

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Identification and Characterization
of Dothistromin Biosynthetic
Genes in the Peanut Pathogen
*Passalora arachidicola***

**A thesis presented in partial fulfillment of the requirements
for the degree of Master of Science in Biochemistry
at Massey University, Palmerston North, New Zealand**

Yanan Guo

2008

Abstract

Dothistromin (DOTH) is a secondary metabolite produced by the fungal peanut pathogen *Passalora arachidicola* and pine needle pathogen *Dothistroma septosporum*. The chemical structure of DOTH is similar to a precursor of aflatoxin (AF) and sterigmatocystin (ST), which are secondary metabolites produced by *Aspergillus* species.

A size fractionated genomic library was made and 11 putative DOTH genes were identified in *P. arachidicola*. The DOTH genes in *P. arachidicola* were compared to DOTH genes in *D. septosporum* as well as to AF and ST genes in *Aspergillus* species. The DOTH gene products in *P. arachidicola* showed 73 - 96% amino acid identity to DOTH genes in *D. septosporum* and 50 - 69% amino acid identity to AF/ST genes in *Aspergillus*. The DOTH biosynthesis genes in *P. arachidicola* had similar gene organization and direction of transcription to DOTH biosynthesis genes in *D. septosporum* and is similar in that 11 putative DOTH genes are separated into three mini-clusters. This differs from the AF/ST clusters in which 25 AF/ST genes are tightly clustered in a 70 kb region.

Identification of transcription factor binding sites upstream of DOTH genes in *P. arachidicola* and *D. septosporum* suggested similar co-regulation of DOTH gene expression in *P. arachidicola* and *D. septosporum*.

Tandem and inverted repeat sequences were identified in intergenic regions in the *P. arachidicola* DOTH gene cluster, but the distribution of those repeats appears to be random. This suggests that the fragmentation of the DOTH biosynthesis gene cluster is not due to retrotransposon activity or recombination between repeat sequences. The DOTH biosynthesis gene clusters in *P. arachidicola* and *D. septosporum* could be ancestral to AF/ST biosynthesis clusters in *Aspergillus* species.

Acknowledgements

I would like to gratefully thank my supervisor Dr Rosie Bradshaw for her constant encouragement and guidance during the course of my Masters. Thanks for your kindness and patience when I have questions. Thanks for your for helping me with my English. You are a great supervisor and I couldn't ask for more.

To my lab colleagues, thanks for Shuguang for helpful technical support and advice, your willingness to listen and sharing your knowledge make it easier for me to start my project. Thanks for Zhilun and Arne for technical assistance and advice. Thanks for Justin, Naydene, Rebecca. It is a pleasure to work with all of you.

I would also like to thank Kathy Hamilton for helping me arrange my course. Thanks for Dr Kathryn Stowell for giving me advice when I start my Master project. Many thanks to other IMBS staff who gave me help over the duration of my studies.

Finally I would like to thank my parents, although it is far from home, you always been so supportive, encouraging and understanding over the duration of my studies. To my husband, thanks for believing in me and helping me with revision before exams.

Table of Contents

Abstract	I
Acknowledgements	II
Table of Contents	III
List of Tables	X
List of Figures	XI
Abbreviations	XIII
Chapter One - Introduction	
1.1 <i>Passalora arachidicola</i> and early leaf spot disease	1
1.1.1 Early leaf spot disease	1
1.1.2 <i>Passalora arachidicola</i>	2
1.1.2.1 The fungus	2
1.1.2.2 <i>Passalora arachidicola</i> produces dothistromin (DOTH) toxin	2
1.2 Dothistromin and biosynthesis gene clusters	3
1.2.1 Structure and biological activity of dothistromin	3
1.2.2 Sterigmatocystin biosynthesis gene cluster	5
1.2.3 Aflatoxin biosynthesis gene clusters	5
1.2.4 Dothistromin biosynthesis gene cluster	8
1.3 Double-stranded-RNA in fungi	11
1.4 Aim and objectives	13
1.4.1 Aim	13
1.4.2 Objectives	13

