period of 30 min, 1 h and 2 h and the average production per h was calculated and presented as a scatter chart (Fig. III. 18 B). The accumulated volatiles were released every 24 h after each measurement.

In conclusion, the AM63 modified yeast strain possess a higher terpene producing capacity, which reach on average ~1.5 fold higher levels compared to EG60, to synthesize cineole from endogenous GPP pool than the parental strain. The maximum production calculated per hour was reached by AM63 after three day of incubation on buffer and started to considerably diminish after 6 days of incubation.

5. Terpene Synthase Genes are Functional in the Yeast

Within the genus *Salvia*, *S. fruticosa*, and *S. pomifera*, each produce a virtually identical range of mono- and sesquiterpenes. Although the qualitative oil composition of both species is similar, quantitatively they are very different with respect to the main compounds. The oil of *Salvia fruticosa* is dominated by high levels of monoterpenes 1,8-cineole, α - and β -pinene, β -myrcene, whereas sabinene is found in low amount. *Salvia pomifera* has high level of sabinene, while 1,8-cineole is found in low amount (Table III. 3).

Table III. 3. Ranges of the quantity of monoterpene compounds in *Salvia* species upper leaves (% of total oil).

	<i>Salvia fruticosa</i>	<i>Salvia pomifera</i>
α-Pinene	2.58 – 6.26	0.42 – 2.40
Sabinene	0.31 – 0.69	1.00 – 2.65
β-Pinene	8.92 – 18.77	0.40 – 1.85
β-Myrcene	1.47 – 5.03	0.48 – 1.38
Limonene	0.06 – 1.28	0.27 – 0.59
1,8-Cineole	22.70 – 49.20	0.09 – 0.20
Terpinene	0.38 – 0.80	0.20 – 0.49

In both species the predominant sesquiterpene is caryophyllene (Table III. 4). α -humulene is found in intermediate levels in *Salvia fruticosa*, while α -cubebene and cadinene are found in intermediate levels in *Salvia pomifera*. Previous studies identified a high degree of variability in quantity and composition of the essential oils from different populations and clones of *Salvia fruticosa* and *Salvia pomifera*. Nevertheless, the qualitative similarity of composition suggests the possibility that both species may contain the same enzymes which can be differentially regulated between the species and within the populations. The terpene synthases

that make these products or their precursors are often ‘promiscuous’ in that, besides their eponymous product they make multiple others at low or trace amounts.

Table III. 4. Ranges of the quantity of sesquiterpene compounds in *Salvia* species upper leaves (% of total oil).

	<i>Salvia fruticosa</i>	<i>Salvia pomifera</i>
α-Cubebene	0.26 – 0.19	1.40 – 4.68
Caryophyllene	4.64 – 12.75	7.82 – 22.75
α-Humulene	2.18 – 4.13	0.43 – 1.49
Curcumene	0.03 – 0.18	0.12 – 5.16
Cadinene	0.03 – 0.70	2.29 – 8.00

in previous work, two cDNA libraries were developed from glandular trichomes of *Salvia fruticosa* and *Salvia pomifera* (D. Ioannidis, 2007; S.Anssour, 2005). EST sequencing of size selected clones (> 500 bp) identified several genes that code for proteins sharing homologies to enzymes of the terpenoid biosynthetic pathway. Five cDNAs resemble monoterpene synthases and twelve of them were identified as putative sesquiterpene synthases (Fig. III. 19 and Fig. III. 20).

Figure III. 19. Phylogenetic tree resulting from the multiple protein alignment of the isolated cDNA clones predicted protein sequences expressed in yeast cells with other terpene synthases previously reported.

		1		50	CS motif	
AAD50304	(1)	MALKVFSVATQMAIPSKLTRCLQPSHLKSSPKLLSSTNSSSRSL		LRVYCS		
AF527416	(1)	-----MAFPRNPTKLLHKPHNKSS-KLISNSRISSYGH		LPLRCS		
clone 215	(1)	-----CSVVIQMAIPSKPTNHLHNSRTKSS-KLSSNSITSVGAR		LRS	SPRC	
AF051899	(1)	----MSSLIMQVVIP-KPAKIFHNNLFSVISKRHRFSTTTITRGGRWAHC				
DQ785793	(1)	----MSSLIMQVVIP-KPAKFFHNNLFSLSKRHRFSTTTTTTRGGRWARC				
AF051900	(1)	----MSIISMNVSILSKPLNCLHNLERRPS-KALLVPCTAPTAR		LRASCS		
AF051901	(1)	-----MSSISINIAMPLNSLHNFERKPS-KAWSTSTCTAPAAR		LRASSS		
DQ785794	(1)	-----MPLNSLHNLERKPS-KAWSTSTCTAPAAR		LQASFS		
clone 1199	(1)	-----				
clone P330	(1)	-----		HMAE	IYASAV	
AAV63785	(1)	-----		MSANCV	SAA	
AY693644	(1)	-----		MTNMF	ASAA	
clone 1025	(1)	-----		EFME	ICSQPI	
ACC66281	(1)	-----		MDSPTT		
AAL79181	(1)	-----			M	
AAX39387	(1)	-----MSTLP		ISSVSFSSSTSP	LVVDDK	
AF374462	(1)	-----		MKDMSIP	LLAAVS	
EEF39510	(1)	-----		MSLQVSAVP	LKTSTQ	
CAA77191	(1)	-----		MASQASQV	LASPH	
Consensus	(1)		RR motif		L AS	
		51			100	
AAD50304	(51)	S--S---QLTTER	RRSGN	YNSPSR	WDVEFIQSLHSDYEE-DKHAIRAS	ELVLT
AF527416	(39)	SQQLPTDEFQVER	RRSGNYS	PSKWDVDYI	QSLHSDYKE-ERHTRRAS	ELIM
clone 215	(45)	SVQLSAGQLQTE	RRSGNYS	PSLWDVDYI	QSLHSDYKE-ERHMRRAS	ELIM
AF051899	(46)	SLQMG-NEIQTG	RRRTGGYQ	PTLWDFSTI	QLFDSEYKE-EKHLMRAAG	MIA
DQ785793	(46)	SLQTG-NEIQTE	RRRTGGYQ	PTLWDFSTI	QSFSEYKE-EKHLMRAAG	MID
AF051900	(46)	SKLQ---EAHQI	RRSGNYQ	PALWDSNYI	QSLNTPYTE-ERHLDRKA	ELIV
AF051901	(43)	LQQE---KPHQI	RRSGDYQ	PSLWDFNYI	QSLNTPYKE-QRHFNRQA	ELIM
DQ785794	(34)	LQQE---EPRQI	RRSGDYQ	PSLWDFNYI	QSLNTPYKE-QRYVNRQA	ELIM
clone 1199	(1)	-----				
clone P330	(11)	PISTKNTNVENT	RRSVTYH	PSVWRDHFL	KYTDDVTKITTAEKQVLE	KEKE
AAV63785	(10)	PTSPKNSDVVEE	IRKSATYHS	SVWGNHFL	SYTSDVTEITAAEKEQLE	KLKE
AY693644	(10)	PISTNNTTVEDM	RRSVTYH	PSVWVDHFL	DYASGITEV---EMEQLQ	KQKE
clone 1025	(11)	PAIKKVKNLDEI	RKSAKFHPSI	WGDYFLQY	SDKTKISDVEQEELAK	QKE
ACC66281	(7)	QRPNME----	IGRA	FVNYHPSI	WGEHFAASP-----DVMRLDAH	KGRG
AAL79181	(2)	SVKEEK----	VIRPI	VHFPPSV	WADQFLIFDDKQAEQ-ANVEQV	VNELRE
AAX39387	(24)	VSTKPD----	VIRHT	MNFNASI	WGDQFLTYDEP-EDL-VMKKQ	LVEELKE
AF374462	(14)	SSTEE-----	TVRPI	ADFHPTL	WGNHFLKSAAD-----VETIDAAT	QEQH
EEF39510	(16)	NATSA-----	VKRHS	STYHPTI	WGDHFLANLSH-----SKI	IDGSI
CAA77191	(14)	PAISS-----	ENRPK	ADFH	PGI	WGD
Consensus	(51)		RRSG	Y PSLW	FI	EL
		101				150
AAD50304	(95)	LVKMELEKE-----	TDH	IRQL	LELIDDLQRMGLSDHFQNEFKEIL	SSIYL
AF527416	(88)	EVKKLLEKE-----	PNPTR	QL	LELIDDLQKLGLSDHFNNEFKEIL	NSVYL
clone 215	(94)	QVKMMLEEE-----	ADP	VRQL	LELIDDLQRLGLGDHFQNEFKEIL	LKSIYL
AF051899	(94)	QVNMLLQEE-----	VDSI	QRLE	LELIDDLRRLGISCHFDREIVEI	LNSKY
DQ785793	(94)	QVKMMLQEE-----	VDSI	RRLE	LELIDDLRRLGISCHFEREIVEI	LNSKY
AF051900	(92)	QVRILLKEK-----	MEPV	QQ	LELIHDLKYLGLSDFQDEI	KEILGVIYN
AF051901	(89)	QVRMLLKVK-----	MEAI	QQ	LELIDDLQYLGLSYFFQDEI	KQILSSIHN
DQ785794	(80)	QVRMLLKVK-----	MEAI	QQ	LELIDDLQYLGLSYFFPDEI	KQILSSIHN
clone 1199	(1)	-----				
clone P330	(61)	DVKKLIAQTP-----	DDSTV	KI	ELIDAIQRLGVAYHFPKEI	DES
AAV63785	(60)	KVKNLLAQTP-----	DESTG	KMELIDAI	QRLGVGYHFTTEI	IQESLRQIHE
AY693644	(57)	RIKTLAQTL-----	DDFV	LKI	ELIDAIQRLGVGYHFEKEI	NHSLRQIYD
clone 1025	(61)	MVKKLLAQTP-----	NNSI	YKMELIDAI	QRLGVEYHFEKEL	DES
ACC66281	(47)	EELKEVVRN-MFSTVND	PL	LLKMN	LIDAIQRLGVAYHFEME	IDKALGQMYD

As it became evident from the sequence comparisons, the two species although residing in the same geographic locations and share highly similar profiles did not express in their transcriptome the same enzymes, at least as far as we failed to identify expression of an ortholog to SfCinS1 in *S. pomifera* and since an ortholog of the sabinene synthase present in *S. pomifera* (homologous to the enzyme from *S. officinalis*) is absent from *S. fruticosa*. Among all the ESTs screened (overall 2,000 for *S. fruticosa* and 500 for *S. pomifera* no actual orthologs between them were identified so far). Beside SfCinS1 thoroughly considered in this study, a novel monoterpene synthase (clone 215) isolated from *S. fruticosa*, distinct from the previously reported SfCinS1, was found in this work to produce mostly cineole. Two other novel terpene synthases, which previously failed to yield any products *in vitro*, when tested as bacterially expressed proteins, were assessed for activity enzymes in yeast cells. They appeared to be sesquiterpene synthases (1025, and P330 clones respectively), the first producing β -farnesene and nerolidol and the second δ -cadinene, α -copaene, trans-caryophyllene and α -cubebene, both of them being enzymes producing multiple products. One last gene tested was a short C-terminal version of a monoterpene synthase gene from *S. fruticosa* (1199 clone) which failed to produce any product.

5.1. Monoterpene synthases are functional in yeast

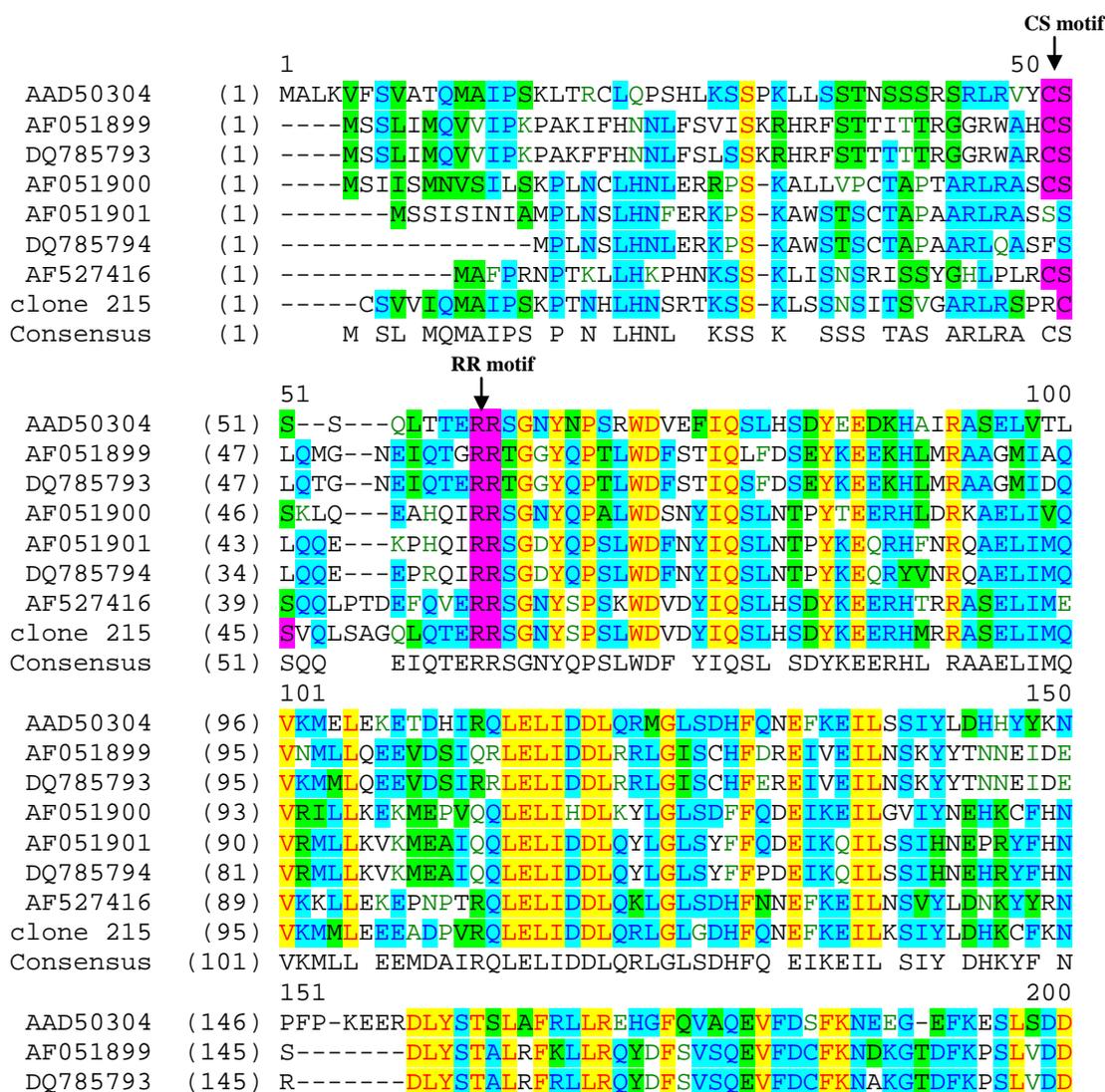
Monoterpene synthases synthesize monoterpenes utilize geranyl pyrophosphate (GPP) as substrate. In yeast the GPP precursors are synthesized by the sterol biosynthetic pathway. Before expressing terpene synthases in yeast cells, it was necessary to assess empty cells for terpene production. Both parental and modified AM63 strains were processed in similar way as the samples. A series of previously characterized as well as and novel monoterpene synthases isolated from *Salvia fruticosa* (Greek sage) were expressed and further evaluated in the AM63 and the parental strain.

5.1.1. Expression and analysis of enzyme activity of Sf 215 ORF from *Salvia fruticosa* in yeast

One of the cDNA inserts isolated from the *Salvia fruticosa* cDNA library and fully sequenced was designated as clone 215 and contained a transcript of 1985 bp. The 215 deduced protein showed 71 % similarity to cineole synthase from *Salvia fruticosa*, SfCinS1, 69 % similarity to 3-carene synthase from *Salvia stenophylla* and 67 % similarity to limonene synthase

from *Mentha longifolia*. These sequences were aligned to each other using the Vector NTI software package and their predicted protein sequences were compared with previously reported monoterpene enzymes showing their close relationship (Fig. III. 21 and Fig. III. 22).

Figure III. 21. Phylogenetic tree resulting from the multiple protein alignment of the isolated 215 clone predicted protein sequences expressed in yeast cells with other monoterpene synthases previously reported.



AF051900	(143)	NEV--EKM	DLYFTALGFRLLRQHGFNISQDVFNCFKNEKGIDFKASLAQD
AF051901	(140)	-----	NDLYFTALGFRILRQHGFNVSEDFVDFCFKIEKCSDFNANLAQD
DQ785794	(131)	-----	NDLYLTALGFRILRQHGFNVSEDFVDFCFKTEKCSDFNANLAQD
AF527416	(139)	GAMKEVER	DLYSTALAFRLLRQHGFQVAQDVLECFKNTKG-EFEPSSDD
clone 215	(145)	----HGEM	DLYSTALAFRLLRQHCFQVAQDVDFCFKNEKG-ELKASLSDD
Consensus	(151)		DLYSTALAFRLLRQHGFNVSQDVDFCFKNEKGSDFKASLADD
		201	250
AAD50304	(194)	TRGLLQLYEASFLLTEGETTLESAREFATKFLERVN--EGGV	DGDLLTR
AF051899	(188)	TRGLLQLYEASFLSAQGEETLHLARDFATKFLHKRVL---	VDKDINLLSS
DQ785793	(188)	TRGLLQLYEASFLSAQGEETLRLARDFATKFLQKRVL---	VDKDINLLSS
AF051900	(191)	TKGMLQLYEASFLLRKGETTLELAREFATKCLQKKLDEGGNE	IDENLLLW
AF051901	(183)	TKGMLQLYEASFLLREGEDTLELARRFSTRSLREKFD	EGGDEIDEDLSSW
DQ785794	(174)	TKGMLQLYEASFLLREGEDTLELARRFSTRSLREKLD	EDGDEIDEDLSSW
AF527416	(188)	TRGLLQLYEASFLLTEGENTLELARDFTTKILEEKL	RN-DEIDDINLVTW
clone 215	(190)	TRGVLQLYEASFLTMEGEKTLDLGREFAAKILEDK	KE-ESSDDL YLLLS
Consensus	(201)	TRGLLQLYEASFLL EGEDTLELARDFATK L EKL E	EIDINLLSW
		251	300
AAD50304	(242)	IAYSLDIPLHWRIKRPNAPAWIEWYRKR	PDMNPVLELAILDNLIVQAQF
AF051899	(235)	IERALELPTHWRVQMPNARSFIDAYKRR	PDMNPTVLELAKLDFNMVQAQF
DQ785793	(235)	IERALELPTHWRVQMPNARSFIDAYKRR	PDMNPTVLELAKLDFNMVQAQF
AF051900	(241)	IRHSLDLPLHWRIQSVEARWFIDAYARR	PDMNPLIFELAKLNFNIIQATH
AF051901	(233)	IRHSLDLPLHWRVQGLEARWFLDAYARR	PDMNPLIFKLAKLNFNIVQATY
DQ785794	(224)	IRHSLDLPLHWRIQGLEARWFLDAYARR	PDMNPLIFKLAKLNFNIVQATY
AF527416	(237)	IRHSLDIPHWRIDRVNTSVWIDVYKRR	PDMNPVLELAVLDSNIVQAQY
clone 215	(239)	IRYALDIPHWRIGRGNASMWIDAYKRR	SDMNPIVLELAILDNIVQAQY
Consensus	(251)	IRHSLDLPLHWRIQ LNAR FIDAYKRR	PDMNPLVLELAKLDFNIVQAQY
		301	350
AAD50304	(292)	QEELKESFRWWRNTGFVEKLPFARDRL	VECYFWNTGII EPRQHASARIMM
AF051899	(285)	QEELKEASRWWNSTGLVHELFPVDR	DRIVECYWTTGVVVERREHGYERIML
DQ785793	(285)	QEELKEASRWWNSTGLVHELFPVDR	DRIVECYWTTGVVERRQHGYERIML
AF051900	(291)	QEELKDLSRWWSRLCFPEKLPFVR	DRLVESFFWAVGMFEPHQHYQRKMA
AF051901	(283)	QEELKDISRWWNSCLA EKLFPVDR	DRIVECFFWAIAAFEPHQYSYQRKMA
DQ785794	(274)	QEELKDVSRWWNSCLA EKLFPVDR	DRIVECFFWAIGAFEPHQYSYQRKMA
AF527416	(287)	QEELKLDLQWWRNTCLA EKLFP	ARDRLVESYFWGVGVVQPRQHG IARMAV
clone 215	(289)	QEELKQDLQWWRNTCIVEKLPF	ARDRLVECYFWTTGIVQPRQHANARITV
Consensus	(301)	QEELKD SRWWNSTCLVEKLPFVR	DRIVECYFWATGVVEPRQHGY RIML
		351	400
AAD50304	(342)	GKVNALITVIDDIYDVYGTLEELE	QFTDLIRRWDINSIDQLPDYMQLCFL
AF051899	(335)	TKINALVTTIDDVFDIYGTLEEL	QLFTTAIQRWDIESMKQLPPYMQICYL
DQ785793	(335)	TKINALVTTIDDVFDIYGTLEEL	QLFTTAIQRWDIESMKQLPPYMQICYL
AF051900	(341)	ATIIVLATVIDDIYDVYGTLELE	LFTDTFKRWDTESITRPLPYMQLCYW
AF051901	(333)	AVIITFITIIDDVYDVYGTLEEL	LELTDMIRRWDNKSISQLPYMQVCYL
DQ785794	(324)	AIIITFVTIIDDVYDVYGTLEEL	LELFTDMIRRWDNISISQLPYMQVCYL
AF527416	(337)	DRSIALITVIDDVYDVYGTLEELE	QFTEAIRRWDISSIDQLPSYMQLCFL
clone 215	(339)	GKVNALITVIDDVYDVYGNLEELE	QFTDVIRRWDMSISIEQLPGYMQLCFL
Consensus	(351)	AKINALITVIDDVYDVYGTLEEL	LELFTDMIRWDI SI QLP YMQLCYL

Figure III. 22. Alignment of deduced amino acid sequences of the 215 monoterpene synthase clone from *Salvia fruticosa* with several reported monoterpene syntases. Where, AAD50304 – limonene synthase [*Mentha longifolia*], AF527416 – 3-carene synthase [*Salvia stenophylla*], AF051899 – 1,8-cineole synthase [*Salvia officinalis*], DQ785793 – cineole synthase CinS1 [*Salvia fruticosa*], AF051900 – (+)-bornyl diphosphate synthase [*Salvia officinalis*], AF051901 – (+)-sabinene synthase [*Salvia officinalis*], DQ785794 – sabinene synthase SabS1 [*Salvia pomifera*]. The alignment was created with the vector NTI software.

Not unexpectedly, phylogenetic data did not provide any real indication as to the products that may be expected from the expressed enzyme, and no homologue of it was identified in the closely related species *S. officinalis* and *S. pomifera*

Previous attempts to obtain activity from this protein expressed in bacterial cells were unsuccessful (S.Ansour, 2005). Although expression of the 215 protein in the bacterial system has been confirmed by western blotting, no detectable activity could be obtained (Mahrez, 2007). This is a rather frequent phenomenon as a large percentage of the isolated cDNAs homologous to terpene synthases reported in the literature fail to yield any detectable products in enzymatic assays. The reasons for this are not very clear.

To functionally identify this putative monoterpene synthase by heterologous expression in yeast cells, the 215 cDNA insert was excised from pCRII-TOPO/215R construct and it was subcloned into pYESmyc plasmid (see Chapter II, paragraph 8.8) into the BamHI and XhoI restriction sites (Fig. III. 23).

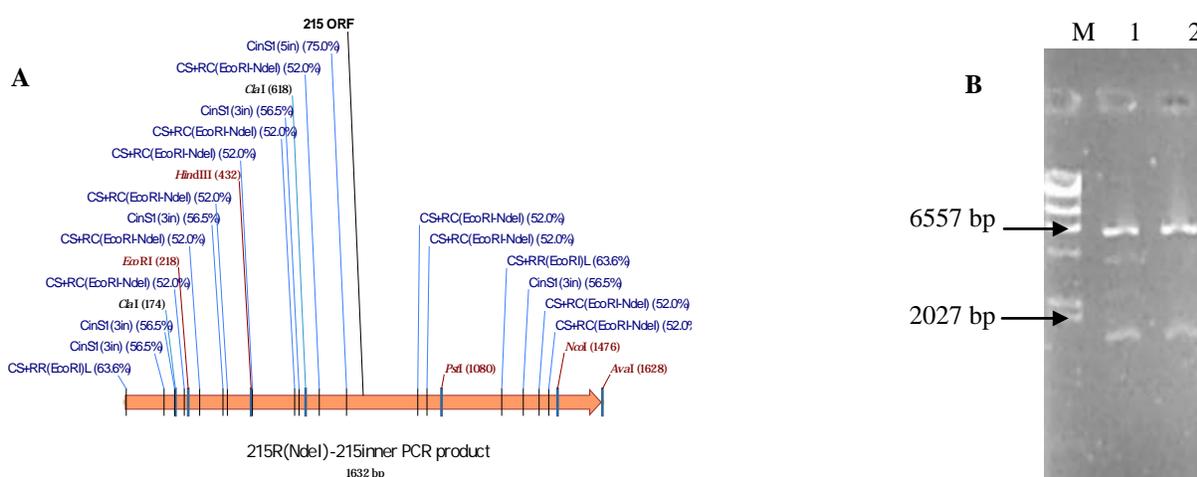


Figure III. 23. Cloning of 215R into pYESmyc vector. A) Linear map of 215R clone B) Agarose gel electrophoresis of pYESmyc/215R construct digested with BamHI-XhoI (1, 2), where M - Gene ruler, λ DNA/*HindIII* marker.

To further assess the activity of 215 monoterpene synthase, the pYESmyc/215R construct or an empty vector were inserted into EG60 and AM63 strains by lithium-acetate transformation. After the two steps of induction, the released volatiles were extracted by SPME method and analysed by GC-MS. The enzyme is shown to catalyze the formation of mostly cineole from

endogenous GPP pool (Fig. III. 24). The yield of cineole from the 215R product was significantly lower than as Sf-CinS1(RR) or as Sf-CinS1(RC) products.

The compounds were identified by comparing the target spectra with reference spectra using the Wiley7 GC-MS spectral library (Appendix 2) with high similarity. The monoterpene profile of the two clones from *Salvia fruticosa* producing cineole, clone 215 and SfCinS1, (Table III. 5) diverge one to each other both qualitative and quantitative.

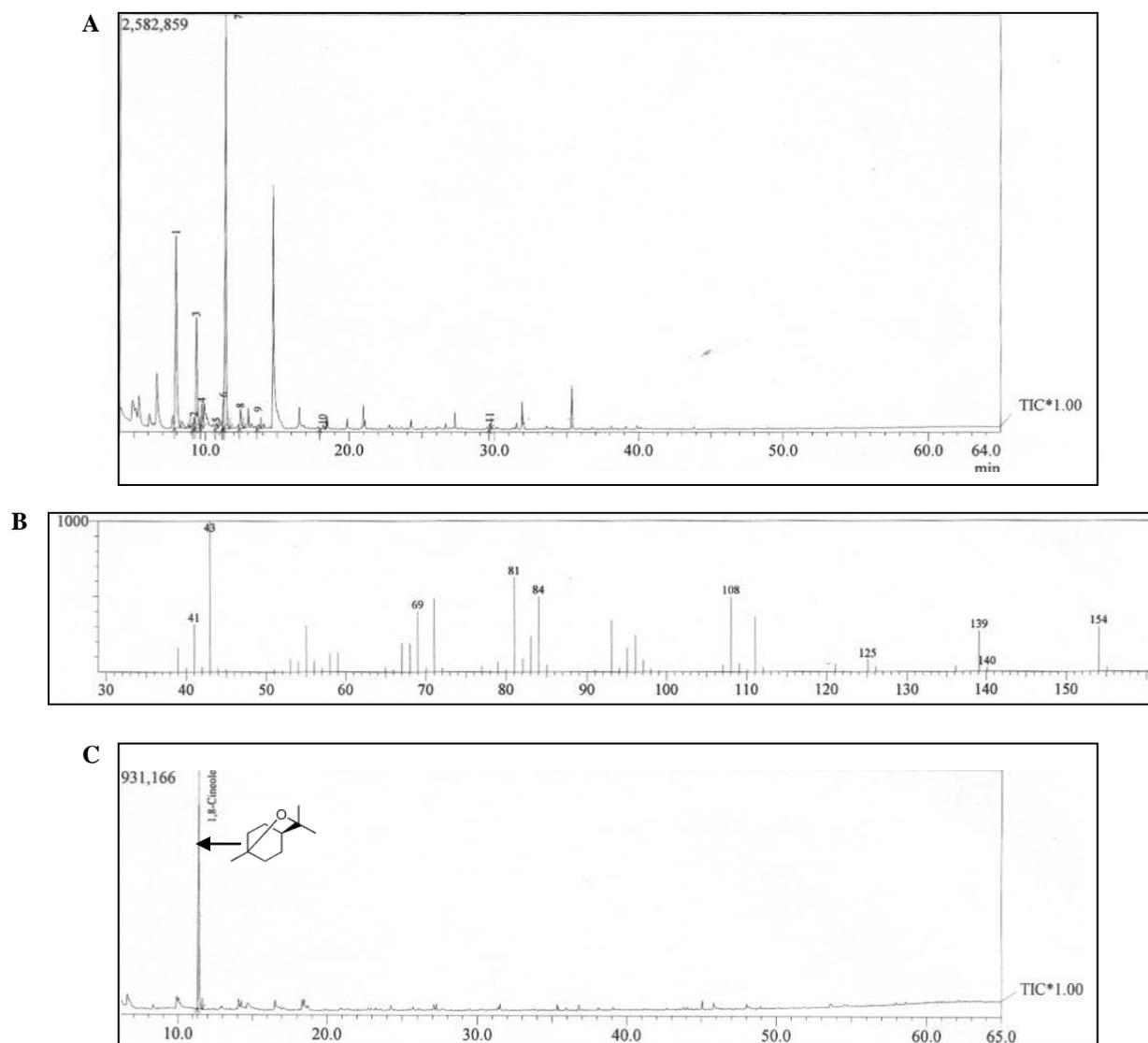


Figure III. 24. The Gas Chromatography/ Mass Spectrometry analysis of volatiles produced by AM63 yeast cells expressing the 215R monoterpene synthase. A) Chromatogram of detected peaks; B) Mass spectrum of 1,8-cineole detected in AM63 yeast cells; C) Chromatogram of 9.11 μ g of standard 1,8-cineole registered the corresponding peak at 11.35 minutes. The yeast cells overexpressing 215R enzyme produced a monoterpene profile of which mostly 1,8-cineole.

Table. III. 5. Monoterpene profile of 215 monoterpene synthase versus SfCinS1 cineole synthase.

Peak number	Compound name	215 Area (%)	SfCinS1 Area (%)
1.	α -Pinene	28.40	7.78
2.	Sabinene	1.06	-
3.	β -Pinene	14.50	8.10
4.	β -Myrcene	2.84	4.9
5.	α -Terpinene	0.73	-
6.	D-Limonene	2.97	-
7.	1,8-Cineole	45.52	75.09
8.	γ -terpinene	2.51	3.57
9.	(+)-2-Carene	0.54	-
10.	α -Terpineol	0.51	0.54

This new enzyme is significantly different from the canonical Cineole synthase 1 which shares 96 % homology to the corresponding enzyme from *Salvia officinalis*. It is noteworthy that a homologue of this gene has not been isolated from *S. officinalis*.

5.2. Sesquiterpene synthases are functional in yeast

5.2.1. Expression of clone P330 from *Salvia pomifera*

The P330 clone is full-length cDNA isolated from the *Salvia pomifera* cDNA library and shares 67 % homology to germacrene D synthase from *Ocimum basilicum*, 58 % homology to selinene synthase from *Ocimum basilicum*, 49 % homology to delta-cadinene synthase from *Ricinus communis*, 42 % homology to beta-caryophyllene synthase from *Artemisia annua*, and 41 % homology to beta-cubebene from *Magnolia grandiflora*. The deduced protein sequence of P330 clone and of previously reported sesquiterpene synthases were analyzed by multiple alignment using Vector NTI software showing close relationship of the P330 enzyme with the other sesquiterpene synthases. Phylogenetic tree of the aligned protein sequences is shown below (Fig. III. 25 and III. 26).

However, this phylogenetic relationship cannot predict the catalytic function of P330 sesquiterpene synthase, whose activity was further analysed by heterologous expression in yeast

cells as phylogenetic closeness of the enzymes reflects rather the closeness of the species rather than the catalytic specificity.

Figure III. 25. Phylogenetic tree resulting from the multiple protein alignment of the isolated P330 clone predicted protein sequences expressed in yeast cells with other sesquiterpene synthases previously reported.

		1		50
AAL79181	(1)	-----MSVKEEKVIRPIVHFPPSVWADQFLIFDDKQAEQANV		
CAA77191	(1)	--MASQASQVLA S PHPAISSEN RPKAD FHPGIWGD MFI ICP--DTDIDAA		
AAV63785	(1)	-MSANCVSAAPT SPKNSDVEEIRKSATYHSSVWGNHFLSYTSDVTEITAA		
AY693644	(1)	-MTNMFASAAPI STNNTTVEDMRRSVTYHPSVWKDHFLDYASGITEV---		
clone P330	(1)	HMAEIIYASAVPI STKNTNVENTRRSVTYHPSVWRDHFLKYTDDVTKITTA		
ACC66281	(1)	-----MDSP T TQRPNMEIGRAFVNYHPSIWGEHFI AASPDVMRLDAH		
EEF39510	(1)	MSLQVSAVPIKT STQNATSAVKRHSSTYHPTIWGDHFLANLSHSKIIDGS		
Consensus	(1)	AS ST N VE R SVTYHPSVWGDHFL Y VTEI AA		
		51		100
AAL79181	(38)	EQVVNELREDVRKDLVSSLDVQTEHTNLLKLIDAIQRLGIAYHFEEIEQ		
CAA77191	(47)	TELYEELKAQVRKMIPEPVD--DSNQKLPFIDAVQRLGVSYHFEKEIED		
AAV63785	(50)	EKEQLEKLKEKVKNLLAQTPD--ESTGKME LIDAIQRLGVGYHFTTEIQE		
AY693644	(47)	EMEQ LQKQKERIKTLLAQTLD--DFVLKIELIDAIQRLGVGYHFEKEINH		
clone P330	(51)	EKQVLEKEKEDVKLLIAQTPD--DSTVKIELIDAIQRLGVAYHFPKEIDE		
ACC66281	(43)	KGRGEE-LKEVVRNMFSTVN---DPLLKMN LIDAIQRLGVAYHFEMEIDK		
EEF39510	(51)	IEQ QF EGLKQKVRKMIIDLNN--EPCKKLGLIDAVQRLGVGYHFKSEIED		
Consensus	(51)	E Q E LKE VK LIA D D KL LIDAIQRLGVAYHFE EIED		
		101		150
AAL79181	(88)	ALQHIYDTYG-----DDWKGRSPSLWFRILRQQGFYVSCDIFKNYKKEDEG		
CAA77191	(95)	ELENIYRDTN NN--DADTDLYTALRFRLLRHGHGFDISCDAFNKFKDEAG		
AAV63785	(98)	SLRQIHEG-QIRN-DDDDVR-VVALRFRLLRQGGYRAPCDVFEKFMDDGG		
AY693644	(95)	SLRQIYDTFQISS-KDNDIR-VVALRFRLLRQHGYPVPSDVFKKFIDNQG		
clone P330	(99)	SLQKIHDYQTQSRKDKDDARVLALRFRLLRQGGYRVTSDVFNGLVDEEG		
ACC66281	(89)	ALGQMYDDHINGK--DDGFDLQTLALQFRLLRQGGYNVSSGVFAKFKDDEG		
EEF39510	(99)	VLQKVYHDYS-----DDEDDLNTVALRFRLLRQHGIKVS CAIFEFKFKDSEG		
Consensus	(101)	SL QIYD YQ DDDDD TVALRFRLLRQ GY VSCDVF KFKDEEG		
		151		200
AAL79181	(133)	SFKESLTNDVEGLLELYEATYLRVQEGVLDDALVFTRTCLEKIAKDL-V		
CAA77191	(143)	NFKASLTSDVQGLLELYEASYMRVHGEDI LDEAISFTTAQLTLALP----		
AAV63785	(145)	NFKESLKKDVEGMLSLEYEASYGIDGEEIMDKALEFSSSHLESMLHNIST		
AY693644	(143)	RLDESVMNNVEGMLSLEYEASNYGMEGEDILDKALEISTSHLEPLASRS--		
clone P330	(149)	NLKEWLISDVEGMLSLEYEASNYGTNEEEILEKILQSTSSHLESLLP----		
ACC66281	(138)	NFSSILSKDTHGLLSLEYEAF LGTHGDDILDEAITFTTVHLKSTLP----		
EEF39510	(145)	NFKTSLINDALGMLSLEYEATHLSIRGEDVLDEALAF TTTNLQSVLP----		
Consensus	(151)	NFKESL DVEGMLSLEYEASYLGI GEDILDEAL FTTSHLESLLP		

		201		250
AAL79181	(182)	HTNPTLSTYIQEALQKPLHKRLTRLEALRYIPMYEQQASHNESLLKLAKL		
CAA77191	(189)	TLHHPQLSEQVGHALKQSIRRGLPRVEARNFISIIYQDLESHNKSLLQFAKI		
AAV63785	(195)	KTNKSLLRRLQEAALDTPISKAARLIGATKFIISTYREDESHNEDLLNFAKL		
AY693644	(191)	-----RRINEALEMPISKTLVRLGARKFISIIYEDESRDEDLLKFAKL		
clone P330	(195)	QMSTLSKRVKEALEMPISKTLMRLGAKKYIPMYQEIESHNELLLNFAKL		
ACC66281	(184)	HVSAPLTKLVELALEIPLHRRMERLQTRFYISIIYEEDRERNDVLLLEFSKL		
EEF39510	(191)	QLNTHLAAQISRALNRPIRKYLPRLLEARNYIDIYATEESYNTLLNFAKL		
Consensus	(201)	L LSK I EALE PI K L RL ARKYISIIY EDESHNE LLNFAKL		
		251		300
AAL79181	(232)	GFNLLQSLHRKELSEVSRWVKGLDVPNNLPYARDRMVECYFWALGVYFEP		
CAA77191	(239)	DFNLLQLLHRKELSEICRWKDLDFTRKLPFARDRVVEGYFWIMGVYFEP		
AAV63785	(245)	DFNILQKMHQE EANYLTRWVEDLDLASKLDFARDRMVESYFWSLGVYFQP		
AY693644	(234)	DFNILQKIHQEE LTHIARWVKELDLGNKLPFARDRVVECYFWILGVYFEP		
clone P330	(245)	DFNIMQKIHQRELNHITRWVDFEFGKKWLLQEIWWSATFGFWETILSH		
ACC66281	(234)	EFRLQSLHQRELRDISLWVKEMDLLAKLPFTRDRVLEGYFWTVGVYFEP		
EEF39510	(241)	DFNMLQELHQKELNVVTKWVKSLDVATKLPYARDRVVECYFWMVGVYFEP		
Consensus	(251)	DFNILQ LHQKEL ITRWVKDLDLA KLPFARDRVVE YFW LGVYFEP		
		301		350
AAL79181	(282)	KYSQARIF-----LAKVISLATVLDPTYDAYGTYEELKIFTEAI		
CAA77191	(289)	QYSLGRKM-----LTKVIAMASIVDDTYDSYATYDELIPYTNAI		
AAV63785	(295)	QYRTSRIY-----LTKIISIVAVIDDYDVYGSFDDLSRFTDVI		
AY693644	(284)	QYNIARRF-----MTKVIAMTSIIDDIYDVHGTLEELQRFDAI		
clone P330	(295)	NMPHQEYS-QKS-L-YPSLMISMFMEL-MNSDA-PMLFEDGILALPMNC		
ACC66281	(284)	HYSRARM-----MTKMIAFATVMDDTYDVYGTLEEELELLTATI		
EEF39510	(291)	QYSFARIM-----MTKIIAITSLDDTYDNYATGEELEILTEAI		
Consensus	(301)	QYS ARI MTKVIAM SVIDD TYD YGT EEL FTDAI		
		351		400
AAL79181	(321)	QRWSITCIDMLPEYLKLLYQGVLDIYIEMEEIMGKEG--K--AHHLSYAK		
CAA77191	(328)	ERWDIKCMNQLPNYMKISYKALLDVYEEEMEQLLANQG--R--QYRVEYAK		
AAV63785	(334)	QSWKISNADELPPYMRICFEALLGIYEDMGDRIG-----APYAI		
AY693644	(323)	RSWDIRAIDELPPYMLCYEALLGMYAEMENEMVKQN--Q--SYRIEYAR		
clone P330	(340)	HRI-EYATKLF-VFMRKWKLKQNEANHIAFNMRKKR--N-WQHTWKRQN		
ACC66281	(323)	ERWNRGDMQLPDYMKVIFIALLDGVDATEDDLTGEG--K--SYRIYYLK		
EEF39510	(330)	ERWDIKAKDALPEYMKIITYTLLDIYNEYEENIAKEE--KSLLYSVYYAK		
Consensus	(351)	RW I ID LP YMKI Y ALLDIY EMED MAK K Y I YAK		
		401		450
AAL79181	(367)	ESMKEFIRSYMMEAKWANEGYVP-----TAEHMSVAFVSSGYSML		
CAA77191	(374)	KAMIRLVQAYLLEAKWTHQNYKP-----TFEEFRDNALPTSGYAML		
AAV63785	(373)	DTMKELVDTYMQEAEWCYTEYVP-----TVDEYMKVALVTGGYLMV		
AY693644	(369)	QEMIKLVTTYMEAKWCYSKYIP-----NMDEYMKLALVSGAYMML		
clone P330	(385)	GVTASMFQQ-RSI-NWHSFPALT-CLQHLL-LEWKTPLPNMILIGSQTNH		
ACC66281	(369)	EAVKDLAKAYLAEARWVSSGYVP-----TSEEYMKVALISAVYPML		
EEF39510	(378)	EVMKRVVRAYLAEVRWRDNCYTP-----TMEEYMQSALLTTCSPML		
Consensus	(401)	E MK LV AYL EAKW YVP T EEYM VALVSSGY ML		
		451		500
AAL79181	(408)	ATTCFVGMG-D-----IVTDEAFKWALT-KPPIIKASCAIARLMDDI		
CAA77191	(415)	AITAFVGMG-E-----VITPETFKWAAS-DPKI IKASTIICRFMDDI		
AAV63785	(414)	ATTFLTGIN-----NITKKDFDWIRN-RPRLLQVAEVLTRLMDDI		
AY693644	(410)	ATTSLVGILGD-----PITKQDFDWITN-EPPILRAASVICRLMDDV		
clone P330	(431)	QFYKLPRLSVD-WTIW-DTGLSKKSQAWIVT-KRMIAQRKLLVNFGE-		
ACC66281	(410)	FVAFVIGMD-E-----VVTKEVLEWAIH-MPTMLRTCIVARLMDDI		
EEF39510	(419)	AIASFVGLK-E-----IATKEAYEWASE-DPKIIRASSIVCRLMDDI		
Consensus	(451)	A T LVGM D IITKE FDWA P IIRASSII RLMDDI		
		501		550
AAL79181	(448)	HSQKEE-----KERIHVASSVESYMKQYDVTTEEHVLKVF		
CAA77191	(455)	AEHKFN-----HRREDDCSAIECYMKQYGVTAQEAAYNEF		

AAV63785	(453)	AGHGTE-----KK---TTAVSCYMKKEYECSEMEASREL
AY693644	(451)	VGHGIE-----QK---ISSVDCYMKENGCSKMEAVGEF
clone P330	(477)	IKHGRI-MRSASSQGRHLCRFSRALSILLASSIYYMLMKMRILIPRLGPK
ACC66281	(450)	PSNKLE-----QERKHVSSSVECYMKEHGTSYHESIQKL
EEF39510	(459)	VSHEFE-----QTRKHVASGVECYIKQYGAEEEEVIKLF
Consensus	(501)	H E R V SSVECYMK YG S EAI F

Figure III. 26. Alignment of deduced amino acid sequences of the P330 sesquiterpene synthase clone from *Salvia pomifera* with several reported sesquiterpene syntases. Where, AAV63785 – selinene synthase [*Ocimum basilicum*], AY693644 – germacrene D synthase [*Ocimum basilicum*], ACC66281 – beta-cubebene synthase [*Magnolia grandiflora*], AAL79181 – beta-caryophyllene synthase QHS1 [*Artemisia annua*], EEF39510 – (+)-delta-cadinene synthase [*Ricinus communis*], CAA77191 – (+)-delta-cadinene synthase [*Gossypium arboreum*]. The alignment was created with the vector NTI software.

To characterize the gene, the P330 clone was digested with EcoRI and XbaI from the pDNR-library vector and subcloned it into the pYES2 vector linearized with the same restriction enzymes (Fig. III. 27) (see Chapter II, paragraph 8.6)

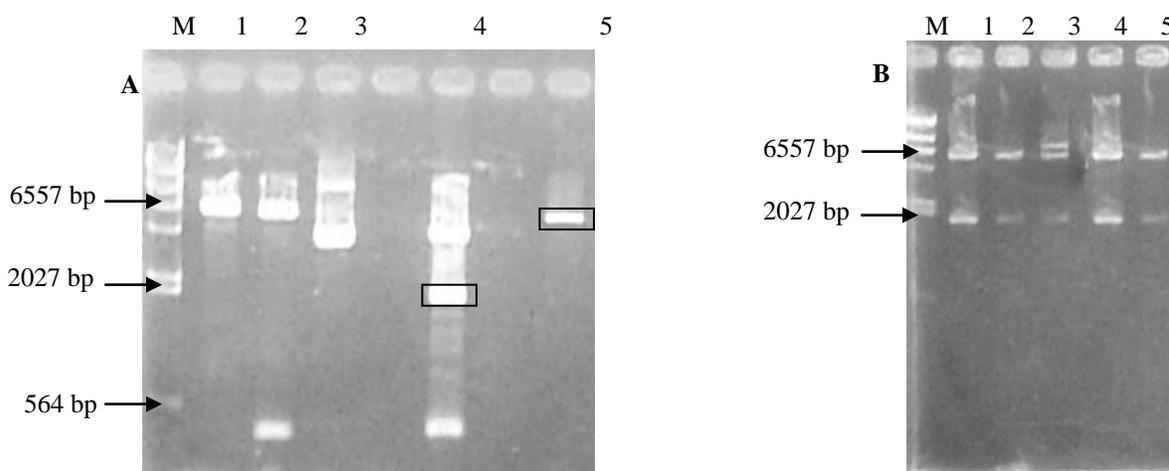


Figure III. 27. Cloning of P330 into yeast vector pYES2: **A)** Restriction digestion of pDNR-lib/P330 and pYES2 with EcoRI and XbaI, where M - Gene ruler, λ DNA/*Hind*III marker; 1- pDNR-lib/P330 cut EcoRI; 2- pDNR-lib/P330 cut XbaI; 3- pDNR-lib/P330 uncut; 4- pDNR-lib/P330 cut EcoRI-XbaI; 5- pYES cut EcoRI-XbaI; **B)** Agarose gel electrophoresis of pYES/P330 digested with EcoRI-XbaI (1-5).

The new construct pYES/P330 was successfully introduced into AM63 and EG60 yeast strains by lithium acetate transformation and induced for expression of P330 protein. The accumulated volatiles were entrapped by SPME and analyzed by GC-MS.

The sesquiterpene synthase homologue encoded by the P330 clone was able to produce a sesquiterpene in yeast cells using the endogenous FPP pool (fig. III. 20) acting as a multi-product

enzyme not unusual for a terpene synthase. The P330 sesquiterpene profile uncovered by GC-MS analysis consist of four major products, such as: 35.52 % δ -cadinene with the retention time at 32.53 min, 12.53 % α -copaene with the retention time at 26.43 min, 9.75 % trans-caryophyllene with the retention time at 28.32 min, and 9.16 % α -cubebene with the retention time at 25.26 min. Nine more minor sesquiterpene products were obtained. All compounds were identified by comparing the target spectra with reference spectra using Wiley7 library (Appendix 2) with high similarity (Fig. III. 28 and III. 29).

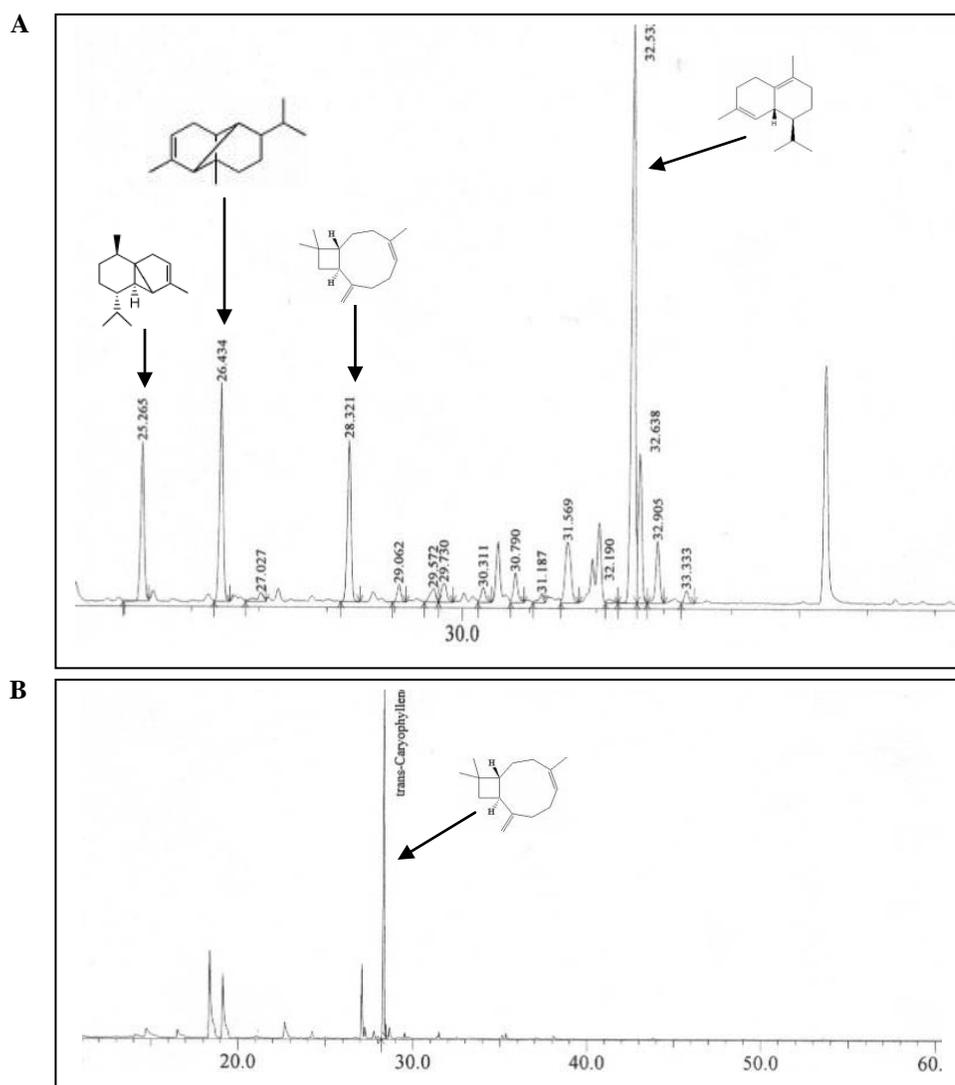


Figure III. 28. GS-MS profile of the sesquiterpenes produced by AM63 yeast strain expressing P330 clone. **A)** The four major peaks corresponde to α -cubebene (25.26 min), α -copaene (26.43 min), trans-caryophyllene (28.32 min), and δ -cadinene (32.53 min); **B)** Chromatogram of 9.02 μ g of standard trans-caryophyllene detected the corresponding peak at 28.31 retention time.

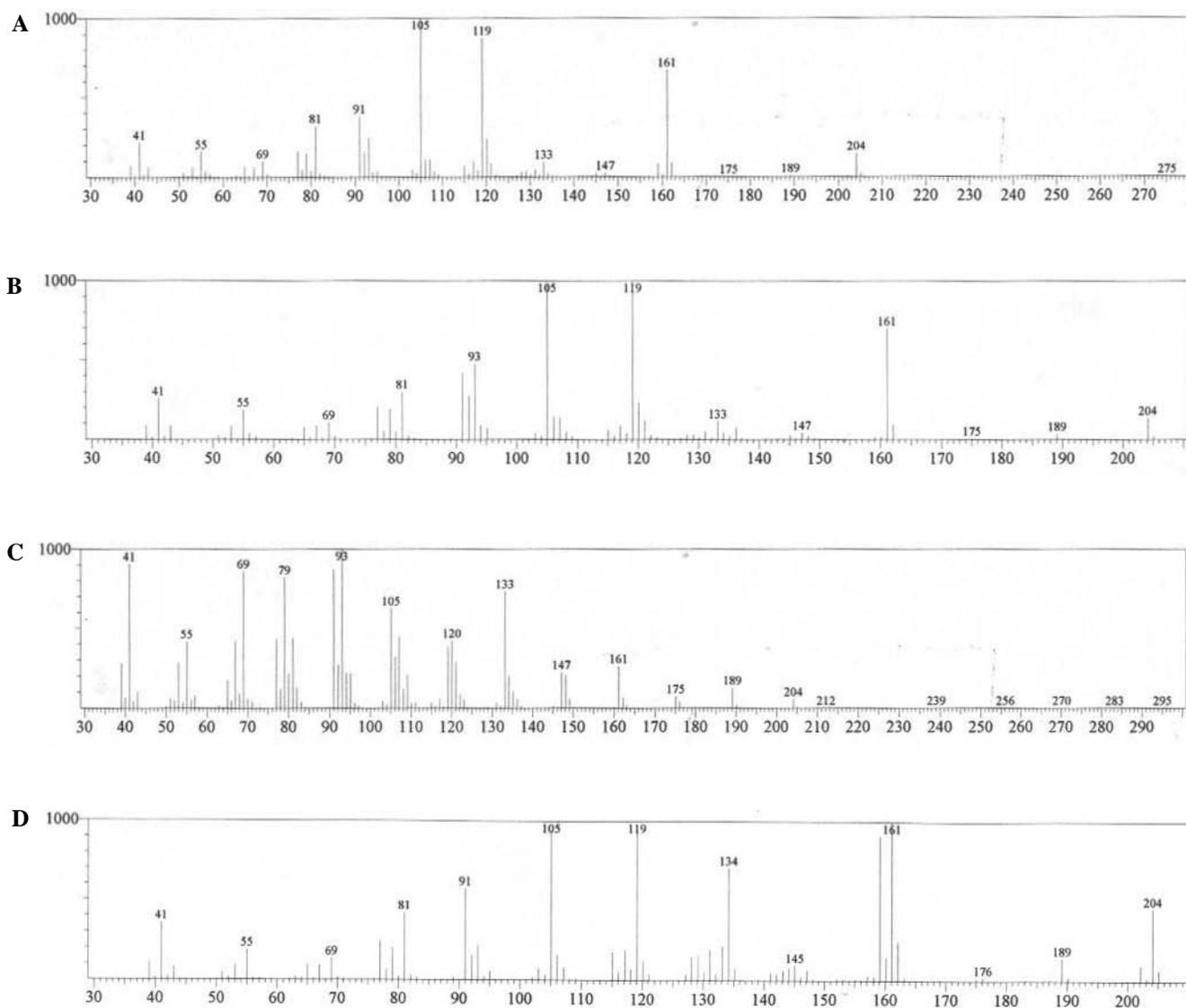


Figure III. 29. Mass spectra of target products. A) Mass spectrum of peak detected at 25.26 min retention time and identified as α -cubebene; B) Mass spectrum of peak detected at 26.43 min retention time and identified as α -copaene; C) Mass spectrum of peak detected at 28.32 min retention time and identified as trans-caryophyllene; D) Mass spectrum of peak detected at 32.53 min retention time and identified as δ -cadinene.