Chapter Two - Materials and Methods

2.1 Growth and maintenance of cultures	14
2.1.1 Growth and maintenance of <i>E. coli</i> cultures	14
2.1.2 Growth, maintenance and harvest of <i>P. arachidicola</i> culture and mycelia	14
2.2 DNA extraction, quantification and analysis	15
2.2.1 Genomic DNA isolation from <i>P. arachidicola</i> by CTAB method	15
2.2.2 Purification of <i>P. arachidicola</i> genomic DNA	16
2.2.2.1 Column purification	16
2.2.2.2 Phenol purification	16
2.2.3 Isolation of plasmid DNA	16
2.2.4 Fluorometric assay to determine DNA concentration	17
2.2.5 Agarose gel electrophoresis	17
2.2.6 Agarose gel purification of size fractionated DNA	18
2.2.7 Restriction endonuclease digestion of DNA	18
2.3 Polymerase chain reaction (PCR)	19
2.3.1 Standard PCR reaction	19
2.3.2 Inverse PCR	19
2.3.3 Touchdown PCR	19
2.3.4 Nested PCR	21
2.3.5 <i>E. coli</i> colony PCR	21
2.4 Ligation and transformation	22
2.4.1 Ligation	22
2.4.1.1 Ligation reaction	22
2.4.1.2 Preparation of the vector	22
2.4.1.3 N-butanol purification of ligation mixture	22
2.4.2 Transformation	23
2.4.2.1 Preparation of <i>E. coli</i> competent cells for electroporation	23
2.4.2.2 Transformation reaction of <i>E. coli</i> cells by electroporation	23
2.5 Hybridization of DIG-labelled probes to blots	24
2.5.1 DIG-labelling of the probes (PCR based)	24
2.5.2 Probe concentration determination	24

2.5.3 Southern blot	25
2.5.4 Hybridisation of DIG labelled probe	25
2.5.5 Immunological detection	25
2.5.6 Stripping blot	26
2.5.7 Screen the library for positive clones	26
2.5.7.1 Colony lifts	26
2.5.7.2 Screen the library for positive clones	27
2.6 DNA sequencing	27
2.6.1 Sequencing	27
2.6.2 Sequence analysis	27

Chapter Three - Results

3.1 Inverse polymerase chain reaction (PCR)	29
3.1.1 Background	29
3.1.2 Results	29
3.1.3 Discussion	31
3.2 RNA that co-purified with genomic DNA	33
3.2.1 Background	33
3.2.2 Results	33
3.2.3 Discussion	35
3.3 Attempts to identify additional dothistromin genes in <i>P. arachidicola</i> using degenerate primers	37
3.3.1 Background	37
3.3.2 Results	37
3.3.3 Discussion	39
3.4 A size fractionated <i>P. arachidicola</i> genomic DNA library	41
3.4.1 Background	41
3.4.2 Results	41
3.4.2.1 Restriction enzyme digestion of purified <i>P. arachidicola</i> 18667 gDNA	41
3.4.2.2 Southern hybridization to identify the size of restriction fragments containing DOTH genes	42

3.4.2.3 A size fractionated <i>EcoRI</i> <i>P. arachidicola</i> genomic DNA library	45
3.4.2.4 Screen the library for positive clones	46
3.4.3 Discussion	49
3.5 Genes identified in plasmids pR284 and pR285 (<i>P. arachidicola</i> <i>pksA</i> cluster)	50
3.5.1 Background	50
3.5.2 Results	51
3.5.2.1 Primary sequence analysis of plasmids pR284 and pR285	51
3.5.2.2 The <i>P. arachidicola</i> <i>Pa-pksA</i> gene	54
3.5.2.3 The <i>P. arachidicola</i> <i>Pa-cypA</i> gene	57
3.5.2.4 The <i>P. arachidicola</i> <i>Pa-avfA</i> gene	58
3.5.2.5 The <i>P. arachidicola</i> <i>Pa-epoA</i> gene	59
3.5.2.6 The <i>P. arachidicola</i> <i>Pa-moxA</i> gene	60
3.5.2.7 The dothistromin biosynthesis genes in <i>P. arachidicola</i> (<i>pksA</i> cluster)	61
3.5.3 Discussion	62
3.6 Genes identified in plasmids pR283 (<i>P. arachidicola</i> <i>vbsA</i> cluster)	66
3.6.1 Background	66
3.6.2 Results	66
3.6.2.1 Primary sequence analysis of plasmids pR283	66
3.6.2.2 The <i>P. arachidicola</i> <i>Pa-hexA</i> gene	68
3.6.2.3 The <i>P. arachidicola</i> <i>Pa11</i> gene	69
3.6.2.4 The <i>P. arachidicola</i> <i>Pa12</i> gene	70
3.6.2.5 The <i>P. arachidicola</i> <i>Pa-vbsA</i> gene	70
3.6.2.6 The dothistromin biosynthesis genes in <i>P. arachidicola</i> (<i>vbsA</i> cluster)	71
3.6.3 Discussion	74
3.7 Genes identified in plasmids pR286 (<i>P. arachidicola</i> <i>dotA</i> cluster)	75
3.7.1 Background	75
3.7.2 Results	75
3.7.2.1 Primary sequence analysis of plasmid pR286	75
3.7.2.2 The <i>P. arachidicola</i> <i>Pa-dotA</i> gene	78

3.7.2.3 The <i>P. arachidicola</i> <i>Pa-dotB</i> gene	79
3.7.2.4 The <i>P. arachidicola</i> <i>Pa-dotC</i> gene	80
3.7.2.5 The <i>P. arachidicola</i> <i>Pa-dotD</i> gene	80
3.7.2.6 The <i>P. arachidicola</i> Partial <i>Pa-mfs</i> gene	81
3.7.2.7 Arrangement of dothistromin biosynthesis genes in <i>P. arachidicola</i> (<i>dotA</i> cluster)	81
3.7.3 Discussion	85
3.8 DOTH biosynthesis gene cluster in <i>P. arachidicola</i>	88
3.8.1 Background	88
3.8.2 Results	88
3.8.2.1 DOTH biosynthesis gene cluster in <i>P. arachidicola</i>	88
3.8.2.2 Regulatory motifs of DOTH biosynthesis genes in <i>P. arachidicola</i>	90
3.8.3 Discussion	97
 Chapter 4 Conclusions and future work	
4.1 The DOTH biosynthesis gene clusters in <i>P. arachidicola</i>	101
4.2 Future works	102
4.2.1 Identification of other DOTH biosynthesis genes in <i>P.</i> <i>arachidicola</i>	102
4.2.2 Identification of DOTH gene functions in <i>P. arachidicola</i>	103
References	105
 Appendix	
Appendix I Media	112
Appendix II Buffers and solutions	113
Appendix III Primers used in this project	116
Appendix IV Nucleotide and amino acid sequence of <i>Pa-pksA</i>	119

Appendix V Alignment of Pa-PksA with other PKSA proteins	124
Appendix VI Nucleotide and amino acid sequence of <i>Pa-cypA</i>	129
Appendix VII Alignment of Pa-CypA with other CypA proteins	131
Appendix VIII Nucleotide and amino acid sequence of <i>Pa-avfA</i>	133
Appendix IX Alignment of Pa-AvfA with other oxidase proteins	134
Appendix X Nucleotide and amino acid sequence of <i>Pa-epoA</i>	135
Appendix XI Alignment of Pa-EpoA with Ds-EpoA protein	135
Appendix XII Nucleotide and amino acid sequence of <i>Pa-moxA</i>	136
Appendix XIII Alignment of Pa-MoxA with other monooxygenase proteins	136
Appendix XIV Nucleotide and amino acid sequence of partial <i>Pa-hexA</i>	137
Appendix XV Alignment of Pa-HexA with other HexA proteins	137
Appendix XVI Nucleotide and amino acid sequence of <i>Pa11</i>	138
Appendix XVII Alignment of Pa11 with Ds11	139
Appendix XVIII Nucleotide and amino acid sequence of <i>Pa12</i>	139
Appendix XIX Alignment of Pa12 with other hypothetical proteins	140
Appendix XX Nucleotide and amino acid sequence of <i>Pa-vbsA</i>	140
Appendix XXI Alignment of Pa-VbsA with other VbsA proteins	143
Appendix XXII Nucleotide and amino acid sequence of partial <i>Pa-dotA</i>	145