The sesquiterpene profile is presented in Table III. 6.

Table. III. 6. Sesquiterpene profile of P330 sesquiterpene synthase.

Peak number	Compound name	Retention time	Area (%)
1.	α -Cubebene	25.265	10.57
2.	α -Copaene	26.434	14.45
3.	β -Cubebene	27.027	1.12
4.	trans-Caryophyllene	28.321	11.25
5.	α -Guaiene	29.062	1.24
6.	α -Humulene	29.730	1.86
7.	γ -Gurjunene	30.790	2.40
8.	β -Selinene	31.187	0.87
9.	γ -Cadinene	32.190	0.45
10.	δ -Cadinene	32.532	40.97
11.	δ -Cadinene	32.638	9.19
12.	Cadina-1,4-diene	32.905	4.58
13.	α -Calacorene	33.333	1.00

5.2.2. Expression of the putative terpene synthase clone 1025 from *Salvia fruticosa*

Among the genes identified as sesquiterpene synthases from *Salvia fruticosa*, the clone 1025 was selected for further bacterial and/or yeast expression studies. The 1025 deduced protein showed 54 % similarity to germacrene D synthase from *Ocimum basilicum*, 48 % similarity to selinene synthase from *Ocimum basilicum*, 43 % similarity to delta-cadinene synthase from *Ricinus communis*, 39 % and 36 % similarity to beta-farnesene synthase from *Citrus junos* and *Artemisia annua*, respectively (Fig. III. 30 and III. 31).

Figure III. 30. Phylogenetic tree resulting from the multiple protein alignment of the isolated 1025 clone predicted protein sequences expressed in yeast cells with other sesquiterpene synthases previously reported.

		1	50
AAV63785	(1)	-MSANCVSAAPTS PKNS-----DVEEIRKSATYHSSVWGNHFLSY	
AY693644	(1)	-MTNMFASAAPISTNNT-----TVEDMRRSVTYHPSVWKDHFLDY	
clone 1025	(1)	EFMEICSQPIPAIKKVK-----NLDEIRKSAKFHPSIWGDYFLQY	
AAX39387	(1)	-MSTLPISVSFSSTSPLVVDDKVKPDVIRHTMNFNASIWGDQFLTY	
AF374462	(1)	----MKDMSIPLLA AVS-----SSTEETVRPIADFHPTLWGNHFLKS	
EEF39510	(1)	--MSLQVSAVPIKTSTQ-----NATS AVKRHSSTYHPTIWGDHFLAN	
Consensus	(1)	MS L VSAIPISTS S TVEEIRKSATFHPSIWGDHFL Y	
		51	100
AAV63785	(40)	TSDVTEITAAEKEQLEKLKEKVNLLAQT PD-----ESTGKMELIDAIQR	
AY693644	(40)	ASGIT EV---EMEQLQKQKERIKTLLAQTLD-----DFVLKIELIDAIQR	
clone 1025	(41)	DSDKTKISDVQEELAKQKEMVKKLLAQT PN-----NSIYKMELIDAIQR	
AAX39387	(50)	DEPEDLV MK--KQLVEELKEEVKKELITIKGSNEPMQHVKLIELIDAVQR	
AF374462	(39)	AADVETIDAAATQEQHAALKQEVRRMITTAN-----KLAQKLHMIIDAVQR	
EEF39510	(41)	LSHSKIIDGSIEQQFEGLKQKVRKMIIDLNN-----EPCKKGLIDAVQR	
Consensus	(51)	SDVT I AAE EQLEKLKEKVNLLAQT N E V KIELIDAIQR	
		101	150
AAV63785	(85)	LGVG YHF TTEIQESLRQIHEG--QIRNDDDDVRVVALRFRLLRQGGYRAP	
AY693644	(82)	LGVG YHF EKEINHSLRQIYDTF-QISSKDNDIRVVALRFRLLRQHGYVPV	
clone 1025	(86)	LGVEYHF EKE LDES LQYIHKH--QNSKDDDDITTV ALRFRLLRQGGYNVP	
AAX39387	(98)	LGIAYHF EEEIEEALQHIHVTYGEQWVDKENLQSISLWFRLLRQGGFNVS	
AF374462	(84)	LGVA YHF EKEIEDELGKVSHDL-----DSDDL YV VSLRFRLFRQGGVKIS	
EEF39510	(86)	LGVG YHF KSEIEDVLQKVYHDYSD---DEDDLNTVALRFRLLRQHGIKVS	
Consensus	(101)	LGVG YHF EKEIEESLQ IH Y Q DDDDL VVALRFRLLRQGGYKVS	
		151	200
AAV63785	(133)	CDVF EK FMD DGGNFKESLKKDVEGMLS LYEASYYGIDGEEIMDKALEFSS	
AY693644	(131)	SDVFKKFIDNQGR LDESVMNNVEGMLS LYEASNYGMGEDILDKALEIST	
clone 1025	(134)	CDVFRKFIDSEGNFVASLKNVVEGLNLNLYEAAYLGTHGEEILERA IQFCS	
AAX39387	(148)	SGVFKDFMDEKGFKESLKNDAQGILALYEA AFMRVEDETILDNALEFTK	
AF374462	(129)	CDVFDKFKDDEGKFESLINDIRGMLS LYEAYLAIRGEDILDEAIVFTT	
EEF39510	(133)	CAIF EK F K DSEGNFKTSLINDALGMLS LYEATHLSIRGEDVLDEALAF TT	
Consensus	(151)	CDVFKKFIDDEGNFKESLINDVEGMLS LYEAYLGIEGEDILDKALEFTT	
		201	250
AAV63785	(183)	SHLESMLHNIS-TKTNKSLLRRLQEQALDTPISKAAIRLGATKFI STYRED	
AY693644	(181)	SHLEPLASRS-----RRINEALEMPI SKTLVRLGARKFISIYEED	
clone 1025	(184)	SHLHTSLHKIT-DVS---LSKR VNEALKMPNRKSLTRLGARKFISVYEED	
AAX39387	(198)	VHLDI IAK---DPSCDSSLRTQIHQALKQPLRRRLARIEALHYMPIYQQE	
AF374462	(179)	THLKSVISISDHS HANSNLAEQIRHSLQIPLRKAARLEARYFLDIYSRD	
EEF39510	(183)	TNLQSVL-----PQLNTHLAAQISRALNRP IRKYLPRL EARNYID IYATE	
Consensus	(201)	SHLESVL NS LARRINEAL MPIRKAL RLGARKFISIIY ED	
		251	300
AAV63785	(232)	ESHNEDILNFAKLDFN ILQKMHQE EANYLTRWWE DL DLASKLDFARDRMV	
AY693644	(221)	ESRDEDLLKFAKLDFN ILQKIHQEELTHIARW WKELDLGNKLPFARDRVV	
clone 1025	(230)	ESHNETLLNFAKLDFN ILVQKMHQRELSDATRWKKLEVANKMPYARDRIV	
AAX39387	(245)	TSHDEVLLKLAKLDFSVLQSMHKKELSHICKWKKDL DLQNKLPYVRDRVV	
AF374462	(229)	DLHDETLLKFAKLDFN ILQAAHQKEASIMTRW WNDLGFPKKVPYARDRII	
EEF39510	(228)	ESYNTTLLNFAKLDFNMLQELHQKELNVVTKW WKS LDVATKLPYARDRVV	
Consensus	(251)	ESHNETLLNFAKLDFN ILQKMHQKELSHITRWKDL DLANKLPYARDRVV	
		301	350
AAV63785	(282)	ESYFWSLG-VYFQPQYRTSRIYLTKIISIVAVIDD IYDVYGSFDDLRSFT	
AY693644	(271)	ECYFWILG-VYFEPQYNIARRFMTKVIAMTSIIDDIYDVHGTLEELQRFT	
clone 1025	(280)	ECFLWVVG-VYFEPQYANARRILVKALSIIASIIDDTY-EYATLHELQILT	
AAX39387	(295)	EGYFWILS-IYYEPQHARTRMFLMKT CMWLVLVDDTFDNYGTYEELIFT	
AF374462	(279)	ETYIWMLLGVSYPEPNLAFGRIFASKVVCMITTIDDTFDAYGTFEELTFT	
EEF39510	(278)	ECYFWMVG-VYFEPQYSFARIMTKIIAITSLLDDTYDNYATGEELEILT	
Consensus	(301)	ECYFWILG VYFEPQYA ARIFLTKIIAIIISIIDDTYD YGTFEEL IFT	
		351	400

AAV63785	(331)	DVIQSWKISN-ADELPPYMRICFEALLGIYEDMGDRIG-----APY
AY693644	(320)	DAIRSWDIRA-IDELPPYMRICYEALLGMYAEMENEMVKQNQ--SYRIEY
clone 1025	(328)	QAVRWVDATMEDSPPYIQMICYRSLIETYVEIEDEMEITGE--SHRVQY
AAX39387	(344)	QAVRWSISC-LDMLPEYMKLIYQELVNLHVEMEESLEKEGK--TYQIHY
AF374462	(329)	EAVTRWDIGL-IDTLPEYMKFIVKALLDIYREAEFEELAKEGR--SYGIPY
EEF39510	(327)	EAIERWDIKA-KDALPEYMKIITYTTLLDIYNEYEENIAKEEKSLLYSVYY
Consensus	(351)	DAIERWDI A IDELPPYMKI IY ALLDIY EMEEEIAKEGK SY I Y 401 450
AAV63785	(371)	AIDTMKELVDTYMQEAEWCYTEYVPTVDEYMKVALVTGGYLMVATFLTG
AY693644	(367)	ARQEMIKLVTTYMEEAKWCYSKYIPNMDEYMKLALVSGAYMMLATTSLVG
clone 1025	(376)	AIQDMKKLGMAFYFEVVKWLYNNYIPTLKEYMKVSLVTSGYMMASSTSVVG
AAX39387	(391)	VKEMAKELVRNLYVEARWLKEGYMPTLEEYMSVSMVTGTYGLMIARSYVG
AF374462	(376)	AKQMMQELIILYFTEAKWLYKGYVPTFDEYKSVALRSIGLRTLAVASFVD
EEF39510	(376)	AKEVMKRVRAYLAEVRWRDNCYTPTMEEYMQSALLTTCSPMLAIASFVG
Consensus	(401)	AKQMMKELV YL EAKWLY YIPTLDEYMKVALVTGGYMLLATTSLVG 451 500
AAV63785	(421)	IN--NITKKDFDWRIRNRPLLQVAEVLTRLMDDIAGHGTEKK----TAV
AY693644	(417)	ILGDPITKQDFDWTINPPILRAASVICRLMDDVVGHGIEQK----ISSV
clone 1025	(426)	MG-DRVRKEDMDWIINEPLIVRASSIICRITDDLGVGDEYEEK----PSWL
AAX39387	(441)	R-GDIVTEDTFKWVSSYPPIIKASCVIVRLMDDIVSHKEEQERGHVASSI
AF374462	(426)	LGDFIATKDNFECILKNAKSLKATETIGRLMDDIAGYKFEQQRGHNP SAV
EEF39510	(426)	L-KEIATKEAYEWASEDPKIIRASSIVCRLMDDIVSHEFEQTRKHVASGV
Consensus	(451)	I DIITKDDFDWISNEPKILRASSVICRLMDDIVGH FEQQR H SAV

Figure III. 31. Alignment of deduced amino acid sequences of the 1025 sesquiterpene synthase clone from *Salvia fruticosa* with several reported sesquiterpene syntases. Where, AAV63785 – selinene synthase [*Ocimum basilicum*], AY693644 – germacrene D synthase [*Ocimum basilicum*], AAX39387 – (E)-beta-farnesene synthase [*Artemisia annua*], AF374462 – (E)-beta-farnesene synthase [*Citrus junos*], EEF39510 – (+)-delta-cadinene synthase [*Ricinus communis*]. The alignment was created with the vector NTI software.

The pJG4-4/1025 construct, kindly provided by Dr. Antonios Makris, was transformed and expressed into AM63 and EG60 yeast strains. Several attempts to test the activity of 1025 clones were not successful. When AM63 yeast cells overexpressing 1025, were incubated for a period of three days in buffer a slight sesquiterpene synthase activity could be detected (Fig. III. 32).

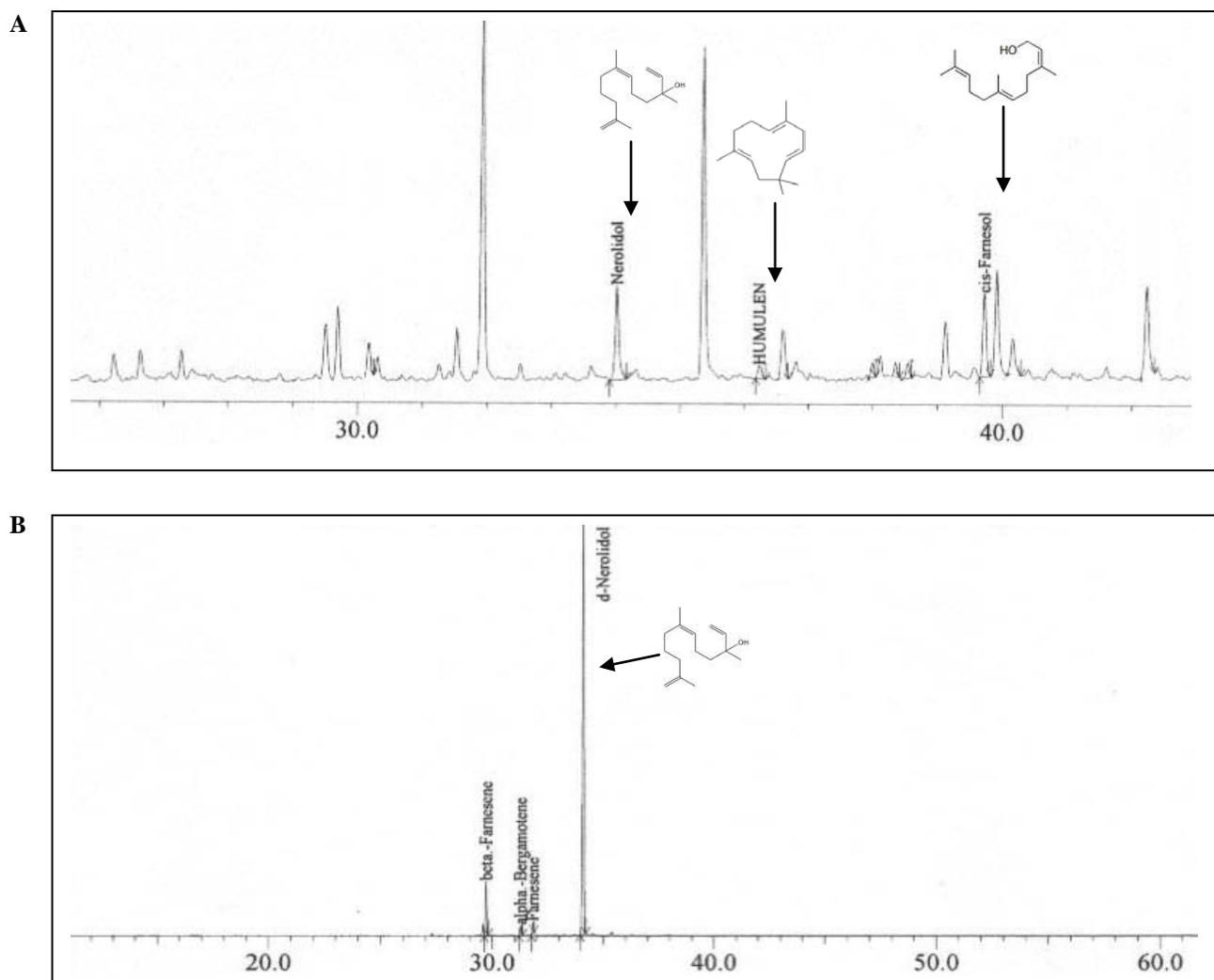


Figure III. 32. The GC-MS profile of the sesquiterpene produced by AM63 strain expressing 1025 clone after three days of incubation. A) The identified peaks correspond to nerolidol (34.01 min), humulene (36.24 min), and cis-farnesol (39.72 min); **B)** Chromatogram of 8.4 μ g of standard d-nerolidol recorded the corresponding peak at 34.09 minute .

The comparison of the target spectra of detected compounds with standard spectra from the Wiley7 GC-MS spectra library (Appendix 2) allowed the identification of three sesquiterpenes with high similarity, which are presented in Table III. 7.

Table. III. 7. Sesquiterpene profile of 1025 sesquiterpene synthase (3rd day of incubation on buffer).

Peak number	Compound name	Retention time	Area (%)
1.	Nerolidol	34.018	51.21
2.	Humulene	36.241	7.39
3.	cis-Farnesol	39.724	41.38

The enzyme activity was further confirmed by conventional enzymatic assay (see Chapter II, subchapter 12). The induced AM63 yeast cultures were disrupted by sonication (see Chapter II, paragraph 11.2.2). Yeast lysates were subsequently incubated with exogenous FPP and the reaction volatile products further evaluated GC-MS (Fig. III. 33). 1025 protein from broken yeast cells was active when exogenous FPP was added, producing small amount of beta-farnesene (29.65 minutes retention time), delta-cadinene (32.47 minutes retention time), and nerolidol (33.97 minutes retention time).

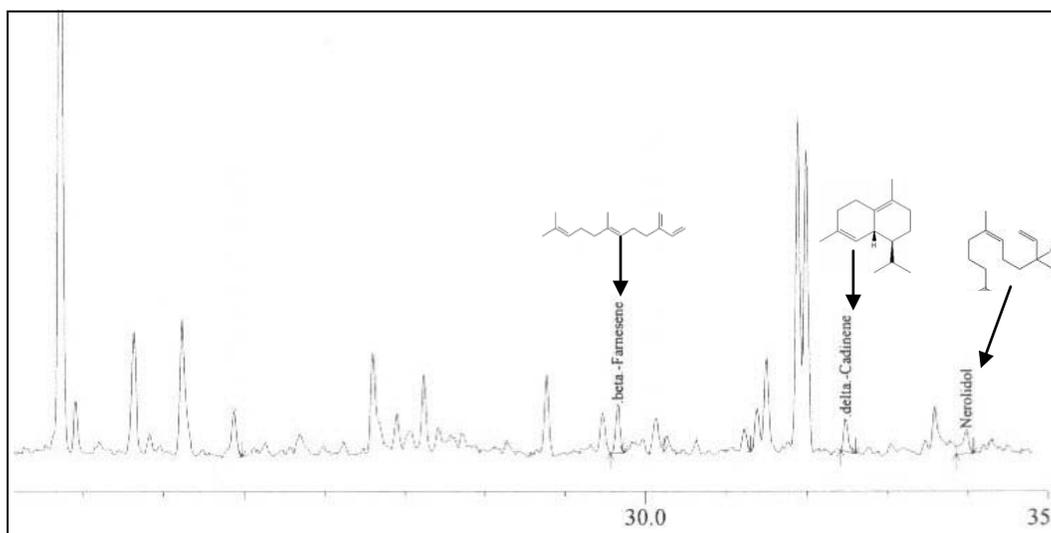


Figure III. 33. The GC-MS analysis of yeast cells expressing 1025 and evaluated by conventional enzymatic assay. Three peaks were identified as beta-farnesene (29.65 min), delta-cadinene (32.47 min) and nerolidol (33.97 min).

When the 1025 clone was expressed into the transformed AM74 and AM75 yeast strains, two out of the three products obtained by conventional enzymatic assay were identified (Fig. III. 34 and III. 35). One of them was beta-farnesene detected at 29.69 minutes retention time and the

other was nerolidol detected at 34.01 minutes retention time. Although both products were produced in small quantity, in the case of nerolidol previously identified in modified AM63 strain a 2 fold increase in the peak area was registered in AM74 yeast strain and 14 fold increase of peak area was registered in AM75 yeast strain, both strains being two steps further modified versions of AM63 (Table III. 8).

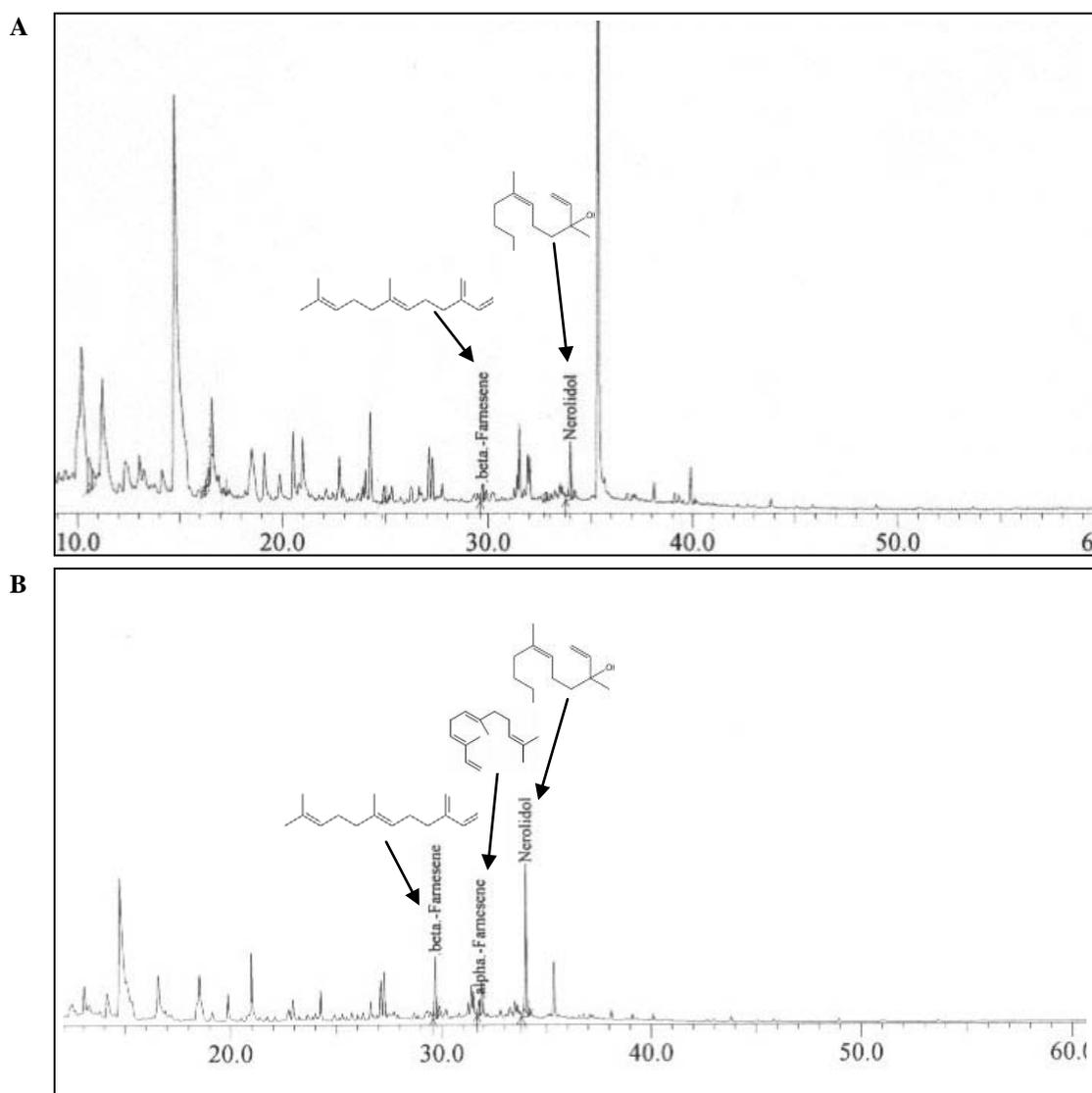


Figure III. 34. The GC-MS profile of the sesquiterpene produced by 1025 clone. A) In strain AM74 expressing the 1025, the identified peaks correspond to beta-farnesene (29.69 min) and d-nerolidol (34.01 min); **B)** In modified strain AM75 expressing 1025, the identified peaks correspond to beta-farnesene (29.58 min), alpha-farnesene (31.66 min), and nerolidol (33.99 min).

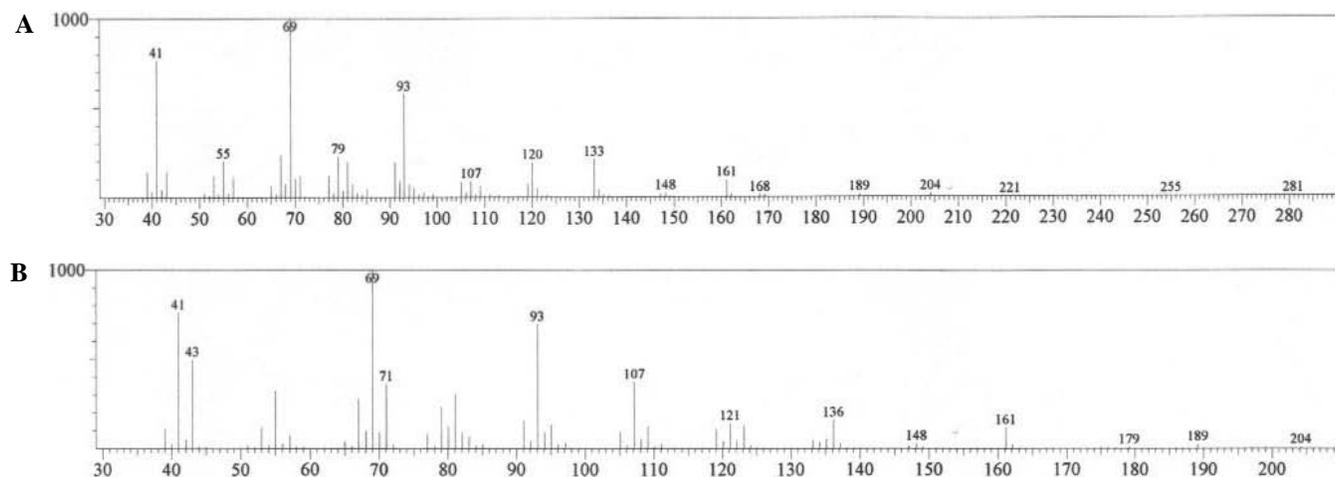


Figure III. 35. Mass spectra of target products detected in AM75 strain expressing 1025 protein. **A)** Mass spectrum of peak detected at 29.58 min retention time and identified as β -farnesene; **B)** Mass spectrum of peak detected at 33.99 min retention time and identified as δ -cadinene.

Table. III. 8. Comparison of sesquiterpene profile of 1025 enzyme *in vivo* expressed in AM63 and AM74.

Peak number	Compound name	Retention time	AM63 Peak area	AM74 Peak area	AM75 Peak area
1.	β -Farnesene	29.69	-	173,695	1,602,942
2.	α -Farnesene	31.77	-	-	536,962
3.	Nerolidol	34.01	265,771	595,032	3,798,631
4.	Humulene	36.24	38,401	-	-
5.	cis-Farnesol	39.72	214,802	-	-

The peaks corresponding to beta-farnesene and nerolidol was detected at very similar retention time not only when the clone samples were examined *in vivo*, but also when it was assessed by conventional enzymatic assay of protein lysate. SPME assay of standard compound analyzed by GC-MS confirmed the appearance of nerolidol peak at 33.97 minute. The target spectra of the peaks registered at 29.69 minutes and at 34.01 minutes shown high similarity with the beta-farnesene and nerolidol, respectively spectra based on the Wiley7 GC-MS spectral library (Appendix 2). The results overall prove that the 1025 protein is as an active enzyme in yeast cells and it possesses beta-farnesene and nerolidol synthase activity.

6. Construction of novel molecular tools for engineering yeast metabolic pathways

To develop further a high yield laboratory strain and with upregulated levels of expression of critical biosynthetic genes leading to high isoprene unit formation, we introduced genetic alterations in the sterol biosynthetic pathway so as to maximize the availability of terpene synthase precursors, GPP and FPP. To this end we developed a set of new molecular tools that enable promoter integration for upregulation of genes involved in the desired pathways or chromosomal insertions of exogenous genes, using in both cases a recyclable cassette system that makes possible the successive insertion of a multitude of genes. The integration scheme employs the loxP-recognition sequence flanking the selection marker, which allows the subsequent excision of the selection marker after the homologous recombination process using the Cre recombinase, and subsequent recycling of the selection marker. At each stage of completion the generated strain can be evaluated for increased terpenoid production and the basic growth properties which ensure a durable strain.

6.1. Generation of a recyclable cassette construct for integrating genes into chromosomal genome of yeast and expressing them under the control of the Galactose promoter

The genes of interest which contribute to enhanced terpene synthesis could become stably integrated in extra copies in the yeast genome under the control of a strong promoter. To this end a cassette containing the Galactose promoter – epitope tag HA – multicloning site – termination sequence *cyc1* was generated as the initial step of construction of the recyclable cassette useful for chromosomal integration of exogenous genes (Fig. III. 36).

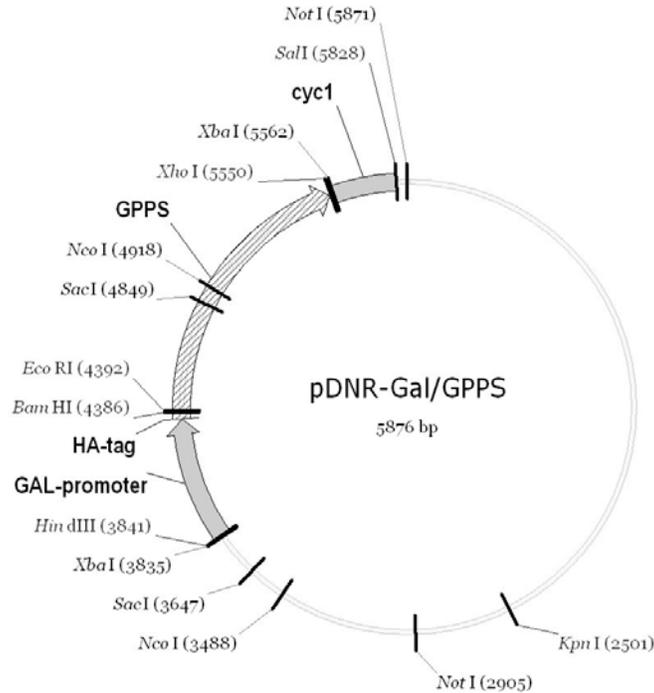


Figure III. 36. The pDNR-GAL-HA-GPPS plasmid map.

First, a pYES-HA plasmid bearing a hemagglutinin tag sequence was developed by ligating into the original pYES2 vector a ds oligonucleotide that contained the methionine codon and the sequence for the short hemagglutinin peptide which is recognized by a commercially available antibody. The complementary tails of the oligos were designed in a way so as to integrate into a HindIII and XhoI tails but to destroy the HindIII upon ligation. The pYES-HA vector was designed to have the BamHI and EcoRI as cloning fusion sites at 0 frame (see Chapter II, paragraph 7.1.1).

Proper insertion was verified by the absence of the HindIII sites. An agarose gel electrophoresis of the pYES-HA plasmid digested with Hind III showed an uncut plasmid, while digested with EcoRI revealed a linearized vector proving in this way the appropriate insertion of HA oligos into pYES vector (Fig. III. 37).

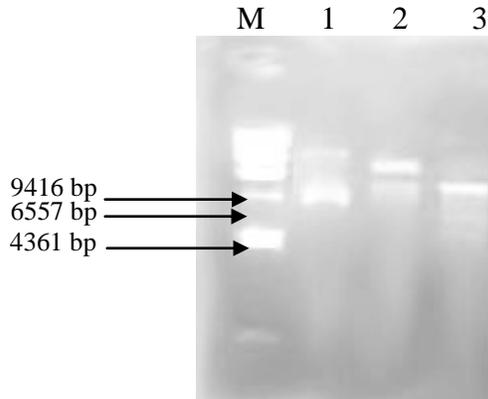


Figure III. 37. Agarose gel electrophoresis of pYES-HA vector, where M - Gene ruler, λ DNA/*Hind*III marker; 1- pYES-HA uncut - control; 2-pYES-HA cut *Hind*III showing uncut plasmid which confirmed the absence of the *Hind* III site; 3- pYES-HA cut *Eco*RI.

Subsequently, pYES-HA was utilized as a template in a PCR reaction to amplify the cassette containing the three elements necessary for generation of the desired integration vector. Specific primers for the 5' region of the GAL promoter and the 3' region of the *cyc1* termination sequence were used and the GALp-HA-*cyc1* fragment of approximately 900 bp was amplified (Fig. III. 38), gel purified and TOPO cloned.

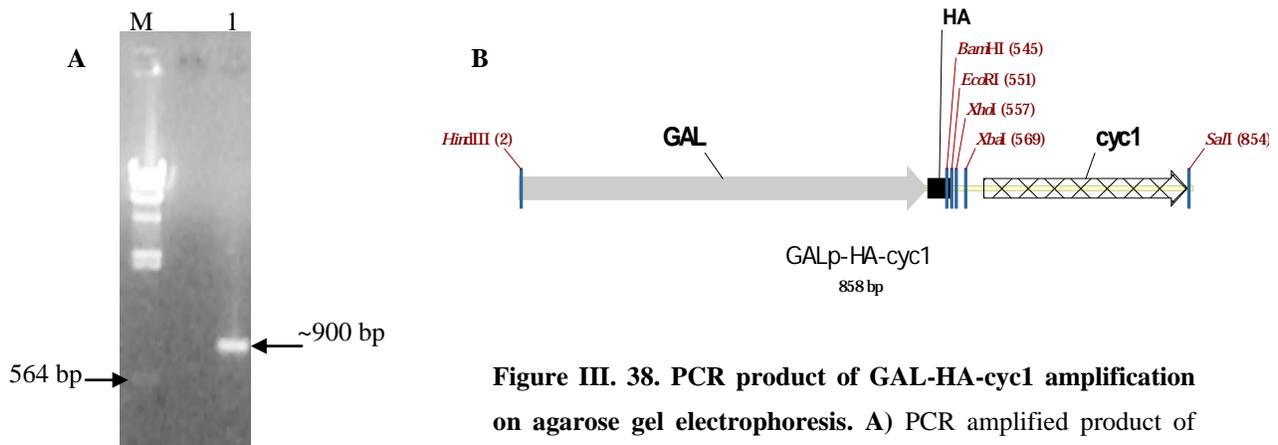


Figure III. 38. PCR product of GAL-HA-*cyc1* amplification on agarose gel electrophoresis. A) PCR amplified product of GAL-HA-*cyc1* (1), where M - Gene ruler, λ DNA/*Hind*III marker; B) Linear map of GALp-HA-*cyc1* cassette.

The presence of the insert for the pCRII-TOPO/GALP-HA-*cyc1* construct was determined by *Hind*III-*Sal*I digestion (Fig. III. 39 A1).

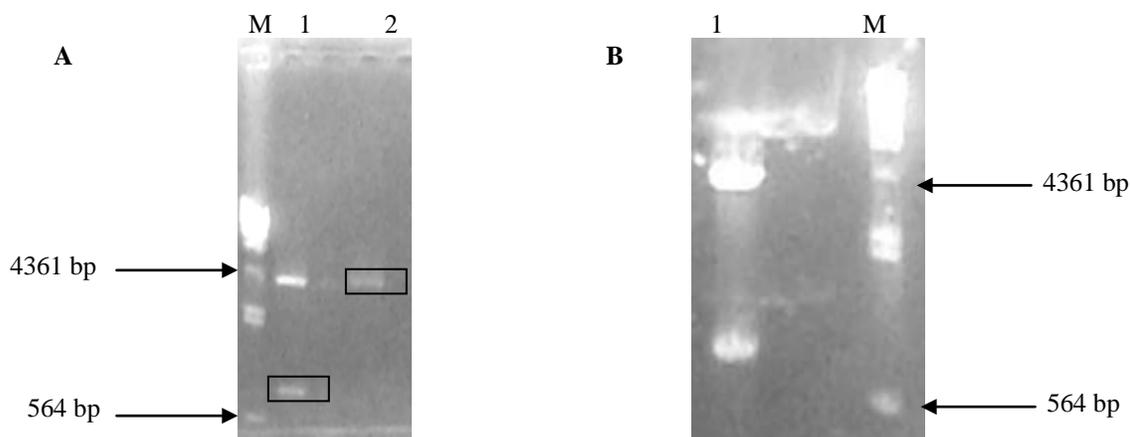


Figure III. 39. Cloning of GALp-HA-cyc1 cassette into pDNR vector. A) Agarose gel electrophoresis of 1- pCRII-TOPO/ GALp-HA-cyc1 cut with HindIII-SalI. The GALp-HA-cyc1 fragment was gel extracted and purified; 2- pDNR cut HindIII-SalI. The plasmid backbone was gel extracted and purified; B) pDNR-GAL digested with HindIII-SalI (1) verify the correctness of the new construct, M - Gene ruler, λ DNA/HindIII marker.

6.2. Generation of a recyclable cassette for promoter integration

To enhance the pool of terpene synthase precursors we overexpressed genes involved in the GPP and FPP biosynthesis by integrating strong promoter elements upstream of open reading frame of genes involved in the biosynthesis of isoprene units.

Integration of these genes in the yeast genome was accomplished by a system that employs a modified loxP-recognition sequence flanking the selection marker and can be used to repeatedly add a promoter upstream of an HA tag fused to different genes. Subsequent Cre-mediated site-specific recombination removes the selection marker and thus creates a stable promoter integration, while the integration cassette is recycled and may be reused.

6.2.1. Developing a GAL promoter-HA tag insertion cassette with a recyclable selection marker

The insertion cassettes uses the loxP-selection marker cassette of the pUG72 plasmid employed for gene deletions, which allows the excision of the selection marker at the end of the homologous recombination process using the Cre recombinase, and subsequent recycling of the selection marker (Fig. III. 40).

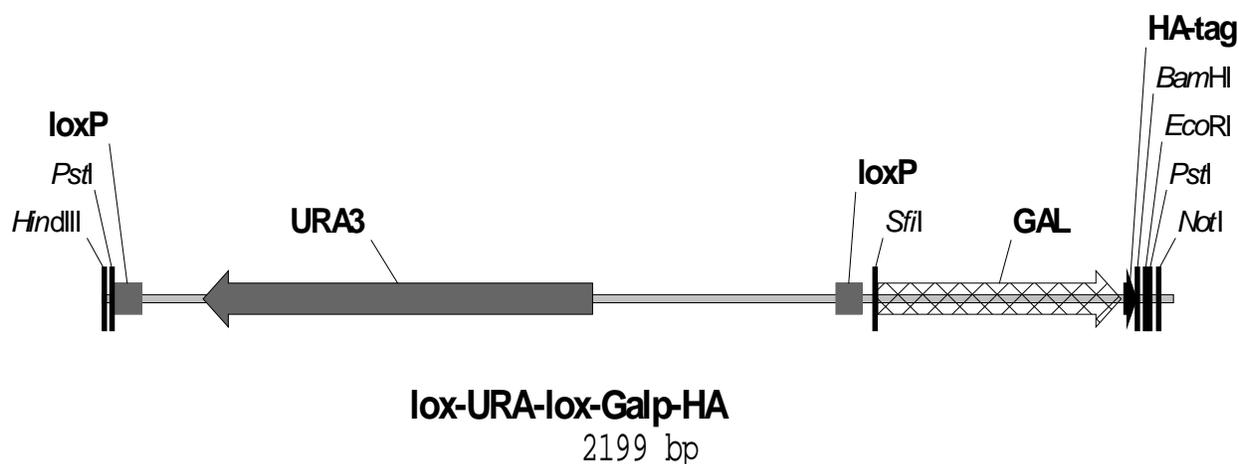


Figure III. 40. loxP-URA-loxP-GALp-HA pCOD2 integration cassette.

The cloning strategy started with the PCR amplification of the GALp-HA fragment of approximately 500bp from the pJG4-6 vector using specific primers (see Chapter II, paragraph 7.1.2) (Fig. III. 41).

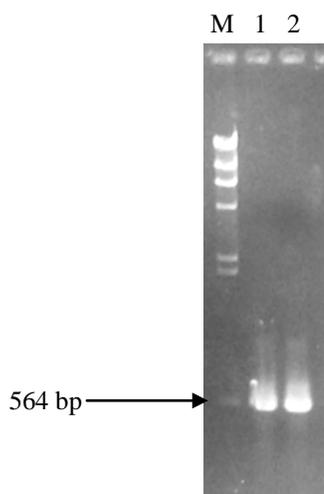


Figure III. 41. PCR product of GALp-HA amplification on agarose gel electrophoresis, where M - Gene ruler, λ DNA/*HindIII* marker; 1-2 PCR product of GALp-HA.

The GALp-HA fragment was TOPO-cloned and randomly extracted minipreps were checked for the presence of the GALp-HA insert in the right orientation. (Fig. III. 42).

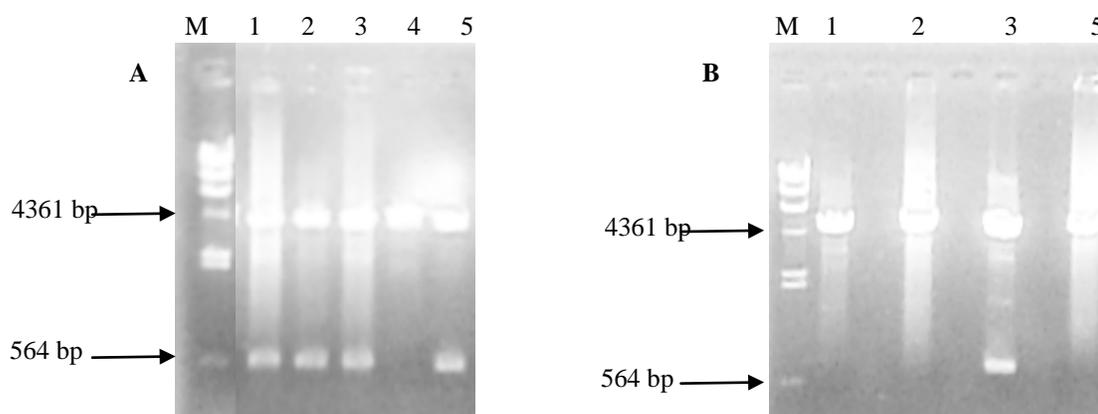


Figure III. 42. Cloning of GALp-HA fragment into pCRII-TOPO vector, where M - Gene ruler, λ DNA/*HindIII* marker; **A)** Digestion of pCRII-TOPO/GALp-HA (1, 2, 3, 4, 5) cut with *EcoRI* revealed the presence GALp-HA insert in constructs 1,2,3,and 5; **B)** The right orientation of the GALp-HA insert was determined after digestion of pCRII-TOPO/GALp-HA (1,2,3,5) with *HindIII-SfiI*. The constructs 1,2, and 5 carried the GALp-HA insert in the right orientation.

Next, the *loxP-URA3-loxP* cassette of approximately 1.7 kb was excised from the pUG72 vector and inserted into *HindIII-SfiI* frame of pTOPO-GAL-HA construct. The verified construct (Fig. III. 43) was named pCOD2 (for details see Chapter II, paragraph 7.1.2):

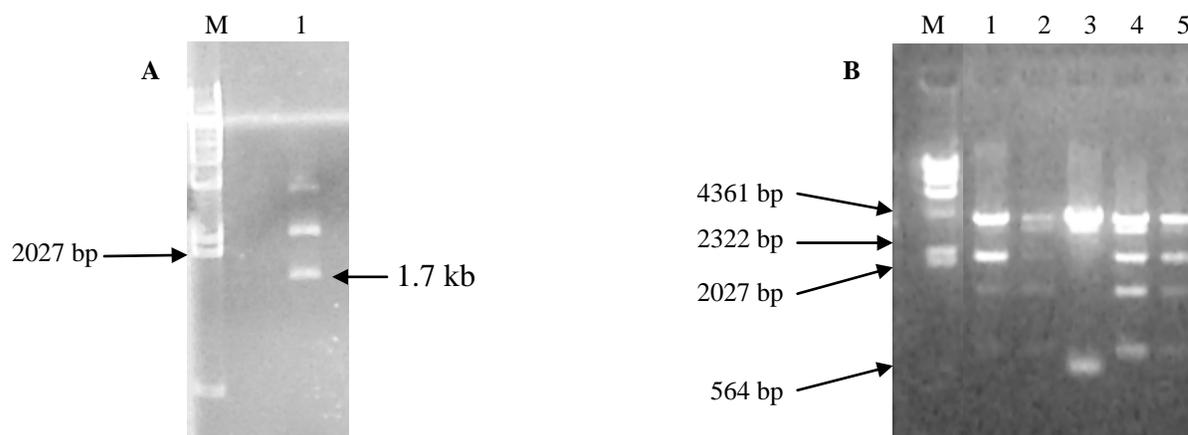


Figure III. 43. Cloning of GALp-HA-cyc1 cassette into pUG27 vector. A) Agarose gel electrophoresis of pUG72 cut with *HindIII-SfiI* (1); B) pCRII-TOPO/GALp-HA-loxP-URA3-loxP cut with *HindIII-EcoRI* (1, 2, 3, 4, 5), where M - Gene ruler, λ DNA/*HindIII* marker. The constructs 1 and 5 are the right constructs.

After size verification, the new construct was confirmed by sequencing, using the M13 reverse and M13 forward. Sequencing was carried out from both directions in order to obtain a

more reliable reading. Contigs were made from the two readings of each clone using the Vector NTI Software package (Appendix 1).

In this paragraph, a new and efficient integration cassette for repeated use by combining the advantages of the selection marker with those of the Cre/loxP recombination system of bacteriophage P1 is described. The integration cassette consists of the URA3 selection marker gene flanked by two 34 bp loxP sites fused to GAL promoter and HA epitope tag. The integration cassette was amplified by PCR using oligonucleotides comprising 25 3'- nucleotides complementary to sequences in the template (pUG72 plasmid) flanking the integration cassette and 45 5'- nucleotides that anneal to sites upstream of the genomic target sequence to be upregulated. After transformation of the linear disruption cassette into yeast cells, selected transformants were checked by PCR for correct integration of the cassette. The verification PCRs were done using combinations of primers complementary to sequences within the cassette and to sequences flanking the target gene. In a diploid yeast cell, one allele of the target gene is upregulated by the GAL promoter integration. Finally, expression of Cre recombinase from pGAL-*cre* plasmid results in removal of the marker gene, leaving behind a single loxP site at the chromosomal integration locus.

6.2.2. Developing a ADH-myc tag insertion cassette with a recyclable selection marker

The completion of the pCOD2 integration cassette facilitated the generation of a similar tool to further improve the molecular engineering of yeast strains. To this end integration of stronger promoters, such as constitutive ADH promoter, and different tags was also desirable for extensive engineering of the sterol biosynthetic pathway in yeast cells (Fig. III. 44).

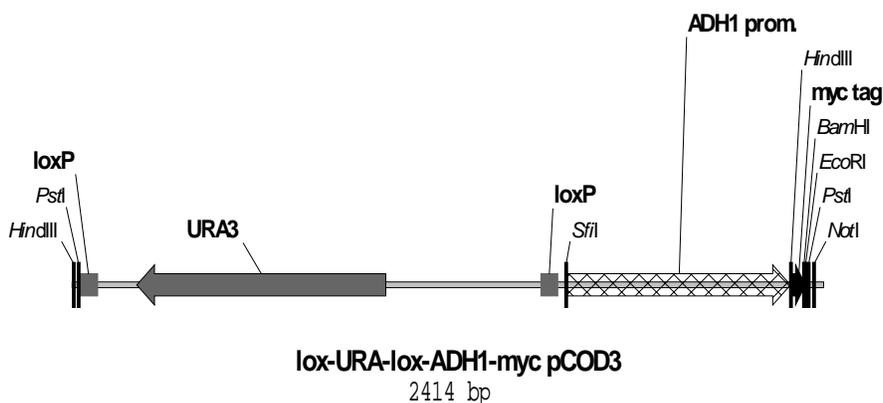


Figure III. 44. The loxP-URA-loxP-ADH1-myc pCOD3 integration cassette.

Following the procedure described above, the loxP-URA3-loxP fragment collected after pUG72 digestion with HindIII-SfiI was used to generate an integration cassette containing the constitutive ADH promoter and the myc tag (Fig. III. 45).

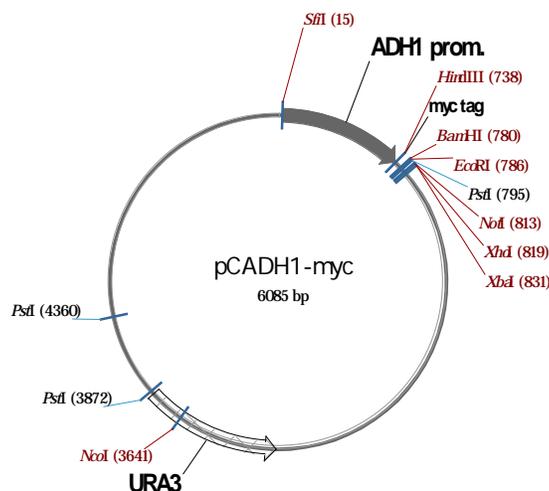


Figure III. 45. The pCADH1-myc plasmid map.

The first step consisted of PCR amplification of the constitutive ADH promoter of approximately 700-800 bp (Fig. III. 46 A) from the pEG202 vector using specific primers (for details see Chapter II, paragraph 7.2.1).

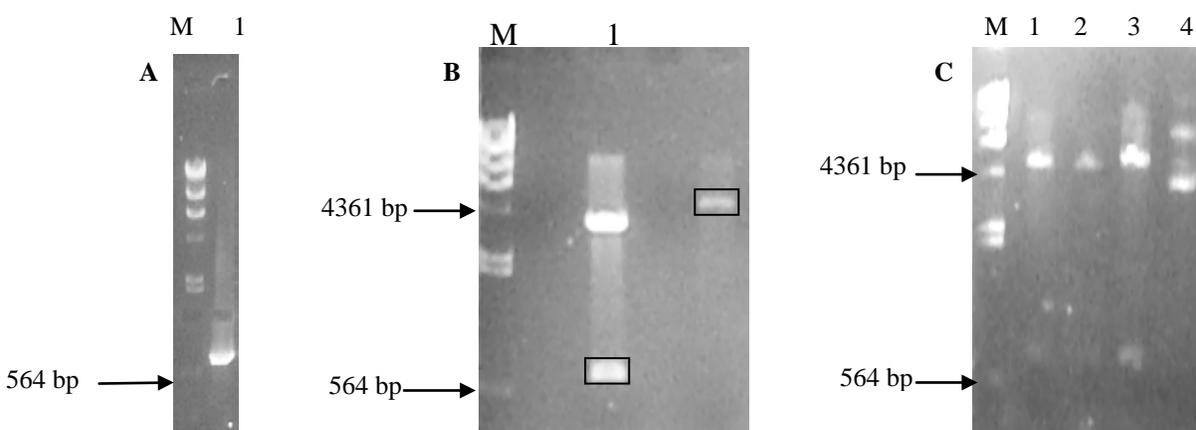


Figure III. 46. Cloning of ADH fragment into pYESmyc vector, where M - Gene ruler, λ DNA/HindIII marker; A) PCR product of ADH amplification; B) Agarose gel electrophoresis of 1- pCRII-TOPO/ADH cut with HindIII-Spe. The ADH fragment was gel extracted and purified; 2- pYESmyc cut HindIII-SpeI. The plasmid backbone was gel extracted and purified; C) pCADH-myc cut with HindIII-SpeI resolved by agarose gel electrophoresis (1, 2, 3, 4). The constructs 1 and 3 are the right constructs.

The amplified ADH promoter was cloned into the pCRII-TOPO vector and additionally subcloned in the pYESmyc vector in the SpeI-HindIII sites to generate a pCADHmyc vector (Fig. III. 46 B, 46 C).

To generate the integration cassette, the ADH-myc fragment was digested from the pCADHmyc vector with SfiI and BamHI, gel purified and ligated into the pCOD2 vector in the same sites (Fig. III. 47). The recombinant plasmid was transformed into *E. coli* MACH1 competent cells, verified by BamHI and SfiI digestion, named pCOD3 (for details see Chapter II, paragraph 7.2.2), and confirmed by sequencing, using the M13 reverse and M13 forward primers. Sequencing was carried out from both directions in order to obtain a more reliable reading. Contigs were made from the two readings of each clone using the Vector NTI Software package (Appendix 1).