Appendix XXIII Alignment of Pa-DotA with other ketoreductase proteins	145
Appendix XXIV Nucleotide and amino acid sequence of <i>Pa-dotB</i>	146
Appendix XXV Alignment of Pa-DotB with other oxidase proteins	147
Appendix XXVI Nucleotide and amino acid sequence of <i>Pa-dotC</i>	148
Appendix XXVII Alignment of Pa-DotC with other MFS transporter proteins	150
Appendix XXVIII Nucleotide and amino acid sequence of <i>Pa-dotD</i>	151
Appendix XXIX Alignment of Pa-DotD with Ds-DotD and C-terminal of Pks proteins	152
Appendix XXX Nucleotide and amino acid sequence of partial <i>Pa-mfs</i>	153
Appendix XXXI Alignment of Pa-Mfs with other Mfs proteins	153

List of Tables

Number	Title	Page
Table 1.1	AF/ST pathway cluster genes and the roles of their gene products	7
Table 2.1	Example of <i>EcoRI</i> digestion of gDNA for Southern blot and fractionated genomic library	18
Table 2.2	Touchdown PCR conditions in this experiment	20
Table 3.1	Degenerate primers used in touchdown PCR	38
Table 3.2	BLASTN result of 1.1 kb fragment	39
Table 3.3	Main BLASTX results of 11.5 kb fragment	53
Table 3.4	Amino acids identity of Pa-PksA to other Pks proteins	55
Table 3.5	Intergenic regions of <i>P. arachidicola</i> and <i>D. septosporum</i> dothistromin genes in 11.5 kb fragment	62
Table 3.6	(A) Inverted repeats (B) Tandem repeats	62
Table 3.7	Main BLASTX results of 5.7 kb fragment	67
Table 3.8	Intergenic regions of <i>P. arachidicola</i> and <i>D. septosporum</i> dothistromin genes in 5.7 kb fragment	73
Table 3.9	Tandem repeats	73
Table 3.10	Main BLASTX results of 5.7 kb fragment	77
Table 3.11	Intergenic regions of <i>P. arachidicola</i> and <i>D. septosporum</i> dothistromin genes in 9.3 kb fragment	83
Table 3.12	Inverted repeats	83
Table 3.13	Tandem repeats	84
Table 3.14	Characterization of genes identified in <i>P. arachidicola</i> DOTH mini-clusters	91

List of Figures

Number	Title	Page
Figure 1.1	Early leaf spot on peanut	1
Figure 1.2	Comparison of the structures of DOTH and AF precursor	3
Figure 1.3	Comparison of dothistromin gene cluster with ST/AF gene clusters	10
Figure 2.1	Model for steps in the evolution of the ST/AF clusters	11
Figure 2.1	An example of probe concentration detection	24
Figure 3.1	Enzyme cutting sites on <i>Pa-cypA</i> and <i>Pa-vbsA</i> fragments and primers used for inverse PCR	30
Figure 3.2	Inverse PCR results	30
Figure 3.3	PCR screens of selected transformants	31
Figure 3.4	The RNA bands	33
Figure 3.5	Restriction digestion of the RNA bands	34
Figure 3.6	The RNA bands in other <i>P. arachidicola</i> strains	34
Figure 3.7	The stability of the RNA bands	35
Figure 3.8	Touchdown PCR using degenerate primers	38
Figure 3.9	PCR using universal primers to check the transformants	39
Figure 3.10	Purified <i>P. arachidicola</i> 18667 gDNA for restriction enzyme digestion	42
Figure 3.11	DIG-labelled probes for Southern blotting	43
Figure 3.12	Working concentration of probes	43
Figure 3.13	Southern blot results	44
Figure 3.14	Checking dephosphorylation of <i>EcoRI</i> digested PIC19H vector	46
Figure 3.15	Colony hybridisation using <i>Pa-cypA</i> probe	47
Figure 3.16	<i>E. coli</i> colony PCR to screen the positive colonies	48
Figure 3.17	<i>EcoRI</i> digestion to check the insert DNA	48
Figure 3.18	Schematic diagram of the 11.5 kb contig showing five ORFs and positions of primers used for sequencing	52
Figure 3.19	Schematic diagram of <i>Pa-pksA</i>	55
Figure 3.20	Alignments of Pa-PksA KS, AT and ACP domains with other PKS proteins	56
Figure 3.21	Schematic diagram of <i>Pa-cypA</i>	57
Figure 3.22	Characteristic cytochrome P450 monooxygenase motifs in Pa-CypA	58
Figure 3.23	Schematic diagram of <i>Pa-avfA</i>	58
Figure 3.24	Comparison of the <i>Pa-epoA</i> gene and the <i>Ds-epoA</i> gene	60
Figure 3.25	Schematic diagram of <i>Pa-moxA</i> gene	60
Figure 3.26	Comparison of <i>D. septosporum</i> and <i>P. arachidicola</i> dothistromin <i>pksA</i> biosynthesis gene orientations	61