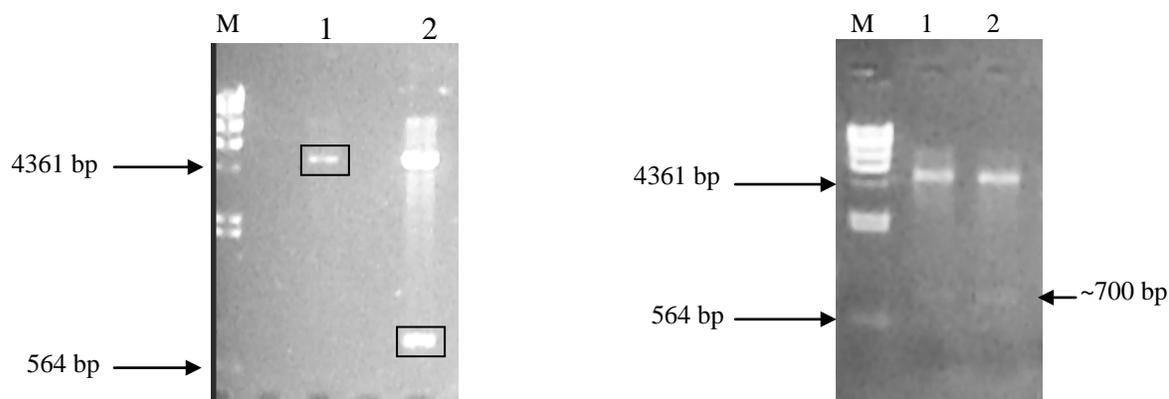


Figure III. 47. Cloning of ADHmyc fragment into the pCOD2 vector. M - Gene ruler, λ DNA/*HindIII* marker; **A)** Agarose gel electrophoresis of 1- pCOD2 cut with SfiI-BamHI. The plasmid backbone was gel extracted and purified; 2- pCADHmyc cut with SfiI-BamHI. The ADHmyc fragment was gel extracted and purified; **B)** pCOD3 cut with BamHI-SfiI resolved by agarose gel electrophoresis (1, 2). Both constructs 1 and 2 are the correct constructs.

7. Yeast Strain Molecular Manipulation.

Yeasts are of invaluable interest as cell factories for the production of compounds that are difficult or impossible to make because of the complexity of structures and stereospecific modifications which are frequently required. Additionally they promise alternative environmentally friendly ways of production, especially for high cost chemicals. Even though there have been a number of reports of metabolic engineering of the carotenoid pathway, the

success in the *in vivo* production of terpenes has been limited. Problems were encountered in the expression and stability of the terpene synthases (in *E. coli*) and the availability of prenyl pyrophosphate precursors. Although, the whole genetic pathway of sterol biosynthesis is present in yeast, it is tightly regulated and the precursors are present in limited quantities. Yeast molecular manipulation could overcome this obstacle by engineering of the biosynthetic pathway to maximize the amounts of GPP and FPP precursors produced intracellularly, which will enhance the production of monoterpenes and sesquiterpenes.

To this end, the expression of yeast genes involved in biosynthesis of geranyl pyrophosphate (GPP) and farnesyl pyrophosphate was modified by integrating strong promoter sequences, by homologous recombination, upstream of the open reading frame of the selected genes, in combination with mutations that prevent the draining of the precursors into competing pathways. The genetic background chosen for molecular manipulation was the high sterol-producing strain EG60 (Mat α , *ura3*, *trp1*, *his3*). All the derivatives of this strain were able to synthesize terpenes but in different quantities.

7.1. Deletion of *erg6* gene in GALp-(K6R)HMG2 strain

The major sterol produced in yeast cells is ergosterol. We reasoned that inhibiting the pathways that drain the accumulating isoprene units could improve isoprene yield by alleviating feedback inhibition of the pathway. Accordingly, the gene *ERG6* encoding an enzyme that converts zymosterol to fecosterol was first chosen for disruption with the rationale that it is the first gene of the pathway, non-essential for yeast viability thus hoping to decrease consumption of terpenoid precursors. Deletion of the *erg6* gene (Fig. III. 48) was generated both in EG60 strain to generate AM62 (Mat α , *ura3*, *trp1*, *his3*, $\Delta erg6:: his5+$), and AM63 strain to generate AM64 (Mat α , GALp-(K6R) HMG2::*URA3*, *trp1*, *his3*, $\Delta erg6:: his5+$)

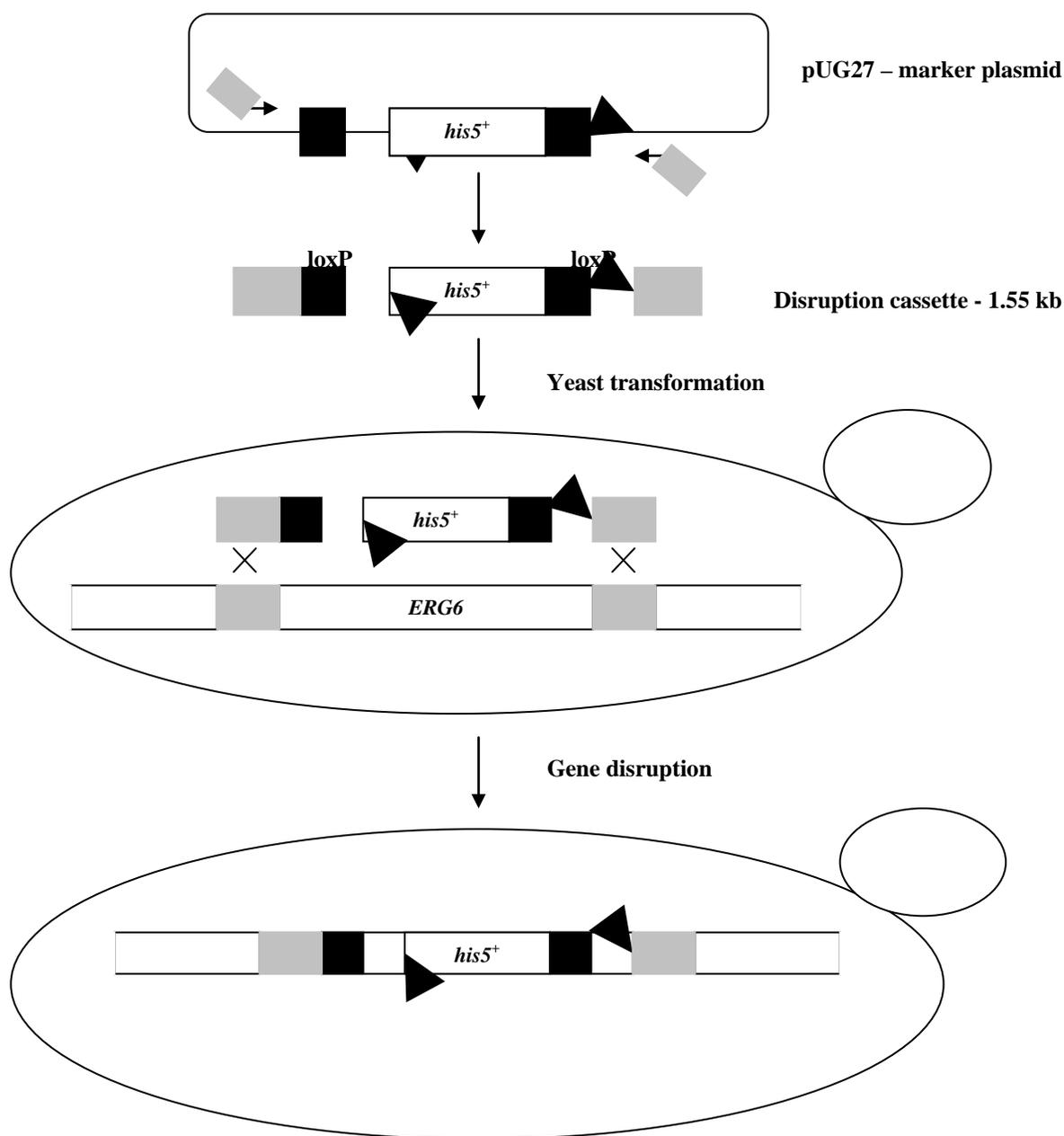


Figure III. 48. Deletion scheme of *ERG6* gene. The gene disruption cassette consists of a *his5⁺* selection marker gene, flanked by two 34 bp *loxP* sequences as direct repeats located adjacent to 45 bp of sequence flanking the chromosomal *ERG6* gene. The disruption cassette is produced by PCR using oligonucleotides comprising complementary nucleotides to sequences in the template (marker plasmid) flanking the disruption cassette and 45 5' nucleotides that anneal to sites upstream or downstream of the *erg6* gene.

The pUG27 disruption cassette (Fig. III. 49) containing *his5+* as a selection marker was PCR amplified using primers comprising nucleotides complementary to sequences in the marker plasmid flanking the disruption cassette and 45 bases of 5' of *ERG6* gene-specific nucleotides (see Chapter II, paragraph 10.5.1.) and 45 bases of 3' sequence from the end of the *ERG6* gene was proceeded.

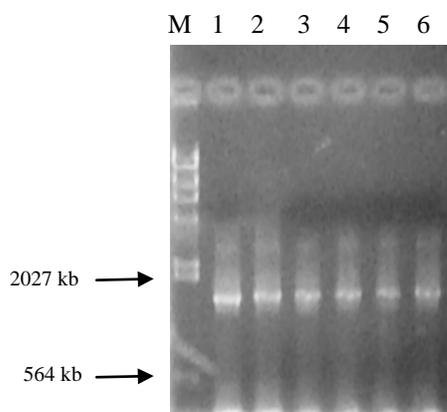
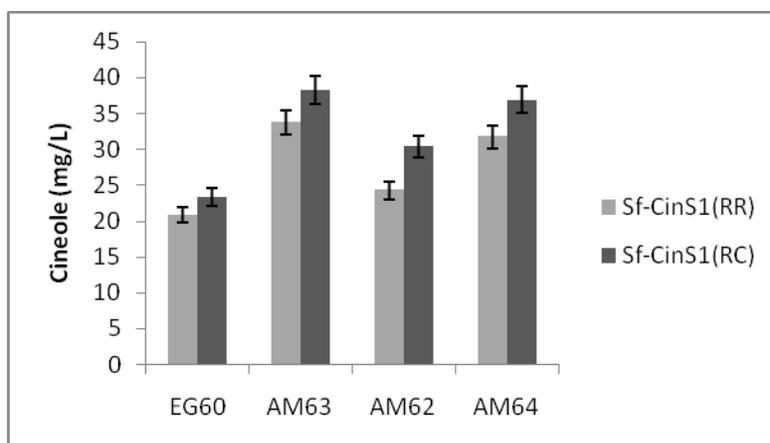


Figure III. 49. PCR amplification of pUG27 disruption cassette, where M - Gene ruler, λ DNA/*Hind*III marker.

Eventhough *ERG6* is not required for normal vegetative growth, meiosis, or sporulation (Garber *et al.*, 1989) the yeast cells carrying the Δ *erg6* deletion only exhibited low transformation rate and sensitivity to normal growth conditions.

To assess the impact of the *erg6* deletion on terpene production, AM62 and AM64 yeast cells were transformed with monoterpene synthases, Sf-CinS1(RR), and Sf-CinS1(RC) and further analyzed for cineole production. In parallel work, strain AM63 and parental strain EG60 were analyzed for cineole production in similar experimental conditions. The double mutant AM64 strain carrying both the GALp-(K6R)-HMG2 cassette and the Δ *erg6* deletion showed higher ability to produce terpenes than the single mutant strain AM62, verifying the anticipated effect of the GALp-(K6R)-HMG2 mutation on the yeast sterol biosynthetic pathway. However, the two strains AM64 and AM62 with in principle improved the genetic background for terpene production than the parental strain, accumulated less cineole than the modified AM63 strain (Fig. III. 50). These outcome may be explained by low rate of transformation and deficient growth observed in Δ *erg6* yeast cells. The effect of the HMG2 mutation on sterol biosynthesis and subsequent terpene production is greater than the effect of the Δ *erg6* mutation, which was

possibly masked in the AM64 strain. Established as the best strain so far for cineole production, the AM63 was selected as starting point for further molecular manipulation.



Yeast strains

EG60 Mat α , *ura3*, *trp1*, *his3*

AM63 Mat α , *ura3*, *trp1*, *his3*, GALp-(K6R)HMG2

AM62 Mat α , Δ *erg6*::HIS5, *ura3*, *trp1*, *his3*

AM64 Mat α , GALp-(K6R)HMG2::*URA3*, *trp1*, *his3* Δ *erg6*::HIS5

Figure III. 50. Cineole production in modified yeast strain. Engineered deletion strains AM62 and AM64, expressing the two variant forms of cineole synthase (Sf-CinS1(RR) and Sf-CinS1(RC)) were monitored for cineole accumulation in parallel with AM63 yeast strain engineered with specific modifications to the MVA pathway (GALp-(K6R)HMG2) and parental strain EG60, within 24 hours. The modified AM63 strain showed higher cineole production when similar expression conditions were respected. Bars indicate the average cineole amount; this experiment was repeated three times with similar results.

Deletion of *erg6* in the EG60 genetic background indeed did lead to an increase in cineole production. However, when we tested the doubly modified strain containing the *erg6* mutation and overexpressing the K6R HMG2 protein, the yields were almost equivalent and the beneficial effects of the deletion were lost. This may be due to that the alleviation of feedback inhibition that could originate from the absence of ergosterol and may influence the HMG2 stability which has already been modified in our case.

Since AM62 and AM64 did not bring any additional significant improvements on terpene production, but conferred growth sensitivity and transformation efficiency defects to yeast cell, the two strains were not pursued further in this work.

7.2. Generation of diploid strain AM66

Most of the biosynthetic enzymes involved in mevalonate pathway are encoded by genes that are essential in yeast. As a consequence, any desirable modifications of these genes in haploid cells could have adverse effects on viability and normal growth parameters. To further address this inconvenience a diploid strain AM66 was generated by mating AM63 (Mat α , GALp-(K6R)HMG2::URA3, *trp1*, *his3*) with KSY10 (Mat a, GALp-(K6R)HMG2::URA3, *trp1*, *his3*), kindly donated from Kristine Stepanian, (MSc student, MAICh). AM66 strain was employed for further molecular strain manipulation targeting one of the two alleles for gene disruption and/or gene over expression by chromosomal promoter integration.

7.3. Upregulation of ERG20 gene by integrating a GALpromoter in one of the two alleles in a GALp-(K6R)HMG2 diploid strain

The gene responsible for GPP synthesis but also for FPP synthesis in sterol biosynthetic pathway, ERG20 was stably upregulated by introducing a galactose inducible promoter upstream of the gene's open reading frame causing an increased production of mRNA transcripts.

A loxP-URA-loxP-GALp-HA cassette was stably integrated in front of the ERG20 gene of AM66 strain (see Chapter II, paragraph 10.4.2) to generate AM67 (Mat α/a , *trp1*, *his3*, GALp-(K6R)HMG2, GALp-HA-ERG20::URA3). The loxP-URA-loxP-GALp-HA pCOD2 cassette was PCR amplified by using gene-integration cassette-specific primers (Fig. III. 51).

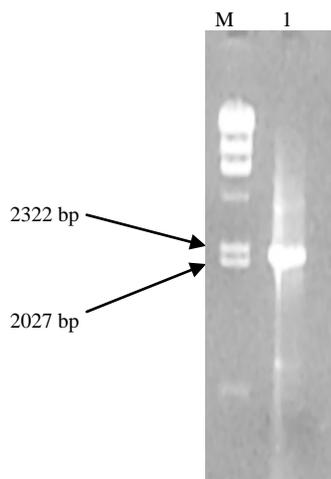


Figure III. 51. PCR product of pCOD2 resolved on agarose gel electrophoresis, where M - Gene ruler, λ DNA/*Hind*III marker; 1-PCR product of pCOD2.

The expression of the ERG20 protein was detected by immunoblotting utilizing the 3-HA tag and anti-hemagglutinin antibody as detection probe (Fig. III. 52).

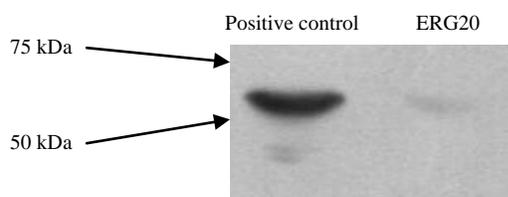


Figure III. 52. Immunodetection of ERG20 protein by HA tag. Total protein lysates from the AM68 yeast cells over-expressing the ERG20 subjected to SDS-PAGE electrophoresis followed by western blot (utilizing the HA tag) to detect and visualize the expression of the protein of interest.

After transformation of the loxP-URA3-loxP-GALp-HA integration cassette into yeast cells, single transformants were selected for genomic DNA extraction by phenol:chlorophorm precipitation. The PCR amplification of ~500 bp fragment utilizing extracted genomic DNAs as template and promoter- and gene-specific primers (Fig. III. 53) revealed an estimated rate of 95% correct integration of the GALp-HA cassette upstream of the *erg20* gene.

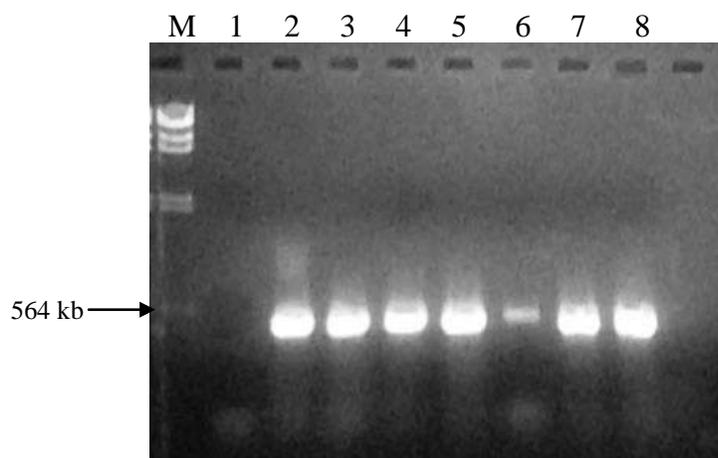


Figure III. 53. PCR confirmation of GALp-HA integration upstream of *erg20* gene in AM67 strain. The extracted genomic DNA from six single colonies that underwent GALp-HA integration upstream of ERG20 were PCR amplified using specific primers to GAL promoter and ERG20 open reading frame (2-8). Five colonies were positive for GALp-HA integration. Sample 1 corresponds to parental yeast cell that do not contain the GALp-HA integration, while M - Gene ruler, λ DNA/*Hind*III marker.

The positive colonies were replated and transformed with plasmids carrying terpene synthase cDNA clones to be assessed for terpene production. Once correctly integrated into the genome the URA3 marker was efficiently excised by transformation with a plasmid carrying the gene for Cre recombinase under the control of the GAL1 promoter and induction of Cre expression by shifting the cells to galactose containing medium (see Chapter II, paragraph 10.3). Cre-induced recombination resulted in excision of the URA3 marker. By plating on FOA media AM67 has been converted to AM68 (Mat α/a , *ura3*, *trp1*, *his3*, GALp-(K6R)HMG2, GALp-HA-ERG20).

7.4. Construction of GALp-(K6R)HMG2, GALp-ERG20, Δ erg9 haploinsufficient strain

In parallel with the efforts to overexpress the genes of the isoprene biosynthetic pathways, work was further focused on inhibition of acting further downstream located pathway which drain the cell's supply of GPP and FPP. For this purpose, the gene ERG9 encoding squalene synthase, an enzyme that converts FPP to squalene was chosen as a suitable target. Recent studies (Song, 2003) have shown that yeast mutants blocked at the ERG9 gene require exogenous ergosterol for normal vegetative growth. To avoid this, partial inhibition of FPP conversion to squalene by interrupting one of the ERG9 alleles of the diploid strains carrying other modifications was generated (Fig. III. 54). This event was expected to reallocate some of the isoprene precursors from sterol biosynthesis to production of terpenes by the introduced terpene synthases, while maintaining the growth rate of yeast cells at acceptable level.

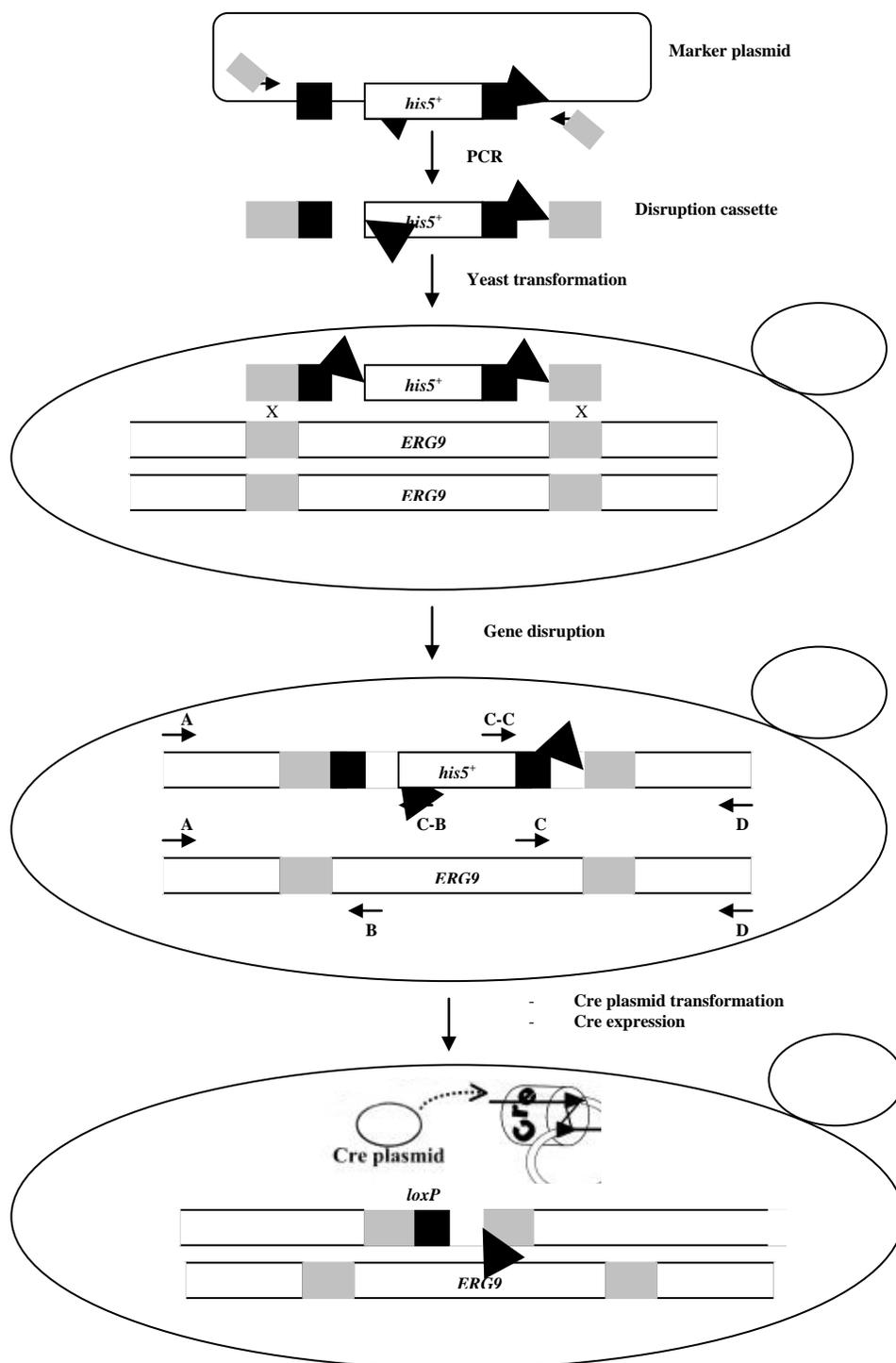


Figure III. 54. Deletion scheme of *ERG9* gene. The gene disruption cassette consists of a *his5⁺* selection marker gene, flanked by two 34 bp *loxP* sequences as direct repeats located adjacent to 45 bp of sequence flanking the chromosomal *ERG9* gene. The disruption cassette is produced by PCR using oligonucleotides comprising of complementary nucleotides to sequences in the template (marker plasmid) flanking the disruption cassette and 45 5' nucleotides that anneal to sites upstream or downstream of the *erg9* gene.

Disruption of *erg9* (see Chapter II, paragraph 10.5.2) was carried out in both AM67 and AM68 diploid strains to generate the haploinsufficient mutants AM69 (Mat α/a , *trp1*, *his3*, GALp-(K6R)HMG2, GALp-HA-ERG20::URA3, ERG9 $\Delta erg9:: his5+$) and AM70 (Mat α/a , *ura3*, *trp1*, *his3*, GALp-(K6R)HMG2, GALp-HA-ERG20, ERG9, $\Delta erg9:: his5+$) respectively, by PCR amplification (Fig. III. 55) of the pUG27 disruption cassette containing *his5+* as a selection marker, using gene-integration cassette-specific primers.

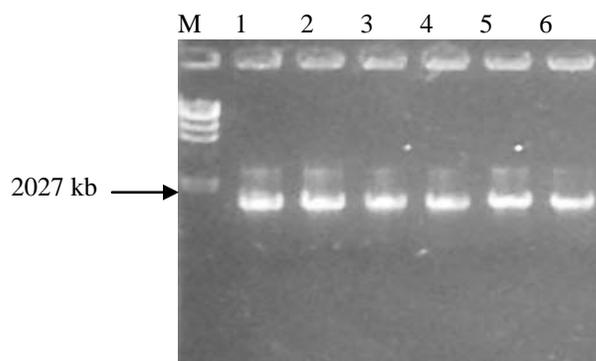


Figure III. 55. PCR product of pUG27 resolved on agarose gel electrophoresis, where M - Gene ruler, λ DNA/*Hind*III marker; 1-6 PCR product of pUG27.

In diploid AM67 and AM68 yeast cells one allele of the target gene was replaced by the disruption cassette generating *erg9* haploinsufficient strains AM69 and AM70, respectively.

After transformation of the linear disruption cassette into yeast cells, genomic DNA extracted from single selected transformants were checked by PCR amplification for correct integration of the cassette and concurrent deletion of the chromosomal target sequence, *erg9*, utilizing primers specific to gene promoter and disruption cassette (Fig. III. 56). The frequency of *erg9* gene disruption was 86 %. The positives colonies were replated and transformed with plasmids carrying terpene synthase cDNA clones to be assessed for terpene production. The *erg9* haploinsufficient yeast cells exhibited low transformation rate and slow growth.

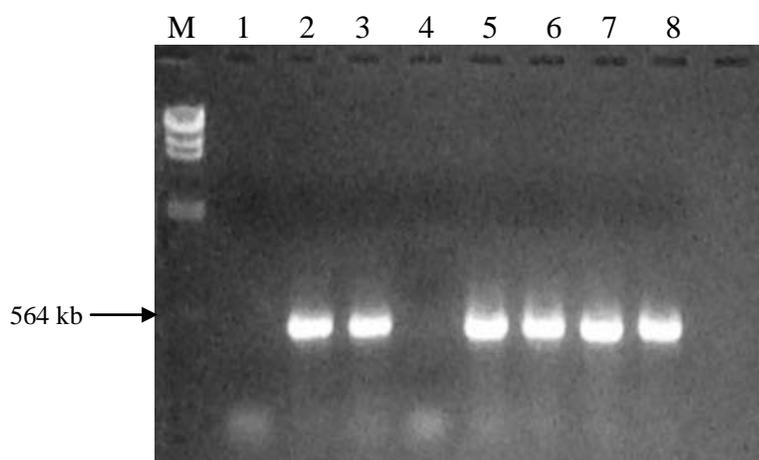


Figure III. 56. PCR confirmation of *erg9* gene disruption in AM70 strain. The extracted genomic DNA from six single colonies that underwent replacement of one *erg9* allele with pUG27 disruption cassette were PCR amplified using specific primers to *erg9* promoter and pUG27 cassette (2-7). Five colonies were positive for *erg9* haploinsufficiency. Sample 1 correspond to parental yeast cell that do not suffer gene disruption, while M - Gene ruler, λ DNA/*Hind*III marker.

7.5. Upregulation and stable truncation of the HMG1 gene by integrating an ADH constitutive promoter in one of the two alleles in different diploid strains

Unlike mammals, yeast express two functional isozymes of HMG-R, Hmg1p and Hmg2p, encoded by the HMG1 and HMG2 genes, respectively. The two proteins are 50% identical in the N-terminal hydrophobic domain and ~93% identical in the C-terminal catalytic domain (Basson *et al.*, 1988). The two are functionally similar in that each can synthesize adequate amounts of mevalonate (Basson *et al.*, 1988), and both appear to be residents of the ER (Koning *et al.*, 1996). In spite of these similarities, the dynamic behavior of the two proteins is strikingly different determined, by the transmembrane regions of the proteins: Hmg1p is extremely stable, whereas Hmg2p is subject to regulated degradation (Hampton and Rine, 1994). Yeast cells respond to increased levels of Hmg1p by proliferating stacked pairs of nuclear-associated membranes called karmellae (Wright *et al.*, 1988). This protein is composed of three distinct domains: an N-terminal transmembrane region, a linker, and the highly conserved C-terminal cytoplasmic domain.

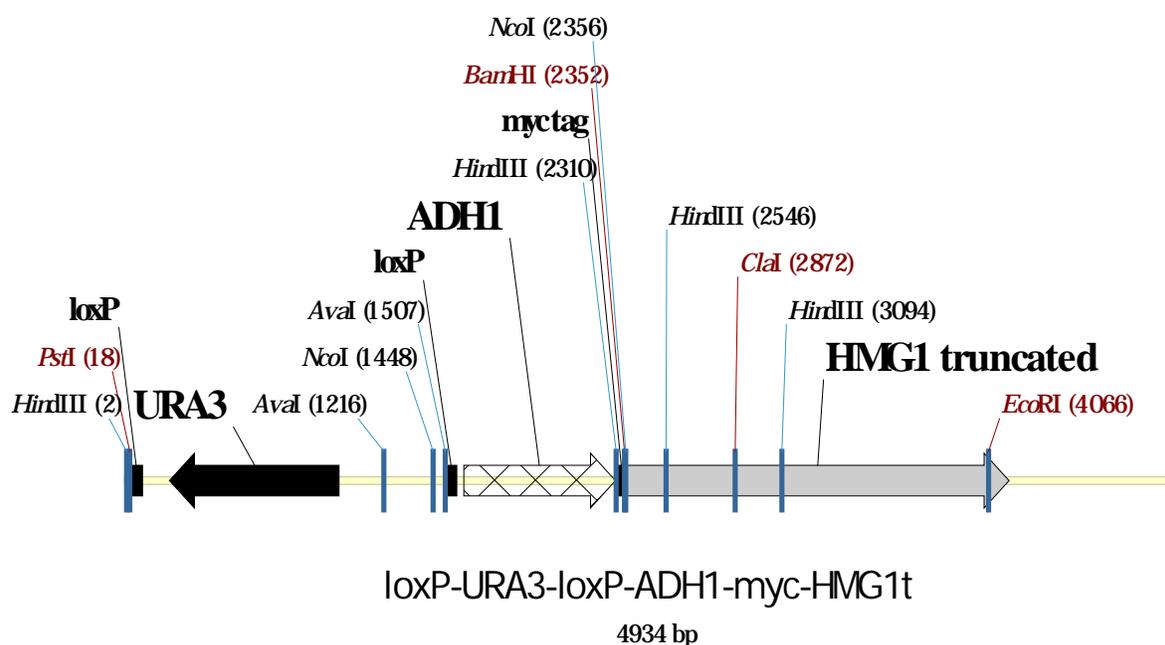


Figure III. 57. The loxP-URA3-loxP-ADH1-myc-HMG1 truncated linear map

Truncations of Hmg1p that shortened the amino terminus containing the membrane domain (Fig. III. 57) were created into the AM68 and AM70 strains by stable integration of the loxP-URA-loxP-ADH1p-myc cassette in the HMG1 promoter, generating AM74 (Mat α/a , GALp-(K6R)HMG2x2::HO, GALp-HA-ERG20, ADH1-myc::HMG1, URA3, *trp1*, *his3*) and AM75 (Mat α/a , GALp-(K6R)HMG2x2::HO, GALp-HA-ERG20, ADH1-myc::HMG1, URA3, *trp1*, *his3*, $\Delta erg9::HIS5$) strains.

The loxP-URA-loxP-ADH1-myc pCOD3 cassette was PCR amplified by using gene-integration cassette-specific primers (Fig. III. 58): HMG1t-COD3-for 5'- ACATAGTGTATCA TTGTCTAATTGTTGATACAAAGTAGATAAATACATAAAAGCTTCGTACGCTGCAGG - 3' and HMG1t-COD3-reverse 5'- GGAGCAGTAAAAGACTTCTTGGTGACTTCAGTTTTTCCCAATTGGTCCATAAGCTTGGAGTTGATTGTATGC - 3'

The amplified linear pCOD3 carrying on the C-terminus an artificially inserted ATG was transformed in the AM68 and AM70 yeast cells where underwent homologous recombination with a region of the N-terminus of HMG1 on one of the two alleles. The outcome is upregulation of a shorter stabilized version of carboxyl terminus HMG1 protein containing the cytosolic catalytic domain. The resulting strains, AM74 and AM75, were analyzed for *in vivo* production of mono- and sesquiterpenes.

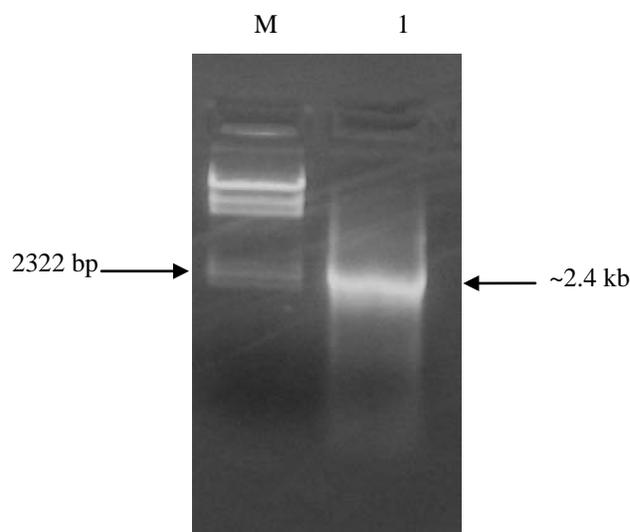


Figure III. 58. PCR product of pCOD3 resolved on agarose gel electrophoresis, where M - Gene ruler, λ DNA/*Hind*III marker; 1-PCR product of pCOD3.

8. Production of terpenes in modified yeast strains

Parental yeast strain EG60 was genetically engineered in the metabolic pathway of sterol biosynthesis. The modifications were designed to expand the endogenous GPP and FPP pool from which terpenes are synthesized by the action of terpene synthases. Four major modifications in sterol biosynthetic pathway were introduced. The first two were aimed at the HMG reductase and its structural genes (Fig. III. 59). The HMG2 protein was stabilized by incorporating a K6R mutation and was upregulated by introducing the GALp inducible promoter upstream of the open reading frame of the HMG2 gene. Both mutations were expected to that supply additional isoprene precursors in the yeast cells. Truncation of HMG1 protein in N-terminus region by ADH promoter incorporation resulted in generation a shorter stabilized C-terminus HMG1 mutant. Further on the endogenous pool of GPP and FPP precursors was more efficiently converted to terpene synthase substrates by upregulation of responsible enzyme encoded by the *erg20* gene. The third modification targeted the *erg9* gene which encodes the enzyme catalyzing the condensation of two molecules of FPP to form squalene, another key compound in the mevalonate pathway and the first sterol synthesized by yeast cells. In this case, squalene synthase expression was downregulated by deletion of its structural gene, *erg9*, in one allele of the two of the diploid cells having a second functional copy of the gene. The modified strains did not produce any terpenes when transformed with empty vector, and were subsequently transformed selected terpene synthases by the lithium acetate method.

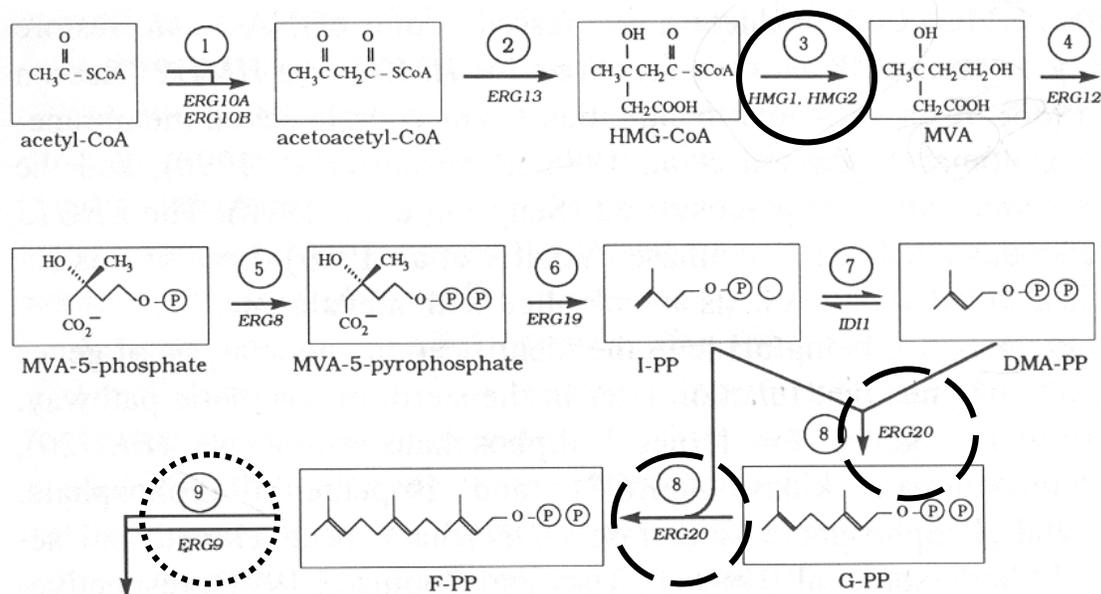
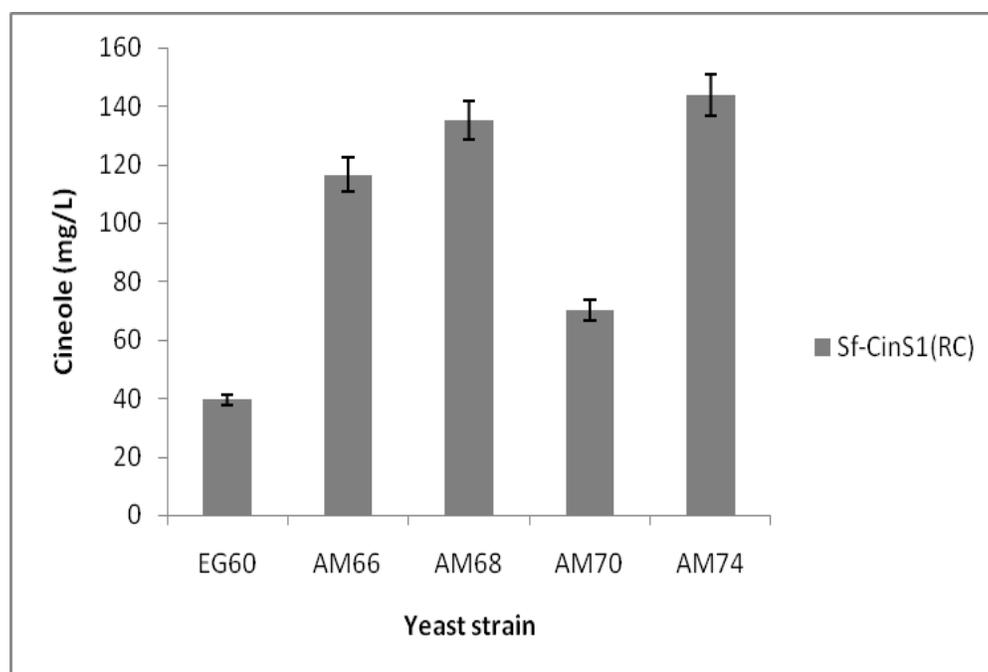


Figure III. 59. Metabolic engineering of sterol biosynthetic pathway. The first genetic modification considered upregulation of K6R HMG2 mutant protein (circle), followed by upregulation of FPP synthase by GAL promoter integration upstream of *erg20* gene (broken line circle), and squalene synthase downregulation achieved by partial deletion *erg9* gene (dotted circle).

8.1. Production of monoterpenes in modified yeast strains

The modified diploid yeast strains AM66 (Mat α/a , GALp-(K6R)HMG2x2::HO, *ura3*, *trp1*, *his3*), AM68 (Mat α/a , GALp-(K6R)HMG2x2::HO, *ura3*, *trp1*, *his3*), and AM70 (Mat α/a , GALp-(K6R)HMG2x2, GALp-HA-ERG20, *ura3*, *trp1*, *his3*, $\Delta erg9::HIS5$) as well as the parental EG60 control strain were transformed with Sf-CinS1(RC) subcloned into appropriate yeast vector and evaluated for cineole production. The cells were induced on galactose containing media for GALp activation and next on enzymatic reaction buffer for a period of three days. The accumulated volatiles were released every 24 hours. The ambient volatiles were trapped on the SPME fiber for a period of time of 30 minutes after three days of yeast culture induction, when the terpene synthases are expected to reach maximum efficiency (see Chapter III, figure III. 13). The quantitative analysis was carried out on the registered peak areas of the detected cineole comparing to the standard 1,8-cineole compound (see Chapter II, paragraph 15.1). All measurements were performed using the same SPME fiber and accomplishing high similarity experimental conditions (Fig. III. 60).



EG60 Mat *a*, *ura3*, *trp1*, *his3*

AM66 Mat *a/a*, GALp-(K6R)HMG2x2::HO, *ura3*, *trp1*, *his3*

AM68 Mat *a/a*, GALp-(K6R)HMG2x2, GALp-HA-ERG20, *ura3*, *trp1*, *his3*

AM70 Mat *a/a*, GALp-(K6R)HMG2x2, GALp-HA-ERG20, *ura3*, *trp1*, *his3*, Δ *erg9::HIS5*

AM74 Mat *a/a*, GALp-(K6R)HMG2x2, GALp-HA-ERG20, ADH-myc-HMG1, URA3, *trp1*, *his3*

Figure III. 60. Cineole production in modified yeast strains. The diploid strains AM66, AM68, AM70, AM74 were assessed for cineole production when transformed with pJG4-6/Sf-Cin(RC) in comparison with the parental strain EG60 carrying the same plasmid, pJG4-6/Sf-Cin(RC), over a period of 3 days, when similar expression levels were expected. AM66, AM68 and AM74 modified strains show ~ 3 fold high cineole production compared with the parental strain, but AM74 exhibit slightly better activity. Lower production of cineole was recorded from AM70 strain comparing to the AM66, AM68, AM70 strains. Bars indicate the average cineole amount. This experiment was repeated two times with similar results.

The Sf-CinS1(RC) enzyme appeared to be active in all tested yeast strains. AM66, AM68 and AM74 carrying *sf-CinS1* subcloned into appropriate yeast vector exhibited 3 fold higher level of cineole synthesis, compared to EG60 expressing the same enzyme. These results confirm the upregulation of the pathway as a combined result of the K6R-HMG2 mutation and the galp-controlled transcription of the *hmg1* and *erg2o* gene expression. The double mutant AM68 bearing the two GALp-upregulated (K6R)HMG2 and GALp-HA-ERG20 genes displayed only slightly better activity than the single mutant AM66, while AM74, bearing in addition the

truncated HMG1 gene under the control of ADH1 constitutive promoter exhibited somewhat better activity than both AM66 and AM68. Contrary to our expectations, squalene synthase downregulation in the AM70 strain, did not increase cineole production over the AM68 levels, but actually reduced the production almost to the EG60 levels. These findings may be explained by assuming that withdrawal of GPP toward FPP synthesis when the ERG20 encoded for a FPP synthase was upregulated, effect that was further worsened when sterol biosynthesis was disturbed further downstream.

8.2. Production of sesquiterpenes in modified yeast strains

During the evaluation of AM67 (Mat α/a , GALp-(K6R)HMG2x2, GALp-HA-ERG20::URA3, *trp1*, *his3*) cells on cineole production, the presence of high values of unexpected products which belong to sesquiterpene class was observed: ~7 fold more caryophyllene than AM67 transformed with empty vector, and ~30 fold more than AM63 (Mat α , GALp-(K6R)HMG2::URA3, *trp1*, *his3*) transformed with cineole synthase Sf-CinS1(RC) subcloned in yeast vector. No caryophyllene was detected in AM63 cells transformed with empty vector (Fig. III. 61).

A plausible explanation for the production of traces of caryophyllene in AM67 cells transformed with empty vector could be the activation of an endogenous yeast enzyme active as a result of FPP synthase upregulation and increased pool of the FPP substrate. The synthesis of caryophyllene in yeast cell expressing cineole synthases could be attributed to enzyme plasticity, as similar observations have been made in plants. It appears that the monoterpene synthases could possibly accommodate in their enzymatic pocket the larger FPP substrate (15 carbons) instead of GPP (10 carbons). To functionally assess the production of sesquiterpenes in the modified yeast strains the P330 novel uncharacterized terpene synthase from *S.pomifera* was chosen for further characterization.

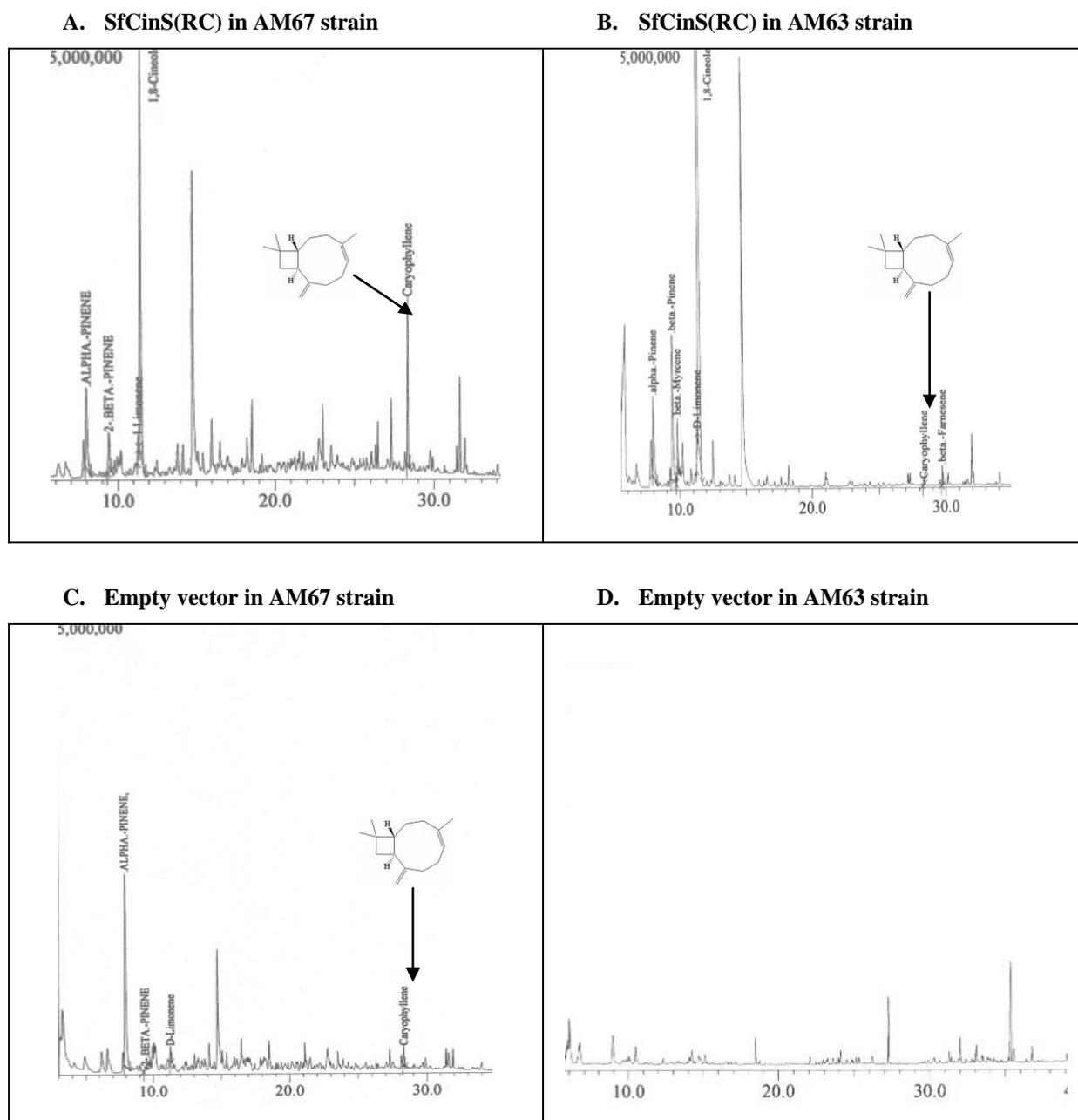
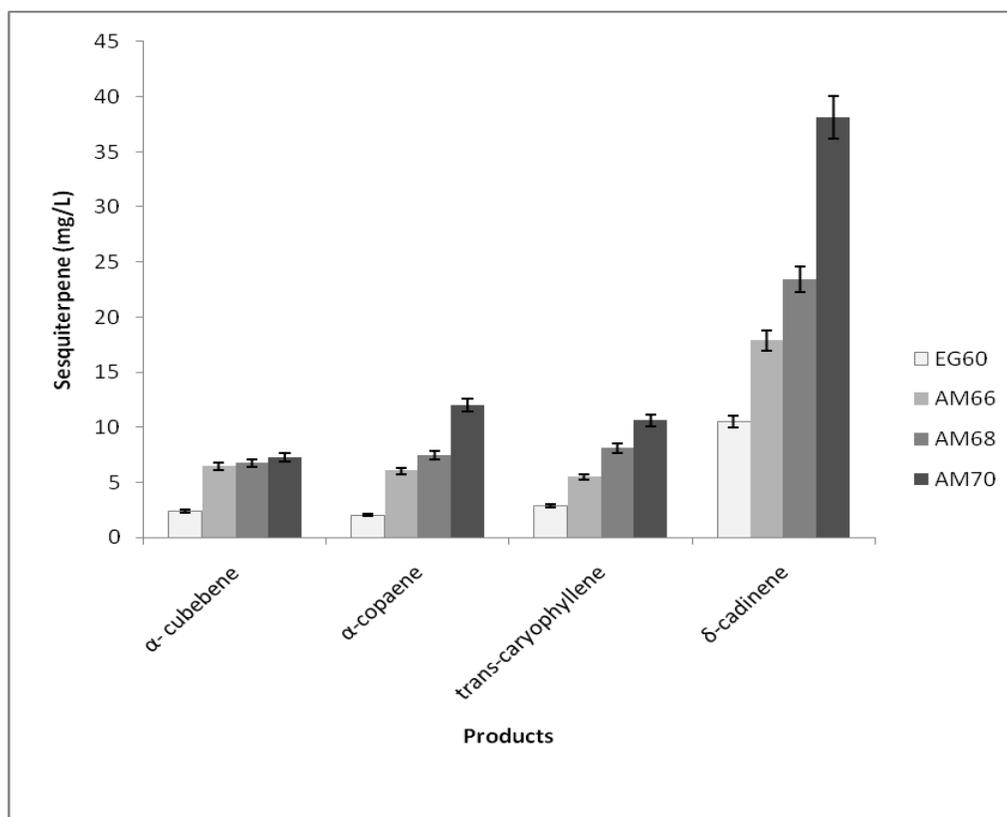


Figure III. 61. Production of caryophyllene in modified yeast strains. A) AM67 diploid strain expressing upregulated K6RHMG and ERG20 transformed with cineole synthase showed appreciable caryophyllene synthase activity B) In AM63 haploid strain expressing upregulated K6R HMG2 and transformed with cineole synthase, produced 30 fold lower amount of caryophyllene; C) AM67 diploid strain expressing upregulated (K6R)HMG2 and ERG20 transformed with empty vector produced 7 fold lower amount of caryophyllene than A; D) In AM63 haploid strain expressing upregulated K6R HMG2 and transformed with empty vector, no caryophyllene synthase activity was registered.

8.2.1. Expression and characterization of P330 enzyme in yeast cells

The diploid yeast strains AM66 (Mat α/a , GALp-(K6R)HMG2x2::HO, *ura3*, *trp1*, *his3*), AM68 (Mat α/a , GALp-(K6R)HMG2x2::HO, *ura3*, *trp1*, *his3*), and AM70 (Mat α/a , GALp-(K6R)HMG2x2, GALp-HA-ERG20, *ura3*, *trp1*, *his3*, $\Delta erg9::HIS5$) as well as the parental EG60 control strain were transformed with P330 sesquiterpene synthase and evaluated for the production of multiple sesquiterpene compounds (Fig. III. 62).



Yeast strain:

EG60 Mat α , *ura3*, *trp1*, *his3*

AM66 Mat α/a , GALp-(K6R)HMG2x2::HO, *ura3*, *trp1*, *his3*

AM68 Mat α/a , GALp-(K6R)HMG2x2, GALp-HA-ERG20, *ura3*, *trp1*, *his3*

AM70 Mat α/a , GALp-(K6R)HMG2x2, GALp-HA-ERG20, *ura3*, *trp1*, *his3*, $\Delta erg9::HIS5$

Figure III. 62. The activity of P330 sesquiterpene synthase in modified yeast strains. The diploid strains AM66, AM68, and AM70 were assessed for cineole production when transformed with P330 sesquiterpene synthase in comparison with the parental strain EG60 within a period of 3 days, when similar expression conditions were respected. The AM70 strain expressing three different genetic modifications in mevalonate pathway produced highest amount of sesquiterpenes, 3-6 folds more than the parental strain. Bars indicate the average cineole amount; this experiment was repeated two times with similar results.

The cells were incubated on buffer subsequent to induction for a period of three days and the accumulated volatiles were released every 24 hours. The ambient volatiles were trapped on the SPME fiber for a period of time of 30 minutes after three days of yeast culture induction. All measurements were performed using the same SPME fiber and accomplishing high similarity experimental conditions. The quantitative analysis was carried out on the registered peak areas of the detected trans-caryophyllene comparing to the standard trans-caryophyllene compound (see Chapter II, paragraph 15.2). Since standard compounds for various other sesquiterpenes produced by P330 enzyme were not available, the quantification of these products was done by estimating the percentage of sesquiterpene produced against the trans-caryophyllene percentage (Fig. III. 62).

The P330 enzyme catalyzed the production of sesquiterpenes utilizing the endogenous FPP pool in all selected strains. Increased sesquiterpene production was registered after each incoming genetic modification as described in Tables III. 9. and III. 10.

Table III. 9. Relative ratio of sesquiterpenes produced by each engineered yeast strains.

Yeast strain	Sesquiterpene products			
	α - cubebene	α -copaene	trans-caryophyllene	δ -cadinene
EG60	13.44	11.28	16.27	58.98
AM66	18	16.88	15.28	49.82
AM68	14.73	16.35	17.75	51.15
AM70	10.67	17.66	15.62	56.03

Table III. 10. Relative yields of sesquiterpenes produced by modified yeast strains (mg/L).

Yeast strain	Sesquiterpene products			
	α - cubebene	α -copaene	trans-caryophyllene	δ -cadinene
EG60	2.405157	2.018614	2.9116	10.55478
AM66	6.453701	6.052138	5.478475	17.86241
AM68	6.740049	7.481317	8.121919	23.40485
AM70	7.258175	12.01306	10.62537	38.11392

The molecular modifications comprising the upregulation of K6R HMG2 mutant and ERG20 synergized well with the downregulation of ERG9 increasing the available pool of endogenous isoprene precursors in yeast cells, opening new venues in overcoming the limitations of *in vivo* production of terpenes. The AM70 strain facilitated the best production of

sesquiterpenes with ~3 fold more α -cubebene, ~6 fold more α -copaene, and ~3.6 more trans-caryophyllene and δ -cadinene than the parental strain. Overall the outcome confirms the anticipated assumption of achieving increased terpene production in yeast cells by metabolic engineering of enzymes involved in mevalonate pathway and points to further improvements that can be made. It is noteworthy that the relative ratios of products do not show any significant alterations between strains, whereas the yield substantially increases. This enables the facile characterization of sesquiterpenes from novel enzymes which are difficult to characterize by in-vitro assays. Furthermore the strains developed having the alterations in to their chromosomal genome are stably modified, easy to use and allow introduction of multiple exogenous genes either in plasmids or into the genome reconstituting highly complex pathways.

9. Co-expression of monoterpene synthases with GPP synthase in yeast

Geranyl diphosphate synthase, which catalyzes the condensation of dimethylallyl diphosphate and isopentenyl diphosphate to geranyl diphosphate, is known only at the enzyme level, appearing to participate primarily in the plastidial biosynthesis of monoterpenes (Somer S. et al., 1995) by supplying the essential precursor of this family of natural products.

Several efforts to improve terpene production in living cells aimed to co-expression of heterologous terpene synthases with a GPP synthase. When cloned in a yeast vector and stably transformed into modified AM63 yeast strain, GPP synthase expressed together with Sf-Cin(RC) has a significant contribution to enhanced production of endogenous GPP and cineole (Fig. III. 63) ~2 fold higher comparing with AM63 yeast cell expressing only Sf-Cin(RC). AM63 cells expressing Sf-Cin(RC) with and without GPP synthase were induced on enzymatic reaction buffer supplemented with 1,85 mM exogenous GPP for 24 hours.

An enhancement of ~1.5 fold of cineole was recorded, indicating an unharnessed potential of the yeast cell to produce even more terpenes when availability of substrates and substrate precursors are not limited.

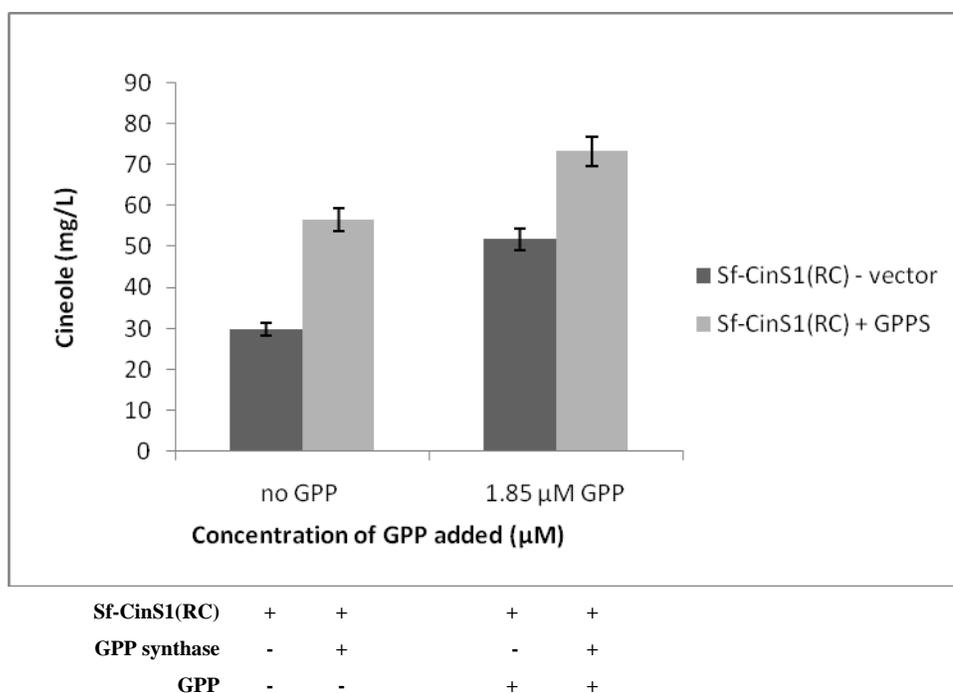


Figure III. 63. Effect of GPP synthase expression and Sf-Cin1 (RC) co-expression on cineole production in AM63 cells. Almost double amount of cineole was produced when Sf-Cin(RC) was co-expressed with GPP synthase. Addition of 1.85 mM exogenous GPP stimulated cineole accumulation with ~1.5 fold more than in crude samples. Bars indicate the average cineole amount; this experiment was repeated three times with similar results.

9.1. Development of a recyclable cassette for recyclable integration of exogenous genes. Insertion of Galp- GPP synthase into the lox-URA pUG27 integration vector

Taking into consideration the effect of GPP synthase on monoterpene production by heterologous genes stably transformed into yeast cell and taking advantage of accessibility to yeast strains a further step in optimizing yeast for maximal efficiency could be the stable integration of the GPP synthase on yeast chromosome. This approach could be used in subsequent steps to insert additional genes at will. To accomplish this aim a molecular tool was developed (see Chapter III, paragraph 3.1) taking advantage of the previously constructed pDNR-Gal-MCS-cyc1 plasmid construct. The GPPS gene was excised from pYESmyc vector with *EcoRI* and *XhoI* and subcloned into pDNR-GAL vector into the same restriction sites (Fig. III. 64 A), and then the GALp-GPPS-cyc1 cassette was removed and subcloned into pUG27 integration vector into *HindIII* and *SalI* restriction sites (Fig. III. 64 B and 64 C).

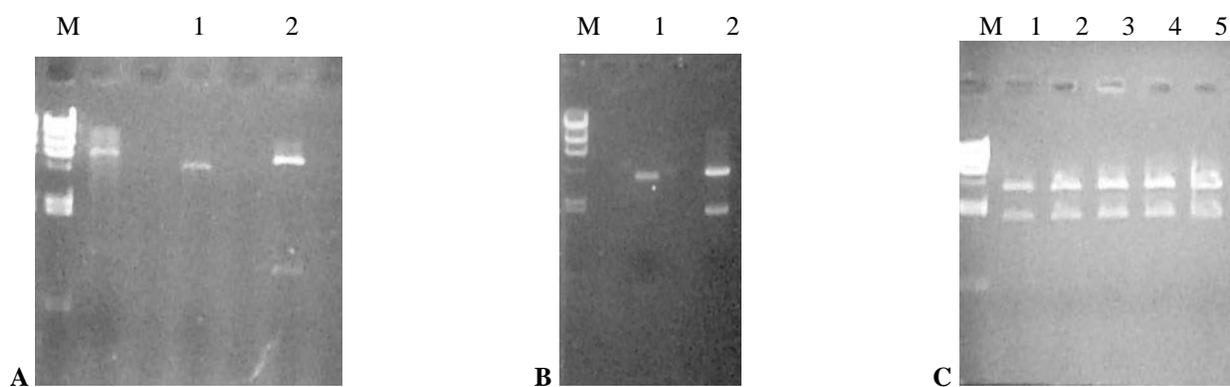


Figure III. 64. Cloning of GPPS fragment into pUG27 integration vector. A) Agarose gel electrophoresis of 1- pDNR-GAL and 2- pYESmyc-GPPS both vectors digested with EcoRI and XhoI. The backbone plasmid pDNR-GAL and the GPPS fragment were gel extracted and purified; B) Agarose gel electrophoresis of 1- pUG27 and 2- pDNR-GAL-GPPS both vectors digested with HindIII and Sall. The backbone plasmid pUG27 and the GAL-GPPS-cyc1 cassette were gel extracted and purified; C) pUG27-GAL-GPPS cut with HindIII-SalI resolved by agarose gel electrophoresis (1, 2, 3, 4, 5). All recombinant DNAs are the right constructs.

An attempt to integrate the GAL-GPPS-cyc1 cassette (Fig. III. 65) into LEU2 gene of AM68 and AM70 chromosomes is currently under investigation by using Hot Star Polymerase and specific primers: **LEU2-pUG F** 5'- ATGTCTGCCCCTAAGAAGATCGTCGTTTTGCCA GGTGACCACGTTGGTCAGCTTACGGATTAGAAAGCCGC - 3' and **LEU2-pUG Rev** 5'- AGCAAGGATTTTCTTAACTTCTTCGGCGACAGCATCACCGACTTCGGTGGATAGG CCACTAGTGGATCTG - 3'.

Stable integration of GPP synthase into AM68 and AM70 chromosome is expected to increase the available endogenous GPP pool restoring the GPP / FPP balance in modified yeast cells and provide a significant boost on terpene production.

Further modifications on sterol biosynthetic pathway that may maximize the production of terpenes in living cell making use of the endogenous substrates precursors pool are under construction .

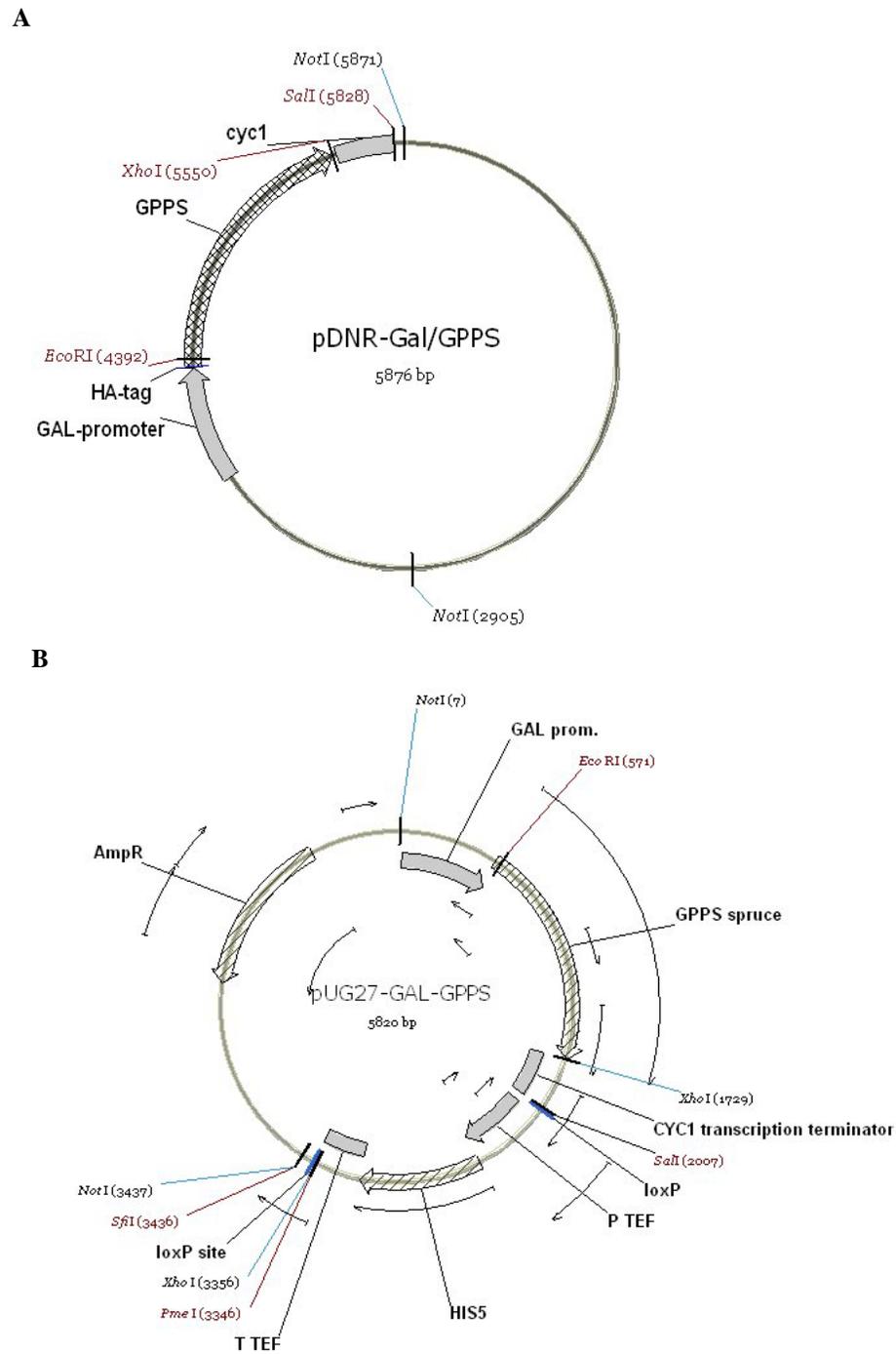


Figure III. 65. Maps of GAL-GPPS-*cyc1* cassette. **A)** pDNR-GAL-GPPS vector; **B)** pUG27-GAL-GPPS vector.

CHAPTER IV

IDENTIFYING INTERACTING PROTEINS OF CINEOLE SYNTHASE

1. Use of two hybrid system as a means of finding interactors of Cineole Synthase full-length *SfCinS1*

To further understand the function of cineole synthase in plants, two yeast two-hybrid screens was performed to identify potential protein interactors for this enzyme and elucidate its functional context in the cell as a part of a likely metabolom. Utilizing the features of the yeast two hybrid system, both the full length and the RC truncated version of cineole synthase were employed as baits during the experimental work. Two different cDNA expression libraries were used: a Tomato library kindly donated by Dr. G. Martin and a *Salvia fruticosa* glandular trichome cDNA library, both of them cloned into pJG4-5 library vector. Screening for cineole synthase *SfCinS* interactors was done using as a reporter system the yeast chromosomal LEU2 gene and selection was based on selecting for the yeast colonies growing on medium lacking leucine in the presence of galactose.

During the first trial, the *SfCinS1* full length based bait, was fused to the heterologous LexA DNA-binding protein in the pGILDA vector, under the control of GAL1 promoter. The fusion takes place on N-terminal end of *SfCinS1*. The Tomato library fused to the B42-activation domain in pJG4-5 plasmid was screened for potential interactors.

In both cases we used EGY48 yeast strain (Erica Golemis Yeast, MAT alpha, trp1, his3, ura3, lexA ops-LEU2). EGY48 is a derivative of U457 in which the endogenous LEU2 gene has been switched by homologous recombination with LEU2 reporters carrying 6 LexA operators.

1.1. Generating the bait construct

The subcloning of *SfCinS1* of 1976 bp bait fragment into pGILDA bait-vector was performed into *EcoRI* and *XhoI* sites (Fig. IV. 1).

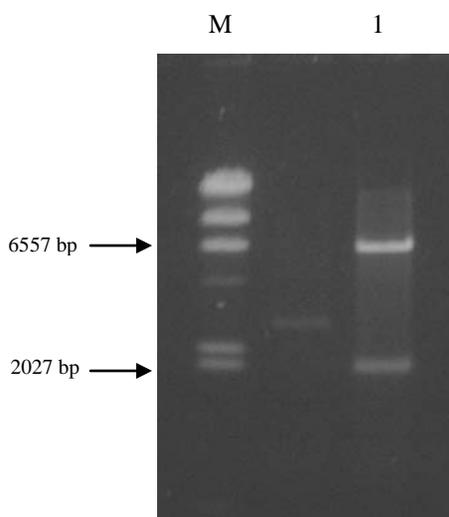


Figure IV. 1. Subcloning of *SfCinS1* into pGILDA, where M - Gene ruler, λ DNA/*Hind*III marker; 1 - pGILDA/*SfCinS1* digested with *Eco*RI and *Xho*I.

1.2. Screening for *SfCinS1* interacting proteins

The introduction of bait plasmid, pGILDA/*SfCinS1* into yeast cells was the first step in the screening. Cells prepared in this way, were transformed with yeast genomic library (cloned into pJG4-5 plasmid). At this point, yeast cells should contain two plasmids. Populations of cells expressing library proteins interacting with the bait are obtained by selecting the colonies growing in the absence of leucine (to maintain selection of EGY48 cells where interaction took place - first reporter system), histidine (to select for pGILDA), and tryptophan (selection for pJG4-5). Positives are selected manually by plating on leucine deficient plates and selecting growing colonies. Finally, the library plasmids are extracted from yeast, transformed into bacteria, propagated, purified and retransformed into the original yeast strain to verify the specific interacting phenotype. Positive library clones were subsequently sent for sequencing.

1.3. Essential control tests before library screening

Before performing library screening, it is necessary to run some control tests, which will ensure, that the colony growth is caused by interacting proteins, but not by any other random events. The controls established that the *SfCinS1* and *SfCinS(RC)* protein was made as stable protein in yeast, capable of entering the nucleus and binding LexA operator sites, and it does not appreciably activate transcription of the LexA operator-based reporter genes. These tests help to limit the frequency of false positives and false negative results.

1.3.1. Testing the autoactivation potential of the bait (LEU2)

The aim of this test is to ensure that EGY48, which contains the LEU2 gene, is an appropriate strain for screening. In order to check this, EGY48 cells were transformed with pGILDA/ *SfCinS1* and later plated on Glu/CM –H, Gall-Raff/CM –H and on Gall-Raff/CM –H – L. On plates where only histidine was absent the colonies were expected to grow, in medium without histidine and leucine, no colonies were expected to appeared. This approach helped confirm the suitability of the bait fusion for further experimentation. The potential problem of background colonies was thus assessed.

Three repetitions with different dilutions of transformants were carried out in order to eliminate any mistakes and help to better evaluate the assay. After plating and two day incubation, colonies on Glu/CM –H and Glu/CM –H, –L appeared (however at a substantially reduced number). There were hardly any colonies on Gall-Raff/CM –H, –L thus proving that there is no need to employ other less sensitive yeast strains. It is worth mentioning that the experiment was done by streaking large numbers of cells which favour the appearance of even trace background under regular conditions (Fig IV. 2).

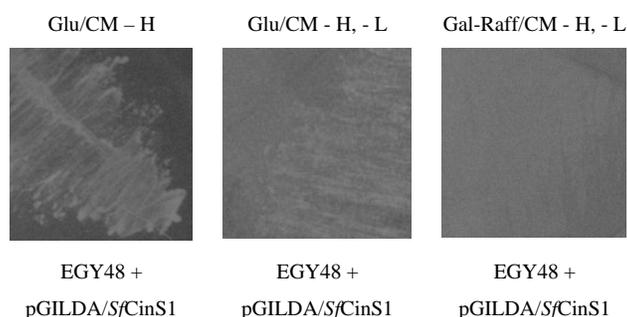


Figure IV. 2. The self activation assay of the bait proteins. Yeast cells carrying the bait were plate CM dropout media lacking leucine to test the ability of the bait itself to turn on the LEU2 reporter gene. No growth present on media lacking leucine can be observed.

1.3.2. Testing expression of bait protein

To assess appropriate expression of the fusion protein we used protein extracts from induced cells and subsequently performed SDS-PAGE and Western blotting with antiLexA antibodies. As a negative control, untransformed yeast EGY48 was used. After developing the film, the visible band with expected size appeared, thus verifying that bait protein expression is carried out in the yeast nucleus. The presence of the target protein, which was detected by

antibody against LexA, showed the presence of a the fused LexA-*SfCinS1* of ~70 kDa molecular weight (Fig. IV. 3).

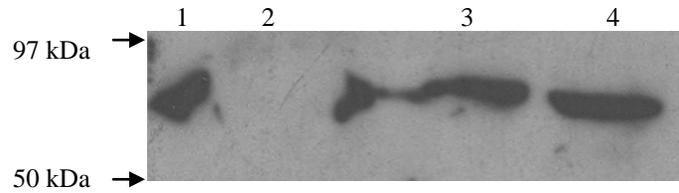


Figure IV. 3. Expression of LexA-*SfCinS1* recombinant protein in yeast cells. Western blot from yeast protein extract using anti-LexA antibody recognized the N-terminal fusion of the *SfCinS1* protein. 1 – positive control; 2 – EGY48 empty cells; 3 & 4 – EGY48 cells carrying pGILDA/ *SfCinS1*.

These controls established that the *SfCinS1* protein is expressed stably in yeast, and does not appreciably activate transcription of the *LexA* operator-based reporter genes.

1.4. Preparation of master plates and selection based on leucine requirements

The first step in interaction hunt was transformation of EGY48 cells with pGilda/*SfCinS1*. After selection of successful clones (based on growing yeast on histidine deficient media) cells were transformed again with the library plasmid pJG4-5/Tomato lib at a large scale (20 transformations). Transformants were plated on Glu/CM - H, - W, large plates, and incubated at 30 °C for >48 hours.

After colonies appeared on the plates, they were switched to 4 °C overnight to harden the agar. The following day, the colonies were scraped off the plates using a sterile microscope slide and placed into a 50ml sterile falcon tube. This cell slurry derived from >10⁶ primary transformants were homogeneously mixed. The cell slurry was subsequently washed with sterile water and centrifuged for 5 minutes at 4000 rpm. The supernatant was discarded. The washing procedure was repeated three times. Resulting pellets were resuspended in an equal volume of 50 % glycerol solution, aliquoted (500 µl per tube) and stored at -80 °C.

An aliquot of the stored slurry was thawed and protein expression was induced on Gall-Raff/CM - H, - W liquid media for 4-6 hours. 500 µl of the aliquots were resuspended into 9,5 ml of corresponding liquid medium containing galactose and raffinose. After 4 hours, serial dilutions were made out of inoculated culture to an optical density at 600nm ~0,5 (in this case 1:40). Subsequently, 200 µl of dilution with necessary optical density were plated on Gal-

Raff/CM - H, - W, - L plates and left at 30 °C until colonies started to appear (from 2-7 days). At this point, colonies and the library plasmid they contain were considered positives if they grew on Gal/Raff based media but not on Glu based media. Grown colonies were carefully picked and streaked on the glucose/CM-H, W master plate to use as starting point for all subsequent processing.

In total, 40 colonies were picked up during the first screening which were further processed (Fig. IV. 4).

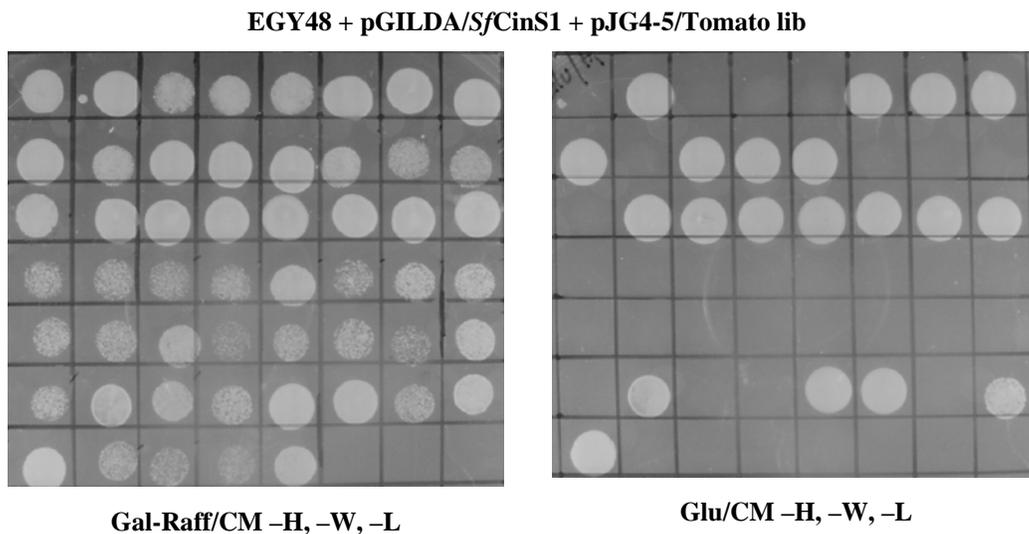


Figure IV. 4. Comparison between colonies grown on solid medium containing galactose and raffinose (where the gene coding for leucine gene was triggered) and glucose plates (where growth on leucine lacking plates was the result of non-specific activation of the reporter gene). Visible differences both in size of colonies as well as in their appearance show clones whose growth was dependent on interactions which take place between the protein of interest encoded by the bait plasmid and the library plasmid.

In order to exclude library clones in which leucine gene was activated due to potential autoactivation of prey, the colonies from master plates were restreaked on two different kinds of solid medium plates (Fig. IV. 4):

a) **Gal-Raff/CM -H, -W, -L**

(on these plates prey is expressed due to activation of promoter by presence of galactose).

b) **Glu/CM -H, -W, -L**

Clones which grow on this medium do not exhibit galactose depending phenotype - prey protein is not expressed, so their growth is not dependent on interaction between bait and prey).

1.5. Plasmid curing

To select the library plasmids that carry putative interacting proteins, the bait plasmid was counterselected by growing the yeast cells three successive times in 5 ml cultures in Glu/-W medium every 24 hours. By providing yeast exogenous source of histidine, we removed the need for yeast to maintain the histidine plasmid and, in effect, after four generations, the histidine plasmid was absent from most of the cells.

1.6. Isolation and characterization of the library interacting clones

The pJG4-5 plasmids, which containing cDNAs for putative interacting proteins with *SfCinS1*, were extracted according to the procedure described in Chapter II, paragraph 16.3.5. The extracted plasmids were subsequently transformed into *E. coli*.

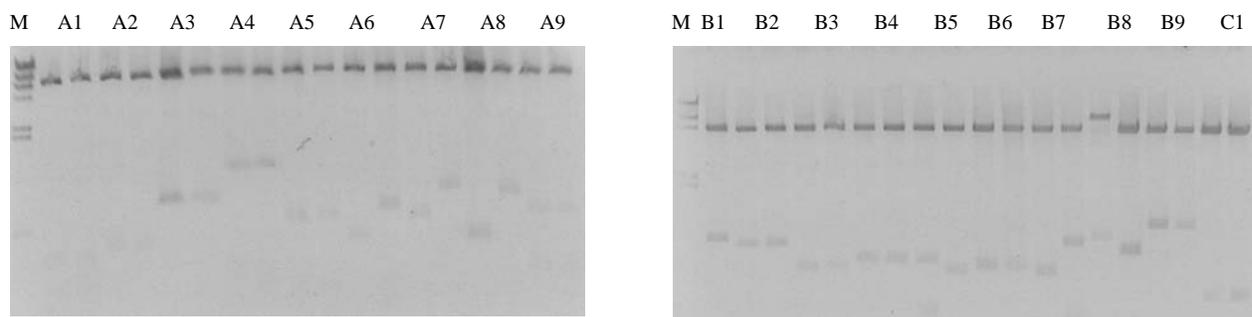


Figure IV. 5. Separation of interacting clones after double digestion with *EcoRI* and *XhoI* by agarose gel electrophoresis. Tomato interactors of *SfCinS1*: clones A3, A5, B9 are confirmed interactors encoding for Putative Phosphatase 2A Inhibitor, HSP90, Glutathione S-transferase.

After plasmid isolation from bacteria, plasmids were digested with restriction enzymes (*EcoRI* and *XhoI*) and resolved by agarose gel electrophoresis for the presence and the size of the insert (Fig. IV. 5).

1.7. Verification of the specificity of the interactors

The final confirmation of specificity was done by retransforming the purified plasmids into “virgin” *LexA*-operator-*LEU2*/*LexA*-operator-*lacZ*/pGILDA/*SfCinS1* – containing strain that has not been subjected to Leu selection and verifying that interaction-dependent phenotypes are still observed. Such false positives could include mutations in the initial EGY48 yeast that favour

growth on Gal medium, library-encoded cDNAs that interact with the LexA DNA-binding domain, or proteins that are sticky and interact with multiple biologically unrelated fusion domains (Serebriiskii et al., 1999).

The purified plasmids were used to transform yeast cells that already contain the following plasmids:

- pGILDA/*SfCinS1*
- pGILDA empty.

Transformation mix was plated on Glu/CM dropout plates lacking tryptophan and histidine. A master dropout plate was created for each library plasmid being tested and incubated 2 to 3 days at 30°C until the colonies appear.

Colonies from this master plate were restreaked to the same series of test plates previously used :

- Glu/CM, -H, -W, -L
- Gal/Raff/CM, -H, -W, -L.

Three positive cDNAs from the first screening have been considered true after this step. They grew on Gal/Raff/CM -Leu plates but not on Glu/CM -Leu plates and only in yeast cell carrying the bait and not in the ones carrying only the empty bait-vector (Fig. IV. 6).

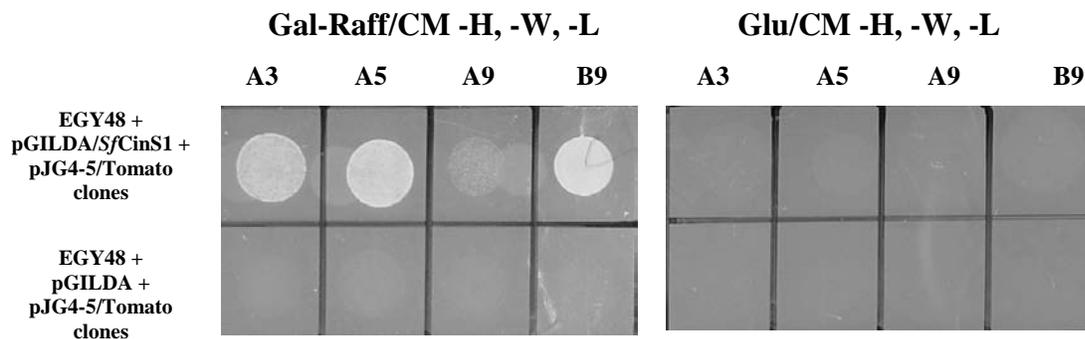


Figure IV. 6. Interaction-dependent phenotype of 4 positive cDNA Tomato library clones and *SfCinS1*. Visible differences in growth on Gall-Raff/CM- HWL plates between transformants resulted from lack of activation of LEU2 gene in control cells. Clones, in which interaction occurred between bait and prey protein showed no growth inhibition on galactose-raffinose, when compared to clones, where bait was not present.

1.8. cDNA clones identification

Three cDNAs clones from tomato library (Table IV. 1) clones from the Tomato library were further analyzed by sequencing in their 5' in order to reveal their identity. The primer used was from the library plasmid proximal to the cloning site. The sequencing data were subjected to nucleotide-nucleotide BLAST to identify the closest homologues. Three genes with different functions were revealed and shown below (Appendix 3).

Table IV. 1. Cineole synthase Tomato interactors

cDNA clone	Accession number	Clone identity	Biological processes	Identity percentage
A5	gi 38154492 gb AY368907.1	Lycopersicon esculentum molecular chaperon Hsp90-2	Heat-shock protein.	99 %
A3	gi 6730705 gb AAFC27100.1	Putative phosphatase 2A inhibitor	Cellular inhibitors for protein phosphatase 2A (PP2A). Important components in the reversible protein phosphorylation events.	73 %
B9	gi 37051105 dbj BAC81649.1	Glutathione S-transferase	Conjugate reduced glutathione with a large number of electrophiles.	68 %

1.9. Description of the identified interactors

1.9.1. Chaperone Hsp90-2

Sequence analysis of cDNA clone A5 suggests that there are two sequences showing high similarity to this clone, *LeHsp90-1*, 96 % identity and *LeHsp90-2*, 99 % identity. The clone encodes open reading frames of 290 amino acids and possess the C-terminal MEEVD pentapeptide that is characteristic to cytoplasmic Hsp90 proteins (Krishna and Gloor, 2001). Comparison of the amino acids sequence of this protein with *Arabidopsis* cytoplasmic Hsp90 protein revealed very high homology.

Heat-shock protein 90 (Hsp90) is an abundant and highly conserved molecular chaperone that is essential for viability in eukaryotes. Despite being a heat-shock protein, Hsp90 is one of the most abundant proteins in unstressed cells (1-2 % of cytosolic protein) where it performs

housekeeping functions contributing to the folding, maintenance of structural integrity and proper regulation of a subset of cytosolic proteins.

Heat-shock proteins (Hsps) act as molecular chaperones, guiding nascent polypeptides through the process of folding and maturation into three-dimensional structures (Pearl LH et al., 2008; Söti C. et al 2005). Chaperones are also responsible for refolding denatured proteins that result from cellular stresses such as nutrient deprivation, abnormal temperature or pH, malignancy, and exposure to various toxins and drugs (Chaudhury S et al., 2006). Heat-shock response is conserved across all species, from prokaryotes to eukaryotes, and provides a mechanism for general upkeep of intercellular processes, including protection against protein aggregation in the cytosol (Frydman J., 2001). It exists in four isoforms: Hsp90 α , Hsp90 β , glucose-regulated protein (GRP94), and Hsp75/tumor necrosis factor receptor associated protein 1 (TRAP-1). Hsp90 α and Hsp90 β can be found in the cytosol, and are the inducible and constitutive forms, respectively.

To date, Hsp90 has been found to interact with over 200 client proteins, as well as ~50 co-chaperones, making it a cornerstone in the cellular protein-folding machinery and an emerging target for the treatment of various disease states (Picard, D. 2008).

1.9.2. Putative phosphatase 2A inhibitor

The clone A3 encodes for a putative phosphatase 2A inhibitor with 73 % identity. PP2A represents a group of highly abundant and ubiquitously expressed Ser/Thr phosphatases in eukaryotes, conserved from yeast to human (Jiang Y, 2006); its activity is found in numerous cellular processes. It was first implicated in cell cycle control by findings showing that its inactivation promoted premature mitotic entry. Since then, its role in mitotic entry has been of a great interest. Type 2A serine/threonine protein phosphatases (PP2A), important components in the reversible protein phosphorylation events in plants and other organisms, are oligomeric complexes constituted by a catalytic subunit and several regulatory subunits that modulate the activity of these phosphatases (Javier Terol et al., 2002). Thus, PP2A proteins can be heterodimers, consisting of a PP2Ac catalytic subunit and a type A regulatory subunit, or heterotrimers that contain an additional regulatory subunit of the B type. PP2Ac subunits are highly conserved in all organisms analyzed, and their activity, specificity, and subcellular

localization depend on the association of this subunit with different A and B regulatory subunits (Hendrix et al., 1993b; Strack et al., 1998).

1.9.3. Glutathione S-transferase

The last clone isolated as real interactor, B9, encodes for a glutathione S-transferase with 68 % identity. These enzymes (GSTs; also known as glutathione S-transferases) are major phase II detoxification enzymes found mainly in the cytosol. In addition to their role in catalyzing the conjugation of electrophilic substrates to glutathione (GSH), they also carry out a range of other functions. They have peroxidase and isomerase activities, they can inhibit the Jun N-terminal kinase (thus protecting cells against H₂O₂-induced cell death), and they are able to bind non-catalytically a wide range of endogenous and exogenous ligands. Cytosolic GSTs of mammals have been particularly well characterized, and were originally classified into Alpha, Mu, Pi and Theta classes on the basis of a combination of criteria such as substrate/inhibitor specificity, primary and tertiary structure similarities and immunological identity (Sheehan et al., 2001).

GSTs are dimeric, mainly cytosolic, enzymes that have extensive ligand binding properties in addition to their catalytic role in detoxification. They have also been implicated in a variety of resistance phenomena involving cancer chemotherapy agents, insecticides, herbicides and microbial antibiotics.

2. Yeast two hybrid screening of a Salvia library to identify interactors of RC truncated cineole synthase *SfCinS*(RC)

Further on a second screening was performed at a pilot scale in collaboration with Lolita Guardari, MSc student. In order to prevent losing potential interactors because of the N-terminal signal peptide of cineole synthase sensitivity which may block the activity of the enzyme, the RC truncated *SfCinS* based bait was used. The LexA-bait fusion was designed in a pYES-LexA engineered vector fusing the LexA protein on the C-terminal end of *SfCinS*(RC). The Salvia library was screened during the second approach.

2.1. Generation of CS (RC)-LexA bait plasmid

A novel recombinant vector was generated to assist cloning the bait sequence in such a way that the LexA fusion will take place at the C-terminal end of the protein.

The LexA gene was PCR amplified using specific primers (see Chapter II, paragraph 16.1.1.3) from pEG202 backbone, and cloned into the pYES2 vector in the *Xho*I and *Xba*I sites (Fig. IV. 7.).

To functionally express the bait protein as N-terminal fusion into the newly generated vector pYES-LexA the bait sequence orf was amplified removing the terminal stop and incorporating an in-frame *Xho*I restriction site to allow the translation of whole bait-LexA recombinant protein. We choose the *Sf*CinS(RC) truncated cineole synthase, which showed slightly higher cineole production than other truncated forms, as bait for the current screening.

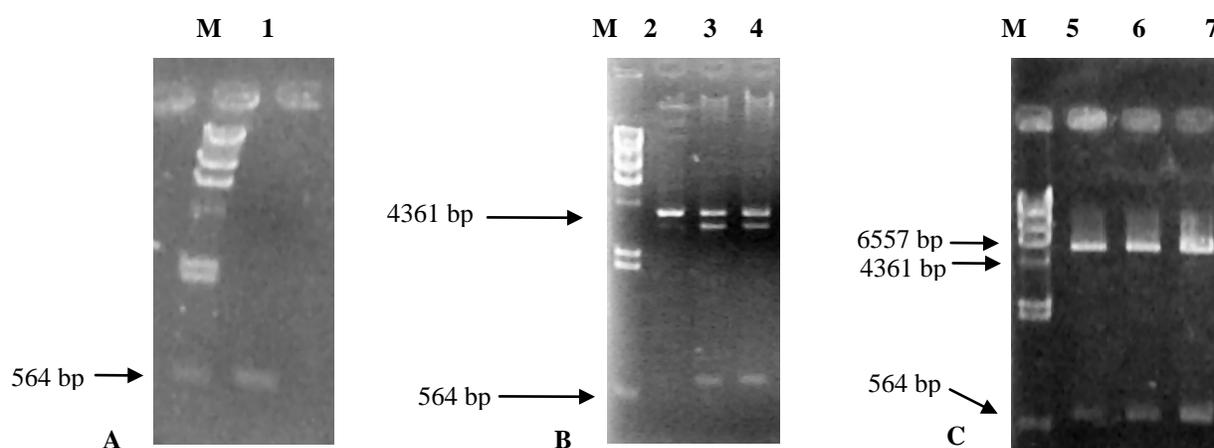


Figure IV. 7. Cloning of LexA fragment into pYES2 vector. A) Agarose gel electrophoresis of PCR amplification of LexA; 1 – PCR product of LexA; B) Digestion with *Eco*RI of pTOPO-LexA vector (2-4); C) Digestion with *Xho*I and *Xba*I of pYES-LexA construct (5-7), where M - Gene ruler, λ DNA/*Hind*III marker. The recombinant DNA - number 3, 4, 6, 7 were validated as right constructs.

The *Sf*CinS(RC) was PCR amplified with specific primers using as template pJG4-6/*Sf*CinS(RC) recombinant vector generated in the frame of this project (see Chapter II, paragraph 8.4) and then cloned into pYES-LexA vector into *Eco*RI and *Xho*I frame (Fig. IV. 8).

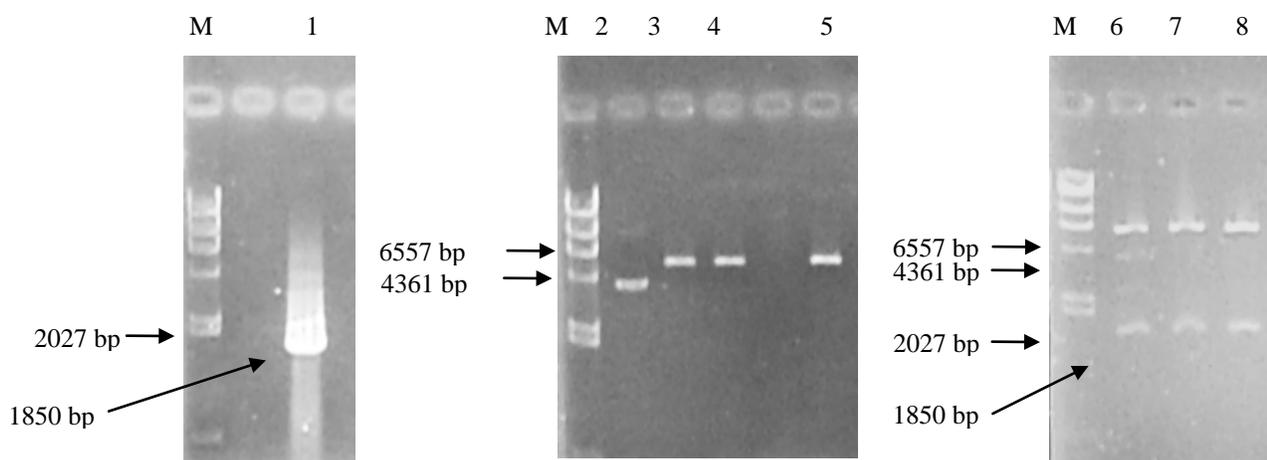


Figure IV. 8. Cloning of *Sf-CinS1* fragment into pYES-LexA vector. A) Agarose gel electrophoresis of PCR amplification of *SfCinS(RC)* no stop codon, where M - Gene ruler, λ DNA/*HindIII* marker; 1 - PCR product of *SfCinS(RC)*; B) Digestion with *EcoRI* and *XhoI* of pYES-LexA vector: 1 - pYES-LexA uncut; 2 - pYES-LexA cut with *EcoRI*; 3 - pYES-LexA cut with *XhoI*; 4 - pYES-LexA cut with *EcoRI* and *XhoI*; C) Digestion with *EcoRI* and *XhoI* of pYES-*SfCinS(RC)*-LexA; 6-8 - pYES-*SfCinS(RC)*-LexA recombinant vectors, all validated as right constructs.

The activity of the bait protein *SfCinS(RC)*-LexA expressed was assessed for cineole production in EGY48 yeast cells by extracting the volatiles released into the head space of the flask by the SPME method and analyzing them by Gas Chromatography. The *SfCinS(RC)* enzyme was able to synthesize cineole using the endogenous GPP pool (Fig. IV. 9).

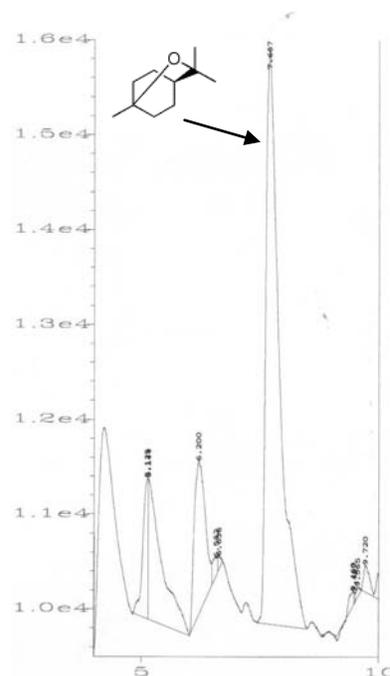


Figure IV. 9. The volatiles produced by yeast cells expressing *Sf-CinS1(RC)* subcloned in pYES-LexA vector were sampled by solid-phase micro-extraction (SPME) and analyzed by Gas Chromatography.

2.2. The pilot scale of the screening

Following the same steps as in previous screening (see Chapter IV, Subchapter 1) a new screening was performed in order to isolate interactors of truncated RC cineole synthase from *Salvia fruticosa* cDNA library. 140 colonies were isolated after the first test as putative interactors and were cured from the bait plasmid, pYES- *SfCinS*(RC)-LexA, by growing on FOA media. 24 colonies were further processed in a pilot screening to assess the efficiency of using the new bait plasmid, pYES-LexA. The rest of the colonies were collected on master plates and stored at 4°C for further exhaustive screening.

Three positives cDNAs for this pilot screening have been considered real after the final test. They grew on Gal/Raff/CM –Leu plates but not on Glu/CM –Leu plates only in yeast cell carrying the bait and not in the ones carrying only the empty bait-vector (Fig. IV. 10)

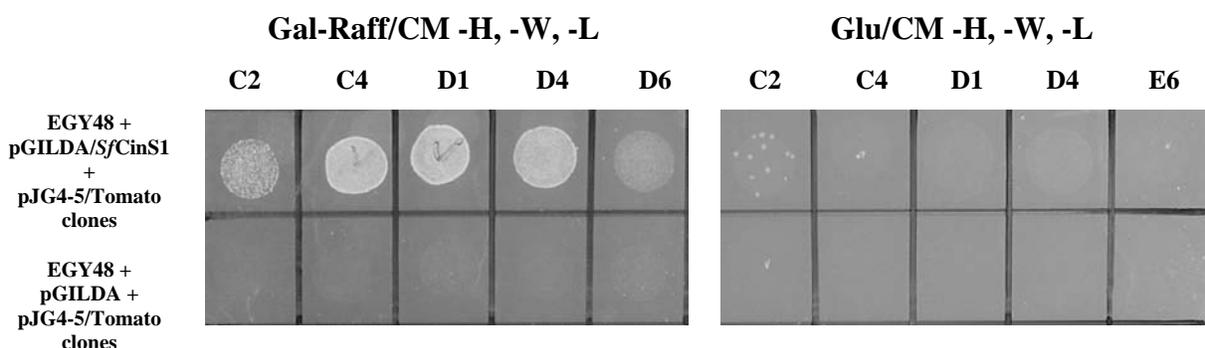


Figure IV. 10. Interaction-dependent phenotype of 5 positive cDNA *Salvia* library clones and *SfCinS*(RC). Visible differences in growth on Gal-Raff/CM HWL plates between transformants resulted from lack of activation of *LEU2* gene in control cells. Clones, in which interaction occurred between bait and prey protein showed no growth inhibition on galactose-raffinose, when compared to clones, where bait was not present.

2.3. cDNA clones identification

3 cDNAs clones (Table IV. 2) from the *Salvia* library were further analyzed by sequencing in order to reveal their identity. The sequencing results were subjected to nucleotide-nucleotide BLAST to identify the closest homologues (Appendix 3). Three genes with different functions were revealed and shown below (Table IV. 2).

Table IV. 2. Cineole synthase *Salvia* interactors

cDNA clone	Accession number	Clone identity	Biological processes	Identity percentage
C4	sp Q03684 BIP4_TOBAC	Luminal-binding protein 4 precursor (BiP4)	Probably plays a role in facilitating the assembly of multimeric protein complexes inside the ER.	96%
D4	gb ACI31551.1	Heat shock protein 90-2	Heat-shock protein.	91%
D1	ref NP_173637.3	SEC14 cytosolic factor family protein	Domain in homologues of a <i>S. cerevisiae</i> phosphatidylinositol transfer protein (Sec14p) and in RhoGAPs, RhoGEFs and the RasGAP, neurofibromin (NF1). Lipid-binding domain. The SEC14 domain of Dbl is known to associate with G protein beta/gamma subunits.	67%

2.4. Description of the identified interactors

2.4.1. Luminal-binding protein 4 precursor (BiP4)

Sequence analysis of cDNA clone C4 of 1048 bp suggests high similarity to Luminal-binding protein 4 precursor (BiP4) 96 % identity and Luminal-binding protein 5 precursor (BiP5) 96 % identity. The clone encodes open reading frames of 188 amino acids. Comparison of the amino acids sequence of this protein with *Nicotiana tabacum* luminal binding protein (BiP) revealed very high homology.

Luminal binding proteins (BiP) have been identified in various mammals (Munro and Pelham, 1986; Ting et al., 1987; Haas and Meo, 1988; Wooden et al., 1988) and yeasts (Normington et al., 1989; Rose et al., 1989), is a member of the HSP70 family that accomplishes its function in the lumen of the endoplasmic reticulum (ER). BiP differs from the other family members in the presence of an N-terminal signal sequence that is required for the cotranslational translocation of proteins through the ER membrane. BiP was shown to bind to newly synthesized, incompletely assembled or malfolded proteins in the lumen of the ER (Haas and Wabl, 1983; Bole et al., 1986; Gething et al., 1986; Kassenbrock et al., 1988). Although constitutively present during normal growth, a sharp induction occurs upon conditions that cause the accumulation of misfolded proteins in the lumen of the ER (Lee, 1987; Kozutsumi et al.,

1988). BiP may prevent the secretion of proteins that have not yet acquired their mature conformation (Bole et al., 1986; Gething et al., 1986), promote protein folding and assembly (Gething et al., 1986), or dissolve protein aggregates that have been formed under normal or stress conditions in the ER (Munro and Pelham, 1986). The recent finding that loss of BiP function blocks translocation of secretory proteins in yeast (Vogel et al., 1990) suggests that the gene product plays a constitutive role both in protein import into the lumen of the ER and in the subsequent maturation steps *in vivo*. Continuous binding of BiP to proteins would only occur so long as proper folding or assembly cannot be achieved, because of either mutations in the transported proteins or external stresses that impair maturation.

Most BiP–protein interaction studies have been performed with synthetic peptides, and only a few model proteins have been analyzed (Vitale et al., 1995; Pedrazzini et al., 1997; Nuttall et al., 2002; Foresti et al., 2003; Mainieri et al., 2004; Randall et al., 2005). BiP–ligands can be coimmunoprecipitated and subsequently released by addition of ATP *in vitro*, suggesting an energy-dependent release mechanism (Munro and Pelham, 1986; Vitale et al., 1995). Consistent with this assumption, dominant-negative BiP ATPase mutants are able to bind to unfolded proteins but compromise proper folding because they fail to release their ligands (Hendershot et al., 1996). When nascent polypeptides emerge in the lumen of the ER during cotranslational translocation, they are still unfolded and interact with BiP (Matlack et al., 1999). Therefore, any secretory protein is a potential BiP–ligand during its life as a folding intermediate. When folding and assembly are complete, hydrophobic regions are no longer displayed and BiP will cease to bind. Permanently misfolded proteins can emerge when physiological conditions are unfavorable or when erroneous proteins are synthesized. In these cases, continuous binding of BiP results eventually in ER-associated protein degradation (ERAD). The current pathway for this event leads via the translocation pore back to the cytosolic proteasome (McCracken and Brodsky, 2003).

BiP is constitutively expressed under normal growth conditions, but transcription can be induced by the accumulation of misfolded proteins in the ER (Kozutsumi et al., 1988; Denecke et al., 1991; Fontes et al., 1991; Nuttall et al., 2002). This process has been termed the unfolded protein response (Rutkowski and Kaufman, 2004; Zhang and Kaufman, 2004). In addition, plant BiP can be induced by unfolded protein response–independent signal transduction pathways (Kalinski et al., 1995; Jelitto-Van Dooren et al., 1999; Wang et al., 2005). However,

transcriptional induction of BiP seldom leads to increased BiP protein levels, even though mRNA concentrations and pulse labeling demonstrate a higher synthesis rate (Leborgne-Castel et al., 1999). This suggests that under ER stress, BiP turnover is increased, but it remains to be shown where BiP degradation occurs.

2.4.2. Heat shock protein 90-2

Sequence analysis of cDNA clone D4 suggests that there are sequences showing high similarity to this clone, Hsp90-2 from *Glycine max* revealed 91% identities, while *LeHsp90-1*, 87% identities and *LeHsp90-2*, 88% identities. The *Salvia* clone encodes open reading frames of 294 amino acids and possess the C-terminal MEEVD pentapeptide that is characteristic to of cytoplasmic Hsp90 proteins as the tomato clone previous isolated (Krishna and Gloor, 2001). Comparison of the amino acids sequence of this protein with *Arabidopsis* cytoplasmic Hsp90 protein revealed high homology.

2.4.3. SEC14 cytosolic factor family protein

Sequence analysis of cDNA clone D1 of 1022 bp suggests similarity to *Arabidopsis* SEC14 cytosolic factor family protein / phosphoglyceride transfer family protein revealed 67% identities. The *Salvia* clone encodes open reading frames of 199 amino acids.

The major phosphatidylinositol transfer proteins (PITP) in yeast *Saccharomyces cerevisiae*, Sec14p, is essential in promoting Golgi secretory function by modulating of its membrane lipid composition. SEC14 is a peripheral membrane protein of the Golgi apparatus where its function is necessary for membrane trafficking from the TGN subcompartment of this organelle. Bioinformatic analyses identify Sec14-like proteins as uniquely eukaryotic proteins that are found in all eukaryotic genomes sequenced to date, and are abundant (>500 members) (Phillips et al., 2005). While this abundance defines a Sec14-protein superfamily, there is remarkably little information regarding their precise physiological functions within cells. It is remarkable that a soluble protein is able to efficiently extract a phospholipid molecule from a membrane bilayer without the need ATP hydrolysis or the action of other cofactors. How release of bound phospholipid is regulated is also mysterious.

As indicated above, this superfamily consists of greater than 500 members, of which the yeast Sec14p is the prototype, and is limited to the *Eukaryota*. The *SEC14* structural gene was

initially defined as a complementation group represented in the original collection of yeast secretory (*sec*) mutants (Phillips et al., 2005). Subsequent molecular characterization of the isolated gene demonstrated it is essential for the viability of yeast cells, and that the protein product is required for membrane trafficking from the yeast TGN. In that regard, there is an interesting specificity to this requirement. Export of the secretory glycoprotein invertase is strongly compromised by Sec14p deficiency whereas trafficking of the vacuolar proteinase carboxypeptidase from the TGN into the endosomal system, and ultimately to the vacuole, is only mildly affected. Thus, while Sec14p controls an essential trafficking pathway from the TGN, it is not required for all trafficking pathways from this Golgi subcompartment.

The prototypical Sec14p from *Saccharomyces cerevisiae* is two-lobed globular protein with a large hydrophobic pocket, which binds the entire PtdIns molecule. While the same overall topology is found in all Sec14p-like proteins, they differ from each other mostly in the region predicted to bind the head group of PtdIns. Other known ligands for these proteins include phosphatidylcholine, trans-retinaldehyde, and PtdIns P3. It seems that Sec14p-like proteins have been adapted during evolution to fulfill a number of different functions that depend on protein–lipid interactions (Mousley C.J. et al., 2007).

3. Assessment of *SfCinS1* activity in the presence of Hsp90 interacting protein.

For further work, HSP90 was chosen to estimate its contribution to terpenoids production in yeast cells. One straightforward way to assess its role in cineole accumulation in yeast cells is to co-express HSP90 in yeast cells and test the production of cineole in increased temperatures which pose a strain on enzyme stability.

3.1. Cloning of Hsp90 into yeast vectors

Full-length Hsp90 of 2359 bp from modified pBluescript1/*AtHsp90* (RALF06-10-K09, RIKEN Arabidopsis) recombinant vector was amplified in a PCR reaction using Hot Star Polymerase and gene-specific primers (see Chapter II, paragraph 8.7).

The amplified PCR product was resolved into 1% agarose gel, purified by gel extraction and cloned into the pCRII vector in a TOPO TA reaction. The Hsp90 fragment was pull out by

restriction digestion with *Bam*HI and *Xho*I and ligated into pYES2 and pYX143HA yeast vectors, linearized with the same enzymes (Fig. IV. 11).

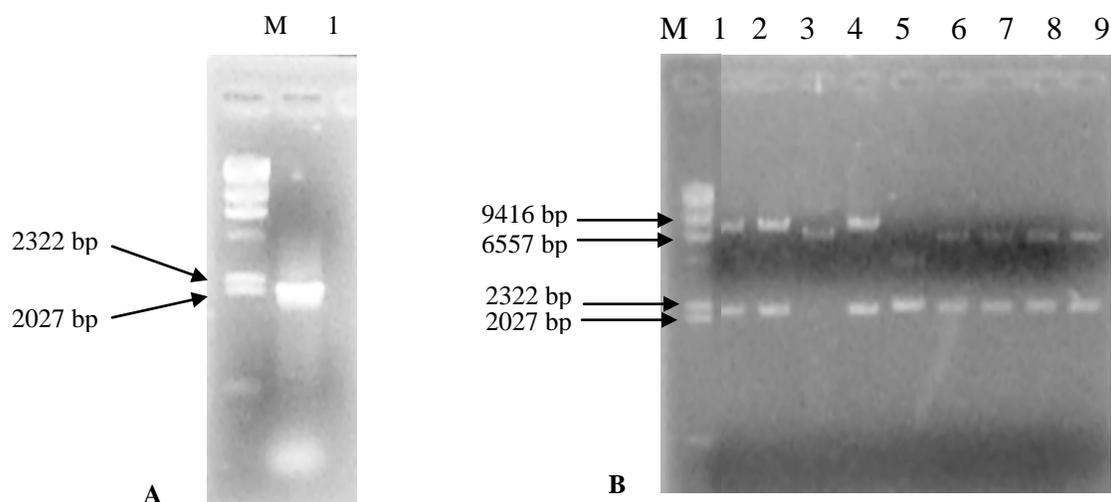


Figure IV. 11. Cloning of AtHSP90-2 into pYES and pYX143HA yeast vectors. A) Agarose gel electrophoresis of PCR amplification of AtHSP90-2 where M - Gene ruler, λ DNA/*Hind*III marker; 1 – PCR product of AtHsp90-2; B) Digestion with *Bam*HI and *Xho*I of pYX143-HA/AtHsp90-2 (1-4) and pYES/AtHsp90-2 (5-9). The recombinant DNA - number 1,2,4,6,7,8,9 were validated as right constructs.

To further evaluate the significance of the interaction of Hsp90 with cineole synthase truncated versions, *Sf*CinS1(RC) and *Sf*CinS1(RR), the two proteins have been co-expressed into AM63 yeast cells, assessed as high terpene producing yeast strain. Co-transformants have been induced according to the protocol described in Chapter II, paragraphs 11.1.1 and 11.1.2 for promoter induction as well as for terpene synthase induction. To assess the effect of Hsp90 protein on enzyme activity, the cineole synthase induction by enzymatic specific buffer (see Chapter II, paragraphs 11.1.2) has been carried out both at 30 °C and 37 °C. Hsp90 protein, isolated twice as a cineole synthase interactor, was shown to influence the activity of these enzymes during temperature-induced stress by increasing cineole production in AM63 yeast cells. When HSP90 was co-expressed with truncated *Sf*CinS1(RC) in AM63 yeast cells at 30 °C the product formation improved by about 30%. Higher temperature, such as 37 °C not only had no effect on cineole production by *Sf*CinS1(RC) co-expressed with HSP90 in AM63 yeast cells, but also brought a gain of 40 % more product (Fig. IV. 12).

The outcome proved that HSP90 acts functionally as a real interactor of cineole synthase that has a valuable contribution on enzyme activity, most probably monitoring proper protein folding, especially during stress conditions.

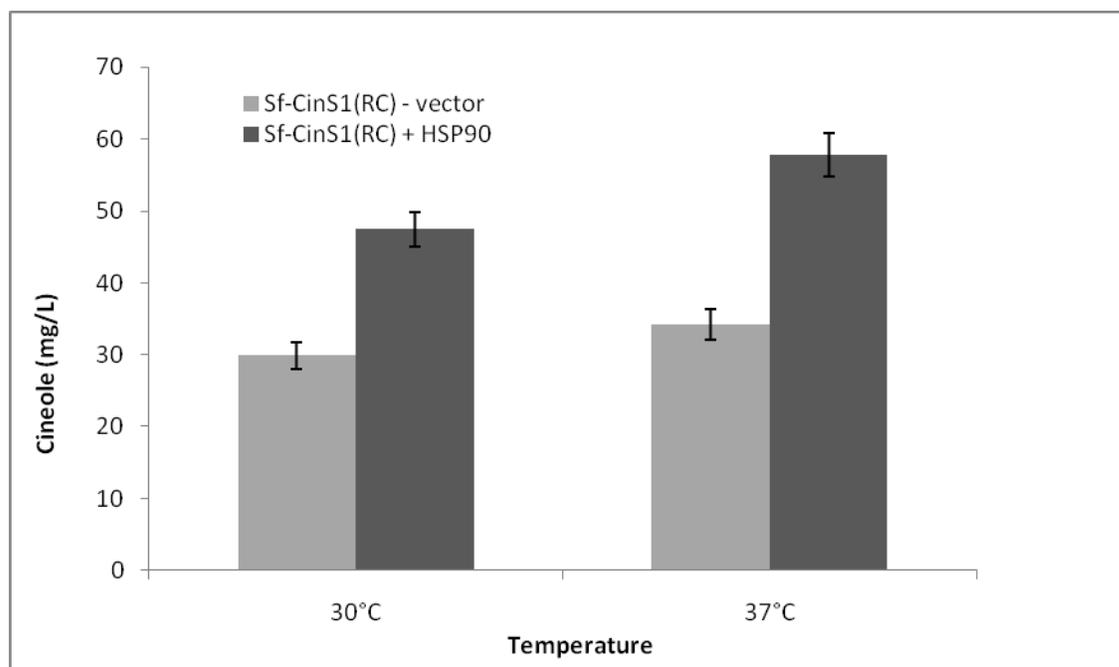


Figure IV. 12. Cineole yield by yeast strain AM63 engineered with *Sf-CinS1(RC)* and HSP90, empty plasmid. The effect of HSP90 interacting protein of cineole synthase on cineole activity at 30°C and 37°C. Bars indicate the average cineole amount; this experiment was repeated three times with similar results.

The two different yeast two hybrid screenings suggest that N-terminus region of monoterpene synthases is particularly sensitive to changes and these are involved significantly in the enzymatic activity of the enzyme. The pilot modified screening employing bait-LexA fusion on C-terminal end of *SfCinS1(RC)* uncovered valuable interactors of cineole synthase which possible play role in protein transport and folding. These findings encourage to further proceed an extensive screening of *Salvia fruticosa* cDNA library. The remaining interactors will be co-expressed in yeast cell with different terpene synthases and assessed for their contribution to terpene production.

CHAPTER V

OXIDATIVE STRESS IN PLANTS

1. Expression of oxidative stress response genes in *Arabidopsis*

In the context of the collaboration in the PENED-GSRT funded program we undertook to examine the *in vivo* function of 3 newly isolated genes in transgenic *Arabidopsis* plants: an isoflavone reductase (IFR) homologue from *Salvia fruticosa*, the glutathione *S*-transferase/peroxidase (BI-GST) and its interacting partner thioredoxin peroxidase 1 (LeTpx1), both genes from *Lycopersicon esculentum*.

Isoflavone reductase (IFR) is an NAD(P)H-dependent oxidoreductase that reduces achiral isoflavones to chiral isoflavanones during the biosynthesis of chiral pterocarpan phytoalexins. A homologue of isoflavone reductase (IFR) was isolated from the *Salvia fruticosa* cDNA library. The protein was assessed for enzymatic activity after bacterial expression using dehydrodiconiferyl alcohol as a substrate and the formation of the product isodihydrodehydrodiconiferyl alcohol was confirmed by LC-MS (Mohamed El Sayed – Master thesis, MAICh 2007). The plant glutathione *S*-transferase BI-GST has been identified as a potent inhibitor of Bax lethality in yeast, a phenotype associated with oxidative stress and disruption of mitochondrial functions (Kilili et al., 2004). LeTpx1 was initially identified as an interacting protein to the glutathione *S*-transferase/peroxidase (BI-GST) (Kampranis et al., 2000) exhibited protection of yeast cells from peroxide-induced cell death, showed atypical glutathione-dependent thioredoxin/peroxidase activity and most likely protecting free cysteines from oxidation.

The *Agrobacterium tumefaciens* AGL-1 strain harboring two binary system recombinant plasmids: (i) the pSoup based helper, plasmid pAL154, and (ii) derivatives of the pGreen0229 d constructed in our lab (see Chapter II, paragraph 3.12.3) were used in *Agrobacterium*-mediated gene transfer. The proteins are expressed under the control of the CaMV 35S promoter. In these experiments, several plasmids were used, all pGREEN300 (35S-MCS-TT) derivatives carrying various inserts (IFRh, BI-GST, LeTpx1, V5-LeTpx1, 51CS-LeTpx1, 76CS-LeTpx1, and 51-

76CS-LeTpx1) which encode for full length isoflavone reductase (IFR) homologue, the glutathione *S*-transferase/peroxidase (BI-GST), and the thioredoxine peroxidase (LeTpx1) homologous proteins, a V5-LeTpx1 fusion protein, and point mutations in the active site of thioredoxin/peroxidase LeTpx1, (51CS)- LeTpx1, (76CS)-LeTpx1, and (51-76CS)-LeTpx1 proteins.

To verify successful transformation of the pGREEN vector constructs carrying the LeTpx1 mutants: pGREEN300/51CS-LeTpx1, pGREEN300/76CS-LeTpx1, pGREEN300/51-76CS-LeTpx1, into the AGL-1 *Agrobacterium* strain, colony PCR reactions were performed. Two colonies of each transformant were picked from the YEP-Km agar plates. The PCR reaction aimed to amplify a ~500 bp segment of the bacterial *bla* phospho resistance gene (BAR) encoding the enzyme phosphotricinacetyl transferase (PAT), which confers resistance to glufosinate ammonium herbicide (Basta) and serves as a marker for selection of transformed plants. The PCR products were analyzed by electrophoresis on 1 % agarose gel (Fig. V. 1).

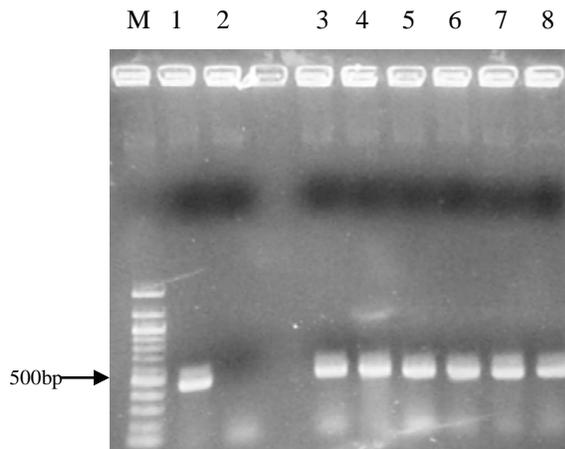


Figure V. 1. Colony PCR products from *Agrobacterium* cells transformed with pGREEN300/1CS-LeTpx1 (3,4), pGREEN300/76CS-LeTpx1 (5,6), pGREEN300/51-76 CS-LeTpx1 (7,8), where M - Gene ruler, λ DNA/*Hind*III marker; 1 - pGREEN300/LeTpx1 vector - positive control; 2 - negative control, empty PCR reaction.

Transformation of *Arabidopsis* plants was performed using a floral dip vacuum infiltration protocol (Bechtold and Pelletier, 1998; Clough and Bent, 1998) using *A. tumefaciens* AGL-1 strain harboring binary plasmid system: (i) the pSoup based helper, plasmid pAL154, and (ii) pGREEN300/IFRh, pGREEN300/BI-GST, pGREEN300/LeTpx1, pGREEN300/V5-LeTpx1, pGREEN300/51CS-LeTpx1, pGREEN300/76CS-LeTpx1, pGREEN300/51-76CS-LeTpx1, and respectively pGREEN300 empty vector. The *Agrobacterium* cells expressing proteins under the control of CAMV35S promoter was used to infect the floral part of healthy wild type *Arabidopsis* Colombia 00 plants, grown to flowering under long day photoperiod in soil in pots.

The first bolts were clipped to encourage proliferation of many secondary bolts. The plants were ready after 8 days of clipping. The plant transformation was achieved according to the protocol described in Chapter II, paragraph 9.5. The plants which endured the transformation procedure were grown until seed maturation and the seeds were collected.

Transformants' selection was based on resistance to glufosinate ammonium herbicide (Basta) to a final concentration of 200-300 μM .

The collected T_0 seeds were replanted in new pots which were placed at 4 °C to break the dormancy. After 4 days the pots were moved to a growing chamber under 16 hours light / 8 hour dark photoperiod. When seedlings reached the stage of 4 leaves, they were sprayed with the required concentration of Basta two rounds with one week interval between, each round consisting of three sprayings with one day interval between. Approximately 8% of plants survived. The seeds of the first generation surviving plants were collected and replanted. The second generation of plants was sprayed again with Basta as described above to remove spontaneous Basta resistant plants. Approximately 90% of T_1 plants survived the Basta treatment, which verified the effectiveness of Basta selection protocol (Fig. V. 2.). Surviving the Basta selection is a first indication of the presence of the transgene of interest in the transformation.

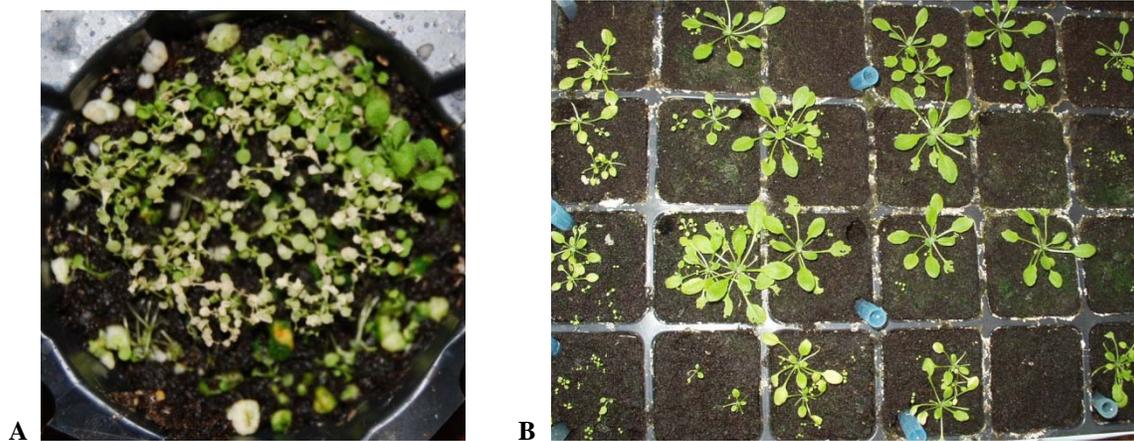


Figure V. 2. Basta selection process. A) *Arabidopsis thaliana* transgenic plants carrying LeTpx1 transgene during selection with Basta; B) Second generation of *Arabidopsis thaliana* transgenic plants carrying BI-GST transgene, Basta resistant.

The effectiveness of the Basta selection method was confirmed by PCR amplification of one of the transgenes. Leaves of surviving plants from T₁ generation considered putative transgenic plants, as well as leaves of wild type plants were collected when fully expanded and used to extract genomic DNAs, by phenol:isoamylalcohol:chlorophorm procedure (see Chapter II, paragraph 4.2.2.). The extracted genomic DNA was checked by agarose gel electrophoresis (Fig. V. 3).

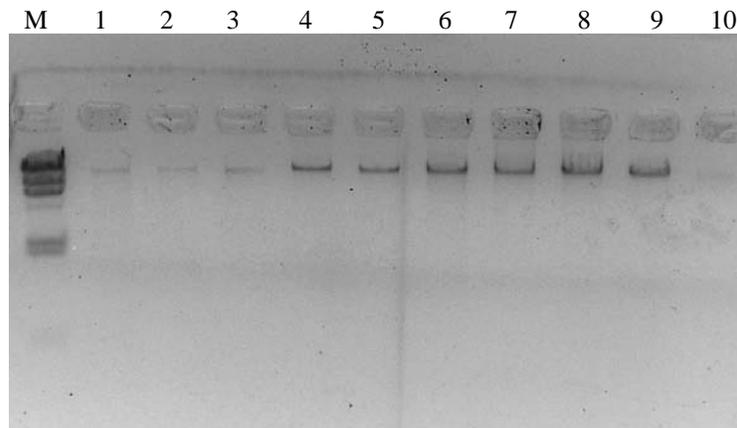


Figure V. 3. *Arabidopsis thaliana* genomic DNA isolated by agarose gel electrophoresis (1-10), where M - Gene ruler, λ DNA/*Hind*III marker.

The extracted genomic DNAs were used as template in a PCR amplification to test the presumed transgenic plants for the presence of the transgene, the bacterial bialophos resistance gene (BAR), by using gene-specific primers and higher annealing temperature, 60 °C, than standard conditions (see Chapter II, paragraph 6.1.). 70 randomly selected putative transformed plants (10 plants with IFR, 10 plants with BI-GST, 10 plants with LeTpx1, 10 plants with LeTpx1-CS51, 10 plants with LeTpx1-CS76, 10 plants with LeTpx1-CS51-76, and 10 plants with empty vector) were examined. The PCR amplification of genomic DNA yielded amplified products at expected size (~500 bp) in 54 examined samples and no bands were observed in control DNA from non-transformed *Arabidopsis* plants. No amplified products were detected in the DNA free control sample. (Fig. V. 4).

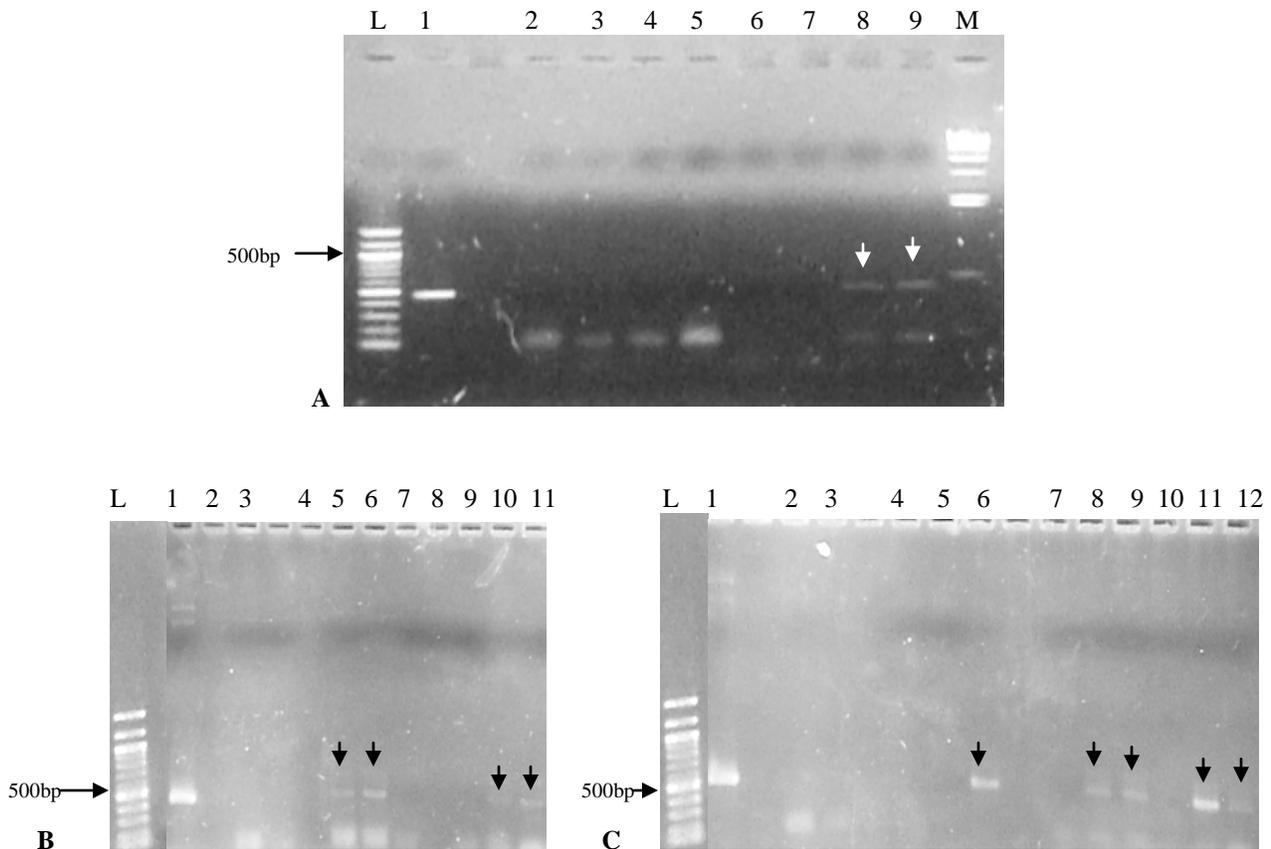


Figure V. 4. PCR amplification of the bacterial bialophos resistance gene (BAR) using *Arabidopsis thaliana* genomic DNAs as a template. A) 1 - pGREEN300/IFR; 2 - At wild type; 3 - negative control; 4 - AtIFR; 5 - AtBI-GST; 6-7 - AtLeTpx-CS76; 8-9 - At pGREEN300 empty; Plant 8 and plant 9 are transgenic plants carrying the pGREEN300 empty vector. **B)** 1 - pGREEN300/IFR; 2 - At wild type; 3 - negative control; 4-9 - AtIFR; 10-11 - AtBI-GST; Plant 5 and plant 6 are transgenic plants carrying the pGREEN300/IFR, while plant 10 and plant 11 are transgenic plants carrying the pGREEN300/BI-GST. **C)** 1- pGREEN300/IFR; 2 - At wild type; 3 - negative control; 4-6 - AtLeTpx-CS76; 7-8 - AtLeTpx; 9-10 - AtLeTpx-CS51-76; 11-12 - AtLeTpx-CS51. Plant 6 is transformed with pGREEN300/LeTpx-CS76, plant 8 is transgenic for pGREEN300/LeTpx, plant 9 is transformed with the double mutant pGREEN300/LeTpx-CS51-76, while plant 11 and plant 12 are transformed with pGREEN300/LeTpx-CS51.

Seeds from PCR positive plants were collected and replanted. Third generation of plants were obtained for further analyses (Fig. V. 5). Transgenic plants were grown under standard conditions monitoring each developmental stage in comparison with non-transformed plants. Morphological, anatomical and other changes were recorded for each independent genotype and developmental stage.



Figure V. 5. Third generation of *Arabidopsis thaliana* transgenic plants: A) Leaf fully expanded LeTpx-wt transgenic plants; B) Inflorescence developmental stage of IFR transgenic plant.

During the selection process, no morphological or fertility phenotypic differences were observed for BI-GST and LeTpx transgenic plants. Expression of BI-GST and LeTPX1 were confirmed by RT-PCR using gene specific primers, by Dr N. Argiriou (data not shown).

2. Effect of acrolein on tomato thioredoxin peroxidase homologous

Plant peroxidases display a conserved catalytic Cys residues, located in the N-terminal region of the protein, which is first transformed into a sulfenic acid after peroxide reduction. For the 2-Cys Prx class of enzyme, the sulfenic acid is reduced via the formation of an intermolecular disulfide bridge using a resolving Cys found on another subunit. Analysis of 2-Cys Prx protein function under stress conditions have shown that activities below the wild-type level have been shown to disturb early shoot development of *Arabidopsis* seedlings and photosynthesis (Baier and Dietz, 1999).

The tomato LeTpx1 protein isolated by the yeast two-hybrid approach as an interactor of BI-GST was shown to share sequence homology to the 2-Cys containing peroxidases. Point mutations on one or both Cys residues on the active site have been generated.

To assess *in-planta* the detoxification capacity of the LeTpx1 wild type and Cys residue mutant proteins, transgenic plants were grown under abiotic pro-oxidant stress conditions. The effect of prooxidants such as acrolein on the rate of germination and seedling development stage of transgenic plants overexpressing the Tpx proteins was examined. Acrolein is a widespread environmental pollutant and it is also an endogenous metabolic product of lipid peroxidation. High-dose acute acrolein toxicity has been suggested as involving oxidative stress subsequent to glutathione (GSH) depletion. To determine the critical concentration of oxidant that inhibits the physiological development, wild type plants were exposed to increasing concentrations of acrolein (Fig. V. 6).

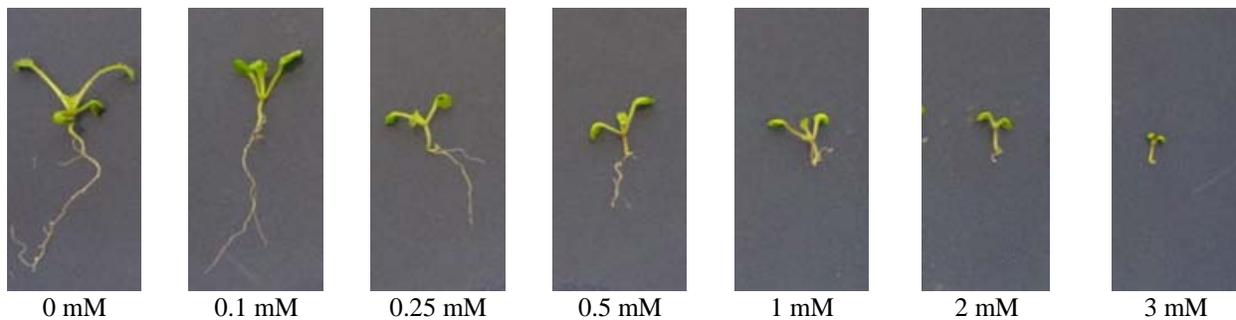


Figure V. 6. *Arabidopsis thaliana* wild type seedlings on the 10th day of growth on MS media supplemented with different concentration of acrolein.

Seeds of wild-type Col-0 were surface sterilized and sown in plates on 50 ml Murashige and Skoog medium containing 2% sucrose and increasing concentrations of acrolein. After 5 days of germination, visible differences on the rate of germination were observed. After 10 days of germination the plate-grown seedlings were compared for morphological differences. Perturbation of primary root development and seedling size appeared when seeds were grown on MS media supplemented with 0.5 mM of acrolein. At the highest concentration of 3 mM, the oxidant completely inhibit the germination of the seeds.

Transgenic *Arabidopsis* seedlings overexpressing LeTpx1 wild type protein and one of the mutants, LeTpx1-76Cys, were assessed for growth in the presence of acrolein at concentrations from 0 mM to 3 mM. Non-transformed seedlings and plants transformed with empty vector were used as controls.

The rate of germination was measured after 5 days of incubation. Approximately 100 seeds were sown per plate. A control plate of MS media without oxidant was sown for each different genotype and the number of germinated seeds on each plate was normalized to the control plate. The resulting data was expressed as scatter charts (Fig. V. 7).

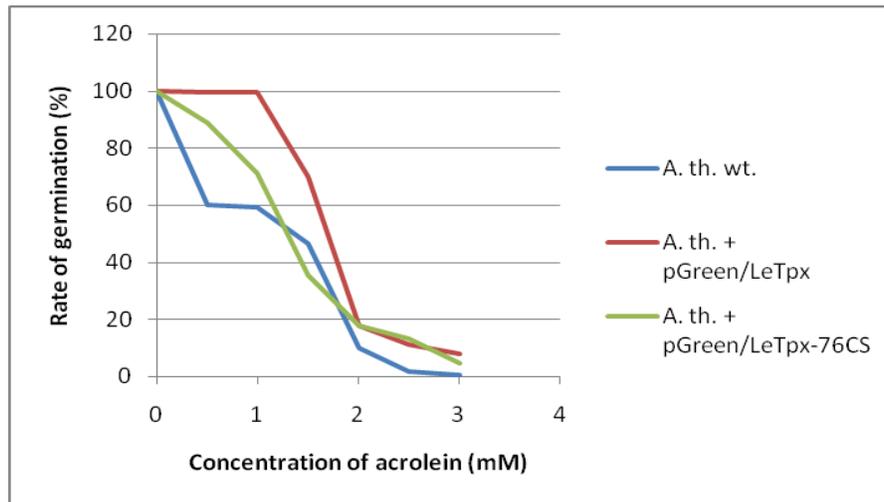


Figure V. 7. Germination rate of *Arabidopsis thaliana* wild type (blue), *Arabidopsis thaliana* LeTpx-wt (red), and *Arabidopsis thaliana* LeTpx-76CS (green) on the 5th day of growth on MS media supplemented with different concentration of oxidants. When acrolein was used as oxidant *Arabidopsis thaliana* LeTpx-wt, transgenic plants showed the highest resistance to oxidative stress regarding the rate of germination.

The results obtained from wild type seeds grown in the presence of acrolein revealed that this oxidant reduced the rate of germination. At 0.5 mM acrolein around 60 % of seeds normally germinated, while 10 % of seeds initiated germination even when 2 mM acrolein was present. Higher concentrations of acrolein inhibited almost completely the seed germination. When transgenic *Arabidopsis* seeds overexpressing the LeTpx1 protein were challenged with acrolein, they exhibited reduced sensitivity in terms of germination than non-transformed seeds, indicating the enhanced protection due to the presence of thioredoxin peroxidase activity (Fig. V. 8). No reduction of the rate of germination was registered for seeds overexpressing LeTpx1 wild type protein when the growing media was supplemented with concentrations of 0.5 mM and 1 mM acrolein. At 1.5 mM acrolein in the growing media the transgenic LeTpx wild type seeds germinated with quite good rate of 70 %. Higher concentrations of acrolein, such as 2 mM or 2.5 mM, could not fully inhibit germination at a percentage of 10 – 20 % of the transformed seeds (Fig. V. 8).

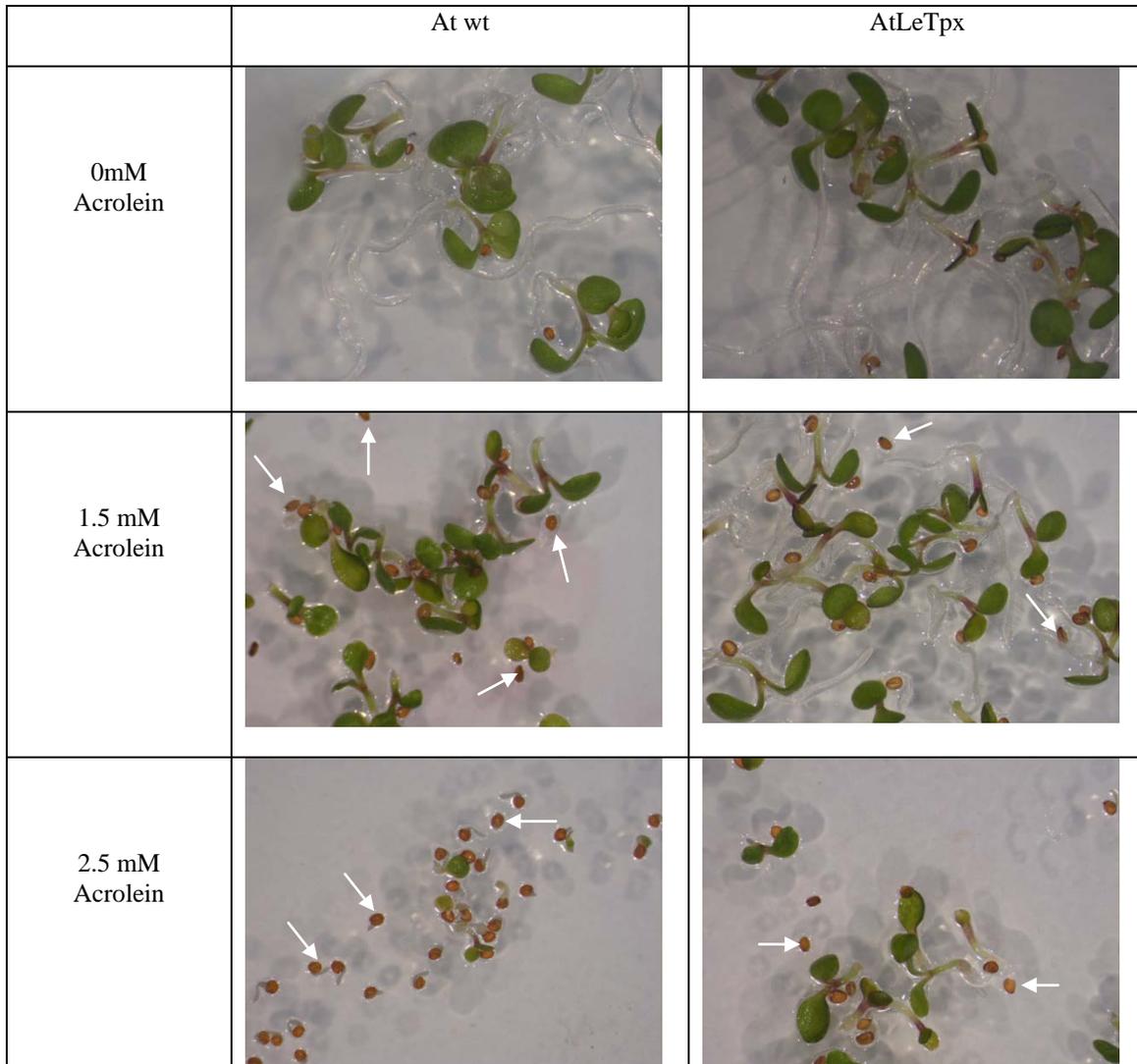


Figure V. 8. The effect of acrolein on the rate of germination of *Arabidopsis thaliana* wild type Col-0, and LeTpx1 transgenic seedlings at 5 days. The arrows indicate the seeds that did not germinate.

When the 76CS mutant form of LeTpx protein was overexpressed and transgenic seeds were exposed to acrolein, germination was affected in a higher proportion than in those overexpressing the LeTpx1 wild type protein, but still less than of wild type seeds, confirming previous data in yeast that showed that despite the deletion of one of the two catalytically active cysteines the protein still maintained some of its protective capacity, possibly as chaperone.

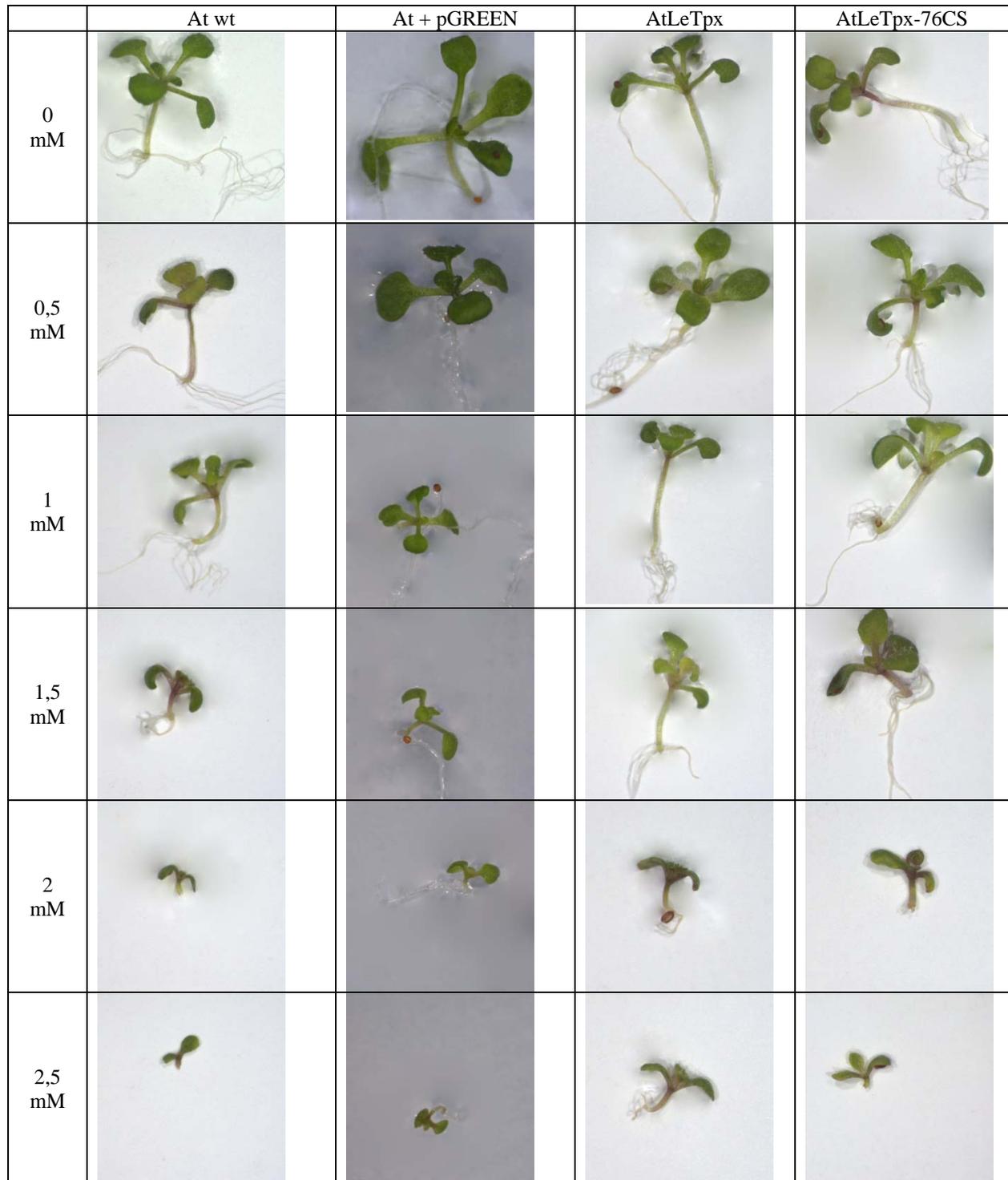


Figure V. 9. The effect of acrolein on transgenic *Arabidopsis* seedlings at 12 days, overexpressing a thioredoxin peroxidase (LeTpx1) and its mutant on 76Cys residue.

The 76CS mutant form of LeTpx1 protein indicates that 76 cysteine mutation in the active site of LeTpx1 enzyme, as anticipated, plays a role and its detoxification capacity, but surprisingly the enzyme still retains some of its protective function.

The effect of acrolein was monitored on plate-grown seedlings in regard to early root development and seedling size in the presence of various concentrations of oxidant for a period of 12 days (Fig. V. 9).

Acrolein visibly disturbed physiological seedling growth of wild type and empty vector transformed plants. The developmental perturbation was most evident at the concentration of 1.5 mM, affecting the size of the seedlings and of the primary roots. This perturbation increased progressively till 3mM acrolein, when germination was completely inhibited. The LeTpx1 wild type and 76CS mutant proteins protected plants from acrolein-induced stress when overexpressed in transgenic plants (fig. V. 9). The seedling overexpressing the LeTpx1 and its mutant form, LeTpx76Cys, were protected from acrolein toxicity, both in terms of seedling size and primary roots formation, but only up to a concentration of 2 mM.

3. Evaluation of the *Arabidopsis* transgenics' resistance to microbial pathogens

Plants respond in a variety of ways to pathogenic microorganisms. *Pseudomonas syringae* is a Gram-negative plant pathogen that defeats plant defenses through effector proteins that are injected into plant cells via the type III secretion system. *P. syringae* strains are assigned to pathovars based largely on their host of origin. *P. syringae* pv. *tomato* DC3000 causes bacterial speck of tomato and has become a model for studying bacterium–plant interactions because it also attacks the experimentally amenable plants *Arabidopsis thaliana*. In susceptible host plants, *P. syringae* can establish large populations in the intercellular spaces of leaves and produce necrotic lesions. In non-host plants and in host plants with R-gene-mediated resistance, *P. syringae* can induce a localized defense response that involves rapid, programmed plant cell death. This hypersensitive response (HR) is associated with cessation of the growth and spread of the bacteria. Genetic analyses of DC3000 have led to the identification of genes essential to pathogenesis and HR induction. Functions of these genes include the production of a type III secretion system that promotes the delivery of various effector proteins into plant cells. Genetic

analyses of *A. thaliana* have profiled genes induced in response to DC3000 and have demonstrated that the plant response is rapid, dynamic, and distinct for susceptibility versus R-gene-specific resistance (de Tores, M. et al., 2003; Quirino, B.F. and Bent, A.F. 2003). During pathogenesis, *P. syringae* pv. *tomato* often attains populations as high as 10^7 to 10^8 colony-forming units (CFU) per cm^2 in *A. thaliana* leaves (Jumbunathan, J. et al., 2001).

Arabidopsis LeTpx1 and BI-GST overexpressing transgenic plants and wild type parental plants (Fig. V. 10) were inoculated with *Pseudomonas syringae* pv. *tomato* DC3000, kindly provided by Dr. P. Sarris, University of Crete, and their susceptibility to pathogen attack was evaluated 8 days after infection..



Figure V. 10. *Arabidopsis* plants at leaf stage fully expanded, prior to infection with *Pseudomonas syringae*.

The infected plants were examined 8 days after inoculation. Wild-type plants exhibited typical symptoms of expanding water-soaked lesions and premature chlorosis (Fig. V. 11 A), whereas transgenic LeTpx1 plants displayed lesions with reduced size (Fig. V. 11 B). Furthermore, inoculation of LeTpx1 plants resulted in dry lesions unable to develop beyond the inoculation site in a significant number of cases. BI-GST plants, overexpressing an antiapoptotic GST protein that possess Bax inhibiting activity in yeast cells were not affected by *Pseudomonas* infection (Fig. V. 11 C). Leaves of infected plants were compared (Fig. V. 11 D). Inoculation with bacterial suspension in wild type plants (Fig. V. 11 D up row) produced the collapse of the whole leaf in several cases, but no visible symptoms appeared in BI-GST plants (Fig. V. 11 D bottom row) and only very mild symptoms in LeTpx1 plants (Fig. V. 11 D middle row).

It appears that, the two proteins LeTpx and BI-GST, from *Lycopersicon esculentum* impart a significant level of resistance in plant defense when expressed in *Arabidopsis thaliana* to *Pseudomonas syringae* pv. *tomato* DC3000.

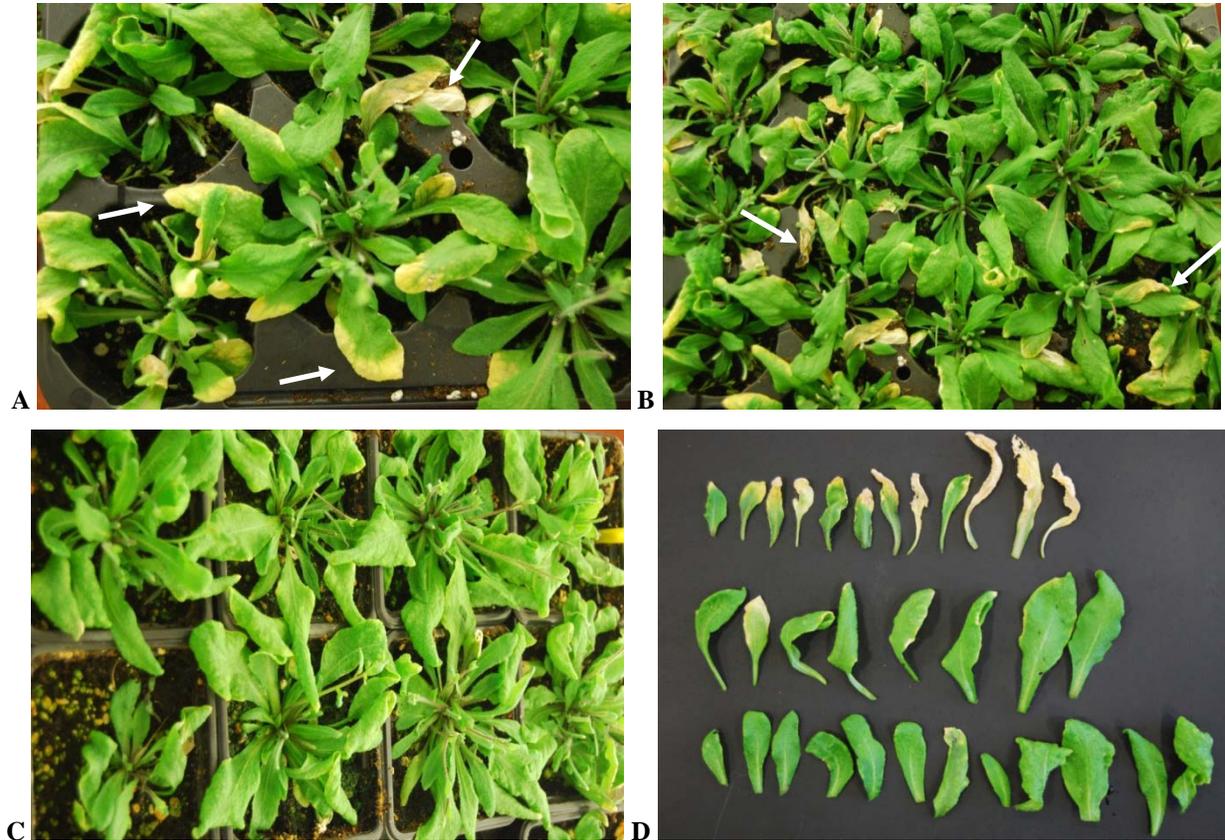


Figure V. 11. *Arabidopsis thaliana* plants infected with *Pseudomonas syringae* pv. *tomato*DC3000, 8 days after inoculation. **A)** Water-soaked progressive lesions of soil-grown wild type plants indicated by white arrows; **B)** LeTpx1 transgenic plants affected by microbial pathogens; **C)** BI-GST transgenic plants resistant to *Pseudomonas* infection; **D)** Leaf symptoms of (from top to bottom) untransformed plants (wt), transgenic LeTpx1, and BI-GST infected with *P. syringae* pv. *tomato*DC3000.

Resistant plants often develop a hypersensitive response (HR), in which necrotic lesions form at the site(s) of pathogen or elicitor entry (Hammond-Kosack and Jones, 1996; Dangl *et al.*, 1996). Just before or concomitant with the appearance of a HR, increased synthesis of several families of pathogenesis-related (PR) proteins in the inoculated leaves takes place (Kombrink and Somssich, 1997). Many of these proteins have been shown to exhibit antimicrobial activity either *in vitro* or *in vivo*. PR proteins subsequently accumulate in the uninoculated portions of the plant, concurrent with the development of a broad-spectrum resistance known as systemic acquired resistance (SAR) (Ryals *et al.*, 1996). Because of this correlation, increased *PR* gene

expression is frequently used as a marker for SAR. Interestingly, although plants lack a circulatory system and do not produce antibodies, SAR shares several common characteristics with the innate immune system of animals (Klessig *et al.*, 2000).

To gain further information on the potential mechanism employed in the increased disease resistance phenotype observed in LeTpx1- and BI-GST-overexpressing *Arabidopsis* plants, the expression levels of PR-5 defense-related gene were analyzed by RT-PCR (Fig. V. 12). This experiment was kindly performed by Dr. Notis Argiriou. Expression was normalised using actin as a control.

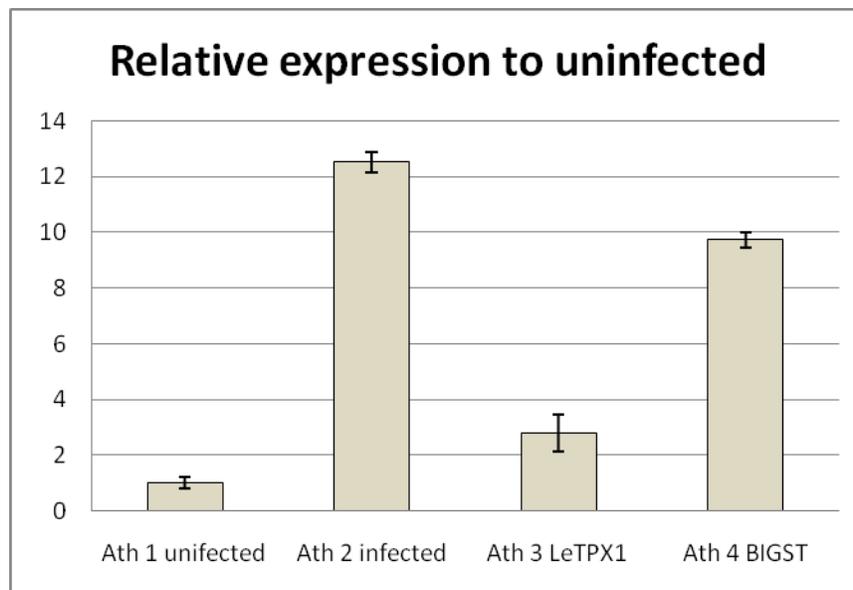


Figure V. 12. Expression analysis of *Arabidopsis* plants infected with *Pseudomonas*. Wild- type *Arabidopsis* plants and LeTpx1 and BI-GST transgenic plants infected with *P. syringae* were considered for analysis of PR-5 gene expression by RT-PCR, using actin as control. Uninfected *Arabidopsis* wild type plants were employed as negative control in this analysis.

In the experiment shown in Fig. V. 12, induced expression of Pathogenesis-related gene PR-5 in uninfected plants was only detected at low levels, suggesting that the growth regime did not cause significant stress. As expected inoculated wild type plants showed a dramatic increase of PR5 >12-fold induction, whereas a reduction in the upregulation of PR5 in transgenic plants is evident. In the case of LeTPX1 transgenics, PR5 accumulation levels almost declined to background levels (uninoculated wild type control), suggesting that the mechanism of resistance between LeTPX1 and BI-GST transgenics may not necessarily be the same.

Future experiments should address the relative PR5 expression levels in all uninfected plants and the potential involvement of the redox sensitive plant defense regulator NDR1 (non race-specific disease resistance 1).

CHAPTER VI

CONCLUSIONS

Over the years, the use of natural products has been the single most successful strategy in the discovery of novel medicines. As many as 60% of successful drugs originate from natural sources. Isoprenoids are among the most diverse groups of compounds synthesized by biological systems. They are constructed from a varying number of isoprene (C₅) units and are classified into classes from the number of C₅ units used. There has been a strong interest in terpenoids as antifungal, antibacterial and anticancer agents. Worldwide sales of terpene based pharmaceuticals are approximately \$ 12 billion. Artemisinin, a sesquiterpene compound produced by *Artemisia annua* which is a novel antimalarial agent and eleutherobin which is a new anticancer agents are such valuable pharmaceutical terpenoids which face the same problems of high production costs. Similarly taxol, a diterpene extracted from the Pacific Yew tree is an extremely potent anticancer agent currently in clinical use.

Metabolic engineering of biologically amenable microorganisms such as *Escherichia coli*, *Streptomyces coelicolor* and *Saccharomyces cerevisiae* aiming to divert their cellular resources and biosynthetic machinery towards the production of desirable heterologous natural products as an environmentally friendlier alternative which can lead to cheaper sources of drugs and industrial chemicals of high value. Previous work on medicinal and aromatic plant terpenoid biosynthesis, identified a series of monoterpene and sesquiterpene synthases encoding genes, as well as cytochrome P450-monooxygenases from *Salvia fruticosa* and *Salvia pomifera* (sage plants).

One of them, a cineole synthase clone was successfully expressed in yeast and was shown to synthesize several monoterpenes in living cells using the endogenous GPP pool, among them a significant amount of 1,8-cineole. During the efforts to produce terpene compounds in yeast it became evident that a major limitation in the production was the restricted amounts of endogenous isoprene precursors. To overcome this, several steps were taken.

Yeast laboratory and industrial strains was shown to possess a great deal of phenotypic variability in the response to adverse stress conditions, the capacity to produce high levels of

heterologous proteins with limited proteolytic degradation, the tolerance to the presence of toxic compounds and growth properties. This variability makes imperative the evaluation of different yeast strains for their capacity to produce high levels of a heterologous terpene synthase and to release increasing quantities of the terpene compound. Initially, a standard laboratory strain of yeast ResGen BY4741, and a high sterol producing yeast strain EG60 were evaluated for terpene production using the endogenous GPP pool. The EG60 strain, showing double cineole production compared to BY4741, have been selected for further modification to enhance terpene production.

The yeast strains were transformed with individual plasmids bearing two truncated forms of cineole synthase, Sf-CinS1(RR), and Sf-CinS1(RC), under control of the GAL promoter. Transformed strains were grown on minimal media with glucose as carbon source, and then induced on corresponding galactose-raffinose media for 12-15 h until cells reached the stationary phase. Further incubation in enzymatic reaction buffer at 30°C aided the removal of non-specific compounds. Volatile compounds were harvested by SPME and analyzed by GS-MS.

An almost complete description of the terpenoid lipid biosynthesis in yeast is available and the genes leading to the formation of GPP and FPP have been identified. Synthesis starts from acetyl-coA and leads in a series of biosynthetic steps to ergosterol. A number of intermediate steps, most of them subsequent of isoprene unit formation are also used for the synthesis of several other essential components such as ubiquinones, dolichols, apocytochrome a and prenylated proteins. Eight biosynthetic steps are involved in the formation of FPP and GPP. Among them, HMG2, the enzyme converting 3-hydroxy-3methylglutaryl (HMG)-CoA into mevalonic acid, has been identified as a rate limiting step in all organisms from yeast to mammals, and the stability of the protein is regulated by the proteasomal system. In this case a K6R point mutation resistant to proteolytic degradation was generated in the cDNAs leading to stabilization of the protein, thus overcoming the proteolytic instability problem. The stabilized K6R HMG2 mutant under the control of GAL promoter was introduced into EG60 genome generating the AM63 strain, that showed better cineole synthesis with 1.5 fold when transformed with Sf-CinS1(RC). The AM63 strain was chosen for further engineering.

Similar work was focused on complete inhibition of pathways located downstream which drain the cell's supply of GPP and FPP. Although the current understanding of the regulation of the final sterol biosynthetic steps, and the effect they can have on shutting down the whole

pathway is limited, it was reasonable to assume that if it exists, it will only reinforce the production of precursors, to compensate for the absence of the final product. The most obvious target is to disrupt the sterol biosynthetic steps that are no longer required for survival. As such, the gene ERG6 encoding an enzyme that converts zymosterol to fecosterol was suitable targets for total deletion generating AM62 and AM64 strains. Eventhough terpene production was higher in last two strains than in the parental EG60 strain this modification did not increase the amount of cineole obtained in AM63 strain, while cellular growth and transformation rate was severe affected.

To functionally evaluate several isolated terpene synthase clones from *Salvia fruticosa* and *Salvia pomifera* cDNA library, the AM63 strain was transformed with four novel terpene synthases which previously failed to yield any products when tested as bacterially expressed proteins. One of them was a monoterpene synthase (215 clone) producing mostly cineole, distinct from the previously identified *Salvia* cineole synthase, while another monoterpene synthase was a short functional fragment of cineole synthase. The other two encoded for sesquiterpene synthases (P330 and 1025), the first producing as major compound alpha-cubebene, alpha-copaene, trans-caryophyllene, and delta-cadinene, while the other possible act as isolongifolene oxide synthase. All the transformed strains were able to synthesize terpenes. In the case of cineole synthase, cineole was clearly the major product. However the sesquiterpene synthases are multiple product enzymes.

Monoterpene synthases are nuclear encoded preproteins that are destined to be imported in the plastids, where they are proteolytically processed into their mature forms. Several investigations have been carried out to identify the exact position of the cleavage of the transit peptide as well as the nature of the modification that blocks the free amino group at the N terminus after processing occurs. Although the precise cleavage site is not yet known for these preproteins, truncation of monterpene synthases exactly before the conserved tandem arginine motif (RRX₈W) has been demonstrated to result in an active and stable enzyme. For monoterpene synthases which contain a transit sequence targeting the enzymes to the chloroplasts, two potential predicted cleavage sites, at the CS and RR motifs were tested to assess the effect on the enzymatic activity of the two versions. The expression and stability of the two truncated cineole synthases, Sf-CinS1(RR) and Sf-CinS1(RC), cloned in the pJG4-6 yeast vector under the GAL promoter and transformed into the AM63 strain were functionally

characterized by Western Blot and assessed their activity by GC-MS for a period of 13 days. Both versions are stable proteins expressed at similar level, but Sf-CinS1(RC) showed a consistent modest improvement in cineole accumulation in comparison to Sf-CinS1(RR).

One step that positively contributed to enhanced terpene synthesis was the modification of the expression of yeast gene involved in the biosynthesis of GPP and FPP, the ERG20 gene encoding for FPP synthase. This was accomplished by introducing strong promoter element, such as the GAL promoter, upstream of the gene's open reading frame, causing increased production of mRNA transcripts. The integration scheme employed the loxP-selection marker cassette, which allows the excision of the selection marker at the end of the homologous recombination process using the cre recombinase, and subsequent recycling of the selection marker. The AM68 generated strain gave 3 fold higher increased cineole production the parental EG60 strain and for basic growth properties which ensure a durable strain.

In parallel with the efforts to overexpress the genes of the isoprene biosynthetic pathways, the inhibition of the pathway that deplete the accumulating isoprene units was further carried on one crucial step immediately downstream of GPP and FPP synthesis. The ERG9 essential gene, encoded for squalene synthase and direct responsible for terpene precursors spending was partial inactivated by deletion of one of the two alleles in diploid AM68 genome generating the AM70 strain. Beside the initial difficulties of cellular growth registered, the AM70 yeast cells behaved as robust enough strain. When AM70 strain was evaluated for terpene production, noticeable reduction of monoterpene production (almost 2 fold), but significant increase of sesquiterpene synthesis up to 3.5 fold compared with the parental EG60 strain. Such opposite effects may be explained by the massive utilization of isoprene precursors resulting from the upregulated ERG20 expression, leading to FPP synthesis and not in favor of GPP synthesis when the sterol flux was altered by the *erg9* haploinsufficiency.

The yeast two-hybrid approach was employed to identify interacting partners of cineole synthase (Sf-CinS1) in the tomato library. Three independent positive clones were isolated after the stringent two-stage screening process. One of them encoding for the HSP90 protein increased cineole accumulation in AM63 yeast cells when co-expressed with Sf-CinS1(RC) to about 30% at 30°C and 40% at 37°C higher yields.

The low efficiency of the two hybrid screening was thought to due to the instability of the N-terminal end of terpene synthases where LexA-Sf-CinS1(RC) fusion occurred. To overcome

this difficulty a novel plasmid bait aiding to the C-terminal LexA-bait fusion was generated. Higher effectiveness was recorded during a pilot screening of the *Salvia* cDNA library with the pYES-Sf-Cin(RC)-LexA bait vector. The HSP90 protein was once again isolated as cineole synthase interactor validating the previous work. Two additional positive clones involved in protein trafficking and folding were isolated. One of them was a luminal-binding protein 4 precursor (BiP4) homologous, and the other a SEC14-like cytosolic factor family protein.

Several novel proteins involved in the plant defence mechanism, LeTpx and BI-GST, were expressed in *Arabidopsis thaliana* transgenic plants and assessed for their biological function during acrolein-induced stress and microbial attack (infection with *Pseudomonas syringe* pv. *tomato* DC3000). The pot soil-grown plants expressing the LeTpx transgene and its mutant variant LeTpx76CS exhibited increased resistance to acrolein-induced stress regarding the rate of growth, seedling size and early root development compared with wild type plants, indicating that the thioredoxin-peroxidase activity of the newly isolated proteins. The *Arabidopsis* plants expressing LeTpx and BI-GST transgenes showed different degrees of resistance to *Pseudomonas* attack. Inoculation of *Pst* DC3000 produced typical leaf lesions 8 days after infection of wild type plants, whereas in the transgenic plants symptom severity was greatly reduced. In keeping with this, infected leaves of transgenic plants did not become chlorotic, indicating that the development of chlorotic symptoms due to the phytotoxin coronatine accompanying pathogen multiplication in the leaf tissue had been abrogated. Thus, the expression of LeTpx and Bi-GST proteins in *Arabidopsis* plants protect them from microbial attack suggesting that they may also be involved in defense against wounding-induced cell death.

Further work will aim to further modifying the yeast strain to maximize terpene productivity without sacrificing cell viability which could hamper the biofermentation process, by incorporation of other strong promoter elements upstream of the genes' open reading frame causing increased production of mRNA transcripts. The integration scheme will employ the loxP-selection marker cassette, which allows the excision of the selection marker at the end of the homologous recombination process using the cre recombinase, and subsequent recycling of the selection marker.. The genes that positively contribute to enhances terpene production, such as GPP synthase, will subsequently become stably integrated in extra copies in the yeast genome under the control of a strong promoter, employing the novel generated integration cassette

pDNR-GAL. The remaining terpene synthases isolated from the EST sequencing approach that do not show significant similarity to any other known activity (five sesquiterpene synthases and three diterpene synthases) will be cloned into appropriate yeast expression vectors and their catalytic activities will be determined. This will not only reveal new enzymatic activities but will also provide insights and material for further protein engineering attempts. Supplementing the yeast culture with substrate analogues, despite the electronic effects or steric bulk introduced by substrate derivatization, such as 2-fluorofarnesyl diphosphate or 4-methylfarnesyl diphosphate, mutants with the ability to transform non-canonical substrates can be selected. This can be further expanded in combination with specific functional assays. Additionally, the facile detection of terpenes as volatiles by SPME / GC-MS will enable the development of large scale targeted mutagenesis approaches to develop libraries of enzymes with novel product specificity. Exhausting screening of the glandular trichome specific *S. pomifera* cDNA library with the bait consisting of Sf-CinS1 with no stop codon fused with the LexA protein on the C-terminal end of the enzyme will aim to uncover more cineole synthase interacting proteins. Details about the protein complexes involved in the biosynthesis of terpenoids and about the role of the N-terminal domain are expected to be elucidated with this approach.

APPENDIX 1

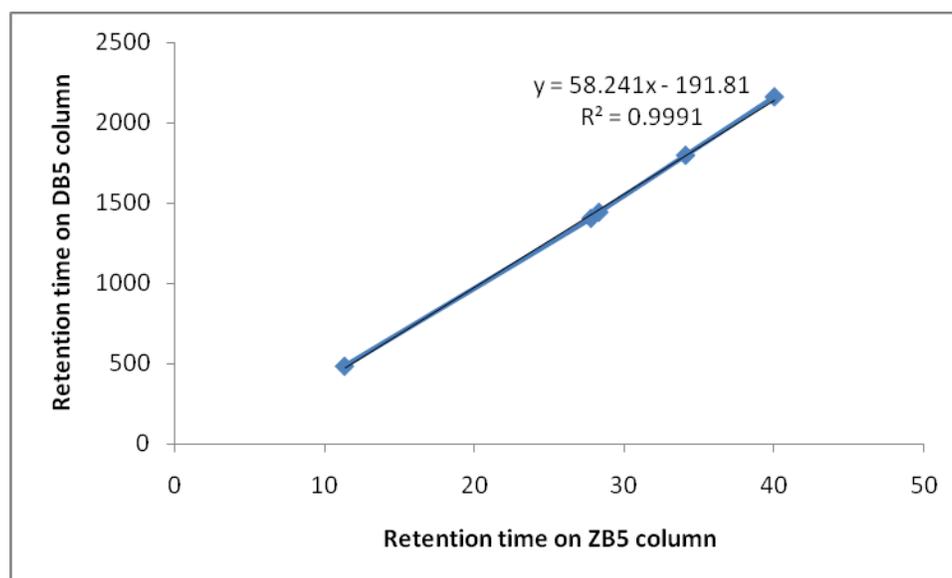
loxP-URA-loxP-GALp-HA pCOD2

1 AAGCTTCGTACGCTGCAGGTCGACAACCCTTAATATAACTTCGTATAATG
 51 TATGCTATACGAAGTTATTAGGTCTAGAGATCCCAATACAACAGATCACG
 101 TGATCTTTTGTAAGATGAAGTTGAAGTGAGTGTTGCACCGTGCCAATGCA
 151 GGTGGCTATTAGATTAAATATGTGATTTGTTCTATTAAGTTTCCTGTATA
 201 ATTAATGGGGAGCGCTGATTCTCTTTTGGTACGCTTCCCATCCAGCATTT
 251 CTGTATCTTTACCTTCAACCTTAGGATCTCTACCCTTGGCGAAAAGTCC
 301 TCTGCCAACAATGATGATATCTGATCCACCCTTACAACCTTCGTGACGG
 351 TTCTGTACTGCTGACCCAATGCATCGCCTTTGTCGTCTAAACCTACACCT
 401 GGGGTCAATGATTAGCCAATCAAACCCTTCTTCTTCTCCTCCCATATCGTT
 451 CTGAGCAATGAACCAATAACGAAATCTTTATCACTCTTTGCAATATCAA
 501 CGGTACCCTTAGTATATTCACCGTGTGCTAGAGAACCCTTGGAAAGACAAT
 551 TCAGCAAGCATCAATAATCCCCTTGGTTCTTTGGTGACCTCTTGCGCACC
 601 TTGTTTCAAGCCAGCAACAATACCAGCACCAGTAACCCCGTGGGCGTTGG
 651 TGATATCAGACCATTCTGCGATACGGTAAACGCCCGATGTATATTGTAAT
 701 TTGACTGTGTTACCGATATCGGCGAATTTTCTGTCTCAATATCAAGAA
 751 CTTGTATTTCTCTGCCAATGCTTTCAATGGAACGACAGTACCCTCATAAC
 801 TGAAATCATCCAAGATATCAACGTGTGTTTTCAAAGGCAATGTATGGA
 851 CCCAACGTTTCAACAAGTTTCAATAGCTCATCAGTCGAACGAACGTCAAG
 901 AGAAGCACACAAATTGGTCTTCTTTTCATCCATTAACCGTAAAAGTTTCG
 951 ATGCAACCGGACTTGATGAGTCTCAGCTCTACTGGTATATGATTTTGTG
 1001 GACATGGTGCAACTAATTGACGGGAGTGTATTGACGCTGGCGTACTGGCT
 1051 TTCACAAAATGGCCCAATCACAACCACATCTTAGATAGTTGAAATGACTT
 1101 TAGATAACATCAATTGAGATGAGCTTAATCATGTCAAAGCTAAAAGTGTC
 1151 ACCATGAACGACAATTCTTAAGCAAATCACGTGATATAGATCCACGAATA
 1201 ACCACCATTGATGCTCGAGGCAAGTAATGTGTGTAAAAAATGCGTTAC
 1251 CACCATCCAATGCAGACCGATCTTCTACCCAGAATCACATATATTTATGT
 1301 ACCGAGTACCTTTTTTCTATCTTCCAATTGCTTCTGCCCATATGATTGTC
 1351 TCCGTAAGCTCGAAATTTCTAAGTTGGATTTTAATCTTCACGCAGGATGA
 1401 CAGTTCGATGAGCTTCTGAGGAGTGTTTAGAACATAATCAGTTTATCCAT
 1451 GGTCTATCTCTTCTTGTGCTTTTTCTCCTCGATAGAACCTAAATAAAAC
 1501 GAGCTCTCGAGAACCCTTAATATAACTTCGTATAATGTATGCTATACGAA
 1551 GTTATTAGGTGATATCAGATCCACTAGTGGCCTATGCGGCCACGGATTAG
 1601 AAGCCGCCGAGCGGGTGACAGCCCTCCGAAGGAAGACTCTCCTCCGTGCG
 1651 TCCTCGTCTTACCAGGTCGCGTTCCTGAAACGCAGATGTGCCTCGCGCCG
 1701 CACTGCTCCGAACAATAAAGATTCTACAATACTAGCTTTTATGGTTATGA
 1751 AGAGGAAAAATTGGCAGTAACCTGGCCCCACAAACCTTCAAATGAACGAA
 1801 TCAAATTNACAACCATAGGATGATAATGCGATTAGTTTTTTAGCCTTATN
 1851 TCTGGGGTAATTAATCAGCGAAGCGATGATTTTTGATCTATTNACAGATA
 1901 TATAAATGCAAAAACCTGCATAACCCTTTAACTAATACTTTCAACATTTT
 1951 CGGTTTGTATTACTTCTTATTCAAATGTAATAAAAGTATCAACAAAAAAT
 2001 TGTTAATATACCTCTATACTTTAACGTCAAGGAGAAAAAACNCCGGATCG
 2051 GACTACTAGCAGCTGTAATACGACTCACTATAGGGAATATTAAGCTCATG

2101 GCACACTACTCTCTAATGAGCAACGGTATACGGCCTTCCTTCCAGTTACT
2151 TGAATTTGAAATAAAAAAAGTTTGCCGCTTTGCTATCAAGTATAAATAGA
2201 CCTGCAATTATTAATCTTTTGTTCCTCGTCATTGTTCTCGTTCCCTTTC
2251 TTCCTTGTTTCTTTTTCTGCACAATATTTCAAGCTATAACCAAGCATACAA
2301 TCAACTCCAAGCTTATGGAACAGAAGCTAATAAGTGAGGAGGATCTGAAT
2351 GGATCCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGC
2401 ATGCTAGAGGGCCT

APPENDIX 2

Calibration of ZB5 retention times according with DB5 retention times



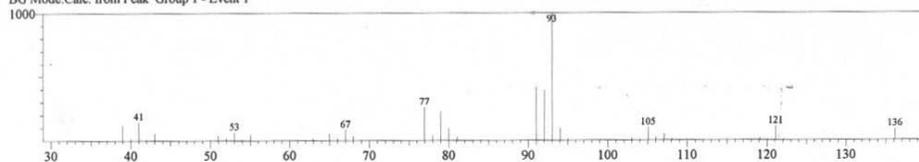
Standard products	Retention time on	
	ZB5 column	DB5 column
cineole	11.356	485
longifolene	27.787	1404
trans-caryophyllene	28.315	1442
trans-nerolidol	34.097	1796
trans-farnesol	40.014	2159

Spectrum comparison of 215 monoterpene synthase products

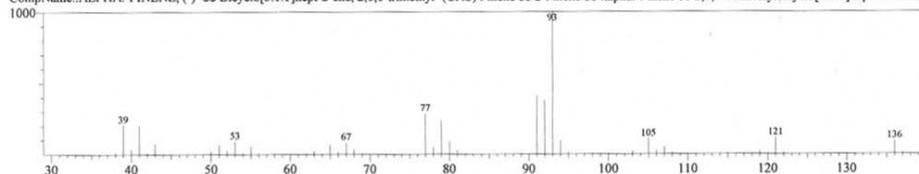
1. α -Pinene

Spectrum Comparison

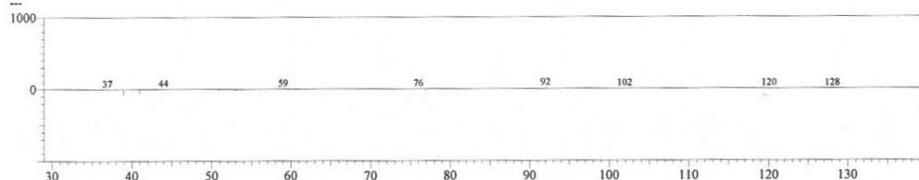
Spectrum1 #Data# 215.QGD R.Time:7.933(Scan#:237)
 MassPeaks:68
 RawMode:Averaged 7.917-7.950(236-238) BasePeak:93.05(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:26447 Formula:C10 H16 CAS:80-56-8 MolWeight:136 RetIndex:0
 MassPeaks:50 BasePeak:93.00(1000)
 CompName:..ALPHA..PINENE, (-)- \$\$ Bicyclo[3.1.1]hept-2-ene, 2,6,6-trimethyl- (CAS) Pinene \$\$ 2-Pinene \$\$ α -Pinene \$\$ 2,6,6-Trimethylbicyclo[3.1.1]hept-2-ene \$\$ a



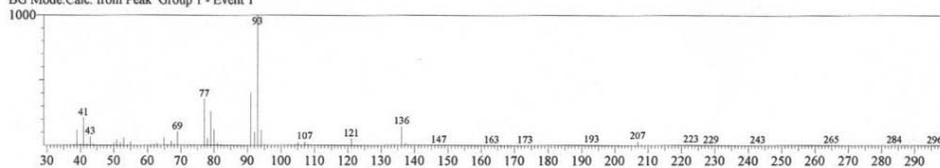
Spectrum3 #Calculation Result#
 MassPeaks:60 BasePeak:92.05(16)



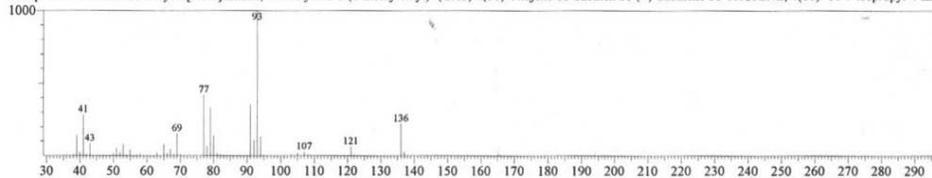
2. Sabinene

Spectrum Comparison

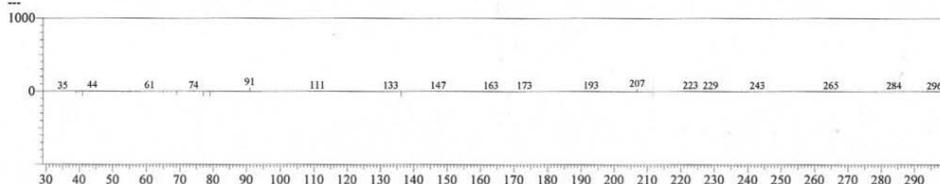
Spectrum1 #Data# 215.QGD R.Time:9.217(Scan#:314)
 MassPeaks:136
 RawMode:Averaged 9.200-9.233(313-315) BasePeak:93.05(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:26425 Formula:C10 H16 CAS:3387-41-5 MolWeight:136 RetIndex:0
 MassPeaks:47 BasePeak:93.00(1000)
 CompName:Sabinene \$\$ Bicyclo[3.1.0]hexane, 4-methylene-1-(1-methylethyl)- (CAS) 4(10)-Thujene \$\$ Sabinen \$\$ (+)-Sabinene \$\$ THUJENE, 4(10)- \$\$ 1-Isopropyl-4-methy



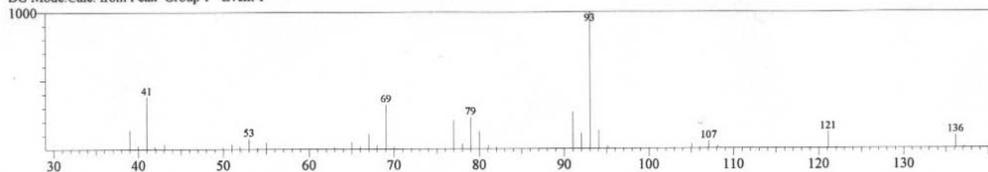
Spectrum3 #Calculation Result#
 MassPeaks:136 BasePeak:91.00(51)



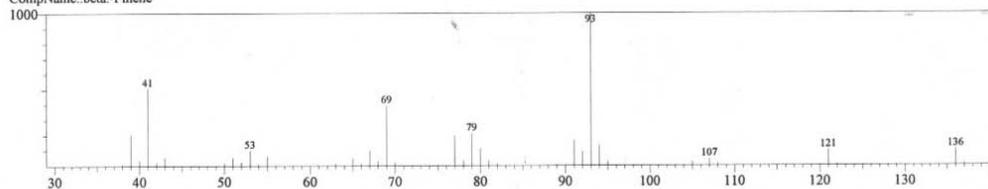
3. β -Pinene

Spectrum Comparison

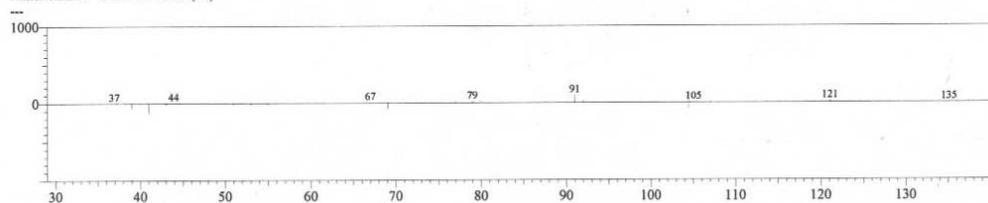
Spectrum1 #Data# 215.QGD R.Time:9.367(Scan#:323)
 MassPeaks:66
 RawMode:Averaged 9.350-9.383(322-324) BasePeak:93.05(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# NIST21.LIB Entry:5429 Formula:C10H16 CAS:127-91-3 MolWeight:136 RetIndex:0
 MassPeaks:50 BasePeak:93.00(1000)
 CompName:beta.-Pinene



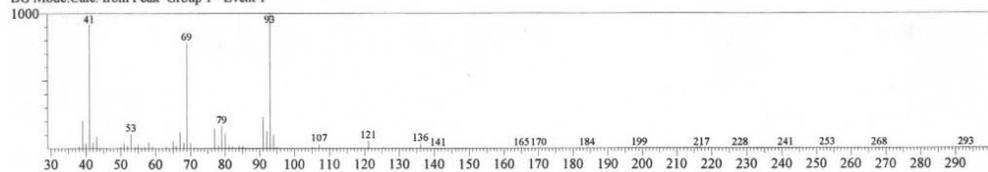
Spectrum3 #Calculation Result#
 MassPeaks:65 BasePeak:91.00(98)



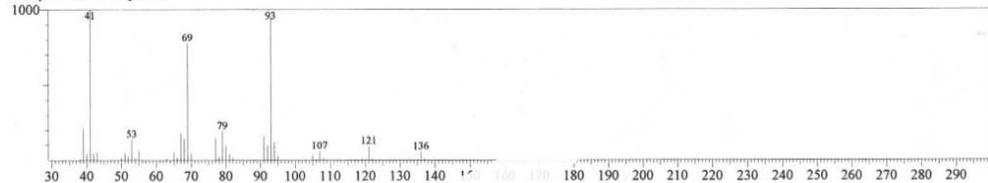
4. β -Myrcene

Spectrum Comparison

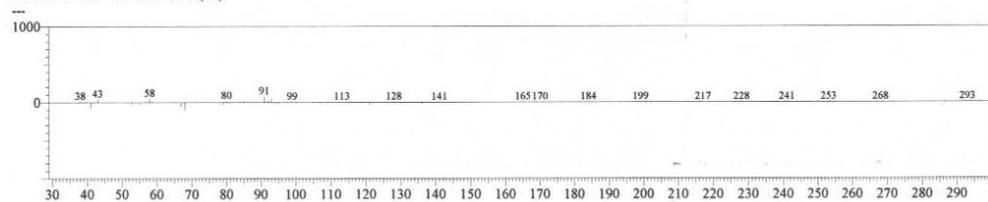
Spectrum1 #Data# 215.QGD R.Time:9.750(Scan#:346)
 MassPeaks:96
 RawMode:Averaged 9.733-9.767(345-347) BasePeak:93.05(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# NIST21.LIB Entry:5442 Formula:C10H16 CAS:123-35-3 MolWeight:136 RetIndex:0
 MassPeaks:50 BasePeak:41.00(1000)
 CompName:beta.-Myrcene



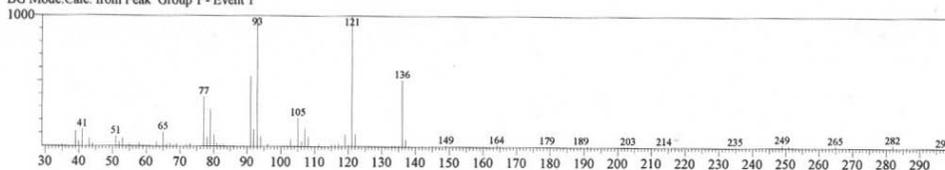
Spectrum3 #Calculation Result#
 MassPeaks:100 BasePeak:91.00(75)



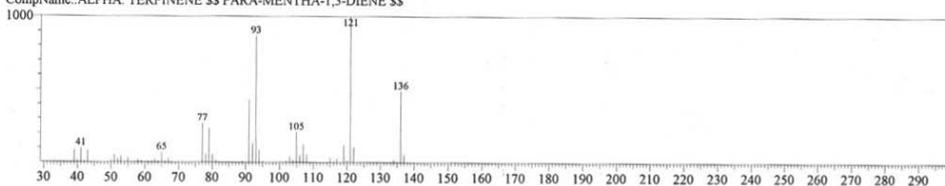
5. α -Terpinene

Spectrum Comparison

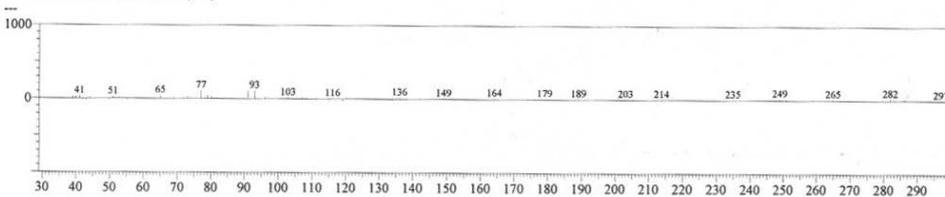
Spectrum1 #Data# 215.QGD R.Time:10.800(Scan#:409)
MassPeaks:134
RawMode:Averaged 10.783-10.817(408-410) BasePeak:121.10(1000)
BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:25507 Formula:C10H16 CAS:99-86-5 MolWeight:136 RetIndex:0
MassPeaks:50 BasePeak:121.00(1000)
CompName:ALPHA. TERPINENE \$\$ PARA-MENTHA-1,3-DIENE \$\$



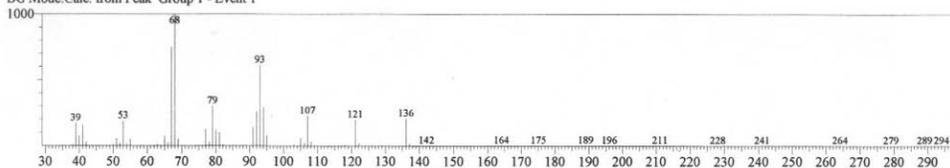
Spectrum3 #Calculation Result#
MassPeaks:136 BasePeak:77.00(110)



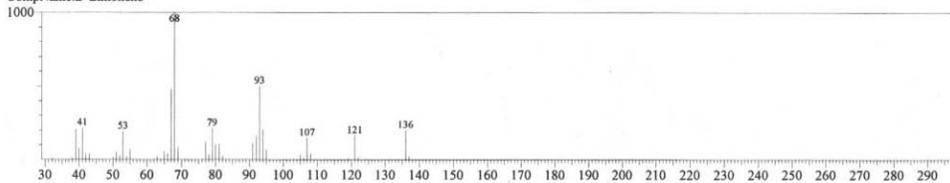
6. D-Limonene

Spectrum Comparison

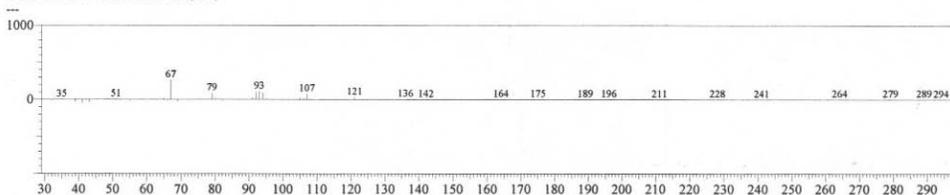
Spectrum1 #Data# 215.QGD R.Time:11.267(Scan#:437)
MassPeaks:113
RawMode:Averaged 11.250-11.283(436-438) BasePeak:68.00(1000)
BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# NIST21.LIB Entry:5415 Formula:C10H16 CAS:5989-27-5 MolWeight:136 RetIndex:0
MassPeaks:50 BasePeak:68.00(1000)
CompName:D-Limonene



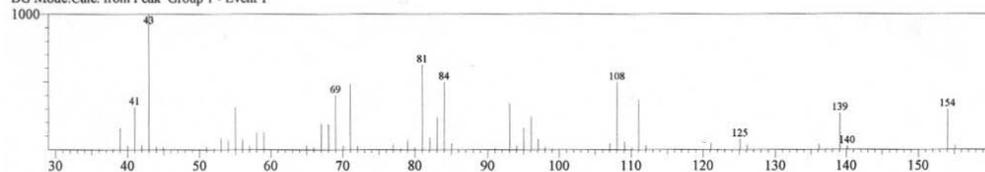
Spectrum3 #Calculation Result#
MassPeaks:119 BasePeak:67.00(263)



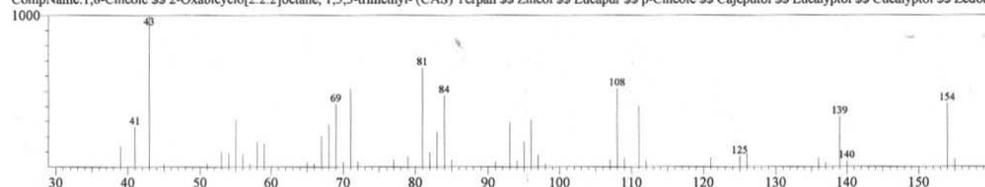
7. 1,8-Cineole

Spectrum Comparison

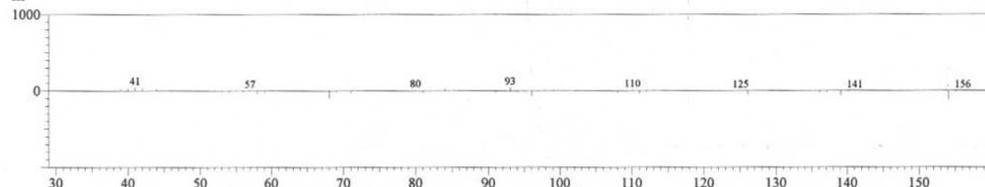
Spectrum1 #Data# 215.QGD R.Time:11.383(Scan#:444)
MassPeaks:81
RawMode:Averaged 11.367-11.400(443-445) BasePeak:43.00(1000)
BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:43991 Formula:C10H18O CAS:470-82-6 MolWeight:154 RetIndex:0
MassPeaks:50 BasePeak:43.00(1000)
CompName:1,8-Cineole SS 2-Oxabicyclo[2.2.2]octane, 1,3,3-trimethyl- (CAS) Terpan SS Zineol SS Eucapur SS p-Cineole SS Cajeputol SS Eucalyptol SS Cucalyptol SS Zedary



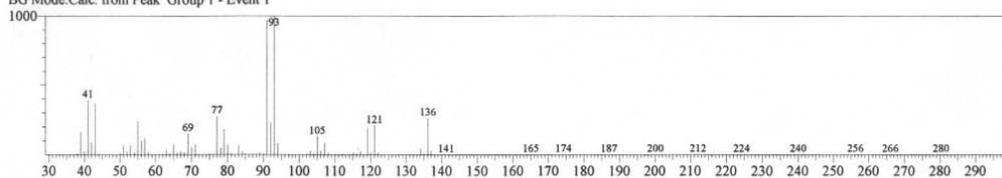
Spectrum3 #Calculation Result#
MassPeaks:79 BasePeak:41.00(49)



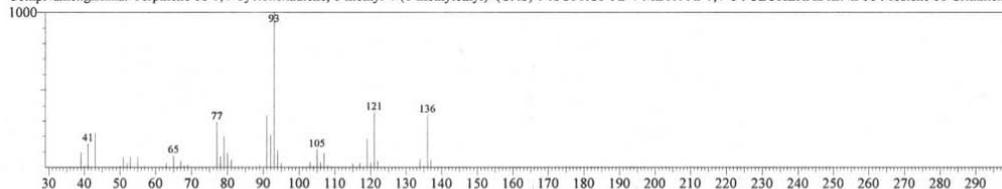
8. γ -Terpinene

Spectrum Comparison

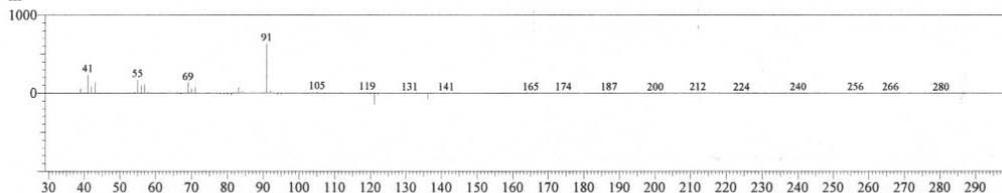
Spectrum1 #Data# 215.QGD R.Time:12.467(Scan#:509)
MassPeaks:127
RawMode:Averaged 12.450-12.483(508-510) BasePeak:93.00(1000)
BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:26282 Formula:C10H16 CAS:99-85-4 MolWeight:136 RetIndex:0
MassPeaks:50 BasePeak:93.00(1000)
CompName:gamma-Terpinene SS 1,4-Cyclohexadiene, 1-methyl-4-(1-methylethyl)- (CAS) 1-ISOPROPYL-4-METHYL-1,4-CYCLOHEXADIENE SS Moslene SS Crithmene



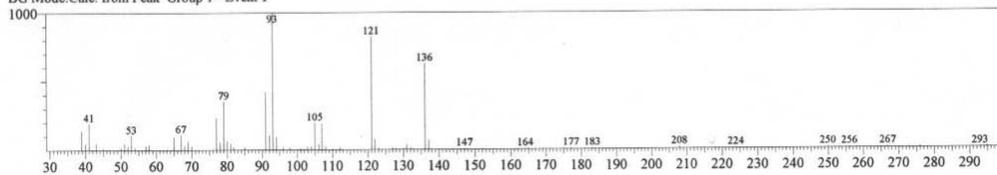
Spectrum3 #Calculation Result#
MassPeaks:126 BasePeak:91.00(634)



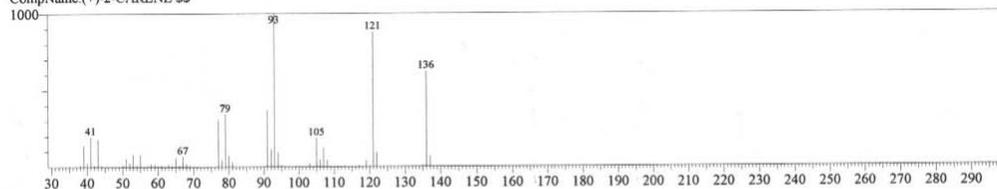
9. (+)-2-Carene

Spectrum Comparison

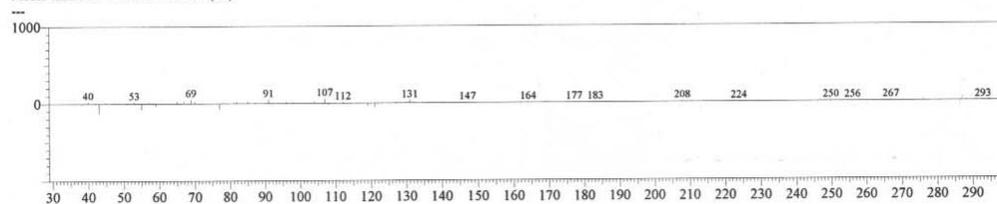
Spectrum1 #Data# 215.QGD R.Time:13.700(Scan#:583)
 MassPeaks:138
 RawMode:Averaged 13.683-13.717(582-584) BasePeak:93.05(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:25524 Formula:C10 H16 CAS:0-00-0 MolWeight:136 RetIndex:0
 MassPeaks:50 BasePeak:93.00(1000)
 CompName:(+)-2-CARENE \$\$

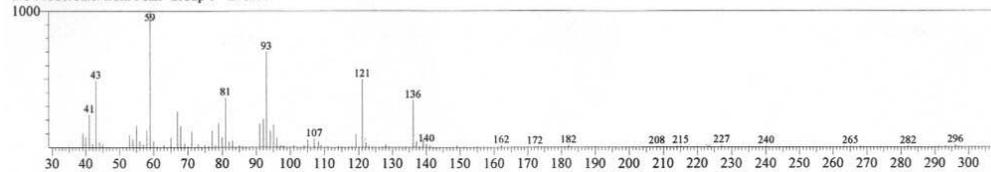


Spectrum3 #Calculation Result#
 MassPeaks:145 BasePeak:107.00(62)

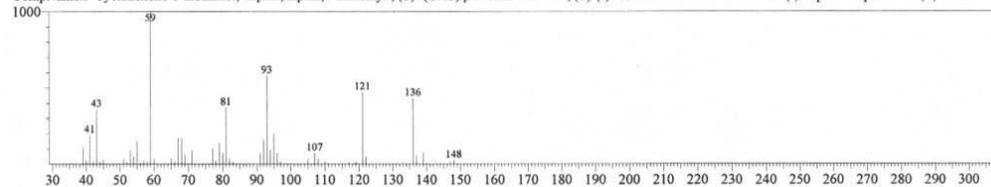
10. α -Terpineol

Spectrum Comparison

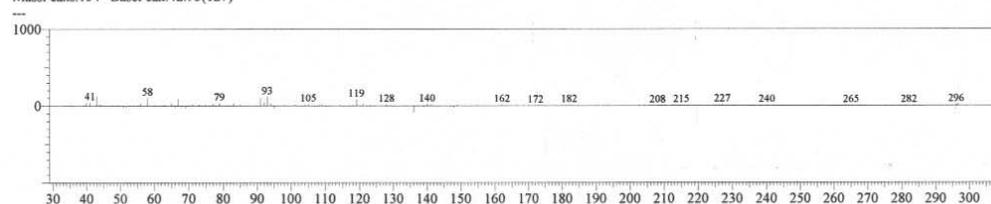
Spectrum1 #Data# 215.QGD R.Time:18.183(Scan#:852)
 MassPeaks:153
 RawMode:Averaged 18.167-18.200(851-853) BasePeak:59.00(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:43792 Formula:C10 H18 O CAS:10482-56-1 MolWeight:154 RetIndex:0
 MassPeaks:50 BasePeak:59.00(1000)
 CompName:3-Cyclohexene-1-methanol, alpha.,alpha.,4-trimethyl-, (S)- (CAS) p-Menth-1-en-8-ol, (S)-(-)- \$\$ ALPHA-TERPINEOL \$\$ (-)-alpha-Terpineol \$\$ (+)-ALPHA-TE



Spectrum3 #Calculation Result#
 MassPeaks:154 BasePeak:42.95(127)

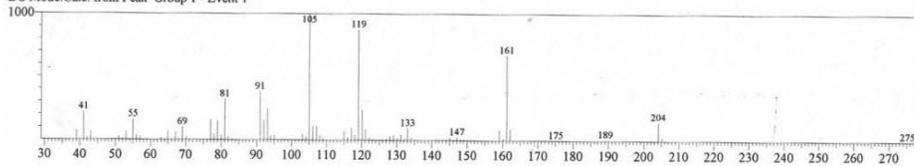


Spectrum comparison of P330 sesquiterpene synthase products

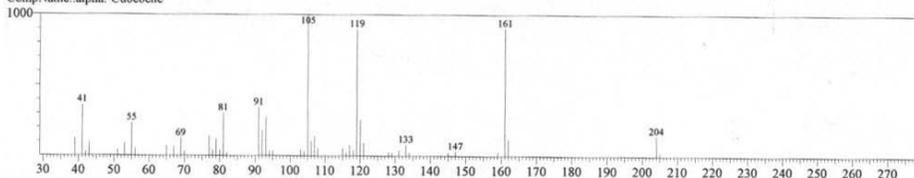
1. α -Cubebene

Spectrum Comparison

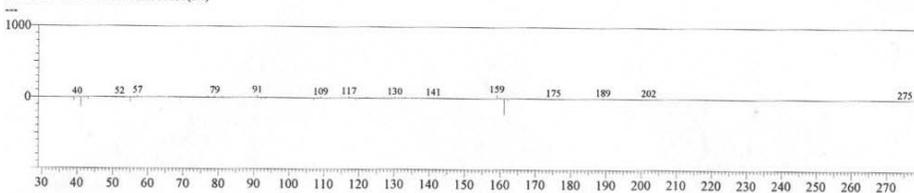
Spectrum1 #Data# p330_140408_R.QGD R.Time:25.267(Scan#:1277)
 MassPeaks:123
 RawMode:Averaged 25.250-25.283(1276-1278) BasePeak:105.05(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# NIST21.LIB Entry:12919 Formula:C15H24 CAS:17699-14-8 MolWeight:204 RetIndex:0
 MassPeaks:50 BasePeak:105.00(1000)
 CompName:alpha-Cubebene



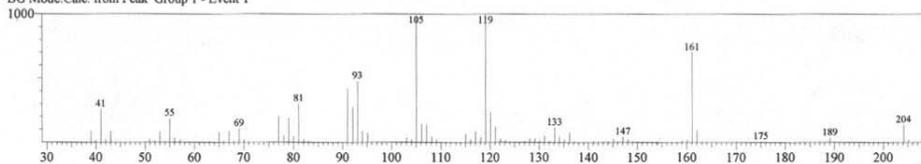
Spectrum3 #Calculation Result#
 MassPeaks:121 BasePeak:159.10(54)



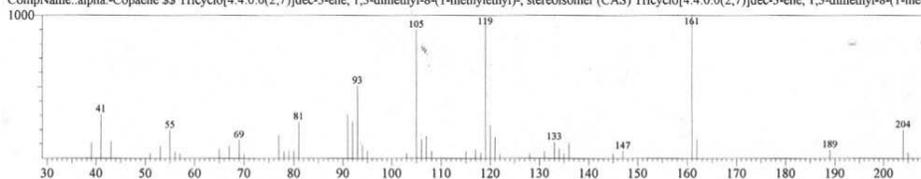
2. α -Copaene

Spectrum Comparison

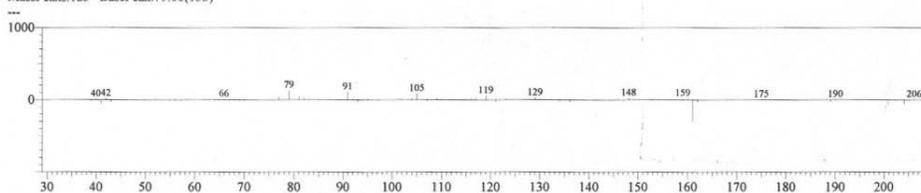
Spectrum1 #Data# p330_140408_R.QGD R.Time:26.433(Scan#:1347)
 MassPeaks:120
 RawMode:Averaged 26.417-26.450(1346-1348) BasePeak:119.10(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:101056 Formula:C15H24 CAS:3856-25-5 MolWeight:204 RetIndex:0
 MassPeaks:50 BasePeak:161.00(1000)
 CompName:alpha-Copaene \$\$ Tricyclo[4.4.0.0(2,7)]dec-3-ene, 1,3-dimethyl-8-(1-methylethyl)-, stereoisomer (CAS) Tricyclo[4.4.0.0(2,7)]dec-3-ene, 1,3-dimethyl-8-(1-methylethyl)-, stereoisomer (CAS)



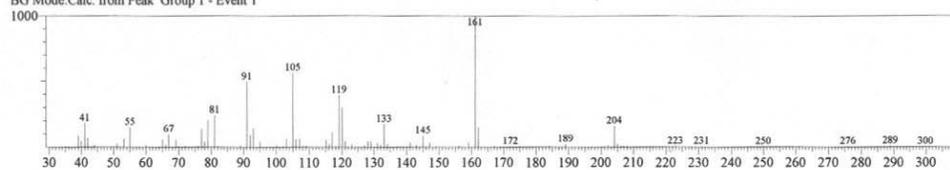
Spectrum3 #Calculation Result#
 MassPeaks:123 BasePeak:79.00(133)



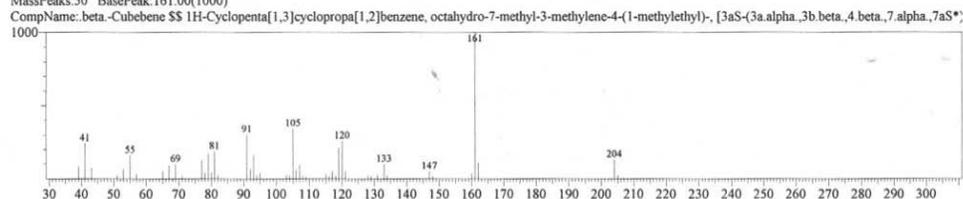
3. β -Cubebene

Spectrum Comparison

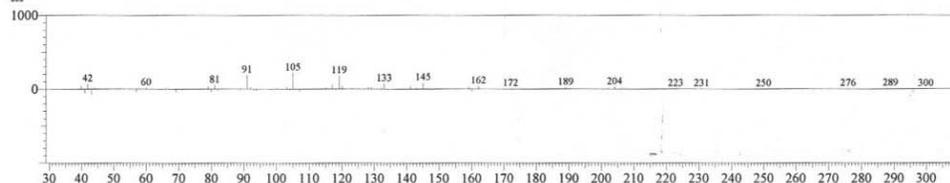
Spectrum1 #Data# p330_140408_R.QGD R.Time:27.033(Scan#:1383)
 MassPeaks:127
 RawMode:Averaged 27.017-27.050(1382-1384) BasePeak:161.15(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:101072 Formula:C15 H24 CAS:13744-15-5 MolWeight:204 RetIndex:0
 MassPeaks:50 BasePeak:161.00(1000)



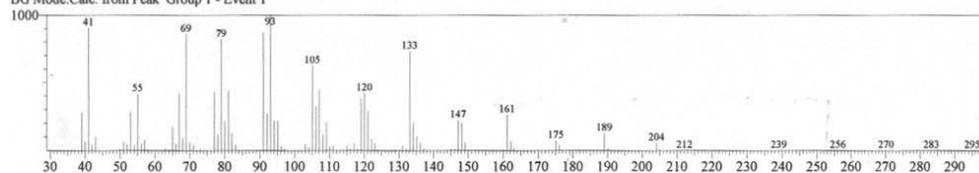
Spectrum3 #Calculation Result#
 MassPeaks:135 BasePeak:105.05(220)



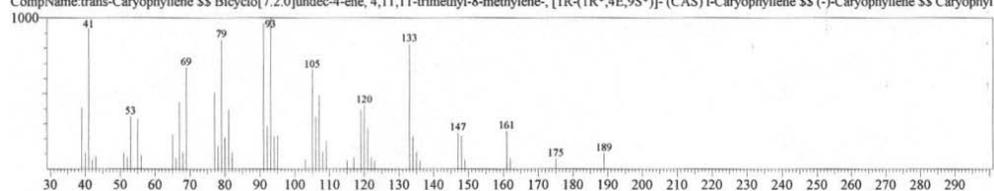
4. trans-Caryophyllene

Spectrum Comparison

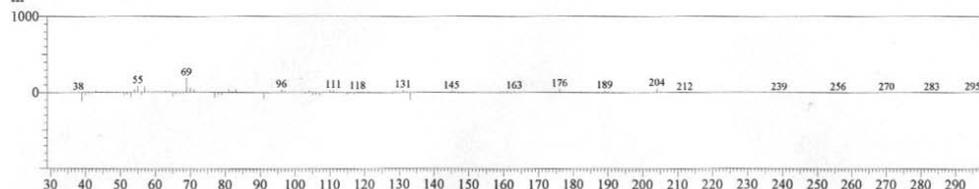
Spectrum1 #Data# p330_140408_R.QGD R.Time:28.317(Scan#:1460)
 MassPeaks:151
 RawMode:Averaged 28.300-28.333(1459-1461) BasePeak:93.05(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:100779 Formula:C15 H24 CAS:87-44-5 MolWeight:204 RetIndex:0
 MassPeaks:50 BasePeak:93.00(1000)



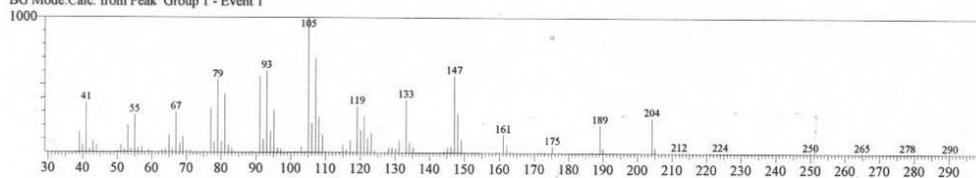
Spectrum3 #Calculation Result#
 MassPeaks:150 BasePeak:69.00(187)



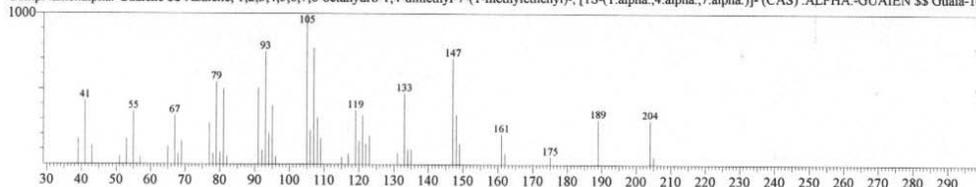
5. α -Guaiene

Spectrum Comparison

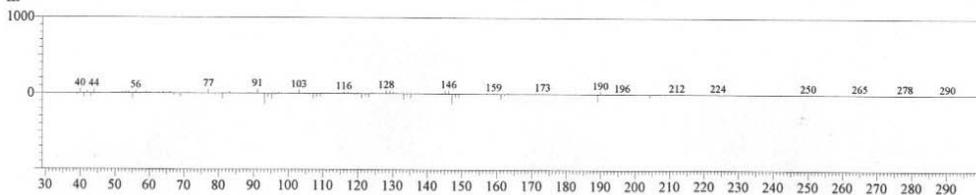
Spectrum1 #Data# p330_140408_R_QGD R.Time:29.067(Scan#:1505)
 MassPeaks:158
 RawMode:Averaged 29.050-29.083(1504-1506) BasePeak:105.05(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:100814 Formula:C15 H24 CAS:3691-12-1 MolWeight:204 RetIndex:0
 MassPeaks:50 BasePeak:105.00(1000)
 CompName: alpha.-Guaiene \$\$ Azulene, 1,2,3,4,5,6,7,8-octahydro-1,4-dimethyl-7-(1-methylethenyl)-, [1S-(1.alpha.,4.alpha.,7.alpha.)]- (CAS) ALPHA.-GUAIEEN \$\$ Guaia-1(5



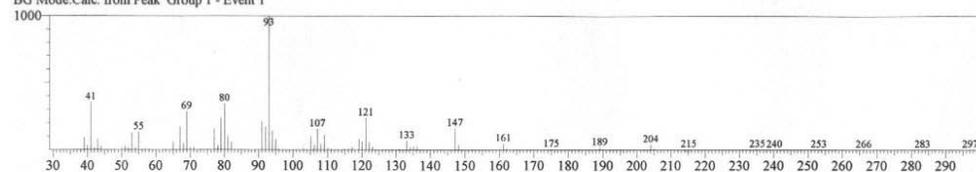
Spectrum3 #Calculation Result#
 MassPeaks:159 BasePeak:91.05(58)



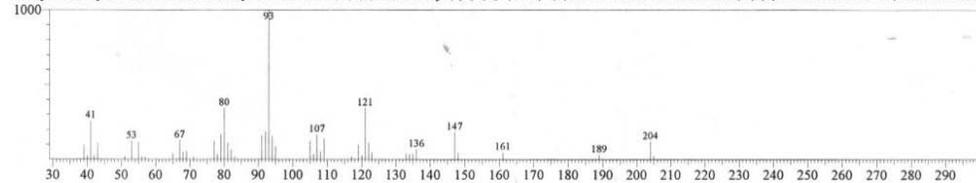
6. α -Humulene

Spectrum Comparison

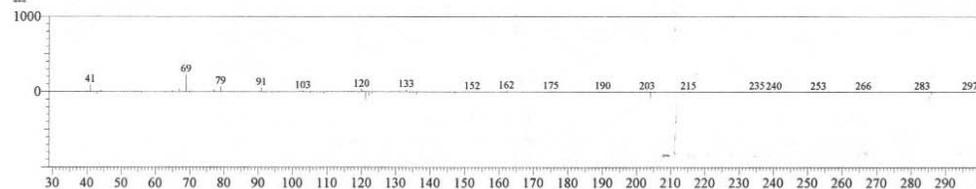
Spectrum1 #Data# p330_140408_R_QGD R.Time:29.733(Scan#:1545)
 MassPeaks:139
 RawMode:Averaged 29.717-29.750(1544-1546) BasePeak:93.05(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:100735 Formula:C15 H24 CAS:6753-98-6 MolWeight:204 RetIndex:0
 MassPeaks:50 BasePeak:93.00(1000)
 CompName: alpha.-Humulene \$\$ 1,4,8-Cycloundecatriene, 2,6,6,9-tetramethyl-, (E,E,E)- (CAS) 4,7,10-CYCLOUNDECATRIENE, 1,1,4,8-TETRAMETHYL-, ALL-CIS \$\$ H



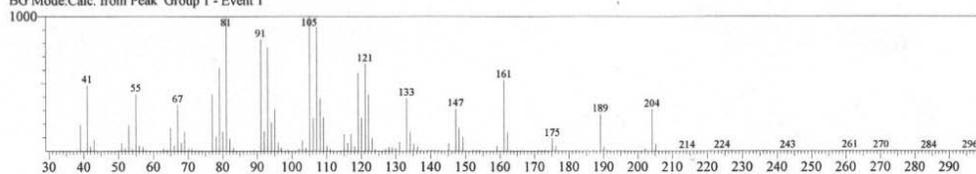
Spectrum3 #Calculation Result#
 MassPeaks:139 BasePeak:69.00(230)



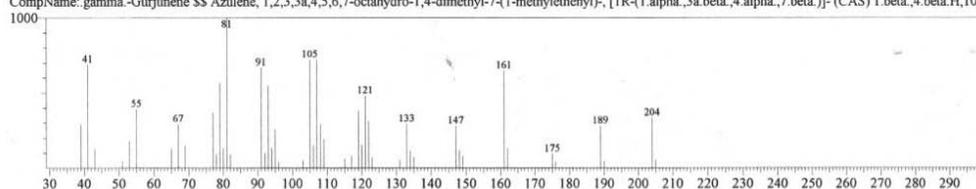
7. γ -Gurjunene

Spectrum Comparison

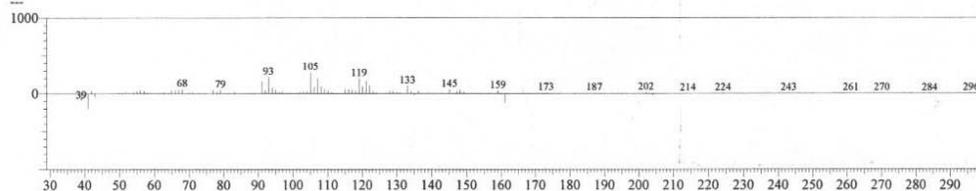
Spectrum1 #Data# p330_140408_R.QGD R.Time:30.783(Scan#:1608)
 MassPeaks:174
 RawMode:Averaged 30.767-30.800(1607-1609) BasePeak:105.05(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:100811 Formula:C15 H24 CAS:22567-17-5 MolWeight:204 RetIndex:0
 MassPeaks:50 BasePeak:81.00(1000)
 CompName: gamma-Gurjunene SS Azulene, 1,2,3,3a,4,5,6,7-octahydro-1,4-dimethyl-7-(1-methylethenyl)-, [1R-(1.alpha.,3a.beta.,4.alpha.,7.beta.)]- (CAS) 1.beta.,4.beta.H,10.b



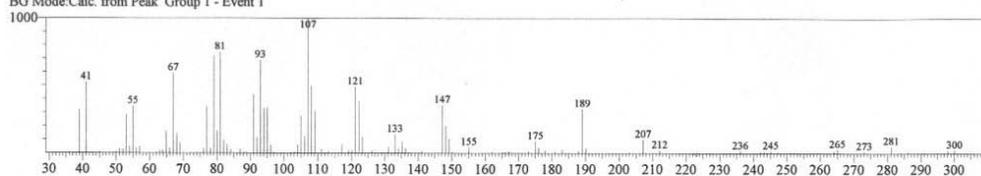
Spectrum3 #Calculation Result#
 MassPeaks:174 BasePeak:105.05(280)



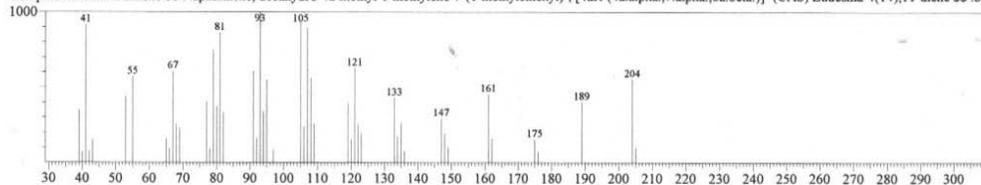
8. β -Selinene

Spectrum Comparison

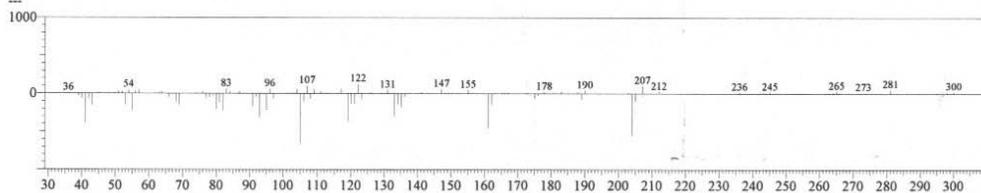
Spectrum1 #Data# p330_140408_R.QGD R.Time:31.183(Scan#:1632)
 MassPeaks:141
 RawMode:Averaged 31.167-31.200(1631-1633) BasePeak:107.05(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:100909 Formula:C15 H24 CAS:17066-67-0 MolWeight:204 RetIndex:0
 MassPeaks:50 BasePeak:93.00(1000)
 CompName: beta-Selinene SS Naphthalene, decahydro-4a-methyl-1-methylene-7-(1-methylethenyl)-, [4aR-(4a.alpha.,7.alpha.,8a.beta.)]- (CAS) Eudesma-4(14),11-diene SS .beta.



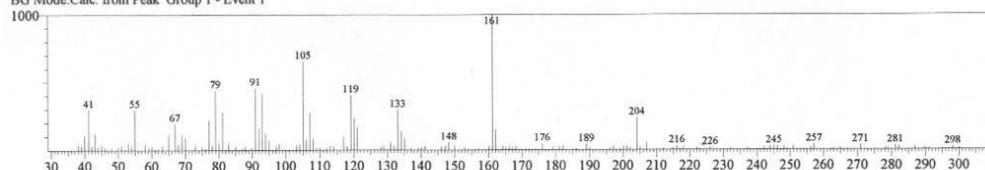
Spectrum3 #Calculation Result#
 MassPeaks:148 BasePeak:122.20(128)



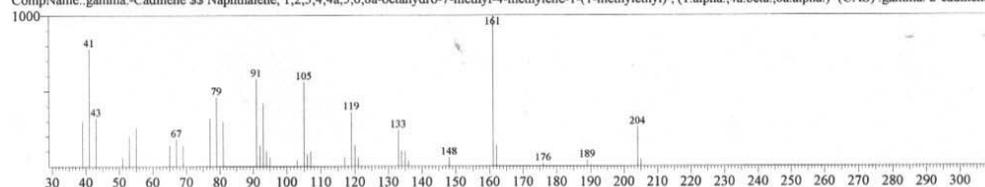
9. γ -Cadinene

Spectrum Comparison

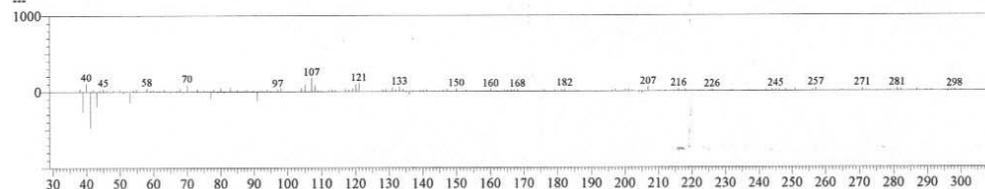
Spectrum1 #Data# p330_140408_R_QGD R.Time:32.183(Scan#:1692)
 MassPeaks:164
 RawMode:Averaged 32.167-32.200(1691-1693) BasePeak:161.15(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:100883 Formula:C15 H24 CAS:39029-41-9 MolWeight:204 RetIndex:0
 MassPeaks:42 BasePeak:161.00(1000)
 CompName:.gamma.-Cadinene SS Naphthalene, 1,2,3,4,4a,5,6,8a-octahydro-7-methyl-4-methylene-1-(1-methylethyl)-, (1.alpha.,4a.beta.,8a.alpha.)-(CAS) .gamma. 2-cadinene !



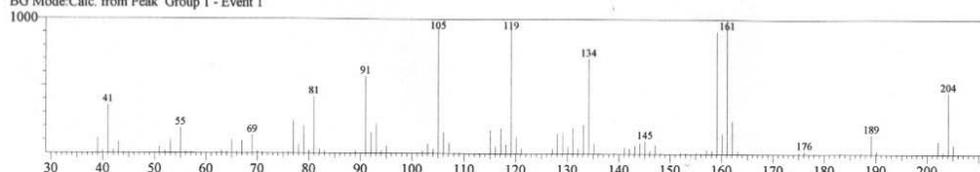
Spectrum3 #Calculation Result#
 MassPeaks:166 BasePeak:107.05(178)



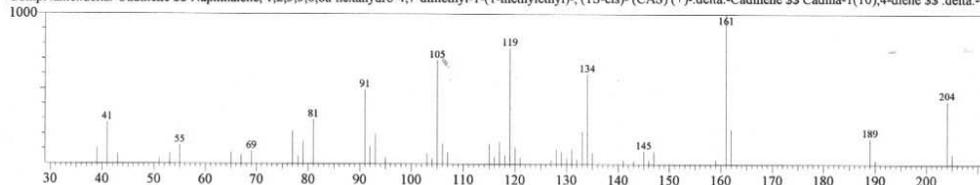
10. δ -Cadinene

Spectrum Comparison

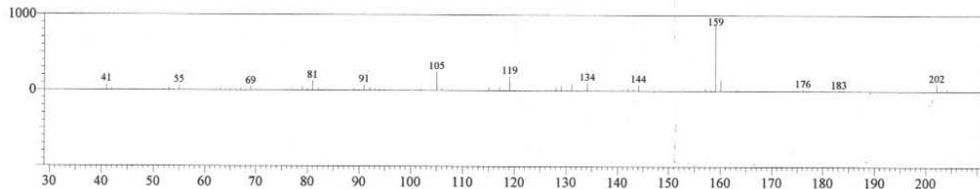
Spectrum1 #Data# p330_140408_R_QGD R.Time:32.533(Scan#:1713)
 MassPeaks:138
 RawMode:Averaged 32.517-32.550(1712-1714) BasePeak:161.10(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:100891 Formula:C15 H24 CAS:483-76-1 MolWeight:204 RetIndex:0
 MassPeaks:50 BasePeak:161.00(1000)
 CompName:.delta.-Cadinene SS Naphthalene, 1,2,3,5,6,8a-hexahydro-4,7-dimethyl-1-(1-methylethyl)-, (1S-cis)- (CAS) (+)-.delta.-Cadinene SS Cadina-1(10),4-diene SS .delta.-



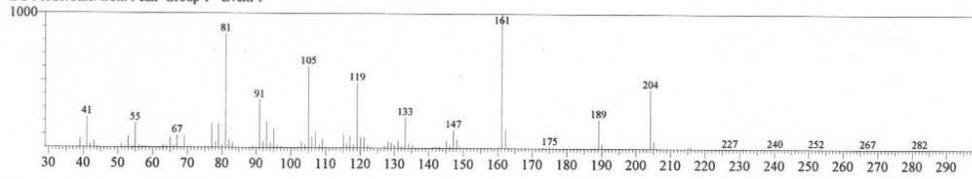
Spectrum3 #Calculation Result#
 MassPeaks:136 BasePeak:159.10(875)



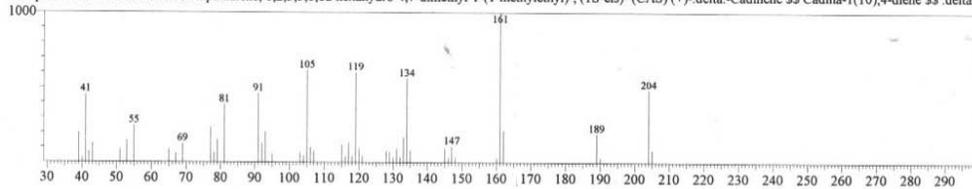
11. δ -Cadinene

Spectrum Comparison

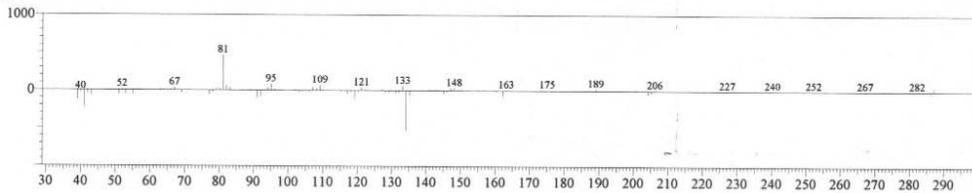
Spectrum1 #Data# p330_140408_R.QGD R.Time:32.633(Scan#:1719)
 MassPeaks:144
 RawMode:Averaged 32.617-32.650(1718-1720) BasePeak:161.10(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:100888 Formula:C15 H24 CAS:483-76-1 MolWeight:204 RetIndex:0
 MassPeaks:50 BasePeak:161.00(1000)
 CompName:delta.-Cadinene SS Naphthalene, 1,2,3,5,6,8a-hexahydro-4,7-dimethyl-1-(1-methylethyl)-, (1S-cis)- (CAS) (+)-delta.-Cadinene SS Cadina-1(10),4-diene SS .delta.-



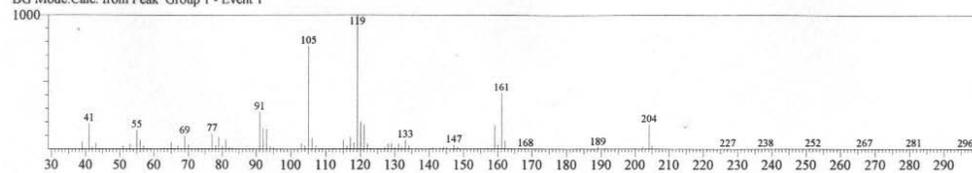
Spectrum3 #Calculation Result#
 MassPeaks:143 BasePeak:81.05(462)



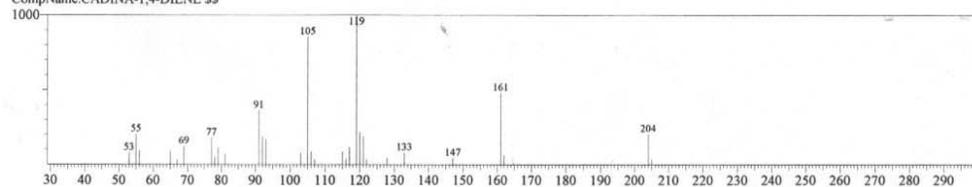
12. Cadina-1,4-diene

Spectrum Comparison

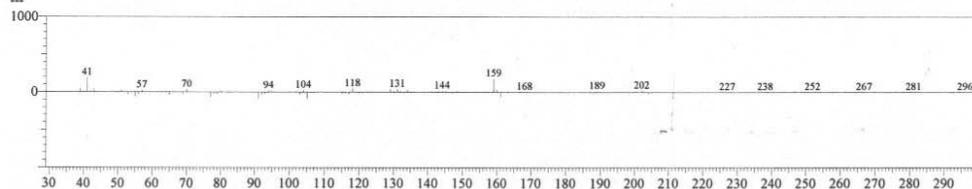
Spectrum1 #Data# p330_140408_R.QGD R.Time:32.900(Scan#:1735)
 MassPeaks:144
 RawMode:Averaged 32.883-32.917(1734-1736) BasePeak:119.10(1000)
 BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:101123 Formula:C15 H24 CAS:0-00-0 MolWeight:204 RetIndex:0
 MassPeaks:33 BasePeak:119.00(1000)
 CompName:CADINA-1,4-DIENE SS



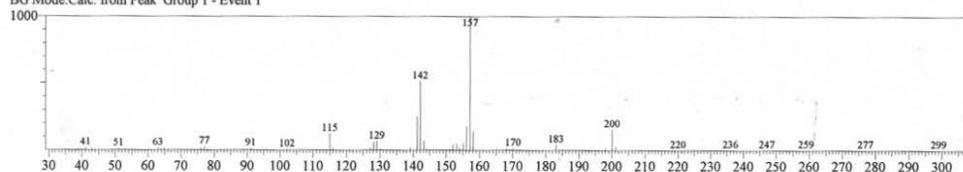
Spectrum3 #Calculation Result#
 MassPeaks:143 BasePeak:41.00(189)



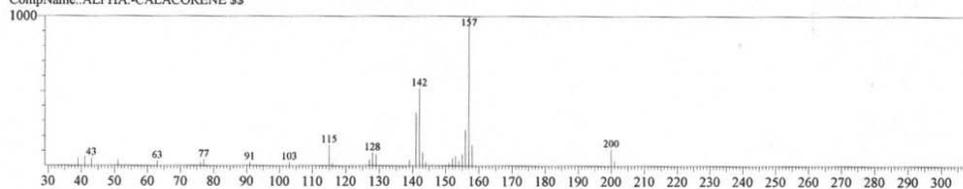
13. α -Calacorene

Spectrum Comparison

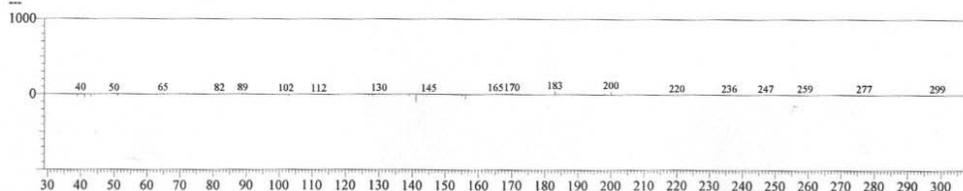
Spectrum1 #Data# p330_140408_R.QGD R.Time:33.333(Scan#:1761)
MassPeaks:140
RawMode:Averaged 33.317-33.350(1760-1762) BasePeak:157.10(1000)
BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:95638 Formula:C15 H20 CAS:21391-99-1 MolWeight:200 RetIndex:0
MassPeaks:32 BasePeak:157.00(1000)
CompName:ALPHA-CALACORENE \$\$



Spectrum3 #Calculation Result#
MassPeaks:141 BasePeak:200.05(45)

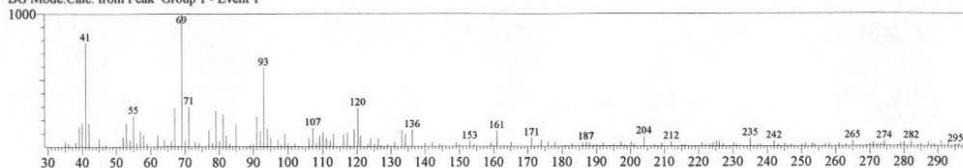


Spectrum comparison of 1025 sesquiterpene synthase products

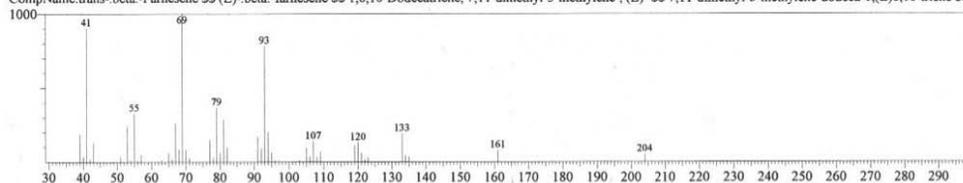
1. β -Farnesene

Spectrum Comparison

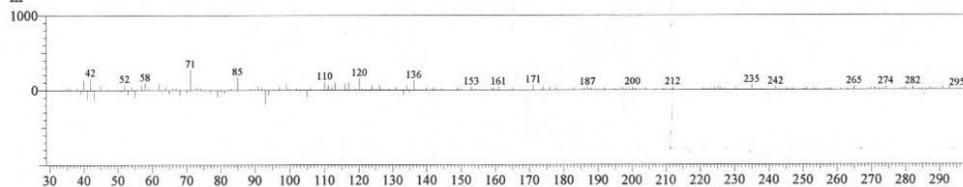
Spectrum1 #Data# 1025.QGD R.Time:29.633(Scan#:1539)
MassPeaks:174
RawMode:Averaged 29.617-29.650(1538-1540) BasePeak:69.10(1000)
BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:100671 Formula:C15 H24 CAS:502-60-3 MolWeight:204 RetIndex:0
MassPeaks:50 BasePeak:69.00(1000)
CompName:trans- β -Farnesene \$\$ (E)- β -farnesene \$\$ 1,6,10-Dodecatriene, 7,11-dimethyl-3-methylene-, (E)- \$\$ 7,11-dimethyl-3-methylene-dodeca-1,(E)6,10-triene \$\$



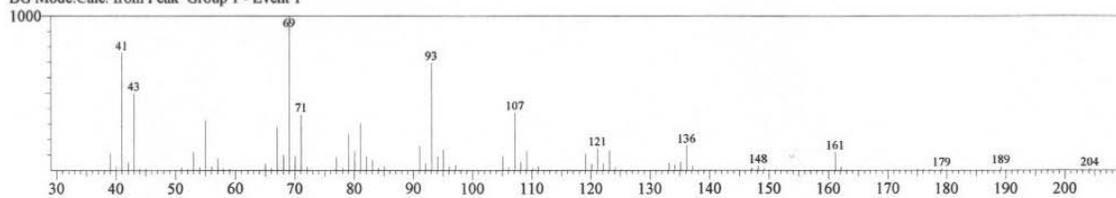
Spectrum3 #Calculation Result#
MassPeaks:180 BasePeak:71.15(272)



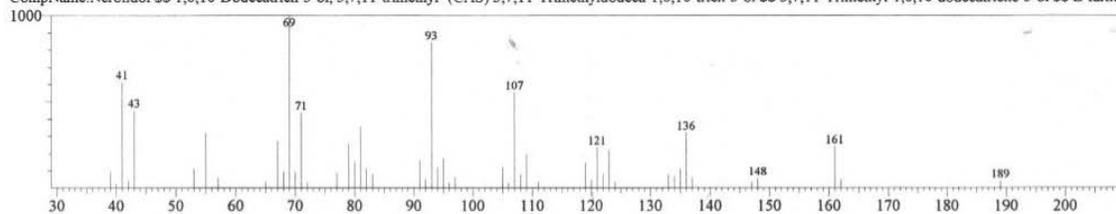
2. Nerolidol

Spectrum Comparison

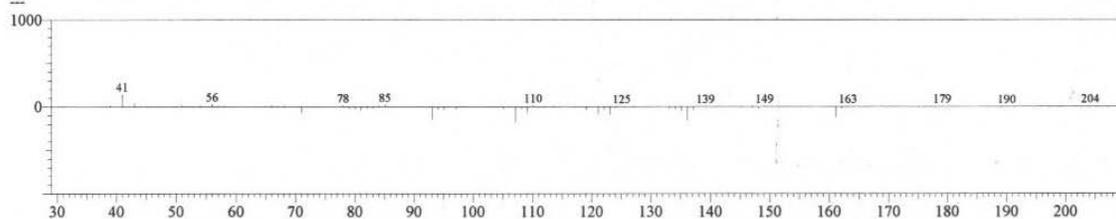
Spectrum1 #Data# AM75B_1025.QGD R.Time:34.000(Scan#:1861)
MassPeaks:124
RawMode:Averaged 33.983-34.017(1860-1862) BasePeak:69.05(1000)
BG Mode:Calc. from Peak Group 1 - Event 1



Spectrum2 #Library# WILEY7.LIB Entry:123919 Formula:C15 H26 O CAS:7212-44-4 MolWeight:222 RetIndex:0
MassPeaks:50 BasePeak:69.00(1000)
CompName:Nerolidol \$\$ 1,6,10-Dodecatrien-3-ol, 3,7,11-trimethyl- (CAS) 3,7,11-Trimethyldodeca-1,6,10-trien-3-ol \$\$ 3,7,11-Trimethyl-1,6,10-dodecatriene-3-ol \$\$ E-farnes



Spectrum3 #Calculation Result#
MassPeaks:123 BasePeak:41.00(146)



APPENDIX 3

Clone A3 – Putative phosphatase 2A inhibitor homologous

1 CCTCTCCCgaMTTCGGCACGAGGCGAAATCAAGAAAATCAGTTAAGCAGC
 51 TCTgTAACTcAGGTGGGAAAAAGGCAAAAAATAATGGTGGTTGACAAAGG
 101 GAAGAAGCAGAAAGTGGAAAGaGGAAAGCTACATTGATGAAAAGCTCATTT
 151 TTTCCATTGAAAAATTGCAAGAAATACAAGACGaCCTTGACAAGATCAAT
 201 GAGAAAGCAAGTGAGGaAGTGTTgGAAATAGAACAGAAGTACAACAAGAT
 251 CCGCAAGCCTGTTTTATGATAAGcGGAATGATATCATTAACTCTATTTCTG
 301 ACTTCTGGTTGACTGCTTTTTTGTAGTCATCCTGTTCTTGGtGACctTCTA
 351 ACTGAAGaGGACCAAAAGATTTTTCAAATTCTTAAGTTCTATTGAAGTGGa
 401 AGACTCGAAAGATGTGAAATTTGGTTACTCAATCACGTTTAACTTTAAGC
 451 CCAATCCTTTCTTTGAAAATTCAAAGCTCTCAAAGACCTATNCCTTCCTT
 501 GAAGaTGGACCTACAAAAATCACAGCTACACCAATAAAATGGAAAGaAGG
 551 cAAAGGcATTCCTAATGGCGTTGCTCAGGaGAAGAAAGGAAACAAGCGAT
 601 CCCaTGCTGAAGAGaGCTTCTTCACCTGGtTCAGTGaaGTCAATAAAAAA
 651 GATGATAGCGATGATGATGAAAATgAGGTTCTGGAGATTCAGGATGAGGT
 701 TGCTGAAATAATCAAGGATGACTTgTGGCCAAACCCTTTAACTTATTTTA
 751 CCAATGaACCTGATGAAGAAGATTTTTGAGGGTGATGAAGGTGGtGATGAG
 801 GGGGAGGACTcTGAAGATGAAGGTGATGAGGAGGAAGAGGAAGACGACGA
 851 AGATGAAGATGACAAATGAACTgTTAATgGACCTCATATTTGATTTgATT
 901 TCTCTTCTTCAATGTTTCAAtTATCATAGTTGGTATCTGTAAAGAAGCTT
 951 AATATTgCAGaTAAATCGATTATATATaGtGGKGRctSCTTTTTTWaCTAC
 1001 AAAAAAAAAAAAAAAAAAaAAAAAAAAAAAAAACTCGGGaGGTTTTTGM

Clone A5 – *Lycopersicon esculentum* HSP90-2 homologous

1 HatgcctMaaAaAaAttCGGcAGNGNtGaGGCGTTCTCTAAAAaCCTCAA
 51 GCTTGGAATCCATGAGGaTTCTCAGAACAGGGCAAAGTTTGCTGAACTGC
 101 TGAGGTACCACtCCACTAAGAGTGGTGATGAGATGACCAGCTTGAAGGAC
 151 TATGTGACCAGAATGAAGGAGGGCCAGAATgATATTTACTACATTACTGG
 201 TGAGAGCAAGAAGGCTGTTGAGAACTCTCCCTTCCtGGAGAACTGAAGA
 251 AGAAGGGATATGAGGTGCTTTACATGGTTGATGCCATTGATGaGTATTCA
 301 ATTGGTCAGCTGAAGGAATTTGAGGGCAAAAAGCTTgTTTCTGCTACCAA
 351 GGAAGGCCTCAAGCTTGctGAGAGTGAAGATGAGAAGAAAAAGCAGGAAG
 401 AATTgAAGGAGAAGTTTGAGGGACTGTGTAAGGTGATGAAGGATGTGCTA
 451 GGaGACAAAGTTGAAAAGGTCATTGTTTCTGACCgTGTYNNTgACTCTCC
 501 CTGCTGTTTGGTCACTGGTGAGTATGGCTGGACyGCTAACATGGaGAGAA
 551 TTATGAAGGCACAGGCACTTAGGGACTCCAGCATGGCTGGATACATGTCT
 601 aGCAAGAAGACCATGGaGATCAACCAGAGAACTCCATCATGGATGAGCT
 651 AAGGAAGAGGGCTgATGCAGACAAGAATGACAAGTCTGTGAAGGACTTGG
 701 TTCTCTTGCTTTTTGAGACTGCCctTCTCACCTCAGGTTTCAGCCTCGAG
 751 GaGCCAAACACCTTTgGCAACAGAATTCACAGGaTGTTGAAACTCGGTTT
 801 GAGCATTGATGAGGAAAGCGGAGATGCTGATGCTGACAtGCCAGCATTGG
 851 AGGATCCtGAAGCTgATGCTGAGGGCAGCAAGATGGAGGAGGTTGATTAA
 901 GtTCATTAATGTTTtATAGTTTTTATgGGTCCC

Clone B8 – Gluthation S-transferase homologous

1 CTcTMCcGAaTTCGGCACGAGAACAAAAATGGCTAACGATGAAGtGATTC
 51 TGTTGGATTTTTGGCCTAGCATGTTTGGAAATGAGGCTAAGGATTgCACTT
 101 GCTGAGAAAGAGATTAAATATGAGTACAGAGACGaGGATTTGAGGAACAA
 151 AAGTCCTCTGCTTTTGCAGATGAATCCTATTCACAAGAAAATCCCTGTTT
 201 TGATTcATAATGGCAAACCAATTTGCGAGTCTATTATTGGAGTTGAGTAT
 251 ATTGATGAAGTGTGGAAGGATAAAGCCCTTTCCCTCCCTTCTGATCCTTA
 301 TGAGAGAGCACAAGCTAGGTTTTGGGCTGATTACATTGATAAGAAGTTGT
 351 ATGATTCTGGGAGGAAGCTATGGACAACAAAGGGAGAAGAGCAAGAGACA
 401 GCTAAAAAAGATTTcATAGAATGCCTTAAGGTGTTAGAGGGAGCACTTGG
 451 AGAGAAGCCTTACTTTGGAGGAGATAATTTTGGTTGTGTGGATATTGCGT
 501 TGATTgGGTACTATAGCTGGTTTTTATGCTTATGAGAGTTATGCTAACATC
 551 AGTGTAGAGGCTGAGTGCCCAAAGTTTgTGGCTTGGGCCAAGAATTgTAT
 601 GCTGAGGGACAGTGTGGCTAAGTCTTTGCCTGATCAACATAAGGtCTGTG
 651 AATTTGTAAAAGTGCTAAGGCAGAAGTTTGGAAATTGAATAAACATATGGa
 701 CCATGATTCTGCTAATAAAGGcGcTTCGCCCTTCTTTTTATTGAAATGTG
 751 TGTTTTGGtCATAGTGTAATTGAGTCTGCTTTTACATGTAATAAAAACTC
 801 TTGGATCTgTTTAAAAAAGAAAGACTTTGGACTTcTTCGcCAGAGG
 851 TTTgGTCAAGTCTCCAATCAAGGTTGTcGGCTTGTCTACCTTGCCAGAAA
 901 TTTACGaAAAGATGGAAAAGGGTCAAATCGTTGGtAGATACGKTGTTGAC
 951 ACTTYTAAATAAGcGAATTTCTTaTGATTTATGTTTTTTW

Clone C4 – Luminal-binding protein 4 homologous

1 AATTcgGCCATTACGGCGGGGAATTATTAACgaCCAACCTGCAGCAGCAAT
 51 TGCATATGGCTTGGATAAGAAAGGTGgTGAAAAGAACATTCTCGTCTTTG
 101 ATCTTgGTGGTGAACCTTTTgATGTcAGTATCTTGACCATTGAtAATGGT
 151 GTGTTTgAGGTTCTTgCAACAAATGGAGACACCCaTTTGGGAGGTGAGGA
 201 TTTTgATCAGAGGGTTATGGAATACTTCATTAATTTGATCAAGAAGAAGC
 251 ACAGTAAGGACATCAGCAAGGATAACAGAGCACTTgGCAAGTTGAGGAGA
 301 GAGTGTGAGCGTGCgAAGAGAGCTTTGAGTAGTCAGCACCAGGTCCGCGT
 351 GGAGATTGAGTCTCTCTTTgACGGTGTGGATTTCTCTgAGCCTCTTACAA
 401 gGGCTCGTTTTGAAGAActGAACAACGATTTATTcAGAAAGACCATGGGT
 451 CCTGTTAAGAAGGCCATGGATGACGCTGGTCTTgAGAAGCGTCAAATTgA
 501 TGAAATTGTTCTTGTGGTGGAAAGTACTAGGATTCCCAAGATTCAACAGC
 551 TTCTCAAGGACTACTTTgATGGCAAGGAGCCCaACAAGGGTGTCAACCCg
 601 ATGAGGCTGTTGCTTTTCGGTGTGCTGTCCAAGGAGGAATTCTGAGTGGC
 651 GAGGGAGGTGATGAGACCAAGATATTCTTCTCTTgGATGTGGCACCATT
 701 GACCCtTGGTATTgAAACTGTTGGAGGAGTGTGACCAAATTGATTCCAA
 751 GGAACACTGTcATTCCAACCAAGAAATCACAAGTCTTcACTACTTACCAG
 801 GaCCaACAAACTACTGTcACAATTCAGGTGTTTgAGGGCGAACGGAGCCT
 851 TACAAAGGACTGTAGGTTGCTTgGGAAATTTGATCTGACAGgAATAGCCC
 901 CAGCCCCAAGGGTACCCACAAATCGAAGTTACATTTGAAGTTGATGCTAA
 951 TGGTATCtggAACGTCAAGGCTGWAGACAAGGCCTCTGGGCGATCAGAGA
 1001 AGATCACAAAtCAMCaACGACAaAGGTCTTTGaGCCAGGAagaGATTD

Clone D1 – SEC14 cytosolic factor family protein homologous

1 SMMTtCgGCCATTACGgCGGGGAGACATTCTCTCATTTTCAGTACTCTTCA
 51 AAAGTATCAAGaaTAATGAGTGCTGGGTCTGAATAAGCCTTCTCTAGACGg
 101 CGAAGAGAATTCCTTATCTCCAGAAACGCAGCAGGGGAAGATTAAGAGG
 151 TAAGAAAGTTGATCGGGCCACTgTCAGGGAAACTAGCCATCTACTGCTCG
 201 GATGCATGTATCGCAAGATACTTGAGAGCACGAAGCTGGAATGTCAAGAA
 251 GGCCgTCAAGATGCTCAAAGCGACTCTAAAGTGGAGACTGGAGTACACGC
 301 CAGAAGAGATACGTTGGGACGACGTTgCAGTGGAAGCAGAGACAGGAAAG
 351 ATTTACAGATCTTcTTACAAAGATAAACATGGGAGGCCGGTTCTTGTAAAT
 401 GAGACCTCGTTGTCAGAATACAAAGTCGATAAAAAGGACAAATTAAGTATT
 451 TGGTTTATTGTCATGGaGAATGCAATTCTGAATCTTgGTGAGGAACAAGAG
 501 CAGATGGTGTGGTTGATTgATTTTCATGGATTTCAGCTTGTCCAATGTATC
 551 AATCAAGGTCACCTAAAGAAACAGCTCATATCCTACAAGACCATTACCCCG
 601 AGAGGCTTgGTGTCGCGATTTTATACGACGCCCCCAAGATTTTCGAACCA
 651 TTTTGGAAAGATGGCGAASctTTCTgGAGCCTAAAACGgCCAACAAAGTCA
 701 AGTTCGTCTACTCCGACGATCCTACCACGAACAAGAtCATGGACGAGTGG
 751 TTCGACATGGAGATGGTGGAGTACGcGTTTGGTGGAAAGGaCGAGGaAGA
 801 TTTTGACATCAGCAAGTACGcGGTGAGGaTGAGGaAGGACGaCGAAAAGC
 851 TCCCCTTCATGTGGCAAACACATCGTAGTTTgATACCACCaCCACCaCCA
 901 CCACCACCACCgCAGAATCCTaCCgCAATCAACTTTACGAGTTCAGACAT
 951 gGATTYcGgAAGCCTCGGAAATAAAAGCAGACAAAGCAGCCgTTTcTATAG
 1001 ATGgAGgWGAAGAGGGaGCMAR

Clone D4 – HSP90-2 homologous

1 GAATTCGGCCATTACGGCCGGGTCTATGAGGCTTTCTCAAAGAACCTcAA
 51 GCTTGGCATCCACGAGGATTTCCAGAACAAGGGAAAGATAGCAGAGTTGC
 101 TGAGGTACCACTCGACCAAGAGTGGTGATGAGATGACCAGCTTGAAGGAC
 151 TATGTGACCAGGATGAAGGAAGGACAGAGCGATATCTTCTACATTACTGG
 201 TGAGAGCAAGAAGGCTGTTGAGAACTCCCCTTTCCTTGAGAGGTTGAAGA
 251 AGAAGGGCTATGAAGTCCTGTACATGGTTGATGCCATTGATGAGTATGCT
 301 GTTGGCCAGTTGAAGGAATTTGAGGGCAAGAAGCTTGTGTCTGCTACCAA
 351 GGAGGGTCTGAAGCTCGATGAGAGCGAGGACGAGAAGGCAAAGAAGGAAG
 401 AGTTGGTGAAGAAATTCGAGGGTCTGTGCAAGGTCATTAAGGATGTGCTT
 451 GGAGACAAGGTTGAGAAGGTTATCGTCTCTGACCgCGTGGTCTGACTCTCC
 501 CTGCTGTTTGGTCACTGGAGAATATGGCCGGACCgCCAACATGGAAGGA
 551 TCATGAAGGCACAAGTTTGGAGGGACTCGAGCATGGCTGGATATATGTCTG
 601 AGCAAGAAGACAGTGGAGATCAACCCTGAGAACCCCATCATGGAAGAGCT
 651 CAGGAAGAGGGCTGATGCTGACAAGAATGACAAGTCTGTCAAGGACTTGG
 701 TTATGCTGCTCTTTGAGACTGCCATGCTCACCTCTGGTTTCAGCCTTGAC
 751 GaGCCCAACACCTTCGGCAACAGGATCCaCAGAATGTTGAAGCTTGGTTT
 801 GAGCATTGACGACGATGCTACTGAAGCTGATGTTGACATGCCACCATTGG
 851 AGGaAGCCGGTCTGAAGCTGAGGGCAGCAAGATGGAGGaAGTCGACTAA
 901 TTCTGCGTCCAGGCAACATATATCTTTTgAAGTTCCCTTgTCTTTACTT
 951 CTGGTATAATGGATTAATATTTCCSBGCTTTTTTGATTTTTTTTTTTTTTa
 1001 CTTTAGgACYGCTTKGCTTATC

REFERENCES

- Ackerman, S. H. and Tzagoloff, A.** (2005). Function, structure, and biogenesis of mitochondrial ATP synthase. *Prog Nucleic Acid Res Mol Biol* **80**, 95-133.
- Adams, J. D., Jr. and Klaidman, L. K.** (1993). Acrolein-induced oxygen radical formation. *Free Radic Biol Med* **15**, 187-93.
- Adams, R. P.**, (1995). Identification of essential oil componenets by GasChromatography / Mass Spectroscopy. Allured Publishing Corporation, Carol Stream, Illinois USA.
- Almeida, J. R., Modig, T., Roder, A., Liden, G. and Gorwa-Grauslund, M. F.** (2008). Pichia stipitis xylose reductase helps detoxifying lignocellulosic hydrolysate by reducing 5-hydroxymethyl-furfural (HMF). *Biotechnol Biofuels* **1**, 12.
- Alper, H., Miyaoku, K. and Stephanopoulos, G.** (2005). Construction of lycopene-overproducing E. coli strains by combining systematic and combinatorial gene knockout targets. *Nat Biotechnol* **23**, 612-6.
- Amoah, B. K., Wu, H., Sparks, C. and Jones, H. D.** (2001). Factors influencing Agrobacterium-mediated transient expression of uidA in wheat inflorescence tissue. *J Exp Bot* **52**, 1135-42.
- Anssour, S.** (2005). Functional expression *in vitro* and biochemical characterization of terpene synthases of *Salvia fruticosa*, Master thesis, MAICH, Greece.
- Arimura, G., Huber, D. P. and Bohlmann, J.** (2004). Forest tent caterpillars (*Malacosoma disstria*) induce local and systemic diurnal emissions of terpenoid volatiles in hybrid poplar (*Populus trichocarpa* x *deltoides*): cDNA cloning, functional characterization, and patterns of gene expression of (-)-germacrene D synthase, PtdTPS1. *Plant J* **37**, 603-16.

Armour, C. D. and Lum, P. Y. (2005). From drug to protein: using yeast genetics for high-throughput target discovery. *Curr Opin Chem Biol* **9**, 20-4.

Atsumi, S., Hanai, T. and Liao, J. C. (2008). Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-9.

Aubourg, S., Lecharny, A. and Bohlmann, J. (2002). Genomic analysis of the terpenoid synthase (AtTPS) gene family of Arabidopsis thaliana. *Mol Genet Genomics* **267**, 730-45.

Back, K. and Chappell, J. (1996). Identifying functional domains within terpene cyclases using a domain-swapping strategy. *Proc Natl Acad Sci U S A* **93**, 6841-5.

Baganz, F., Hayes, A., Marren, D., Gardner, D. C. and Oliver, S. G. (1997). Suitability of replacement markers for functional analysis studies in *Saccharomyces cerevisiae*. *Yeast* **13**, 1563-73.

Bakhrat, A., Jurica, M. S., Stoddard, B. L. and Raveh, D. (2004). Homology modeling and mutational analysis of Ho endonuclease of yeast. *Genetics* **166**, 721-8.

Barr, M. M. (2003). Super models. *Physiol Genomics* **13**, 15-24.

Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F. and Cullin, C. (1993). A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **21**, 3329-30.

Bayne, M. L., Cascieri, M. A., Kelder, B., Applebaum, J., Chicchi, G., Shapiro, J. A., Pasleau, F. and Kopchick, J. J. (1987). Expression of a synthetic gene encoding human insulin-like growth factor I in cultured mouse fibroblasts. *Proc Natl Acad Sci U S A* **84**, 2638-42.

Beggs, J. D. (1978). Transformation of yeast by a replicating hybrid plasmid. *Nature* **275**, 104-9.

- Bertea, C. M., Freije, J. R., van der Woude, H., Verstappen, F. W., Perk, L., Marquez, V., De Kraker, J. W., Posthumus, M. A., Jansen, B. J., de Groot, A. et al.** (2005). Identification of intermediates and enzymes involved in the early steps of artemisinin biosynthesis in *Artemisia annua*. *Planta Med* **71**, 40-7.
- Bertram, J. S. and Vine, A. L.** (2005). Cancer prevention by retinoids and carotenoids: independent action on a common target. *Biochim Biophys Acta* **1740**, 170-8.
- Beukers, M. W. and Ijzerman, A. P.** (2005). Techniques: how to boost GPCR mutagenesis studies using yeast. *Trends Pharmacol Sci* **26**, 533-9.
- Bicchi, C., Cordero, C., Liberto, E., Sgorbini, B. and Rubiolo, P.** (2008). Headspace sampling of the volatile fraction of vegetable matrices. *J Chromatogr A* **1184**, 220-33.
- Bick, J. A. and Lange, B. M.** (2003). Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane. *Arch Biochem Biophys* **415**, 146-54.
- Bino, R. J., Ric de Vos, C. H., Lieberman, M., Hall, R. D., Bovy, A., Jonker, H. H., Tikunov, Y., Lommen, A., Moco, S. and Levin, I.** (2005). The light-hyperresponsive high pigment-2dg mutation of tomato: alterations in the fruit metabolome. *New Phytol* **166**, 427-38.
- Blanch, H. W., Adams, P. D., Andrews-Cramer, K. M., Frommer, W. B., Simmons, B. A. and Keasling, J. D.** (2008). Addressing the need for alternative transportation fuels: the Joint BioEnergy Institute. *ACS Chem Biol* **3**, 17-20.
- Bohlmann, J. and Keeling, C. I.** (2008). Terpenoid biomaterials. *Plant J* **54**, 656-69.
- Bohlmann, J., Meyer-Gauen, G. and Croteau, R.** (1998). Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proc Natl Acad Sci U S A* **95**, 4126-33.

- Bohlmann, J., Stauber, E. J., Krock, B., Oldham, N. J., Gershenzon, J. and Baldwin, I. T.** (2002). Gene expression of 5-epi-aristolochene synthase and formation of capsidiol in roots of *Nicotiana attenuata* and *N. sylvestris*. *Phytochemistry* **60**, 109-16.
- Bohlmann, J., Steele, C. L. and Croteau, R.** (1997). Monoterpene synthases from grand fir (*Abies grandis*). cDNA isolation, characterization, and functional expression of myrcene synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase. *J Biol Chem* **272**, 21784-92.
- Bouvier, F., Rahier, A. and Camara, B.** (2005). Biogenesis, molecular regulation and function of plant isoprenoids. *Prog Lipid Res* **44**, 357-429.
- Breitmaier, E.** (2006). Terpenes : flavors, fragrances, pharmaca, pheromones. Weinheim: Wiley-VCH.
- Brent, R. and Ptashne, M.** (1985). A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43**, 729-36.
- Bro, C., Regenberg, B., Forster, J. and Nielsen, J.** (2006). In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metab Eng* **8**, 102-11.
- Broach, J. R.** (1982). The yeast plasmid 2 mu circle. *Cell* **28**, 203-4.
- Broun, P. and Somerville, C.** (2001). Progress in plant metabolic engineering. *Proc Natl Acad Sci U S A* **98**, 8925-7.
- Buerstedde, J. M., Reynaud, C. A., Humphries, E. H., Olson, W., Ewert, D. L. and Weill, J. C.** (1990). Light chain gene conversion continues at high rate in an ALV-induced cell line. *Embo J* **9**, 921-7.

- Busch, T. and Kirschning, A.** (2008). Recent advances in the total synthesis of pharmaceutically relevant diterpenes. *Nat Prod Rep* **25**, 318-41.
- Butcher, R. A. and Schreiber, S. L.** (2004). Identification of Ald6p as the target of a class of small-molecule suppressors of FK506 and their use in network dissection. *Proc Natl Acad Sci U S A* **101**, 7868-73.
- Cakar, Z. P., Seker, U. O., Tamerler, C., Sonderegger, M. and Sauer, U.** (2005). Evolutionary engineering of multiple-stress resistant *Saccharomyces cerevisiae*. *FEMS Yeast Res* **5**, 569-78.
- Cane, D. E. and Bowser, T. E.** (1999). Trichodiene synthase: mechanism-based inhibition of a sesquiterpene cyclase. *Bioorg Med Chem Lett* **9**, 1127-32.
- Cannell, R. J. P.** (1998). Natural products isolation. Totowa, N.J.: Humana.
- Caruthers, J. M., Kang, I., Rynkiewicz, M. J., Cane, D. E. and Christianson, D. W.** (2000). Crystal structure determination of aristolochene synthase from the blue cheese mold, *Penicillium roqueforti*. *J Biol Chem* **275**, 25533-9.
- Chaudhuri, B., Steube, K. and Stephan, C.** (1992). The pro-region of the yeast prepro-alpha-factor is essential for membrane translocation of human insulin-like growth factor 1 in vivo. *Eur J Biochem* **206**, 793-800.
- Chen, K. S. and Gould, M. N.** (2004). Development of a universal gap repair vector for yeast-based screening of knockout rodents. *Biotechniques* **37**, 383-8.
- Christianson, D. W.** (2006). Structural biology and chemistry of the terpenoid cyclases. *Chem Rev* **106**, 3412-42.
- Chu, B. C. and Lee, H.** (2007). Genetic improvement of *Saccharomyces cerevisiae* for xylose fermentation. *Biotechnol Adv* **25**, 425-41.

- Colas, P. and Brent, R.** (1998). The impact of two-hybrid and related methods on biotechnology. *Trends Biotechnol* **16**, 355-63.
- Colby, S. M., Alonso, W. R., Katahira, E. J., McGarvey, D. J. and Croteau, R.** (1993). 4S-limonene synthase from the oil glands of spearmint (*Mentha spicata*). cDNA isolation, characterization, and bacterial expression of the catalytically active monoterpene cyclase. *J Biol Chem* **268**, 23016-24.
- Connolly, J. D. and Hill, R.** (1991). Dictionary of terpenoids. London: Chapman & Hall.
- Connolly, J. D. and Hill, R. A.** (2008). Triterpenoids. *Nat Prod Rep* **25**, 794-830.
- Crock, J., Wildung, M. and Croteau, R.** (1997). Isolation and bacterial expression of a sesquiterpene synthase cDNA clone from peppermint (*Mentha x piperita*, L.) that produces the aphid alarm pheromone (E)-beta-farnesene. *Proc Natl Acad Sci U S A* **94**, 12833-8.
- Croteau, R., Alonso, W. R., Koeppe, A. E. and Johnson, M. A.** (1994). Biosynthesis of monoterpenes: partial purification, characterization, and mechanism of action of 1,8-cineole synthase. *Arch Biochem Biophys* **309**, 184-92.
- Croteau, R., El-Bialy, H. and Dehal, S. S.** (1987). Metabolism of Monoterpenes : Metabolic Fate of (+)-Camphor in Sage (*Salvia officinalis*). *Plant Physiol* **84**, 643-648.
- Cuppett, S. L. and Hall, C. A., 3rd.** (1998). Antioxidant activity of the Labiatae. *Adv Food Nutr Res* **42**, 245-71.
- Czabany, T., Athenstaedt, K. and Daum, G.** (2007). Synthesis, storage and degradation of neutral lipids in yeast. *Biochim Biophys Acta* **1771**, 299-309.

- Daum, G., Wagner, A., Czabany, T. and Athenstaedt, K.** (2007). Dynamics of neutral lipid storage and mobilization in yeast. *Biochimie* **89**, 243-8.
- Davies, B. S. and Rine, J.** (2006). A role for sterol levels in oxygen sensing in *Saccharomyces cerevisiae*. *Genetics* **174**, 191-201.
- Dejong, J. M., Liu, Y., Bollon, A. P., Long, R. M., Jennewein, S., Williams, D. and Croteau, R. B.** (2006). Genetic engineering of taxol biosynthetic genes in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* **93**, 212-24.
- Dewick, P. M.** (2002). The biosynthesis of C5-C25 terpenoid compounds. *Nat Prod Rep* **19**, 181-222.
- Dietz, K. J., Jacob, S., Oelze, M. L., Laxa, M., Tognetti, V., de Miranda, S. M., Baier, M. and Finkemeier, I.** (2006). The function of peroxiredoxins in plant organelle redox metabolism. *J Exp Bot* **57**, 1697-709.
- Dimster-Denk, D., Thorsness, M. K. and Rine, J.** (1994). Feedback regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase in *Saccharomyces cerevisiae*. *Mol Biol Cell* **5**, 655-65.
- Duchen, M. R.** (2004). Roles of mitochondria in health and disease. *Diabetes* **53 Suppl 1**, S96-102.
- Duester, G.** (2008). Retinoic acid synthesis and signaling during early organogenesis. *Cell* **134**, 921-31.
- Edwards, R. and Gatehouse, J. A.** (1999). Secondary metabolism. In: *Plant biochemistry and molecular biology*. New York: John Wiley & Sons. 193-218.
- Eisenreich, W., Rohdich, F. and Bacher, A.** (2001). Deoxyxylulose phosphate pathway to terpenoids. *Trends Plant Sci* **6**, 78-84.

El Sayed, M. (2005). Studies on the biological and biochemical properties of a *Salvia fruticosa* isoflavone reductase homologue. Master thesis, MAICh, Greece.

El Tamer, M. K., Lucker, J., Bosch, D., Verhoeven, H. A., Verstappen, F. W., Schwab, W., van Tunen, A. J., Voragen, A. G., de Maagd, R. A. and Bouwmeester, H. J. (2003). Domain swapping of Citrus limon monoterpene synthases: impact on enzymatic activity and product specificity. *Arch Biochem Biophys* **411**, 196-203.

Engels, B., Dahm, P. and Jennewein, S. (2008). Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. *Metab Eng* **10**, 201-6.

Ennaifer, M. S. (2005). Environmental and ontogenic variation of terpene biosynthesis. *MAICh. Natural products and Biotechnology Department*.

Ernst, J. F. (1986). Improved secretion of heterologous proteins by *Saccharomyces cerevisiae*: effects of promoter substitution in alpha-factor fusions. *Dna* **5**, 483-91.

Evans, E., Sugawara, N., Haber, J. E. and Alani, E. (2000). The *Saccharomyces cerevisiae* Msh2 mismatch repair protein localizes to recombination intermediates in vivo. *Mol Cell* **5**, 789-99.

Facchini, P. J., Huber-Allanach, K. L. and Tari, L. W. (2000). Plant aromatic L-amino acid decarboxylases: evolution, biochemistry, regulation, and metabolic engineering applications. *Phytochemistry* **54**, 121-38.

Farhat, G. N., Affara, N. I. and Gali-Muhtasib, H. U. (2001). Seasonal changes in the composition of the essential oil extract of East Mediterranean sage (*Salvia libanotica*) and its toxicity in mice. *Toxicol* **39**, 1601-5.

Fields, S. and Sternglanz, R. (1994). The two-hybrid system: an assay for protein-protein interactions. *Trends Genet* **10**, 286-92.

- Fischer, C. R., Klein-Marcuschamer, D. and Stephanopoulos, G.** (2008). Selection and optimization of microbial hosts for biofuels production. *Metab Eng* **10**, 295-304.
- Ford, S. K. and Pringle, J. R.** (1991). Cellular morphogenesis in the *Saccharomyces cerevisiae* cell cycle: localization of the CDC11 gene product and the timing of events at the budding site. *Dev Genet* **12**, 281-92.
- Forkmann, G. and Martens, S.** (2001). Metabolic engineering and applications of flavonoids. *Curr Opin Biotechnol* **12**, 155-60.
- Fortman, J. L., Chhabra, S., Mukhopadhyay, A., Chou, H., Lee, T. S., Steen, E. and Keasling, J. D.** (2008). Biofuel alternatives to ethanol: pumping the microbial well. *Trends Biotechnol* **26**, 375-81.
- Fox, L. R. and Morrow, P. A.** (1981). Specialization: Species Property or Local Phenomenon? *Science* **211**, 887-893.
- Fraiser, L. H., Kanekal, S. and Kehrer, J. P.** (1991). Cyclophosphamide toxicity. Characterising and avoiding the problem. *Drugs* **42**, 781-95.
- Friesen, J. A. and Rodwell, V. W.** (2004). The 3-hydroxy-3-methylglutaryl coenzyme-A (HMG-CoA) reductases. *Genome Biol* **5**, 248.
- Furr, H. C.** (2004). Analysis of retinoids and carotenoids: problems resolved and unsolved. *J Nutr* **134**, 281S-285S.
- Gaber, R. F., Copple, D. M., Kennedy, B. K., Vidal, M. and Bard, M.** (1989). The yeast gene ERG6 is required for normal membrane function but is not essential for biosynthesis of the cell-cycle-sparking sterol. *Mol Cell Biol* **9**, 3447-56.

- Gardner, R. G. and Hampton, R. Y.** (1999). A 'distributed degron' allows regulated entry into the ER degradation pathway. *Embo J* **18**, 5994-6004.
- Gelhaye, E., Rouhier, N., Navrot, N. and Jacquot, J. P.** (2005). The plant thioredoxin system. *Cell Mol Life Sci* **62**, 24-35.
- Gelvin, S. B.** (2006). Agrobacterium transformation of Arabidopsis thaliana roots: a quantitative assay. *Methods Mol Biol* **343**, 105-13.
- Gelvin, S. B.** (2006). Agrobacterium virulence gene induction. *Methods Mol Biol* **343**, 77-84.
- Gersbach, P. V.** (2002). The essential oil secretory structures of Prostanthera ovalifolia (Lamiaceae). *Ann Bot (Lond)* **89**, 255-60.
- Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Luca-Danila, A., Anderson, K., Andre, B. et al.** (2002). Functional profiling of the Saccharomyces cerevisiae genome. *Nature* **418**, 387-91.
- Gietz, R. D. and Schiestl, R. H.** (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* **2**, 31-4.
- Gietz, R. D. and Sugino, A.** (1988). New yeast-Escherichia coli shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites. *Gene* **74**, 527-34.
- Gil, G., Faust, J. R., Chin, D. J., Goldstein, J. L. and Brown, M. S.** (1985). Membrane-bound domain of HMG CoA reductase is required for sterol-enhanced degradation of the enzyme. *Cell* **41**, 249-58.
- Glick, B. S. and Pon, L. A.** (1995). Isolation of highly purified mitochondria from Saccharomyces cerevisiae. *Methods Enzymol* **260**, 213-23.

- Goldstein, J. L. and Brown, M. S.** (1990). Regulation of the mevalonate pathway. *Nature* **343**, 425-30.
- Goldstein, J. L., DeBose-Boyd, R. A. and Brown, M. S.** (2006). Protein sensors for membrane sterols. *Cell* **124**, 35-46.
- Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Quorollo, B., Goldman, B. S., Cao, Y., Askenazi, M., Halling, C. et al.** (2001). Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science* **294**, 2323-8.
- Grabinska, K. and Palamarczyk, G.** (2002). Dolichol biosynthesis in the yeast *Saccharomyces cerevisiae*: an insight into the regulatory role of farnesyl diphosphate synthase. *FEMS Yeast Res* **2**, 259-65.
- Grochowski, L. L., Xu, H. and White, R. H.** (2006). *Methanocaldococcus jannaschii* uses a modified mevalonate pathway for biosynthesis of isopentenyl diphosphate. *J Bacteriol* **188**, 3192-8.
- Guo, X. and Mitra, S.** (2000). On-line membrane extraction liquid chromatography for monitoring semi-volatile organics in aqueous matrices. *J Chromatogr A* **904**, 189-96.
- Guo, Z., Severson, R. F. and Wagner, G. J.** (1994). Biosynthesis of the diterpene cis-abienol in cell-free extracts of tobacco trichomes. *Arch Biochem Biophys* **308**, 103-8.
- Haber, J. E.** (1998). Mating-type gene switching in *Saccharomyces cerevisiae*. *Annu Rev Genet* **32**, 561-99.
- Hahn, F. M., Eubanks, L. M., Testa, C. A., Blagg, B. S., Baker, J. A. and Poulter, C. D.** (2001). 1-Deoxy-D-xylulose 5-phosphate synthase, the gene product of open reading frame (ORF) 2816 and ORF 2895 in *Rhodobacter capsulatus*. *J Bacteriol* **183**, 1-11.

- Hamano, Y., Dairi, T., Yamamoto, M., Kuzuyama, T., Itoh, N. and Seto, H.** (2002). Growth-phase dependent expression of the mevalonate pathway in a terpenoid antibiotic-producing *Streptomyces* strain. *Biosci Biotechnol Biochem* **66**, 808-19.
- Hampton, R., Dimster-Denk, D. and Rine, J.** (1996). The biology of HMG-CoA reductase: the pros of contra-regulation. *Trends Biochem Sci* **21**, 140-5.
- Hampton, R. Y.** (1998). Genetic analysis of hydroxymethylglutaryl-coenzyme A reductase regulated degradation. *Curr Opin Lipidol* **9**, 93-7.
- Hampton, R. Y.** (2002). Proteolysis and sterol regulation. *Annu Rev Cell Dev Biol* **18**, 345-78.
- Hampton, R. Y. and Bhakta, H.** (1997). Ubiquitin-mediated regulation of 3-hydroxy-3-methylglutaryl-CoA reductase. *Proc Natl Acad Sci U S A* **94**, 12944-8.
- Hampton, R. Y. and Rine, J.** (1994). Regulated degradation of HMG-CoA reductase, an integral membrane protein of the endoplasmic reticulum, in yeast. *J Cell Biol* **125**, 299-312.
- Harbison, C. T., Gordon, D. B., Lee, T. I., Rinaldi, N. J., Macisaac, K. D., Danford, T. W., Hannett, N. M., Tagne, J. B., Reynolds, D. B., Yoo, J. et al.** (2004). Transcriptional regulatory code of a eukaryotic genome. *Nature* **431**, 99-104.
- Harborne, J. B.** (1990). Role of secondary metabolites in chemical defence mechanisms in plants. *Ciba Found Symp* **154**, 126-34; discussion 135-9.
- Hayes, F.** (2003). Transposon-based strategies for microbial functional genomics and proteomics. *Annu Rev Genet* **37**, 3-29.
- Heiderpriem, R. W., Livant, P. D., Parish, E. J., Barbuch, R. J., Broaddus, M. G. and Bard, M.** (1992). A simple method for the isolation of zymosterol from a sterol mutant of *Saccharomyces cerevisiae*. *J Steroid Biochem Mol Biol* **43**, 741-3.

- Hemmerlin, A., Hoeffler, J. F., Meyer, O., Tritsch, D., Kagan, I. A., Grosdemange-Billiard, C., Rohmer, M. and Bach, T. J.** (2003). Cross-talk between the cytosolic mevalonate and the plastidial methylerythritol phosphate pathways in tobacco bright yellow-2 cells. *J Biol Chem* **278**, 26666-76.
- Herzberg, K., Bashkirov, V. I., Rolfsmeier, M., Haghazari, E., McDonald, W. H., Anderson, S., Bashkirova, E. V., Yates, J. R., 3rd and Heyer, W. D.** (2006). Phosphorylation of Rad55 on serines 2, 8, and 14 is required for efficient homologous recombination in the recovery of stalled replication forks. *Mol Cell Biol* **26**, 8396-409.
- Heyer, W. D., Li, X., Rolfsmeier, M. and Zhang, X. P.** (2006). Rad54: the Swiss Army knife of homologous recombination? *Nucleic Acids Res* **34**, 4115-25.
- Hezari, M. and Croteau, R.** (1997). Taxol biosynthesis: an update. *Planta Med* **63**, 291-5.
- Hitzeman, R. A., Hagie, F. E., Levine, H. L., Goeddel, D. V., Ammerer, G. and Hall, B. D.** (1981). Expression of a human gene for interferon in yeast. *Nature* **293**, 717-22.
- Hoekema, A., Roelvink, P. W., Hooykaas, P. J. and Schilperoort, R. A.** (1984). Delivery of T-DNA from the *Agrobacterium tumefaciens* chromosome into plant cells. *Embo J* **3**, 2485-2490.
- Hoffmann, G. R.** (1994). Induction of genetic recombination: consequences and model systems. *Environ Mol Mutagen* **23 Suppl 24**, 59-66.
- Holmes, A. M. and Haber, J. E.** (1999). Double-strand break repair in yeast requires both leading and lagging strand DNA polymerases. *Cell* **96**, 415-24.
- Holmgren, A.** (2000). Antioxidant function of thioredoxin and glutaredoxin systems. *Antioxid Redox Signal* **2**, 811-20.

- Holstein, S. A. and Hohl, R. J.** (2004). Isoprenoids: remarkable diversity of form and function. *Lipids* **39**, 293-309.
- Hoog, J. L. and Antony, C.** (2007). Whole-cell investigation of microtubule cytoskeleton architecture by electron tomography. *Methods Cell Biol* **79**, 145-67.
- Hope, I. A. and Struhl, K.** (1986). Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* **46**, 885-94.
- Houston, P. L. and Broach, J. R.** (2006). The dynamics of homologous pairing during mating type interconversion in budding yeast. *PLoS Genet* **2**, e98.
- Hua, S. B., Qiu, M., Chan, E., Zhu, L. and Luo, Y.** (1997). Minimum length of sequence homology required for in vivo cloning by homologous recombination in yeast. *Plasmid* **38**, 91-6.
- Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S. and O'Shea, E. K.** (2003). Global analysis of protein localization in budding yeast. *Nature* **425**, 686-91.
- Iijima, Y., Davidovich-Rikanati, R., Fridman, E., Gang, D. R., Bar, E., Lewinsohn, E. and Pichersky, E.** (2004). The biochemical and molecular basis for the divergent patterns in the biosynthesis of terpenes and phenylpropenes in the peltate glands of three cultivars of basil. *Plant Physiol* **136**, 3724-36.
- Itoh, H., Ueguchi-Tanaka, M. and Matsuoka, M.** (2008). Molecular biology of gibberellins signaling in higher plants. *Int Rev Cell Mol Biol* **268**, 191-221.
- Jackson, B. E., Hart-Wells, E. A. and Matsuda, S. P.** (2003). Metabolic engineering to produce sesquiterpenes in yeast. *Org Lett* **5**, 1629-32.
- Jeffries, T. W.** (2006). Engineering yeasts for xylose metabolism. *Curr Opin Biotechnol* **17**, 320-6.

Jenke-Kodama, H., Muller, R. and Dittmann, E. (2008). Evolutionary mechanisms underlying secondary metabolite diversity. *Prog Drug Res* **65**, 119, 121-40.

Jennewein, S. and Croteau, R. (2001). Taxol: biosynthesis, molecular genetics, and biotechnological applications. *Appl Microbiol Biotechnol* **57**, 13-9.

Jennewein, S., Park, H., DeJong, J. M., Long, R. M., Bollon, A. P. and Croteau, R. B. (2005). Coexpression in yeast of Taxus cytochrome P450 reductase with cytochrome P450 oxygenases involved in Taxol biosynthesis. *Biotechnol Bioeng* **89**, 588-98.

Jennewein, S., Wildung, M. R., Chau, M., Walker, K. and Croteau, R. (2004). Random sequencing of an induced Taxus cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proc Natl Acad Sci U S A* **101**, 9149-54.

Jensen-Pergakes, K. L., Kennedy, M. A., Lees, N. D., Barbuch, R., Koegel, C. and Bard, M. (1998). Sequencing, disruption, and characterization of the *Candida albicans* sterol methyltransferase (ERG6) gene: drug susceptibility studies in *erg6* mutants. *Antimicrob Agents Chemother* **42**, 1160-7.

Jewett, M. C., Hofmann, G. and Nielsen, J. (2006). Fungal metabolite analysis in genomics and phenomics. *Curr Opin Biotechnol* **17**, 191-7.

Jiang, H., Wood, K. V. and Morgan, J. A. (2005). Metabolic engineering of the phenylpropanoid pathway in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **71**, 2962-9.

Jin, Y. S. and Stephanopoulos, G. (2007). Multi-dimensional gene target search for improving lycopene biosynthesis in *Escherichia coli*. *Metab Eng* **9**, 337-47.

- Joly, J. C., Leung, W. S. and Swartz, J. R.** (1998). Overexpression of Escherichia coli oxidoreductases increases recombinant insulin-like growth factor-I accumulation. *Proc Natl Acad Sci U S A* **95**, 2773-7.
- Julsing, M. K., Koulman, A., Woerdenbag, H. J., Quax, W. J. and Kayser, O.** (2006). Combinatorial biosynthesis of medicinal plant secondary metabolites. *Biomol Eng* **23**, 265-79.
- Kampranis, S. C., Ioannidis, D., Purvis, A., Mahrez, W., Ninga, E., Katerelos, N. A., Anssour, S., Dunwell, J. M., Degenhardt, J., Makris, A. M. et al.** (2007). Rational conversion of substrate and product specificity in a Salvia monoterpene synthase: structural insights into the evolution of terpene synthase function. *Plant Cell* **19**, 1994-2005.
- Kang, M. J., Lee, Y. M., Yoon, S. H., Kim, J. H., Ock, S. W., Jung, K. H., Shin, Y. C., Keasling, J. D. and Kim, S. W.** (2005). Identification of genes affecting lycopene accumulation in Escherichia coli using a shot-gun method. *Biotechnol Bioeng* **91**, 636-42.
- Karousou, R., Balta, M., Hanlidou, E. and Kokkini, S.** (2007). "Mints", smells and traditional uses in Thessaloniki (Greece) and other Mediterranean countries. *J Ethnopharmacol* **109**, 248-57.
- Karousou, R., Koureas, D. N. and Kokkini, S.** (2005). Essential oil composition is related to the natural habitats: Coridothymus capitatus and Satureja thymbra in NATURA 2000 sites of Crete. *Phytochemistry* **66**, 2668-73.
- Karst, F. and Lacroute, F.** (1977). Ergosterol biosynthesis in Saccharomyces cerevisiae: mutants deficient in the early steps of the pathway. *Mol Gen Genet* **154**, 269-77.
- Keegan, L., Gill, G. and Ptashne, M.** (1986). Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science* **231**, 699-704.

- Keeling, C. I., Weisshaar, S., Lin, R. P. and Bohlmann, J.** (2008). Functional plasticity of paralogous diterpene synthases involved in conifer defense. *Proc Natl Acad Sci U S A* **105**, 1085-90.
- Kehrer, J. P. and Biswal, S. S.** (2000). The molecular effects of acrolein. *Toxicol Sci* **57**, 6-15.
- Kennedy, M. A., Barbuch, R. and Bard, M.** (1999). Transcriptional regulation of the squalene synthase gene (ERG9) in the yeast *Saccharomyces cerevisiae*. *Biochim Biophys Acta* **1445**, 110-22.
- Ketchum, R. E., Horiguchi, T., Qiu, D., Williams, R. M. and Croteau, R. B.** (2007). Administering cultured *Taxus* cells with early precursors reveals bifurcations in the taxoid biosynthetic pathway. *Phytochemistry* **68**, 335-41.
- Khosla, C. and Keasling, J. D.** (2003). Metabolic engineering for drug discovery and development. *Nat Rev Drug Discov* **2**, 1019-25.
- Kind, T. and Fiehn, O.** (2007). Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics* **8**, 105.
- Kingston, D. G.** (2000). Recent advances in the chemistry of taxol. *J Nat Prod* **63**, 726-34.
- Kintzios, S. E.** (2000). Sage : the genus *Salvia*. Australia: Harwood Academic ; Abingdon : Marston.
- Kjeldsen, T.** (2000). Yeast secretory expression of insulin precursors. *Appl Microbiol Biotechnol* **54**, 277-86.
- Kleinhans, F. W., Lees, N. D., Bard, M., Haak, R. A. and Woods, R. A.** (1979). ESR determinations of membrane permeability in a yeast sterol mutant. *Chem Phys Lipids* **23**, 143-54.

Koehn, F. E. and Carter, G. T. (2005). The evolving role of natural products in drug discovery. *Nat Rev Drug Discov* **4**, 206-20.

Konig, J., Baier, M., Horling, F., Kahmann, U., Harris, G., Schurmann, P. and Dietz, K. J. (2002). The plant-specific function of 2-Cys peroxiredoxin-mediated detoxification of peroxides in the redox-hierarchy of photosynthetic electron flux. *Proc Natl Acad Sci U S A* **99**, 5738-43.

Kumar, A. and Snyder, M. (2001). Emerging technologies in yeast genomics. *Nat Rev Genet* **2**, 302-12.

Lange, B. M. and Croteau, R. (1999). Genetic engineering of essential oil production in mint. *Curr Opin Plant Biol* **2**, 139-44.

Lange, B. M., Wildung, M. R., Stauber, E. J., Sanchez, C., Pouchnik, D. and Croteau, R. (2000). Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes. *Proc Natl Acad Sci U S A* **97**, 2934-9.

Lashkari, D. A., DeRisi, J. L., McCusker, J. H., Namath, A. F., Gentile, C., Hwang, S. Y., Brown, P. O. and Davis, R. W. (1997). Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc Natl Acad Sci U S A* **94**, 13057-62.

Laule, O., Furholz, A., Chang, H. S., Zhu, T., Wang, X., Heifetz, P. B., Grisse, W. and Lange, M. (2003). Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* **100**, 6866-71.

Lee, P. C., Petri, R., Mijts, B. N., Watts, K. T. and Schmidt-Dannert, C. (2005). Directed evolution of *Escherichia coli* farnesyl diphosphate synthase (IspA) reveals novel structural determinants of chain length specificity. *Metab Eng* **7**, 18-26.

Lee, P. C., Salomon, C., Mijts, B. and Schmidt-Dannert, C. (2008). Biosynthesis of ubiquinone compounds with conjugated prenyl side chains. *Appl Environ Microbiol* **74**, 6908-17.

- Lee, P. C. and Schmidt-Dannert, C.** (2002). Metabolic engineering towards biotechnological production of carotenoids in microorganisms. *Appl Microbiol Biotechnol* **60**, 1-11.
- Lee, S. K., Chou, H., Ham, T. S., Lee, T. S. and Keasling, J. D.** (2008). Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. *Curr Opin Biotechnol* **19**, 556-63.
- Lees, N. D., Bard, M. and Kirsch, D. R.** (1999). Biochemistry and molecular biology of sterol synthesis in *Saccharomyces cerevisiae*. *Crit Rev Biochem Mol Biol* **34**, 33-47.
- Legrain, P. and Selig, L.** (2000). Genome-wide protein interaction maps using two-hybrid systems. *FEBS Lett* **480**, 32-6.
- Lehmann, T., Sager, F. and Brenneisen, R.** (1997). Excretion of cannabinoids in urine after ingestion of cannabis seed oil. *J Anal Toxicol* **21**, 373-5.
- Leonard, E., Yan, Y., Lim, K. H. and Koffas, M. A.** (2005). Investigation of two distinct flavone synthases for plant-specific flavone biosynthesis in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **71**, 8241-8.
- Lesburg, C. A., Caruthers, J. M., Paschall, C. M. and Christianson, D. W.** (1998). Managing and manipulating carbocations in biology: terpenoid cyclase structure and mechanism. *Curr Opin Struct Biol* **8**, 695-703.
- Lesburg, C. A., Zhai, G., Cane, D. E. and Christianson, D. W.** (1997). Crystal structure of pentalenene synthase: mechanistic insights on terpenoid cyclization reactions in biology. *Science* **277**, 1820-4.
- Lindahl, A. L., Olsson, M. E., Mercke, P., Tollbom, O., Schelin, J., Brodelius, M. and Brodelius, P. E.** (2006). Production of the artemisinin precursor amorpha-4,11-diene by engineered *Saccharomyces cerevisiae*. *Biotechnol Lett* **28**, 571-80.

- Lionis, C., Faresjo, A., Skoula, M., Kapsokefalou, M. and Faresjo, T.** (1998). Antioxidant effects of herbs in Crete. *Lancet* **352**, 1987-8.
- Little, D. B. and Croteau, R. B.** (2002). Alteration of product formation by directed mutagenesis and truncation of the multiple-product sesquiterpene synthases delta-selinene synthase and gamma-humulene synthase. *Arch Biochem Biophys* **402**, 120-35.
- Lommen, A., van der Weg, G., van Engelen, M. C., Bor, G., Hoogenboom, L. A. and Nielen, M. W.** (2007). An untargeted metabolomics approach to contaminant analysis: pinpointing potential unknown compounds. *Anal Chim Acta* **584**, 43-9.
- Lorence, A. and Verpoorte, R.** (2004). Gene transfer and expression in plants. *Methods Mol Biol* **267**, 329-50.
- Lotan, T. and Hirschberg, J.** (1995). Cloning and expression in Escherichia coli of the gene encoding beta-C-4-oxygenase, that converts beta-carotene to the ketocarotenoid canthaxanthin in Haematococcus pluvialis. *FEBS Lett* **364**, 125-8.
- Lucker, J., El Tamer, M. K., Schwab, W., Verstappen, F. W., van der Plas, L. H., Bouwmeester, H. J. and Verhoeven, H. A.** (2002). Monoterpene biosynthesis in lemon (Citrus limon). cDNA isolation and functional analysis of four monoterpene synthases. *Eur J Biochem* **269**, 3160-71.
- Luttgen, H., Rohdich, F., Herz, S., Wungsintaweekul, J., Hecht, S., Schuhr, C. A., Fellermeier, M., Sagner, S., Zenk, M. H., Bacher, A. et al.** (2000). Biosynthesis of terpenoids: YchB protein of Escherichia coli phosphorylates the 2-hydroxy group of 4-diphosphocytidyl-2C-methyl-D-erythritol. *Proc Natl Acad Sci U S A* **97**, 1062-7.
- Lynd, L. R., van Zyl, W. H., McBride, J. E. and Laser, M.** (2005). Consolidated bioprocessing of cellulosic biomass: an update. *Curr Opin Biotechnol* **16**, 577-83.
- Ma, J. and Ptashne, M.** (1987). A new class of yeast transcriptional activators. *Cell* **51**, 113-9.

Ma, J. and Ptashne, M. (1988). Converting a eukaryotic transcriptional inhibitor into an activator. *Cell* **55**, 443-6.

Mahmoud, S. S. and Croteau, R. B. (2003). Menthofuran regulates essential oil biosynthesis in peppermint by controlling a downstream monoterpene reductase. *Proc Natl Acad Sci U S A* **100**, 14481-6.

Mahrez, W. (2007). Structural, functional, and mechanistic studies of terpene synthases. Master thesis, MAICh, Greece.

Mann, J. (1987). Secondary metabolism. Oxford: Clarendon Press.

Mansour, M. P., Volkman, J. K. and Blackburn, S. I. (2003). The effect of growth phase on the lipid class, fatty acid and sterol composition in the marine dinoflagellate, *Gymnodinium* sp. in batch culture. *Phytochemistry* **63**, 145-53.

Maraz, A. (2002). From yeast genetics to biotechnology. *Acta Microbiol Immunol Hung* **49**, 483-91.

Marrero, P. F., Poulter, C. D. and Edwards, P. A. (1992). Effects of site-directed mutagenesis of the highly conserved aspartate residues in domain II of farnesyl diphosphate synthase activity. *J Biol Chem* **267**, 21873-8.

Martin, D. M., Faldt, J. and Bohlmann, J. (2004). Functional characterization of nine Norway Spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. *Plant Physiol* **135**, 1908-27.

Martin, V. J., Pitera, D. J., Withers, S. T., Newman, J. D. and Keasling, J. D. (2003). Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat Biotechnol* **21**, 796-802.

- Mas, S., Villas-Boas, S. G., Hansen, M. E., Akesson, M. and Nielsen, J.** (2007). A comparison of direct infusion MS and GC-MS for metabolic footprinting of yeast mutants. *Biotechnol Bioeng* **96**, 1014-22.
- McCaskill, D. and Croteau, R.** (1999). Strategies for bioengineering the development and metabolism of glandular tissues in plants. *Nat Biotechnol* **17**, 31-6.
- McGarvey, D. J. and Croteau, R.** (1995). Terpenoid metabolism. *Plant Cell* **7**, 1015-26.
- McGlynn, P. and Lloyd, R. G.** (2002). Recombinational repair and restart of damaged replication forks. *Nat Rev Mol Cell Biol* **3**, 859-70.
- Meganathan, R.** (2001). Biosynthesis of menaquinone (vitamin K2) and ubiquinone (coenzyme Q): a perspective on enzymatic mechanisms. *Vitam Horm* **61**, 173-218.
- Mercke, P., Bengtsson, M., Bouwmeester, H. J., Posthumus, M. A. and Brodelius, P. E.** (2000). Molecular cloning, expression, and characterization of amorpho-4,11-diene synthase, a key enzyme of artemisinin biosynthesis in *Artemisia annua* L. *Arch Biochem Biophys* **381**, 173-80.
- Mijts, B. N., Lee, P. C. and Schmidt-Dannert, C.** (2004). Engineering carotenoid biosynthetic pathways. *Methods Enzymol* **388**, 315-29.
- Mills, G. A. and Walker, V.** (2001). Headspace solid-phase microextraction profiling of volatile compounds in urine: application to metabolic investigations. *J Chromatogr B Biomed Sci Appl* **753**, 259-68.
- Miura, Y., Kondo, K., Shimada, H., Saito, T., Nakamura, K. and Misawa, N.** (1998). Production of lycopene by the food yeast, *Candida utilis* that does not naturally synthesize carotenoid. *Biotechnol Bioeng* **58**, 306-8.

Morimoto, M., Cantrell, C. L., Libous-Bailey, L. and Duke, S. O. (2008). Phytotoxicity of constituents of glandular trichomes and the leaf surface of camphorweed, *Heterotheca subaxillaris*. *Phytochemistry*.

Mukhopadhyay, A., Redding, A. M., Rutherford, B. J. and Keasling, J. D. (2008). Importance of systems biology in engineering microbes for biofuel production. *Curr Opin Biotechnol* **19**, 228-34.

Mustacchi, R., Hohmann, S. and Nielsen, J. (2006). Yeast systems biology to unravel the network of life. *Yeast* **23**, 227-38.

Newman, D. J., Cragg, G. M. and Snader, K. M. (2003). Natural products as sources of new drugs over the period 1981-2002. *J Nat Prod* **66**, 1022-37.

Nielsen, J. and Jewett, M. C. (2008). Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*. *FEMS Yeast Res* **8**, 122-31.

O'Brien, P. J. and Herschlag, D. (1999). Catalytic promiscuity and the evolution of new enzymatic activities. *Chem Biol* **6**, R91-R105.

Oehlen, L. and Cross, F. R. (1998). The mating factor response pathway regulates transcription of *TEC1*, a gene involved in pseudohyphal differentiation of *Saccharomyces cerevisiae*. *FEBS Lett* **429**, 83-8.

Ohloff, G. (1990). *Riechstoffe und Geruchssinn : die molekulare Welt der Düfte ; mit 16 Tabellen*. Berlin {[u.a.]}: Springer.

Ohnuma, S., Nakazawa, T., Hemmi, H., Hallberg, A. M., Koyama, T., Ogura, K. and Nishino, T. (1996). Conversion from farnesyl diphosphate synthase to geranylgeranyl diphosphate synthase by random chemical mutagenesis. *J Biol Chem* **271**, 10087-95.

- Ohnuma, S., Narita, K., Nakazawa, T., Ishida, C., Takeuchi, Y., Ohto, C. and Nishino, T.** (1996). A role of the amino acid residue located on the fifth position before the first aspartate-rich motif of farnesyl diphosphate synthase on determination of the final product. *J Biol Chem* **271**, 30748-54.
- Oliver, S. G., van der Aart, Q. J., Agostoni-Carbone, M. L., Aigle, M., Alberghina, L., Alexandraki, D., Antoine, G., Anwar, R., Ballesta, J. P., Benit, P. et al.** (1992). The complete DNA sequence of yeast chromosome III. *Nature* **357**, 38-46.
- Olsson, L. and Nielsen, J.** (2000). The role of metabolic engineering in the improvement of *Saccharomyces cerevisiae*: utilization of industrial media. *Enzyme Microb Technol* **26**, 785-792.
- O'Neil, M. J. and Merck & Co.** (2001). The Merck index : an encyclopedia of chemicals, drugs, and biologicals. xvi, 1818, [746] p.
- Opitz, S., Kunert, G. and Gershenzon, J.** (2008). Increased terpenoid accumulation in cotton (*Gossypium hirsutum*) foliage is a general wound response. *J Chem Ecol* **34**, 508-22.
- Ostergaard, S., Olsson, L. and Nielsen, J.** (2000). Metabolic engineering of *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* **64**, 34-50.
- Pan, X., Yuan, D. S., Xiang, D., Wang, X., Sookhai-Mahadeo, S., Bader, J. S., Hieter, P., Spencer, F. and Boeke, J. D.** (2004). A robust toolkit for functional profiling of the yeast genome. *Mol Cell* **16**, 487-96.
- Paques, F. and Haber, J. E.** (1999). Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* **63**, 349-404.
- Parks, L. W. and Casey, W. M.** (1995). Physiological implications of sterol biosynthesis in yeast. *Annu Rev Microbiol* **49**, 95-116.

- Parks, L. W., McLean-Bowen, C., Bottema, C. K., Taylor, F. R., Gonzales, R., Jensen, B. W. and Ramp, J. R.** (1982). Aspects of sterol metabolism in the yeast *Saccharomyces cerevisiae* and in *Phytophthora*. *Lipids* **17**, 187-96.
- Pauli, G. F., Pro, S. M. and Friesen, J. B.** (2008). Countercurrent separation of natural products. *J Nat Prod* **71**, 1489-508.
- Pelham, H. R.** (2001). Traffic through the Golgi apparatus. *J Cell Biol* **155**, 1099-101.
- Perfumi, M., Arnold, N. and Tacconi, R.** (1991). Hypoglycemic activity of *Salvia fruticosa* Mill. from Cyprus. *J Ethnopharmacol* **34**, 135-40.
- Peters, R. J. and Croteau, R. B.** (2003). Alternative termination chemistries utilized by monoterpene cyclases: chimeric analysis of bornyl diphosphate, 1,8-cineole, and sabinene synthases. *Arch Biochem Biophys* **417**, 203-11.
- Petranovic, D. and Nielsen, J.** (2008). Can yeast systems biology contribute to the understanding of human disease? *Trends Biotechnol* **26**, 584-90.
- Polakowski, T., Bastl, R., Stahl, U. and Lang, C.** (1999). Enhanced sterol-acyl transferase activity promotes sterol accumulation in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* **53**, 30-5.
- Polakowski, T., Stahl, U. and Lang, C.** (1998). Overexpression of a cytosolic hydroxymethylglutaryl-CoA reductase leads to squalene accumulation in yeast. *Appl Microbiol Biotechnol* **49**, 66-71.
- Raal, A., Orav, A. and Arak, E.** (2007). Composition of the essential oil of *Salvia officinalis* L. from various European countries. *Nat Prod Res* **21**, 406-11.
- Rajakumari, S., Grillitsch, K. and Daum, G.** (2008). Synthesis and turnover of non-polar lipids in yeast. *Prog Lipid Res* **47**, 157-71.

- Ralston, L., Subramanian, S., Matsuno, M. and Yu, O.** (2005). Partial reconstruction of flavonoid and isoflavonoid biosynthesis in yeast using soybean type I and type II chalcone isomerases. *Plant Physiol* **137**, 1375-88.
- Rao, A. V. and Agarwal, S.** (2000). Role of antioxidant lycopene in cancer and heart disease. *J Am Coll Nutr* **19**, 563-9.
- Rao, A. V. and Rao, L. G.** (2007). Carotenoids and human health. *Pharmacol Res* **55**, 207-16.
- Rasmann, S., Kollner, T. G., Degenhardt, J., Hiltbold, I., Toepfer, S., Kuhlmann, U., Gershenzon, J. and Turlings, T. C.** (2005). Recruitment of entomopathogenic nematodes by insect-damaged maize roots. *Nature* **434**, 732-7.
- Ratushny, V. and Golemis, E.** (2008). Resolving the network of cell signaling pathways using the evolving yeast two-hybrid system. *Biotechniques* **44**, 655-62.
- Ravn, M. M., Coates, R. M., Flory, J. E., Peters, R. J. and Croteau, R.** (2000). Stereochemistry of the cyclization-rearrangement of (+)-copalyl diphosphate to (-)-abietadiene catalyzed by recombinant abietadiene synthase from *Abies grandis*. *Org Lett* **2**, 573-6.
- Reiling, K. K., Yoshikuni, Y., Martin, V. J., Newman, J., Bohlmann, J. and Keasling, J. D.** (2004). Mono and diterpene production in *Escherichia coli*. *Biotechnol Bioeng* **87**, 200-12.
- Rivera, D., Obon, C., Inocencio, C., Heinrich, M., Verde, A., Fajardo, J. and Llorach, R.** (2005). The ethnobotanical study of local Mediterranean food plants as medicinal resources in Southern Spain. *J Physiol Pharmacol* **56 Suppl 1**, 97-114.
- Ro, D. K. and Douglas, C. J.** (2004). Reconstitution of the entry point of plant phenylpropanoid metabolism in yeast (*Saccharomyces cerevisiae*): implications for control of metabolic flux into the phenylpropanoid pathway. *J Biol Chem* **279**, 2600-7.

Ro, D. K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S., Kirby, J. et al. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940-3.

Robinson, G. W., Tsay, Y. H., Kienzle, B. K., Smith-Monroy, C. A. and Bishop, R. W. (1993). Conservation between human and fungal squalene synthetases: similarities in structure, function, and regulation. *Mol Cell Biol* **13**, 2706-17.

Rodriguez-Concepcion, M. (2004). The MEP pathway: a new target for the development of herbicides, antibiotics and antimalarial drugs. *Curr Pharm Des* **10**, 2391-400.

Rodriguez-Concepcion, M. and Boronat, A. (2002). Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. *Plant Physiol* **130**, 1079-89.

Rohdich, F., Hecht, S., Gartner, K., Adam, P., Krieger, C., Amslinger, S., Arigoni, D., Bacher, A. and Eisenreich, W. (2002). Studies on the nonmevalonate terpene biosynthetic pathway: metabolic role of IspH (LytB) protein. *Proc Natl Acad Sci U S A* **99**, 1158-63.

Rohdich, F., Wungsintaweekul, J., Fellermeier, M., Sagner, S., Herz, S., Kis, K., Eisenreich, W., Bacher, A. and Zenk, M. H. (1999). Cytidine 5'-triphosphate-dependent biosynthesis of isoprenoids: YgbP protein of *Escherichia coli* catalyzes the formation of 4-diphosphocytidyl-2-C-methylerythritol. *Proc Natl Acad Sci U S A* **96**, 11758-63.

Rohmer, M., Knani, M., Simonin, P., Sutter, B. and Sahn, H. (1993). Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate. *Biochem J* **295** (Pt 2), 517-24.

Rontein, D., Onillon, S., Herbette, G., Lesot, A., Werck-Reichhart, D., Sallaud, C. and Tissier, A. (2008). CYP725A4 from yew catalyzes complex structural rearrangement of taxadiene into the cyclic ether 5(12)-oxa-3(11)-cyclotaxane. *J Biol Chem* **283**, 6067-75.

- Ross, S. J., Findlay, V. J., Malakasi, P. and Morgan, B. A.** (2000). Thioredoxin peroxidase is required for the transcriptional response to oxidative stress in budding yeast. *Mol Biol Cell* **11**, 2631-42.
- Roth, R. J. and Acton, N.** (1989). A simple conversion of artemisinic acid into artemisinin. *J Nat Prod* **52**, 1183-5.
- Rouhier, N. and Jacquot, J. P.** (2005). The plant multigenic family of thiol peroxidases. *Free Radic Biol Med* **38**, 1413-21.
- Runquist, M., Ericsson, J., Thelin, A., Chojnacki, T. and Dallner, G.** (1992). Biosynthesis of trans,trans,trans-geranylgeranyl diphosphate by the cytosolic fraction from rat tissues. *Biochem Biophys Res Commun* **186**, 157-65.
- Ruzicka, L.** (1953). The isoprene rule and the biogenesis of terpenic compounds. *Experientia* **9**, 357-67.
- Rynkiewicz, M. J., Cane, D. E. and Christianson, D. W.** (2001). Structure of trichodiene synthase from *Fusarium sporotrichioides* provides mechanistic inferences on the terpene cyclization cascade. *Proc Natl Acad Sci U S A* **98**, 13543-8.
- Rynkiewicz, M. J., Cane, D. E. and Christianson, D. W.** (2002). X-ray crystal structures of D100E trichodiene synthase and its pyrophosphate complex reveal the basis for terpene product diversity. *Biochemistry* **41**, 1732-41.
- Santos-Gomes, P. C. and Fernandes-Ferreira, M.** (2003). Essential oils produced by in vitro shoots of sage (*Salvia officinalis* L.). *J Agric Food Chem* **51**, 2260-6.
- Sauer, U.** (2006). Metabolic networks in motion: ¹³C-based flux analysis. *Mol Syst Biol* **2**, 62.

- Savage, T. J., Hatch, M. W. and Croteau, R.** (1994). Monoterpene synthases of *Pinus contorta* and related conifers. A new class of terpenoid cyclase. *J Biol Chem* **269**, 4012-20.
- Schiestl, R. H. and Gietz, R. D.** (1989). High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet* **16**, 339-46.
- Schwab, W., Williams, D. C., Davis, E. M. and Croteau, R.** (2001). Mechanism of monoterpene cyclization: stereochemical aspects of the transformation of noncyclizable substrate analogs by recombinant (-)-limonene synthase, (+)-bornyl diphosphate synthase, and (-)-pinene synthase. *Arch Biochem Biophys* **392**, 123-36.
- Shaw, J. M. and Nunnari, J.** (2002). Mitochondrial dynamics and division in budding yeast. *Trends Cell Biol* **12**, 178-84.
- Shiba, Y., Paradise, E. M., Kirby, J., Ro, D. K. and Keasling, J. D.** (2007). Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab Eng* **9**, 160-8.
- Sies, H.** (1997). Oxidative stress: oxidants and antioxidants. *Exp Physiol* **82**, 291-5.
- Skalnik, D. G., Narita, H., Kent, C. and Simoni, R. D.** (1988). The membrane domain of 3-hydroxy-3-methylglutaryl-coenzyme A reductase confers endoplasmic reticulum localization and sterol-regulated degradation onto beta-galactosidase. *J Biol Chem* **263**, 6836-41.
- Sopko, R., Huang, D., Preston, N., Chua, G., Papp, B., Kafadar, K., Snyder, M., Oliver, S. G., Cyert, M., Hughes, T. R. et al.** (2006). Mapping pathways and phenotypes by systematic gene overexpression. *Mol Cell* **21**, 319-30.
- Sprague, G. F., Jr. and Winans, S. C.** (2006). Eukaryotes learn how to count: quorum sensing by yeast. *Genes Dev* **20**, 1045-9.

Spring, D. R. (2005). Chemical genetics to chemical genomics: small molecules offer big insights. *Chem Soc Rev* **34**, 472-82.

Stamellos, K. D., Shackelford, J. E., Shechter, I., Jiang, G., Conrad, D., Keller, G. A. and Krisans, S. K. (1993). Subcellular localization of squalene synthase in rat hepatic cells. Biochemical and immunochemical evidence. *J Biol Chem* **268**, 12825-36.

Starks, C. M., Back, K., Chappell, J. and Noel, J. P. (1997). Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. *Science* **277**, 1815-20.

Stashenko, E. E., Jaramillo, B. E. and Martinez, J. R. (2004). Analysis of volatile secondary metabolites from Colombian *Xylopiya aromatica* (Lamarck) by different extraction and headspace methods and gas chromatography. *J Chromatogr A* **1025**, 105-13.

Stashenko, E. E., Jaramillo, B. E. and Martinez, J. R. (2004). Comparison of different extraction methods for the analysis of volatile secondary metabolites of *Lippia alba* (Mill.) N.E. Brown, grown in Colombia, and evaluation of its in vitro antioxidant activity. *J Chromatogr A* **1025**, 93-103.

Stashenko, E. E. and Martinez, J. R. (2008). Sampling flower scent for chromatographic analysis. *J Sep Sci* **31**, 2022-31.

Steele, C. L., Katoh, S., Bohlmann, J. and Croteau, R. (1998). Regulation of oleoresinosis in grand fir (*Abies grandis*). Differential transcriptional control of monoterpene, sesquiterpene, and diterpene synthase genes in response to wounding. *Plant Physiol* **116**, 1497-504.

Stephanopoulos, G. (2007). Challenges in engineering microbes for biofuels production. *Science* **315**, 801-4.

Stephanopoulos, G. (2008). Metabolic engineering: enabling technology for biofuels production. *Metab Eng* **10**, 293-4.

- Suzuki, K., Hattori, Y., Uraji, M., Ohta, N., Iwata, K., Murata, K., Kato, A. and Yoshida, K.** (2000). Complete nucleotide sequence of a plant tumor-inducing Ti plasmid. *Gene* **242**, 331-6.
- Swayne, T. C., Gay, A. C. and Pon, L. A.** (2007). Fluorescence imaging of mitochondria in yeast. *Methods Mol Biol* **372**, 433-59.
- Szkopinska, A., Grabinska, K., Delourme, D., Karst, F., Rytka, J. and Palamarczyk, G.** (1997). Polyprenol formation in the yeast *Saccharomyces cerevisiae*: effect of farnesyl diphosphate synthase overexpression. *J Lipid Res* **38**, 962-8.
- Szkopinska, A., Swiezewska, E. and Karst, F.** (2000). The regulation of activity of main mevalonic acid pathway enzymes: farnesyl diphosphate synthase, 3-hydroxy-3-methylglutaryl-CoA reductase, and squalene synthase in yeast *Saccharomyces cerevisiae*. *Biochem Biophys Res Commun* **267**, 473-7.
- Takahashi, S., Kuzuyama, T., Watanabe, H. and Seto, H.** (1998). A 1-deoxy-D-xylulose 5-phosphate reductoisomerase catalyzing the formation of 2-C-methyl-D-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. *Proc Natl Acad Sci U S A* **95**, 9879-84.
- Tanaka, K., Urbanczyk, H., Matsui, H., Sawada, H. and Suzuki, K.** (2006). Construction of physical map and mapping of chromosomal virulence genes of the biovar 3 *Agrobacterium* (*Rhizobium vitis*) strain K-Ag-1. *Genes Genet Syst* **81**, 373-80.
- Tarshis, L. C., Yan, M., Poulter, C. D. and Sacchettini, J. C.** (1994). Crystal structure of recombinant farnesyl diphosphate synthase at 2.6-Å resolution. *Biochemistry* **33**, 10871-7.
- Teoh, K. H., Polichuk, D. R., Reed, D. W., Nowak, G. and Covello, P. S.** (2006). *Artemisia annua* L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a

key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. *FEBS Lett* **580**, 1411-6.

Tholl, D. (2006). Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Curr Opin Plant Biol* **9**, 297-304.

Thulasiram, H. V. and Poulter, C. D. (2006). Farnesyl diphosphate synthase: the art of compromise between substrate selectivity and stereoselectivity. *J Am Chem Soc* **128**, 15819-23.

Tkacz, J. S. and MacKay, V. L. (1979). Sexual conjugation in yeast. Cell surface changes in response to the action of mating hormones. *J Cell Biol* **80**, 326-33.

Tong, F., Black, P. N., Bivins, L., Quackenbush, S., Ctrnacta, V. and DiRusso, C. C. (2006). Direct interaction of *Saccharomyces cerevisiae* Faa1p with the Omi/HtrA protease orthologue Ynm3p alters lipid homeostasis. *Mol Genet Genomics* **275**, 330-43.

Trachtulcova, P., Frydlova, I., Janatova, I. and Hasek, J. (2004). The absence of the Isw2p-Itc1p chromatin-remodelling complex induces mating type-specific and Flo11p-independent invasive growth of *Saccharomyces cerevisiae*. *Yeast* **21**, 389-401.

Trantas, E., Panopoulos, N. and Ververidis, F. (2009). Metabolic engineering of the complete pathway leading to heterologous biosynthesis of various flavonoids and stilbenoids in *Saccharomyces cerevisiae*. *Metab Eng*.

Trapp, S. and Croteau, R. (2001). Defensive Resin Biosynthesis in Conifers. *Annu Rev Plant Physiol Plant Mol Biol* **52**, 689-724.

Turnbull, J. J., Nakajima, J., Welford, R. W., Yamazaki, M., Saito, K. and Schofield, C. J. (2004). Mechanistic studies on three 2-oxoglutarate-dependent oxygenases of flavonoid biosynthesis: anthocyanidin synthase, flavonol synthase, and flavanone 3beta-hydroxylase. *J Biol Chem* **279**, 1206-16.

- Uchida, K.** (2000). Cellular response to bioactive lipid peroxidation products. *Free Radic Res* **33**, 731-7.
- Uetz, P. and Stagljar, I.** (2006). The interactome of human EGF/ErbB receptors. *Mol Syst Biol* **2**, 2006 0006.
- Usaitė, R., Wohlschlegel, J., Venable, J. D., Park, S. K., Nielsen, J., Olsson, L. and Yates Iii, J. R.** (2008). Characterization of global yeast quantitative proteome data generated from the wild-type and glucose repression *Saccharomyces cerevisiae* strains: the comparison of two quantitative methods. *J Proteome Res* **7**, 266-75.
- Valenzuela, P., Medina, A., Rutter, W. J., Ammerer, G. and Hall, B. D.** (1982). Synthesis and assembly of hepatitis B virus surface antigen particles in yeast. *Nature* **298**, 347-50.
- Vandamme, E. J.** (1992). Microbial production of vitamins and biofactors: an overview. *J Nutr Sci Vitaminol (Tokyo) Spec No*, 244-7.
- Vassarotti, A. and Goffeau, A.** (1992). Sequencing the yeast genome: the European effort. *Trends Biotechnol* **10**, 15-8.
- Veen, M. and Lang, C.** (2004). Production of lipid compounds in the yeast *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* **63**, 635-46.
- Veen, M., Stahl, U. and Lang, C.** (2003). Combined overexpression of genes of the ergosterol biosynthetic pathway leads to accumulation of sterols in *Saccharomyces cerevisiae*. *FEMS Yeast Res* **4**, 87-95.
- Verdonk, J. C., Ric de Vos, C. H., Verhoeven, H. A., Haring, M. A., van Tunen, A. J. and Schuurink, R. C.** (2003). Regulation of floral scent production in petunia revealed by targeted metabolomics. *Phytochemistry* **62**, 997-1008.

- Verdoucq, L., Vignols, F., Jacquot, J. P., Chartier, Y. and Meyer, Y.** (1999). In vivo characterization of a thioredoxin h target protein defines a new peroxiredoxin family. *J Biol Chem* **274**, 19714-22.
- Vidan, S. and Snyder, M.** (2001). Large-scale mutagenesis: yeast genetics in the genome era. *Curr Opin Biotechnol* **12**, 28-34.
- Villas-Boas, S. G., Moxley, J. F., Akesson, M., Stephanopoulos, G. and Nielsen, J.** (2005). High-throughput metabolic state analysis: the missing link in integrated functional genomics of yeasts. *Biochem J* **388**, 669-77.
- Wahl, H. G., Hoffmann, A., Luft, D. and Liebich, H. M.** (1999). Analysis of volatile organic compounds in human urine by headspace gas chromatography-mass spectrometry with a multipurpose sampler. *J Chromatogr A* **847**, 117-25.
- Walker, G. M.** (1999). Synchronization of yeast cell populations. *Methods Cell Sci* **21**, 87-93.
- Wallaart, T. E., Bouwmeester, H. J., Hille, J., Poppinga, L. and Maijers, N. C.** (2001). Amorpha-4,11-diene synthase: cloning and functional expression of a key enzyme in the biosynthetic pathway of the novel antimalarial drug artemisinin. *Planta* **212**, 460-5.
- Wang, L., Kao, R., Ivey, F. D. and Hoffman, C. S.** (2004). Strategies for gene disruptions and plasmid constructions in fission yeast. *Methods* **33**, 199-205.
- Wang, X., Ira, G., Tercero, J. A., Holmes, A. M., Diffley, J. F. and Haber, J. E.** (2004). Role of DNA replication proteins in double-strand break-induced recombination in *Saccharomyces cerevisiae*. *Mol Cell Biol* **24**, 6891-9.
- Wang, Y. H., Davila-Huerta, G. and Essenberg, M.** (2003). 8-Hydroxy-(+)-delta-cadinene is a precursor to hemigossypol in *Gossypium hirsutum*. *Phytochemistry* **64**, 219-25.

Wendland, B., Emr, S. D. and Riezman, H. (1998). Protein traffic in the yeast endocytic and vacuolar protein sorting pathways. *Curr Opin Cell Biol* **10**, 513-22.

Wheeler, C. J. and Croteau, R. (1987). Monoterpene cyclases. Stereoelectronic requirements for substrate binding and ionization. *J Biol Chem* **262**, 8213-9.

White, N. J. (2008). Qinghaosu (artemisinin): the price of success. *Science* **320**, 330-4.

Whittington, D. A., Wise, M. L., Urbansky, M., Coates, R. M., Croteau, R. B. and Christianson, D. W. (2002). Bornyl diphosphate synthase: structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proc Natl Acad Sci U S A* **99**, 15375-80.

Wickner, R. B. (1995). Prions of yeast and heat-shock protein 104: 'coprion' and cure. *Trends Microbiol* **3**, 367-9.

Wijeratne, S. S. and Cuppett, S. L. (2007). Potential of rosemary (*Rosemarinus officinalis* L.) diterpenes in preventing lipid hydroperoxide-mediated oxidative stress in Caco-2 cells. *J Agric Food Chem* **55**, 1193-9.

Wilkinson, B. and Bachmann, B. O. (2006). Biocatalysis in pharmaceutical preparation and alteration. *Curr Opin Chem Biol* **10**, 169-76.

Williams, D. C., Carroll, B. J., Jin, Q., Rithner, C. D., Lenger, S. R., Floss, H. G., Coates, R. M., Williams, R. M. and Croteau, R. (2000). Intramolecular proton transfer in the cyclization of geranylgeranyl diphosphate to the taxadiene precursor of taxol catalyzed by recombinant taxadiene synthase. *Chem Biol* **7**, 969-77.

Williams, D. C., McGarvey, D. J., Katahira, E. J. and Croteau, R. (1998). Truncation of limonene synthase preprotein provides a fully active 'pseudomature' form of this monoterpene cyclase and reveals the function of the amino-terminal arginine pair. *Biochemistry* **37**, 12213-20.

Winde, J. H. d. (2003). Functional genetics of industrial yeasts. Berlin {[u.a.]: Springer.

Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J. D., Bussey, H. et al. (1999). Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901-6.

Wise, M. L., Savage, T. J., Katahira, E. and Croteau, R. (1998). Monoterpene synthases from common sage (*Salvia officinalis*). cDNA isolation, characterization, and functional expression of (+)-sabinene synthase, 1,8-cineole synthase, and (+)-bornyl diphosphate synthase. *J Biol Chem* **273**, 14891-9.

Wise, M. L., Urbansky, M., Helms, G. L., Coates, R. M. and Croteau, R. (2002). Syn stereochemistry of cyclic ether formation in 1,8-cineole biosynthesis catalyzed by recombinant synthase from *Salvia officinalis*. *J Am Chem Soc* **124**, 8546-7.

Wodicka, L., Dong, H., Mittmann, M., Ho, M. H. and Lockhart, D. J. (1997). Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat Biotechnol* **15**, 1359-67.

Wong, S. L., Zhang, L. V., Tong, A. H., Li, Z., Goldberg, D. S., King, O. D., Lesage, G., Vidal, M., Andrews, B., Bussey, H. et al. (2004). Combining biological networks to predict genetic interactions. *Proc Natl Acad Sci U S A* **101**, 15682-7.

Wood, D. W., Setubal, J. C., Kaul, R., Monks, D. E., Kitajima, J. P., Okura, V. K., Zhou, Y., Chen, L., Wood, G. E., Almeida, N. F., Jr. et al. (2001). The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**, 2317-23.

Wu, S., Schalk, M., Clark, A., Miles, R. B., Coates, R. and Chappell, J. (2006). Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. *Nat Biotechnol* **24**, 1441-7.

- Wungsintaweekul, J., Herz, S., Hecht, S., Eisenreich, W., Feicht, R., Rohdich, F., Bacher, A. and Zenk, M. H. (2001). Phosphorylation of 1-deoxy-D-xylulose by D-xylulokinase of *Escherichia coli*. *Eur J Biochem* **268**, 310-6.
- Yamano, S., Ishii, T., Nakagawa, M., Ikenaga, H. and Misawa, N. (1994). Metabolic engineering for production of beta-carotene and lycopene in *Saccharomyces cerevisiae*. *Biosci Biotechnol Biochem* **58**, 1112-4.
- Yan, Y., Chemler, J., Huang, L., Martens, S. and Koffas, M. A. (2005). Metabolic engineering of anthocyanin biosynthesis in *Escherichia coli*. *Appl Environ Microbiol* **71**, 3617-23.
- Yan, Y., Kohli, A. and Koffas, M. A. (2005). Biosynthesis of natural flavanones in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **71**, 5610-3.
- Yang, C. S., Landau, J. M., Huang, M. T. and Newmark, H. L. (2001). Inhibition of carcinogenesis by dietary polyphenolic compounds. *Annu Rev Nutr* **21**, 381-406.
- Yoshikuni, Y., Ferrin, T. E. and Keasling, J. D. (2006). Designed divergent evolution of enzyme function. *Nature* **440**, 1078-82.
- Yoshikuni, Y., Martin, V. J., Ferrin, T. E. and Keasling, J. D. (2006). Engineering cotton (+)-delta-cadinene synthase to an altered function: germacrene D-4-ol synthase. *Chem Biol* **13**, 91-8.
- Zhang, D., Jennings, S. M., Robinson, G. W. and Poulter, C. D. (1993). Yeast squalene synthase: expression, purification, and characterization of soluble recombinant enzyme. *Arch Biochem Biophys* **304**, 133-43.
- Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T. et al. (2001). Global analysis of protein activities using proteome chips. *Science* **293**, 2101-5.

Zinser, E. and Daum, G. (1995). Isolation and biochemical characterization of organelles from the yeast, *Saccharomyces cerevisiae*. *Yeast* **11**, 493-536.

Zinser, E., Paltauf, F. and Daum, G. (1993). Sterol composition of yeast organelle membranes and subcellular distribution of enzymes involved in sterol metabolism. *J Bacteriol* **175**, 2853-8.