Figure 3.27	Schematic diagram of the 5.7 kb <i>P. arachidicola</i> contig showing four ORFs and positions of primers used for sequencing	67
Figure 3.28	Schematic diagram of <i>Pa-hexA</i>	68
Figure 3.29	Alignments of Pa-HexA ACP domain with other HexA proteins	69
Figure 3.30	Schematic diagram of <i>PaI1</i>	70
Figure 3.31	Schematic diagram of <i>Pa-vbsA</i>	71
Figure 3.32	Alignments of Pa-VbsA flavin binding domain with other flavin binding domains	71
Figure 3.33	Comparison of <i>D. septosporum</i> and <i>P. arachidicola</i> dothistromin <i>vbsA</i> biosynthesis gene clusters	72
Figure 3.34	Schematic diagram of the 9.3 kb <i>P. arachidicola</i> contig showing four ORFs and positions of primers used for sequencing	77
Figure 3.35	Schematic diagram of <i>Pa-dotA</i>	78
Figure 3.36	Alignments of Pa-DotA adenine binding motif with other adenine binding motifs	79
Figure 3.37	Alignments of Pa-DotB heme-binding domain with other oxidase proteins	79
Figure 3.38	Schematic diagram of <i>Pa-dotC</i>	80
Figure 3.39	Alignment of Pa-DotD with other thioesterase and polyketide synthase proteins	81
Figure 3.40	Comparison of <i>D. septosporum</i> and <i>P. arachidicola</i> dothistromin <i>dotA</i> biosynthesis gene cluster	82
Figure 3.41	Amino acid and nucleotide identity of genes and gene gaps between <i>P. arachidicola</i> and <i>D. septosporum</i>	92
Figure 3.42	The GC content of three DOTH gene clusters in <i>P. arachidicola</i>	93
Figure 3.43	Comparison of the DOTH, AF and ST biosynthesis clusters in <i>P. arachidicola</i> , <i>D. septosporum</i> and <i>Aspergillus</i> species	94
Figure 3.44	Compare the intergenic region between <i>Pa-dotD-Pa-mfs</i> and <i>Ds-dotD-Ds9</i>	89
Figure 3.45	Proposed DOTH biosynthesis pathway in <i>P. arachidicola</i> compared to the AF biosynthesis pathway in <i>A. parasiticus</i>	95
Figure 3.46	AflR binding sites of DOTH genes in <i>P. arachidicola</i> and <i>D. septosporum</i>	96

Abbreviations

Abbreviation	Meaning
AF	aflatoxin
AVF	averufin
AVN	averantin
bp	base pair
cDNA	complementary deoxyribonucleic acid
°C	degree Celsius
DMST	demethylsterigmatocystin
Dnase	deoxyribonuclease
dNTP	deoxynucleotide triphosphate
DOTH	dothistromin
G	gram
HAVN	5'-hydroxy-averantin
ITS	internal transcribed spacer region
IPTG	isopropyl- β -D-thiogalactoside
kb	kilobase pair
L	litre
M	molar
Mb	megabase
MFS transporter	major facilitator superfamily transporter
ml	milliliter
mM	millimolar
NOR	norsolorinic acid
OAVN	oxoaverantin
OMST	o-methylsterigmatocystin
ORF	open reading frame
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
SDS	sodium dodecyl sulfate
ST	sterigmatocystin
μ l	microlitre
μ M	micromolar
μ g	microgram
UV	ultraviolet
VAL	versiconal
VER	versicolorin
VHA	versiconal hemiacetal acetate
X-Gal	5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside