Genetic Potential of Lichen-Forming Fungi in Polyketide Biosynthesis

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Yit Heng Chooi
B.Sc. (Hons)

School of Applied Sciences
Science, Engineering and Technology Portfolio
RMIT University

March 2008
Declaration

The work presented in this thesis was completed in the period of August 2004 to March 2008 under the co-supervision of Assoc. Professor Ann Lawrie and Professor David Stalker at School of Applied Sciences, RMIT University, and Dr. Simone Louwhoff at Royal Botanical Gardens, Victoria.

In compliance with the university regulations, I declare that:

I. except where due acknowledgement has been made; the work is that of the author alone
II. the work has not been submitted previously, in whole or in part, to qualify for any other academic award;
III. the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program;
IV. ethics procedures and guidelines have been followed.

_________________
Yit Heng Chooi
27th March 2008
Dedication

To my parents and
my beloved wife Lee Ngoh
Acknowledgement

There are many individuals without whom the work described in this thesis might not have been possible, and to whom I am greatly indebted.

I would like to thank my supervisor Associate Professor Ann Lawrie for, firstly, the opportunity to undertake this research topic of my own interest, and for her guidance, patience, and encouragement in the most challenging times. I would also like to express my sincere thanks to both of my co-supervisors. To Professor David Stalker who has taught me the true spirit of research, and to Dr. Simone Louwhoff who has introduced me to the beautiful world of lichenology. I am very grateful to RMIT University for providing the RMIT Malaysia-Australia Colombo plan Commemoration (MACC) Scholarship.

I am deeply indebted to Emeritus Professor Jack Elix at the Department of Chemistry, Australian National University for his assistance in chemical analysis, for proofreading the introductory part of my thesis and for having been an extremely supportive and generous mentor. His insights into the mechanisms of organic reactions and his wealth of knowledge on lichens and their secondary metabolites have been crucial in my research, and his enthusiasm and passion in science and teaching has been truly inspirational.

I would like to express my gratitude to Professor Isao Fujii and Professor Yutaka Ebizuka at the Graduate School of Pharmaceutical Sciences, The University of Tokyo for the opportunity to work with a group of vibrant scientists at the forefront of natural product research, for the valuable contributions of time and materials, and for their generous advice and guidance (Prof. Fujii is now at School of Pharmacy, Iwate Medical University). Although the research visit has only lasted for five weeks, I have learnt a lot from the people in the lab and the experience was invaluable for my scientific career development. A special thanks to Dr. Tetsuo Kushiro and Dr. Yasuyo Seshime, Takayuki Itoh, Tomohiro Moriya, Tomomi Moriguchi, Natsuki Nambu and Hiroki Sano, for their assistance in my experiments and made my stay in Tokyo enjoyable. I am also very grateful to T. Itoh, T. Kushiro, I. Fujii and Y. Ebizuka for communicating their unpublished results.
Similarly, I am very grateful to Associate Professor Meryl Davis at the Genetics Department, University of Melbourne for her contribution of ideas and encouragement and for providing the Aspergillus nidulans strain, plasmid and other resources. I would also like to thank David Clarke at University of Melbourne for his assistance in A. nidulans transformation. Associate Professor Daniele Armaleo at the Department of Biology, Duke University is also thanked for his constructive suggestions and for communicating unpublished results.

Those who aided me in the lab; I thank Dr. Gregory Nugent, who was then a post-doc but is now a lecturer in the department, and Dr. Brian Meehan, for always being there to give me the appropriate advice and encouragement. I would also like to thank my friends and colleagues in the lab and department for their company, friendship and support. In particular to Kavitha who is now a proud mother, for the long hours spent together over failed Southern blotting and difficult cloning, and made the journey fun.

I would like to thank my parents for believing in me and for being proud of me. Their unconditional love and support has given me the strength and courage while I am away from home.

Most importantly, I would like to thank my wife Lee Ngoh, for being the most patient and supportive witness to my academic journey over the past three and a half years. Without her support, love and encouragement, this study would not have been possible.
# TABLE OF CONTENTS

Title Page I  
Declaration II  
Dedication III  
Acknowledgement IV  
Table of Contents VI  
List of Figures XI  
List of Tables XIX  
List of Abbreviations XXI  
Publications XXIII  
Abstract XXIV  

## CHAPTER 1 INTRODUCTION AND BACKGROUND 1

1.1 Introduction 2  
1.2 Background 5  
1.2.1 The lichens 5  
1.2.2 Chemistry and lichens 7  
1.2.3 Lichen secondary metabolites 9  
1.2.3.1 The biological significance of lichen metabolites 13  
1.2.3.2 Chemotaxonomy of lichens 13  
1.2.4 Prospects and applications of lichen natural products 14  
1.2.4.1 Pharmacology and medicine 14  
1.2.4.2 Potential applications in agriculture 18  
1.2.4.3 Cosmetics and perfume industry 19  
1.2.5 The polyketide factories 20  
1.2.6 The unique lichen compounds and the ubiquitous PKSs 27  
1.2.6.1 Lichen polyketides common to other fungi 28  
1.2.6.2 Polyketide compounds common in lichens 30  
1.2.6.2.1 Basic monoaromatic units that form lichen coupled phenolics 33  
1.2.6.2.2 The formation of various linkages in lichen coupled phenolics 37  
1.2.6.2.3 Derivatization of the basic monoaromatic units 39  
1.2.6.2.4 Depsides 43  
1.2.6.2.5 Depsidones 50  
1.2.6.2.6 Diphenyl ethers 56  
1.2.6.2.7 Depsones 61  
1.2.6.2.8 Dibenzofurans and usnic acid homologs 63  
1.2.6.2.9 Biphenyls 67  
1.2.6.3 Lichens as combinatorial chemists 68  
1.3 Research scope and objectives 75
CHAPTER 2  COMMON MATERIALS AND METHODS  

2.1  Introduction  

2.2  Molecular biology techniques  

2.2.1  Small-scale lichen DNA extraction  

2.2.2  Large-scale lichen DNA extraction  

2.2.3  Polymerase chain reaction (PCR)  

2.2.4  Gel electrophoresis (GE)  

2.2.5  DNA purification and gel extraction  

2.2.6  DNA sequencing  

2.2.7  Plasmid miniprep  

2.2.8  Transformation of *Escherichia coli* by electroporation  

2.2.9  Colony-direct PCR  

2.2.10  Cloning of PCR products  

2.2.11  Restriction digestion and ligation of DNA  

2.2.12  Southern blotting  

2.3  Bioinformatic analysis  

2.3.1  Raw sequence editing  

2.3.2  Multiple sequence alignment and phylogenetic analysis  

2.3.3  Gene annotation and plasmid map drawing  

2.3.4  Drawing of chemical structures  

CHAPTER 3  DIVERSITY OF POLYKETIDE SYNTHASE GENES IN LICHENS  

3.1  Introduction  

3.1.1  Exploring the genetic diversity of fungal PKSs  

3.1.2  Correlating the structure of polyketides and PKS genes with phylogenetic analysis  

3.1.3  Research goals  

3.2  Lichen identification and chemical analysis  

3.2.1  Materials and methods  

3.2.1.1  Lichen collection and identification  

3.2.1.2  Thin layer chromatography (TLC)  

3.2.1.3  High performance liquid chromatography (HPLC)  

3.2.1.4  PCR amplification of internal transcribed spacer regions  

3.2.2  Results  

3.2.2.1  Lichen specimens and identification  

3.2.2.2  Chemical analysis of lichens  

3.2.2.3  Internal transcribed spacer (ITS) regions  

3.2.3  Discussion  

VII
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3 NR- and PR-type KS domain-specific degenerate primers</td>
<td>113</td>
</tr>
<tr>
<td>3.3.1 Materials and methods</td>
<td>114</td>
</tr>
<tr>
<td>3.3.1.1 Design of degenerate primers</td>
<td>114</td>
</tr>
<tr>
<td>3.3.1.2 PCR amplification of KS domain</td>
<td>114</td>
</tr>
<tr>
<td>3.3.1.3 KS domain sequence analysis</td>
<td>114</td>
</tr>
<tr>
<td>3.3.2 Results</td>
<td>115</td>
</tr>
<tr>
<td>3.3.2.1 NRKS degenerate primers</td>
<td>115</td>
</tr>
<tr>
<td>3.3.2.2 PRKS degenerate primers</td>
<td>115</td>
</tr>
<tr>
<td>3.3.2.3 PCR results and KS domain sequence analyses</td>
<td>118</td>
</tr>
<tr>
<td>3.3.3 Discussion</td>
<td>120</td>
</tr>
<tr>
<td>3.3.3.1 NRKS amplified KS domains</td>
<td>121</td>
</tr>
<tr>
<td>3.3.3.2 PRKS amplified KS domain</td>
<td>121</td>
</tr>
<tr>
<td>3.4 Clade III NR-type KS domain-specific degenerate primers</td>
<td>124</td>
</tr>
<tr>
<td>3.4.1 Materials and methods</td>
<td>125</td>
</tr>
<tr>
<td>3.4.1.1 Design of degenerate primers</td>
<td>125</td>
</tr>
<tr>
<td>3.4.1.2 PCR amplification of KS domain and sequence analysis</td>
<td>125</td>
</tr>
<tr>
<td>3.4.2 Results</td>
<td>125</td>
</tr>
<tr>
<td>3.4.2.1 NR3KS degenerate primers</td>
<td>125</td>
</tr>
<tr>
<td>3.4.2.2 PCR results and KS domain sequence analyses</td>
<td>127</td>
</tr>
<tr>
<td>3.4.3 Discussion</td>
<td>128</td>
</tr>
<tr>
<td>3.5 Phylogenetic analysis of lichen KS domains</td>
<td>130</td>
</tr>
<tr>
<td>3.5.1 Materials and methods</td>
<td>131</td>
</tr>
<tr>
<td>3.5.1.1 Multiple sequence alignment</td>
<td>131</td>
</tr>
<tr>
<td>3.5.1.2 Construction of phylogenetic tree</td>
<td>131</td>
</tr>
<tr>
<td>3.5.2 Results</td>
<td>132</td>
</tr>
<tr>
<td>3.5.3 Discussion</td>
<td>134</td>
</tr>
<tr>
<td>3.5.3.1 Non-reducing PKS Clade I</td>
<td>134</td>
</tr>
<tr>
<td>3.5.3.2 Non-reducing PKS Clade II</td>
<td>139</td>
</tr>
<tr>
<td>3.5.3.3 Non-reducing PKS basal to Clades I and II</td>
<td>142</td>
</tr>
<tr>
<td>3.5.3.4 Non-reducing PKS Clade III</td>
<td>143</td>
</tr>
<tr>
<td>3.5.3.5 Partial-reducing PKSs</td>
<td>147</td>
</tr>
<tr>
<td>3.5.3.6 Highly-reducing PKSs</td>
<td>148</td>
</tr>
<tr>
<td>3.6 Conclusions and overview</td>
<td>149</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>155</td>
</tr>
<tr>
<td>4.1.1 Rationale and research goal</td>
<td>160</td>
</tr>
<tr>
<td>4.2 Locating and cloning of partial PKS genes in <em>Flavoparmelia caperatulla</em></td>
<td>161</td>
</tr>
<tr>
<td>4.2.1 Materials and methods</td>
<td>162</td>
</tr>
<tr>
<td>4.2.1.1 Southern blotting</td>
<td>162</td>
</tr>
<tr>
<td>4.2.1.2 Construction of partial genomic library</td>
<td>162</td>
</tr>
</tbody>
</table>
4.2.1.3 Screening of partial genomic library

4.2.2 Results
4.2.2.1 Southern blotting
4.2.2.2 Construction and screening of partial genomic library

4.2.3 Discussion
4.2.3.1 High molecular weight lichen DNA for Southern and cloning
4.2.3.2 Southern-guided library construction and PCR screening

4.3 Locating and cloning of partial PKS genes in *Xanthoparmelia semiviridis*

4.3.1 Materials and methods
4.3.1.1 Southern blotting
4.3.1.2 Construction of *X. semiviridis* partial genomic library
4.3.1.3 Screening of partial genomic library

4.3.2 Results
4.3.2.1 Southern blotting
4.3.2.2 Construction and screening of partial genomic library
4.3.2.3 Bioinformatic analysis

4.3.3 Discussion
4.3.3.1 Southern-guided PKS gene cloning
4.3.3.2 Bioinformatic analysis

4.4 Detection of CMeT domain in NR-PKS genes

4.4.1 Materials and methods
4.4.1.1 Design of NR-type CMeT domain-specific primers
4.4.1.2 PCR amplification of CMeT domain
4.4.1.3 Southern blotting with CMeT domain probe
4.4.1.4 KS-CMeT domain-hopping PCR

4.4.2 Results
4.4.2.1 NRMeT degenerate primers
4.4.2.2 PCR amplification of CMeT domains
4.4.2.3 Southern blotting with CMeT domain probe
4.4.2.4 Extending the PKS gene sequence by domain-hopping PCR

4.4.3 Discussion
4.4.3.1 Clade III NR-type CMeT domain
4.4.3.2 *X. semiviridis* clade III NR-PKS gene (*xsepks1*)

4.5 Cloning of a full-length PKS gene from *X. semiviridis*

4.5.1 Materials and methods
4.5.1.1 Southern blotting with the *xsepks1* KS-CMeT probe
4.5.1.2 Construction and screening of partial genomic library with large inserts

4.5.2 Results
4.5.2.1 Southern blotting with the *xsepks1* KS-CMeT probe
4.5.2.2 Cloning of the full-length PKS gene
4.5.2.3 Sequence and PKS domain analysis

4.5.3 Discussion
4.5.3.1 A clade III NR-PKS from *X. semiviridis* (*xsepks1*)
4.5.3.2 Regions surrounding the *xsepks1*
<table>
<thead>
<tr>
<th>CHAPTER 5</th>
<th>HETEROLOGOUS EXPRESSION OF LICHEN PKS GENE IN ASPERGILLUS</th>
<th>205</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>206</td>
</tr>
<tr>
<td>5.2</td>
<td>Transformation and transcription analysis of partial xsepks1 in <em>Aspergillus nidulans</em></td>
<td>210</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Materials and methods</td>
<td>210</td>
</tr>
<tr>
<td>5.2.1.1</td>
<td>Strain and plasmids</td>
<td>210</td>
</tr>
<tr>
<td>5.2.1.2</td>
<td><em>A. nidulans</em> transformation</td>
<td>210</td>
</tr>
<tr>
<td>5.2.1.3</td>
<td>Selection and PCR screening of transformants</td>
<td>211</td>
</tr>
<tr>
<td>5.2.1.4</td>
<td>Transcription detection by RT-PCR</td>
<td>211</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Results</td>
<td>212</td>
</tr>
<tr>
<td>5.2.2.1</td>
<td><em>A. nidulans</em> transformation and PCR screening</td>
<td>212</td>
</tr>
<tr>
<td>5.2.2.2</td>
<td>Transcriptional analysis of the partial PKS gene in <em>A. nidulans</em></td>
<td>213</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Discussion</td>
<td>214</td>
</tr>
<tr>
<td>5.3</td>
<td>Heterologous expression of full-length xsepks1 in <em>Aspergillus oryzae</em></td>
<td>215</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Materials and methods</td>
<td>218</td>
</tr>
<tr>
<td>5.3.1.1</td>
<td>Construction of entry clone by BP reaction</td>
<td>218</td>
</tr>
<tr>
<td>5.3.1.2</td>
<td>Construction of entry clone by LR reaction</td>
<td>219</td>
</tr>
<tr>
<td>5.3.1.3</td>
<td><em>A. oryzae</em> transformation</td>
<td>219</td>
</tr>
<tr>
<td>5.3.1.4</td>
<td>PCR screening of <em>A. oryzae</em> transformants</td>
<td>220</td>
</tr>
<tr>
<td>5.3.1.5</td>
<td>Transcription analysis of <em>A. oryzae</em> transformants</td>
<td>220</td>
</tr>
<tr>
<td>5.3.1.6</td>
<td>Fermentation and organic extraction</td>
<td>222</td>
</tr>
<tr>
<td>5.3.1.7</td>
<td>HPLC analysis</td>
<td>222</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Results</td>
<td>223</td>
</tr>
<tr>
<td>5.3.2.1</td>
<td><em>A. oryzae</em> transformation and screening</td>
<td>223</td>
</tr>
<tr>
<td>5.3.2.2</td>
<td><em>A. oryzae</em> transformants transcription analysis</td>
<td>223</td>
</tr>
<tr>
<td>5.3.2.3</td>
<td>Chemical analysis of <em>A. oryzae</em> transformants</td>
<td>228</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Discussion</td>
<td>237</td>
</tr>
<tr>
<td>5.3.3.1</td>
<td>Transformation and expression of xsepks1 in <em>A. oryzae</em></td>
<td>237</td>
</tr>
<tr>
<td>5.3.3.2</td>
<td>Functionality of XsePKS1 protein in <em>A. oryzae</em></td>
<td>241</td>
</tr>
<tr>
<td>5.3.3.3</td>
<td>Possible functions of clade III NR-PKSs in lichens</td>
<td>244</td>
</tr>
</tbody>
</table>

| CHAPTER 6 | CONCLUSIONS AND FUTURE DIRECTIONS | 246 |

References 249
List of Figures

CHAPTER 1

Figure 1.1 The leaf-like lichen Xanthoria parietina. A, Laminal view. B, Detail of a vertical cross-section: uc, conglutinate upper cortex; ph, photobiont layer harboring the globose cells of the green alga Trebouxia arboricola; m, gas-filled medullary layer built up by aerial hyphae; lc, conglutinate lower cortex. Adapted from Honegger (1996).

Figure 1.2 Probable pathways leading to the major groups of lichen metabolites. Adapted from Elix (1996).

Figure 1.3 The basic pathway of fatty acid and polyketide biosynthesis, showing the roles of the various activities carried out by the subunits or domains of the fatty acid or polyketide synthase (box). A-D represent the alternative versions of the reductive cycle that lead to keto, hydroxy, enoyl, or methylene functionality, respectively, at specific α-carbons during the assembly of reduced polyketides. Figure from Simpson (1995).

Figure 1.4 Type I, II and III polyketide synthases. KS, ketosynthase, AT, acyltransferase; ACP, acyl-carrier-protein; KR, ketoreductase; DH, dehydratase. Figure from Shen (2002).

Figure 1.5 Biosynthesis of 6-methylsalicylic acid. Adapted from Cox (2007).

Figure 1.6 Examples of typical domain architectures of A) non-reducing, B) partially reducing, and C) highly reducing PKSs and their corresponding products. KS, ketosynthase; AT, acyltransferase, DH, dehydratase; KR, ketoreductase, ACP, acyl carrier protein; TE, thioesterase; CYC, Claisen cyclase; CMeT, C-methyltransferase. Related references: Colletotrichum lagenarium PKS1 (Fujii et al. 2000); Penicillium patulum MSAS (Beck et al. 1990); Aspergillus terreus LDKS (Hendrickson et al. 1999; Kennedy et al. 1999).

Figure 1.7 Biosynthesis of orsellinic acid.

Figure 1.8 Examples of fungal metabolite biosynthesis related to some lichen metabolites. A, biosynthesis and folding pattern of some well-characterized fungal polyketides; B, some related lichen polyaromatic metabolites.

Figure 1.9 Various groups of coupled phenolics present in lichens where monoaromatic units were bonded by ester, ether and/or carbon-carbon linkages. The approximate number of compounds known [based on Huneck & Yoshimura (1996)] is indicated in parentheses. New compounds have been discovered more recently and were reviewed in Huneck (2001).

Figure 1.10 Proposed biosynthetic routes to orsellinic acid [1], β-orsellinic acid [2], and methylphloroacetophenone [3] via the polyketide pathway. PKS domains involved in the biosynthesis are KS, ketosynthase; AT, acyltransferase; ACP, acyl carrier protein, and CYC, Claisen cyclase. The three monoaromatic units formed the basis of various coupled phenolics present in lichens.

Figure 1.11 Proposed biosynthesis of tetraketide [4] and pentaketide [5] alkyl-resorcylic acids. Adapted from (Funa et al. 2007).

Figure 1.12 Examples of lichen depside/depsozones with o xoalkylresorcylic acid moieties, in which the keto group in the alkyl chain has not originated from the polyketide pathway. A) Depside (miriquidic acid) with a 3'-oxoresorcylic acid moiety and depsozone (loxodin) with a 1'-oxoresorcylic acid moiety. B) A biosynthetic step in aflatoxin biosynthesis involving a monooxygenase and a reductase according to Yabe & Nakajima (2004).

Figure 1.13 Proposed biosynthesis steps for ester bond formation in depsides.
Figure 1.14 Coupling of phenols via one-electron oxidation to form various types of biophenyls and diphenyl ethers. Box: formation of phenoxy radical (B) and its mesomeric forms (A and C) via one-electron oxidation. Adapted from Dewick (2002).

Figure 1.15 Common β-orsellinic acid derivatives incorporated in depsides and depsidones with the methyl groups at C-3 and C-6 subjected to increasing level of oxidation. In grey, theoretically possible β-orsellinic acid derivatives but seldom or yet to be found in lichen depsides and depsidones. Adapted from Culberson (1969).

Figure 1.16 Biosynthesis of haematommmic acid and its incorporation into atranorin. Adapted from Vicente et al. (2003).

Figure 1.17 Post-PKS modifications of the aromatic units of chloroatranorin (A) and fumarprotocetraric acid (B).

Figure 1.18 Labelling pattern in A) glyphoric acid and B) atranorin. 14C from malonyl-CoA (●) and formic acid (△). (Mosbach 1964; Yamazaki & Shibata 1966).

Figure 1.19 Proposed biosynthetic routes to meta- and para-scrobiculin. Route A), hydroxylation of nordivaricatic acid at C-3’ leads to para-scrobiculin; route B) hydroxylation of divaric acid followed by intermolecular esterification to form scrobiculin. Acyl migration results in interconversion of para- and meta-scrobiculin.

Figure 1.20 Basic routes of enzymatic hydroxylation of aromatic compounds via dioxygenase (DO), monoxygenase (MO), and peroxidase (PO). DH, dehydrogenase; Re, rearrangement. (1) aromatic substrate, (2) cis-dihydrodiol, (3) catecholic product, (4) epoxide intermediate, (5) phenolic product, (6) cis,trans-dihydrodiol, (7) hypothetical dioxetane intermediate. Adapted from Ullrich and Hofrichter (2007).

Figure 1.21 Hydroxylation of a phenolic substrate (8) by A) tyrosinase (Tyr) and by B) a laccase (Lacc) or phenol-oxidizing peroxidases (PO) indirectly. (3) catechol product, (9) o-benzoquinone product, (10) two mesomeric forms of the phenoxy radical that can disproportionate to a cyclodienone cation (11) and a phenolic molecule (8), (12) unstable hydroxycyclohexadienone tautomer rearranging to a hydroquinone (13) or p-benzoquinone (14) product. Adapted from Ullrich and Hofrichter (2007).

Figure 1.22 Two structurally related depsides isolated from plants – Papaver rhoeas, Myrciaria cauliflora and Vaccinium marocarpon.

Figure 1.23 Proposed phenol oxidative coupling of lichen depside (olivetoric acid) to form depsidone (physodic acid). Based on Culberson (1964).

Figure 1.24 Oxidation of a synthetic depside to diploicin with manganese dioxide. Box, tumidulin is a depside structurally related to diploicin. Based on Brown et al. (1961).

Figure 1.25 Proposed depsidone (diploicin) biosynthesis via a benzophenone intermediate. Adapted from Elix et al. (1984b).

Figure 1.26 Different polyketide folding patterns that would produce a benzophenone moiety but not the benzophenone that is crossed. Box, a carboxyl group at the benzophenone moiety that would be required for production of most depsidone carboxylic acids via the mechanism illustrated in Figure 1.25.

Figure 1.27 Proposed biosynthesis of divaronic acid (depsidone) via a meta-depside intermediate. Based on Elix et al. (1987).

Figure 1.28 Proposed mechanism for biosynthesis of depsidones from a para-depside precursor mediated by a dioxygenase. NADH is thought to act as a hydrogen donor (in litt. to J.A. Elix).

Figure 1.29 Excelsione and the structurally related lichen depsidones – variolaric acid and
cryptostictic acid.

**Figure 1.30** Proposed biosynthetic relationship between the depsidones – loxodin [1], norlobaridone [3], and the diphenyl ethers – loxodinol [2], isonorlobaridone [4] and norlobariol methylpsuedoester [5]. [5] could be obtained directly from norlobaridone by methanolysis.

**Figure 1.31** Proposed Smiles rearrangement of prasinic acid to micareic acid.

**Figure 1.32** Proposed phenol oxidative coupling of methylphloroacetophenone to form leprolomin (diphenyl ether), usnic acid (dibenzofuran-like) and contortin (biphenyl).

**Figure 1.33** Proposed biosynthetic relationship between picrolichenic acid (depsone), colensoic acid (depsidone), epiphorelic acid I and perlatolic acid (depside). 2-O-methylanziaic acid is the likely precursor of picrolichenic acid, while anziaic acid is a possible precursor of colensoic acid and epiphorelic acid I.

**Figure 1.34** Structure of friesiic acid and 2'-O-methylfriesic acid.

**Figure 1.35** Biosynthesis of usnic acid via phenol oxidative coupling follows by a nucleophilic attack of a phenol group on to an enone system (Dewick 2002).

**Figure 1.36** Alternative nucleophilic attack of the two phenol groups to the two available enone systems (Taguchi et al 1966). Route (a) lead to formation of usnic acid; (b) lead to (B) iso usnic acid; (c) and (d) lead to (C) and (D) respectively. Box: numbering of the methylphloroacetophenone unit used in this figure.

**Figure 1.37** Proposed biosynthesis for A) porphyric acid and B) strepsilin. Adapted from Mosbach (1969).

**Figure 1.38** Lichen biphenyl (contortin) and diphenylmethane [bis-(2,4-dihydroxy-6-n-propylpheny)methane].

**Figure 1.39** The hierarchy of variations introduced at each stage of the biosynthesis of lichen coupled phenolic compounds. The variations introduced at each level increase the diversity by several folds.

**Figure 1.40** Comparison of the sterigmatocystin (ST), dothistromin (DOT) and aflatoxin (AF) gene clusters and their corresponding metabolic products. Genes predicted to be homologous and have similar functions are in the same colour. Putative AflR binding sites in promoter regions of the genes are indicated by stemmed balls. Figure reproduced from Zhang (2000).

**Figure 1.41** Biosynthesis of sterigmatocystin, aflatoxin B1, and dothistromin via the common precursor – norsolorinic acid. The homologous PKS genes involve in the biosynthesis of norsolorinic acid are stcA (sterigmatocystin), pksA (dothistromin), and aflC (aflatoxin B1).

---

**CHAPTER 2**

**Figure 2.1** Standard touchdown PCR cycle conditions. Ta, optimal annealing temperature.

---

**CHAPTER 3**

**Figure 3.1** Genealogy of type I PKSs, inferred by maximum parsimony analysis of the KS domain. Taken from Kroken et al. (2003). KS, ketosynthase; AT, acyl transferase; DH, dehydratase; ER, enoyl reductase; KR, ketoreductase; ME, C-methyltransferase; PP, phosphopanteinylation site (acyl carrier protein).

**Figure 3.2** The locations where the lichen specimens were collected across Victoria, Australia. For a more detailed map of the collecting locations, see Google Map link:
Figure 3.3  The lichens collected from Mount Waverley (MW), Wilsons Promontory National Park (WP), Wyperfeld National Park (WF), and Melbourne central business district (CBD). Figure continues next page.

Figure 3.4  The chemical structures of compounds detected in the lichen specimens (Table 3.1).

Figure 3.5  Multiple sequence alignment of NR-PKSs (in bold) and the conserved regions corresponding to NRKS-F/R and LC1/2c primers. Col, Colletotrichum lagenarium; Pep, Penicillium patulum; Asn, Aspergillus nidulans; Asf, A. fumigatus; Asp, A. parasiticus; Ast, A. terreus. 6MSAS (a PR-PKS) and LNKS (a HR-PKS) are outgroups for comparison.

Figure 3.6  Multiple sequence alignment of PR-PKSs (in bold) and the conserved regions corresponding to PRKS-F/R and LC3/5c primers. Stv, Streptomyces viridochromogenes; Mie, Micromonospora echinospora. WA (a NR-PKS) and LNKS (a HR-PKS) are outgroups for comparison.

Figure 3.7  Gel electrophoreses of the PCR products amplified using A) NRKS-F/R primers and B) PRKS-F/R primers. The numbering of the wells corresponds to the collection number for the lichens specimens in Figure 3.3 and Table 3.3. Lane 1 = F. caperatulla, 2 = T. scutellatum, 6 = P. jackii, 7 = F. soredians, 8 = F. caperatulla, 11 = X. semiviridis, 12 = P. chinense, 14 = P. reticulatum, 15 = P. borreri, 16 = P. cunninghamii, 21 = H. revoluta, 23 = F. caperatulla, 27=28 = U. oncodeoides, 30 = Cladonia sp..

Figure 3.8  Multiple sequence alignment of the clade III NR-type KS domains and the corresponding sites of NR3KS-F/R primers and other KS primers (NRKS-F/R, PRKS-F/R, LC1). Representative sequences from the HR-, PR- and non-clade III NR-PKSs are shown as outgroups to highlight the differences in sequences and the selectivity of the KS degenerate primers.

Figure 3.9  KS domain phylogeny inferred from Maximum Evolution (ME) method (bootstrap = 1000 replicates). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test were shown next to the branches. The evolutionary distances are in the units of the number of amino acid substitutions per site. Classification of PKSs are based on Kroken et al. (2003).

Figure 3.10  Proposed biosynthesis of zearalenone by Gibberella zeae PKS13 (a NR-PKS) and PKS4 (a HR-PKS). The highly-reduced polypeptide chain produced by PKS4 is in grey. Box: aldol cyclization of the tetrapetide moiety that is analogous to orsellinic acid biosynthesis.

Figure 3.11  Examples of lichen compounds formed from highly-reduced polyketides. * A homodimeric compound with two identical polyketide chains; ** A homotrimeric compound with three identical polyketide chains.

CHAPTER 4

Figure 4.1  Southern blotting using FcNRKS1 probe. Left, digested (BamHI, EcoRI, HindIII) and undigested (lane unct) genomic DNA electrophoresed on a 0.7% agarose gel along with the FcNRKS1 PCR product (lane PCR) as positive control. Right, chemiluminescent detection of the corresponding blot.

Figure 4.2  Southern blotting using FcNR3KS1 probe. Left, same as Figure 4.1 (left). Right, chemiluminescent detection of the corresponding blot (striped and reprobed with FcNR3KS1 probe).

Figure 4.3  PCR screening of the F. caperatulla subgenomic library with NR3KS-F/R primers: A) PCR products from pooled plasmids of 12 colonies each lane (IA-IH, IIA-IID), B) PCR products from the individual 12 colonies from group G (lane 1-12).
Figure 4.4  The approximate location of FcNR3KS1-iF primer binding site on the *F. caperatulla* NR3KS1 KS domain.

Figure 4.5  Southern blots of *X. semiviridis* DNA hybridized with XsNR3KS1 and XsNR3KS2 probes. A) Digested genomic DNA electrophoresed on a 0.7 % agarose gel. B) blot probed with XsNR3KS1, C-D) blot probed with XsNR3KS2, band targeted for cloning/library construction is circled. Ba, BamHI; Bg, BglII; Ec, EcoRI; H, HindIII; P, PstI; X, XhoI; BaH, BamHI+HindIII, BgX, BglII+XhoI; EcP, EcoRI+PstI.

Figure 4.6  The plasmid map of pUC52KS containing a partial PKS gene (xsepks1) from *X. semiviridis*. The small arrows indicate the approximate binding sites for the specific primers XsNR3KS2-F/R. Bent arrows indicate the putative translation start sites and transcriptional directions.

Figure 4.7  Potential promoter elements found in the intergenic region between xsepks1 and a putative oxydoreductase. Numbering started from the BamHI site of the pUC52KS insert.

Figure 4.8  Design and relative binding positions of the NR3MeT-F and -R primers.

Figure 4.9  CMeT domain PCR products of *X. semiviridis* amplified with NR3MeT-F and -R primers. Lane 1-2 are duplicates. The arrow indicate the size of the PCR products (~380 bp).

Figure 4.10  Southern blots of *X. semiviridis* hybridized with A) XsNRMeT mixed probe, and B) XsNR3KS2 probe. Grey dotted arrows indicate the bands of similar sizes detected in both blots. Ba, BamHI; Bg, BglII; Ec, EcoRI; H, HindIII; P, PstI; X, XhoI; BaH, BamHI+HindIII, BgX, BglII+XhoI; EcP, EcoRI+PstI.

Figure 4.11  Overlapping regions of 5’xsepks1 (pUC52KS insert), KS-CMeT PCR product (4.5 kb), and XsNRMeT1 PCR product/probe.

Figure 4.12  Comparison of the KS (left) and CMeT (right) domain phylogeny of clade III NR-PKSs. The evolutionary relationships were inferred using the Minimum Evolution (ME) method described in Section 3.5.1.2. Dotted lines linked the KS and CMeT domains that are part of the same PKS protein. Scales indicate the number of amino acid substitutions per site.

Figure 4.13  Southern blot of *X. semiviridis* DNA hybridized with KS-CMeT probe. NE, NotI-EcoRI; XE, XbaI-EcoRI; XE, XbaI-HindIII.

Figure 4.14  Overlapping region of 5’xsepks1 in pUC52KS and KS-CMeTprobe, and a NotI site predicted at about 12 kb downstream of the EcoRI site (2144).

Figure 4.15  Plasmid pJpks-ID5 isolated from a *X. semiviridis* partial genomic library. The small black arrows represent the approximate binding sites of the primers used for sequencing the pJpks-ID5. // represents region that has not been sequenced.

Figure 4.16  Multiple sequence alignment of the catalytic cores of the minimum PKS functional domains - β-ketoacyl synthase (KS), acyltransferase (AT), and acyl carrier protein (ACP). The proposed key amino acids in the active sites are shaded.

Figure 4.17  Multiple sequence alignment of the CMeT domains in XsePKS1 and other known PKSs showing the conserved motifs as defined by (Kagan & Clarke 1994; Miller et al. 2003). Proposed key amino acids are indicated by (*)

Figure 4.18  Multiple sequence alignment of the SAT domains in XsePKS1 and other known PKSs showing the catalytic triads (*) as proposed by Crawford et al. (2006).

Figure 4.19  Comparison of the PKS domain architecture of XsePKS1 and other PKSs. The first four PKSs belong to the clade III NR-type. PKS domains: SAT, starter unit-ACP transacylase; KS, β-ketoacyl synthase; AT, acyltransferase, ACP, acyl carrier protein; CMeT, C-methyltransferase; R, reductase; CYC,
Claisen cyclase/thioesterase; Est?, esterase/lipase (COG0657); DH, dehydratase; KR, ketoreductase; ER, enoyl reductase; C, condensation.

Figure 4.20  Assembled nucleotide sequence combining pUC52KS insert and pJpks-ID5. ORFs that showed homology to protein sequences in the GenBank database are indicated as blocks. Ambiguous or incomplete ORFs are shaded in grey. A gap of 700 bp estimated size is remained unsequenced. RT, reverse transcriptase.

Figure 4.21  The three-frames (positive) translational map of the 3’ end region of xsepks1; the shaded amino acids indicate regions that showed homology to the reductase (R) domain of PKSCT at different frames (based on BLASTx result). Position 1 in the map corresponds to 6220 bp downstream from the putative start codon of xsepks1. The putative stop codon (underlined) is located at position 337 in this map.

Figure 4.22  Phylogenetic analysis of individual PT domains. NSAS, norsolorinic acid synthase; STS, sterigmatocystin synthase; THNS, tetrahydroxynaphthalene synthase; WAS, YWA1 naphthopyrone synthase. Unmodified figure taken from Cox (2007).

Figure 4.23  The PT domain phylogeny inferred using the Maximum Parsimony method (Eck & Dayhoff 1966). The bootstrap consensus tree (mid-point rooted) inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed (Felsenstein 1985). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (Felsenstein 1985). The MP tree was obtained using the Close-Neighbor-Interchange algorithm (Nei & Kumar 2000) (pg. 128) with search level 3 in which the initial trees were obtained with the random addition of sequences (10 replicates). The tree is drawn to scale, with branch lengths calculated using the average pathway method (Nei & Kumar 2000) (pg. 132) and are in the units of the number of changes over the whole sequence. All positions containing gaps and missing data were eliminated from the dataset (Complete Deletion option). There was a total of 276 positions in the final dataset, out of which 260 were parsimony informative. Phylogenetic analyses were conducted in MEGA4 (Tamura et al. 2007).

Figure 4.24  Citrinin biosynthetic gene cluster in Monascus purpureus (AB243687) (Shimizu et al. 2007) and its parrellel to xsepks1 and associated genes.

CHAPTER 5

Figure 5.1  The plasmid map of pUC52KS and the approximate location of the binding sites for M13 forward, Gap1-F, RT52KS-F and RT52KS-R primers.

Figure 5.2  A. nidulans transformants growing on glufosinate-containing selective medium. Two transformants (a10 and a22) in the first plate carried the 5’ xsepks1.

Figure 5.3  PCR analysis of A. nidulans transformants a10 (lanes 1-2) and a22 (lanes 3-4) using the universal M13/pUC forward primer (lanes 1 & 3) or GAP-1F forward primer (lanes 2 & 4) with RT52KS-R as reverse primer.

Figure 5.4  Transcription analysis of a10 and a22 transformants with RT52KS-F/R primers flanking the putative intron. Lanes 1-3, transformant a10; lanes 4-6, transformant a22; lanes 1 & 4, cDNA; lanes 2 & 5, genomic DNA; lanes 3 & 6, total RNA as negative control.

Figure 5.5  Conversion of pTAex3 vector into a GATEWAY destination vector, pTAex3R.

Figure 5.6  The plasmid map of pTA-Cse3 containing the xsepks1 ORF. Small arrows indicate the binding sites of the primers used in 3’ RACE.
Figure 5.7  The differences between the one-step and two-steps RT-PCR. 1) PCR from genomic DNA: intron present in the PCR product, 2) one-step RT-PCR: both poly-A tailed mRNA and truncated mRNA were amplified, 3) two-steps RT-PCR: 3’-CDS A primer selectively reverse transcribed poly-A tailed mRNA; PCR in the second step with gene-specific primers amplified only from the xsepks1 first strand cDNA.

Figure 5.8  A. oryzae transformants growing on Czapek-Dox (CD) agar medium.

Figure 5.9  Gel electrophoresis of 3’ RACE products of transformants 5 (T5) and 6 (T6).

Figure 5.10  Electrophoreograms of the 3.1 kb 3´ RACE products: A) transformant 5, B) transformant 6. Mixed peaks were observed in both sequences after “TCCTAG”.

Figure 5.11  Gel electrophoresis of nested PCR products of 3´ cDNA ends.

Figure 5.12  Sequencing of 3´cDNA ends (1.05 kb 3´RACE product, Figure 5.11) showing splicing of an intron. (Top): pairwise alignment of xsepks1 3´ cDNA end sequence to genomic sequence; (Bottom): electrophoretogram showing no mixed sequence was observed after the “TCCTAG” region.

Figure 5.13  Pairwise alignment of the XsePKS1 transcripts: 1) with and 2) without the intron2 removed. Arrows and (*) indicates the position of stop codons.

Figure 5.14  HPLC analyses of medium extracts from A. oryzae transformants. Bottom: UV spectra of the peaks at Rt A) 10.0, B) 12.6 and C) 24.0 min monitored at 240 – 350 nm.

Figure 5.15  HPLC chromatograms of extracts from transformant 1 (a, 1st fermentation; b, 2nd fermentation) and transformant 5, analyzed in J.A. Elix’s lab.

Figure 5.16  UV spectra of compounds B, C and D in the extracts of A. oryzae transformants (Figure 5.15) and the library search results.* The structure of methylisoplacodiolic was not resolved, but a CH2 is known to present in addition to the isoplacodiolic acid based on mass spectroscopy (J.A. Elix, pers. comm.).

Figure 5.17  UV spectra of compounds E and F found in transformants 3, 4 and 5. Bottom: HPLC chromatogram of the extract from the A. oryzae control strain carrying the empty pTAex3 plasmid; compounds E and F were also detected in the extract.

Figure 5.18  HPLC analyses of extracts from the medium and mycelia of transformant 6 culture. Bottom: UV spectrum and library search results of compound G.

Figure 5.19  Orange pigmentation appeared in the mycelium of A. oryzae transformant 1.

Figure 5.20  HPLC analyses of the methanol extract from the X. semiviridis thalli and the structural formula of the three major polyketide compounds detected - succinprotocetraric acid [1], fumarprotocetraric acid [2], and usnic acid [3]. The retention time (Rt) for [1], [2] and [3] were 20.6 min, 22.4 min and 28.2 min respectively. [1] and [2] are originate from β-orsellinic acid, while [3] is from methylphloroacetophenone.

Figure 5.21  Summary of results for expression of xsepks1 in A. oryzae and factors that might have affected the outcome.

Figure 5.22  Structural formula of Ochratoxin A. Box, citrinin-like moiety; in grey, phenylalanine moiety.

Figure 5.23  A PKS gene cloned from the lichen Cladonia grayi. The small triangles indicate the intron sites (D. Armaleo, unpublished data). KS, ketosynthase; AT, acyltransferase; ACP, acyl carrier protein; MT, methyltransferase; CYC, Claisen cyclase, P-450, cytochrome p450 protein. Instead of producing β- orcinol or methylphloroacetophenone derivatives, C. grayi produced the orcinol depside (4-O-demethylsphaerophorin) and orcinol depsidones (grayanic and 4-O-demethylgrayanic acids) in its natural thalli and mycobiont.

XVII
Since a CMeT domain is not required for biosynthesis of orcinol depsides and depsidones, this PKS might be responsible for the biosynthesis of other undetected metabolites.

Figure 5.24 Proposed biosynthetic relationship between β-orsellinic acid, methylphloroacetophenone, 3-methylorcinaldehyde and 3,5-dimethylorsellinic acid. Note: The methylation by CMeT might occurs after cyclization as proposed by Cox (2007).
# List of Tables

## CHAPTER 1

<table>
<thead>
<tr>
<th>Table 1.1</th>
<th>The major classes of lichen secondary metabolites based on biosynthetic pathways. Classified according to (Elix 1996).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.2</td>
<td>Antibiotic and antifungal activities of some lichen compounds.</td>
</tr>
<tr>
<td>Table 1.3</td>
<td>Antiviral activities of some lichen compounds.</td>
</tr>
<tr>
<td>Table 1.4</td>
<td>Antitumour and antimutagenic activities of some lichen compounds.</td>
</tr>
<tr>
<td>Table 1.5</td>
<td>Enzyme inhibitory activities</td>
</tr>
<tr>
<td>Table 1.6</td>
<td>The distribution of β-orsellinic acid derivatives from Figure 1.15 into various depsides and depsidones. The assignment of A and B rings is according to Figure 1.9. Decarboxylated units are indicated by (-CO2) after the number of the unit. Table adapted from Culberson (1969).</td>
</tr>
<tr>
<td>Table 1.7</td>
<td>Examples of lichen depsides from coupling of various aromatic units via ester bond. (continue next page)</td>
</tr>
<tr>
<td>Table 1.8</td>
<td>Lichen depsidones. Coupling of various aromatic units via an ester bond and an ether bond.</td>
</tr>
<tr>
<td>Table 1.9</td>
<td>Lichen diphenyl ethers. Coupling of aromatic units via ether bond.</td>
</tr>
<tr>
<td>Table 1.10</td>
<td>Dibenzofurans and usnic acid homologs. Coupling of aromatic units via carbon-carbon and ether linkages.</td>
</tr>
</tbody>
</table>

## CHAPTER 2

<table>
<thead>
<tr>
<th>Table 2.1</th>
<th>The list of primers used in this study, with their corresponding sequences, in the order of appearance in this thesis.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.2</td>
<td>GenBank accession number for reference sequences used in this study.</td>
</tr>
</tbody>
</table>

## CHAPTER 3

<table>
<thead>
<tr>
<th>Table 3.1</th>
<th>Compounds detected in the lichens collected by TLC and/or HPLC (*). Colour indicates the compound classes: monoaromatics, depsides, depsidones, dibenzofurans, anthraquinones. Major compounds detected in HPLC are in bold.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3.2</td>
<td>Closest BLASTn matches of the internal transcribed spacer (ITS) 1, 5.8s ribosomal RNA gene, and ITS 2 of selected lichens. *The expected (E) value for all the closest Blast hit = 0.0</td>
</tr>
<tr>
<td>Table 3.3</td>
<td>PCR products obtained from the lichen DNA using LC1/2c, LC3/5c, NRKS-F/R and PRKS-F/R primers. PCR products that yielded clean sequences are shaded; those that yielded mixed sequences or failed the sequencing are unshaded. (-) indicates no PCR products were obtained; blank indicates no PCR reaction was carried out.</td>
</tr>
<tr>
<td>Table 3.4</td>
<td>PCR products obtained using the NR3KS-F/R primers and the closest homolog among the PKSs analysed in Kroken et al. (2003).</td>
</tr>
<tr>
<td>Table 3.5</td>
<td>Characterised polyketide synthases in the non-reducing (NR) clade I and their corresponding products.</td>
</tr>
<tr>
<td>Table 3.6</td>
<td>Lichen KS domains that grouped with clade I NR-PKSs and compounds detected in the corresponding lichens.</td>
</tr>
<tr>
<td>Table 3.7</td>
<td>Characterised polyketide synthases in the non-reducing (NR) clade II and their corresponding products.</td>
</tr>
<tr>
<td>Table 3.8</td>
<td>Lichen KS domains that grouped with clade II NR-PKSs and compounds detected in the corresponding lichens.</td>
</tr>
<tr>
<td>Table 3.9</td>
<td>Characterised polyketide synthases in the non-reducing (NR) clade III and their corresponding products.</td>
</tr>
<tr>
<td>Table 3.10</td>
<td>Lichen KS domains that grouped with clade III NR-PKSs and compounds detected in the corresponding lichens.</td>
</tr>
</tbody>
</table>

### CHAPTER 4

| Table 4.1 | The various approaches used previously in locating and cloning of fungal PKS genes. | 156 |
| Table 4.2 | Comparison of lichen KS and CMet domains PCR results. | 187 |
List of Abbreviations

ACD  Advanced Chemistry Development Inc.
ACP  acyl carrier protein (domain)
AMP  adenylation (domain)
AT   acyltransferase (domain)
BLAST/Blast basic local alignment search tool
cDNA complementary DNA
CMet C-methyltransferase (domain)
CoA  coenzyme A
COG  cluster of orthologous genes
CON  condensation (domain)
CYC  Claisen cyclase (domain)
DH   dehydratase (domain)
DHN  1,8-dihydroxynapthalene
DIG  digoxygenin
DNA  deoxyribonucleic acid
EBI  European Bioinformatics Institute
EBV  Epstein-bar virus
ER   enoyl reductase (domain)
Est  predicted esterase (domain)
FAS  fatty acid synthase
GE   gel electrophoresis
HIV  human immunodeficiency virus
HPLC high performance liquid chromatography
HR-PKS highly-reducing polyketide synthase
ITS  internal transcribed spacer
JGI  Joint Genome Institute
KR   ketoreductase (domain)
KS   $\beta$-ketoacyl synthase/ketosynthase (domain)
MEGA Molecular Evolutionary Genetic Analysis software
MIT  Massachusetts Institute of Technology
mRNA messenger RNA
MSAS 6-methylsalicylic acid synthase
NAD  nicotinamide adenine dinucleotide
NADP nicotinamide adenine dinucleotide phosphate
NCBI National Center for Biotechnology Information
NR-PKS | non-reducing polyketide synthase  
NRPS | non-ribosomal peptide synthase  
NUP | Nested Universal Primers  
OAS | orsellinic acid synthase  
ORF | open reading frame  
PAGE | polyacrylamide gel electrophoresis  
PCR | polymerase chain reaction  
PFGE | pulse-field gel electrophoresis  
PKS | polyketide synthase  
PPTase | phosphopanteinyl transferase  
PR-PKS | partially-reducing polyketide synthase  
PT | product template (domain)  
R/R-domain | reductase (domain)  
RMIT | Royal Melbourne Institute of Technology  
RACE | rapid amplification cDNA ends  
RNA | ribonucleic acid  
RT-PCR | reverse transcriptase-polymerase chain reaction  
SAT | Starter-unit acyltransferase (domain)  
SDS | sodium dodecyl sulphate  
SMART | Switching Mechanism At 5’ end of the RNA Transcript  
TE | thioesterase (domain)  
THN | tetrahydroxynaphthalene  
TLC | thin layer chromatography  
UPM | Universal Primer Mix  
UV | Ultraviolet  
WGS | whole genome sequencing

Genetic and amino acid codes is according to the IUPAC-IUB. International System (SI) of Units are adopted. Standard symbols are used for chemical drawings and names.
Publications

Parts of the work described in this thesis appeared in the following publications:

1. Journal publication by invitation – Mycological Research special issue “Fungal Secondary Metabolite Research”


2. Conference abstract/poster

Abstract

Lichens produce a diverse array of bioactive secondary metabolites, many of which are unique to the organisms. Their potential applications, however, are limited by their finite sources and the slow-growing nature of the organisms in both laboratory and environmental conditions. This thesis set out to investigate polyketide synthase genes in lichens, with the ultimate goal of providing a sustainable source of lichen natural products to support these applications.

To expand the diversity of PKS genes that could be detected in lichens, new degenerate primers targeting ketoacylsynthase (KS) domains of specific clades of PKS genes have been developed and tested on various lichen samples. Using these primers, 19 KS domains from various lichens were obtained. Phylogenetic analysis of the KS domains was used to infer the function of the PKS genes based on the predicted PKS domain architecture and chemical analysis by TLC and/or HPLC. KS domains from PKS clades not previously known in lichens were identified; this included the clade III NR (non-reducing)-PKSs, PR (partially reducing)-PKSs and HR (highly reducing)-PKSs. The discovery of clade III NR-PKSs with C-methyltransferase (CMeT) domain and their wide occurrence in lichens was especially significant. Based on the KS domain phylogenetic analysis and compounds detected in the individual lichens, the clade III NR-PKSs were hypothesized to be involved in the biosynthesis of β-orsellinic acid and methylphloroacetophenone – the monoaromatic precursors for many lichen coupled phenolic compounds, such as β-orcinol depsides/depsidones and usnic acid.

A strategy has been developed to isolate clade III NR-PKSs directly from environmental lichen DNA using clade III NR-type KS amplified from the degenerate primers (NR3KS-F/R) as homologous probes. Another pair of degenerate primers specific to the CMeT domain of NR-PKSs has also been developed to facilitate the cloning and probing of new clade III NR-PKS genes in lichens. A clade III NR-PKS gene (xsepks1) from X. semiviridis was cloned successfully. This is the first report of the isolation of a full-length PKS gene from environmental lichen DNA. The domain architecture of xsepks1 is KS-AT-ACP-CMeT, as expected for a clade III NR-PKS, suggesting that the newly developed clade-specific primers are useful for cloning
new clade III NR-PKS genes and that KS domain phylogenetic analysis could predict the functional domains in PKSs.

Attempts were made to characterize the function of xsepks1 by heterologous expression in Aspergillus species. Both A. nidulans (transformed with 5´partial xsepks1 including native promoter) and A. oryzae (transformed with full-length xsepks1 under the regulation of starch-inducible amyB promoter) were tested as potential hosts for the expression of lichen PKS genes. Transcriptional analysis showed that A. nidulans could potentially utilize the lichen PKS gene promoter and both fungal hosts could splice the introns of a lichen PKS gene. Several compounds unique to the A. oryzae transformants carrying xsepks1 were detected, but they could not be reproduced in subsequent fermentations even though the gene was transcribed into mRNA. None of the expected products (β-orsellinic acid, methylphloroacetophenone or similar methylated monoaromatic compounds) was detected in A. oryzae transformants, and the function of xsepks1 remains to be determined. The other clade III NR-PKS genes detected in X. semiviridis could also be responsible for the biosynthesis of β-orsellinic acid or methylphloroacetophenone, as precursors of the major secondary metabolites detected in X. semiviridis (i.e. fumarprotocetraric acid, succinprotocetraric acid and usnic acid).

Overall, the work in this thesis demonstrated the prospect of using a molecular approach to access the lichen biosynthetic potential without going through the cumbersome culturing stage.
CHAPTER 1 INTRODUCTION AND BACKGROUND

1.1 Introduction .............................................................................................................. 2

1.2 Background .............................................................................................................. 5

1.2.1 The lichens ........................................................................................................... 5

1.2.2 Chemistry and lichens ......................................................................................... 7

1.2.3 Lichen secondary metabolites .............................................................................. 9

1.2.3.1 The biological significance of lichen metabolites ........................................... 13

1.2.3.2 Chemotaxonomy of lichens ............................................................................. 13

1.2.4 Prospects and applications of lichen natural products ...................................... 14

1.2.4.1 Pharmacology and medicine .......................................................................... 14

1.2.4.2 Potential applications in agriculture ............................................................... 18

1.2.4.3 Cosmetics and perfume industry ...................................................................... 19

1.2.5 The polyketide factories ....................................................................................... 20

1.2.6 The unique lichen compounds and the ubiquitous PKSs .................................... 27

1.2.6.1 Lichen polyketides common to other fungi ...................................................... 27

1.2.6.2 Polyketide compounds common in lichens ....................................................... 30

1.2.6.2.1 Basic monoaromatic units that form lichen coupled phenolics .................. 33

1.2.6.2.2 The formation of various linkages in lichen coupled phenolics .................. 37

1.2.6.2.3 Derivatization of the basic monoaromatic units ........................................... 39

1.2.6.2.4 Depsides ...................................................................................................... 43

1.2.6.2.5 Depsidones .................................................................................................. 50

1.2.6.2.6 Diphenyl ethers .......................................................................................... 56

1.2.6.2.7 Depsones ..................................................................................................... 61

1.2.6.2.8 Dibenzofurans and usnic acid homologs ..................................................... 63

1.2.6.2.9 Biphenyls ..................................................................................................... 67

1.2.6.3 Lichens as combinatorial chemists .................................................................... 68

1.3 Research scope and objectives ................................................................................. 75
1.1 Introduction

Natural products have played an important role in medicine since the 1940s (e.g. penicillins) and continually serve as an important source and inspiration for new drugs (Newman et al. 2003). Continuous searches for novel chemical entities, by expanding the natural products searched, are essential to combat the adaptability of infectious bacteria and to keep pace of the ever-increasing need for new drugs to cure various diseases. The diverse structures of natural products have been selected over millions of years of chemical evolution and interact with various biological molecules. Consequently, they offer a potentially infinite source of chemical diversity unmatched by many synthetic chemical collections or combinatorial libraries, which often occupied less ‘biologically relevant chemical space’ (Dobson 2004). The research undertaken in this thesis was motivated by such a goal – to expand the natural product resources for pharmaceutical and industrial applications.

Lichen metabolites are some of the potential natural product resources that exhibit manifold bioactivities including antibiotic, antimycobacterial, antiviral, antiinflammatory, analgesic, antipyretic, antiproliferative and cytotoxic effects (Huneck 1999; Muller 2001). Over 1000 lichen metabolites with various biological actions have been reported (Huneck 1999, 2001). Despite the large number of unique lichen compounds with diverse structures and biological activities, their therapeutic potential remains pharmaceutically unexploited (Muller 2001). This is especially true for Australia, where a large diversity of lichens, many unique to the continent, is found in assorted environments and climates, from the alps and rainforest to the deserts. According to Miao (2001), there are two reasons for lack of exploitation of the resources: 1) lichens are slow-growing in nature, and so to harvest large amounts of lichen from nature is neither sustainable nor ecologically sensible; 2) lichens are neither easy to grow in laboratory culture nor to cultivate as crop. Although there have been successful attempts to grow lichen mycobionts in axenic culture, the cultures often do not produce metabolites that resemble compounds in the natural thalli (Crittenden & Porter 1991; Culberson & Armaleo 1992; Huneck 1999; Stocker-Worgotter & Elix 2002; Yamamoto et al. 1994).
Hence, transferring the lichen metabolite-producing genes into an easily grown heterologous host is becoming an attractive alternative approach (Muller 2001), as this might provide a sustainable source of useful lichen metabolites. The recent improved understanding of the secondary metabolic pathways of various microorganisms and advances in molecular genetic techniques has opened up the possibility of accessing the lichen genomes directly for their potential in production of useful secondary metabolites. Cloning of metabolite-producing genes into heterologous hosts would not only enable expression of secondary metabolites that are normally found in lichen thalli but could also reveal silent or cryptic metabolic pathways, such as substances only produced under stress or extreme conditions (Huneck 1999). Production of lichen metabolites in a fast-growing and well-characterized heterologous host also makes fermentation scaling-up and titre improvement easier to achieve when potential therapeutic molecules are identified or more compounds are needed for further screening.

Among the various secondary metabolic pathways in microbes, the genetics and biochemistry of polyketide biosynthesis are especially well-studied. As prior knowledge, especially from other fungi, should be transferable to lichens, it would be reasonable to focus the initial explorations on polyketide pathways in lichens. Furthermore, most unique lichen compounds, including the potential inhibitors of HIV-1 integrase, i.e. depsides and depsidones, which are rarely reported elsewhere (Neamati et al., 1997), are formed via polyketide pathways (Crittenden & Porter 1991; Fahselt 1994). The tendency of polyketide synthase (PKS) genes, which code for the key enzymes involved in polyketide biosynthesis, to cluster with other genes within the same pathway (Keller & Hohn 1997) also eases the cloning and expression of all the biosynthetic genes required to produce the final products in the pathway.

In this direction, the scientists at the Cubist Pharmaceuticals proposed a genetic screening approach employing the polymerase chain reaction (PCR) with degenerate primers to amplify the conserved domain of the PKS genes (Miao et al. 2001). The sequence of the PCR-amplified gene fragment could be used as a “gene tag” to facilitate the discrimination and selection of unique enzymes and avoid duplications by clustering analysis. The candidate genes and the associated biosynthetic gene cluster could be recovered from the respective lichen DNA libraries using the “gene
tags” as probes. This would be followed by expression of the biosynthetic genes in a surrogate host for metabolite production. A patent regarding this approach for isolating lichen biosynthetic genes has been filed (US patent 6297007). Preliminary investigations reported by Miao et al. (2001) suggested that the approach is highly feasible.

Despite the potential benefits and commercial interests underlying the investigation, the progress of research on the molecular biology of lichen secondary metabolism (based on the public information available) has been relatively slow compared with that in other free-living fungi. At the inception of this project in late 2004, there was no complete PKS gene sequence in the public domain (e.g. GenBank database), except for a study on the phylogeny of some lichen polyketide synthase genes inferred from the conserved β-ketoacyl synthase (KS) domain sequences (Grube & Blaha 2003). Understandably, progress was hampered by the difficulties mentioned earlier, i.e. lichens are slow-growing and difficult to culture. There might also be hurdles in obtaining high quality DNA from environmental lichen samples for genetic manipulation. The lack of an established lichen transformation system was also a major limiting factor, as it prevented the characterization of lichen biosynthetic genes using conventional gene disruption and complementation techniques.

This project will investigate the genetics of polyketide biosynthesis in lichenized fungi in an attempt to overcome these difficulties. The ultimate aim is to use molecular genetic technologies to provide a sustainable source of lichen metabolites for natural product discovery and industrial applications. The following sections in this chapter provide the background information to the investigations, including the organisms (lichens), the history of lichen chemistry, the lichen secondary metabolites, and the applications and prospects of the lichen metabolites. The function of polyketide synthases and their role in biosynthesis of characteristic lichen compounds will also be discussed. Lastly, the chapter ends with the research scope and objectives of this thesis.
1.2 Background

1.2.1 The lichens

Before S. Schwendener (1829-1919) discovered the dual nature of lichens, most biologists mistakenly treated lichens as plants and thought they were one organism (Honneger 2000). Hawksworth (1984) described a lichen as a stable, self-supporting, mutualistic symbiosis involving a fungus (the mycobiont) and a microalga and/or cyanobacterium (the photobiont). Like mycorrhizae and plant-associated fungi, lichenization should be appreciated as another successful nutritional strategy of the fungal kingdom.

Lichens are probably the earliest colonizers of terrestrial habitats on the earth, with fossil records tracing back to 400-600 million years ago (Taylor et al. 1995; Yuan et al. 2005). These fossils indicate that fungi developed symbiotic partnerships with photoautotrophs before the evolution of vascular plants. There are about 300 genera and 18000 species of presently recognised lichens, and they account for about 20% of all fungi (Galun 1988). The lichen mycobiont lacks photosynthetic capabilities and obtains carbon sources from the photobiont while manipulating its growth. In return, the mycobiont, with its highly differentiated morphological structures, secures adequate illumination, water, mineral salts and gas exchange for the photobiont. Some lichenologists prefer to treat the relationship as a kind of controlled parasitism by fungi of the algal partner rather than as a mutualistic symbiosis. This is partly because the fungi form haustoria, structures frequently found in pathogenic fungi, which penetrate the photobiont cells and adsorb nutrients (Nash 1996).

The lichenized ‘life-style’ of fungi is now widely accepted as a source of evolutionary innovation that has stimulated an enormous morphological radiation in fungi, mostly in the ascomycetes (Margulis & Fester 1991). Occasionally, and mainly in tropics, the mycobiont may belong to the basidiomycetes (Margulis & Fester 1991). The nomenclature of lichens is derived from the fungal component, since many lichens share the same photobiont and the morphological variation is largely due to the mycobiont.
Figure 1.1. The leaf-like lichen *Xanthoria parietina*. A, Laminal view. B, Detail of a vertical cross-section: uc, conglutinate upper cortex; ph, photobiont layer harboring the globose cells of the green alga *Trebouxia arboricola*; m, gas-filled medullary layer built up by aerial hyphae; lc, conglutinate lower cortex. Adapted from Honegger (1996).

On the basis of their forms and habitats, lichens are traditionally divided into three main morphological groups: crustose, foliose and fructicose (Büdel & Scheidegger 1996). However, this classification is phylogenetically irrelevant and members of a particular genus may belong to more than one of these growth forms. The best-known growth form are the foliose lichens, where the mycobiont is highly organized and has a leaf-like thallus formed by flattened lobes that show various degrees of branching and are only partially attached to the substrate (Figure 1.1A). The mycobiont is differentiated into three morphologically distinct layers analogous to a plant leaf - the upper cortex, medulla, and lower cortex (Figure 1.1B). In contrast to what is commonly perceived, most lichens are crustose, which are less differentiated, lack a cortex at the lower surface and are attached tightly to the substrate. Fructicose lichens, on the other hand, display highly differentiated forms with stems, small branches and leaf-like structures expanded to the third dimensional plane.
Overall the lichen symbiosis is a very successful one, as lichens are found in almost all terrestrial habitats from the tropics and deserts to polar regions (Galloway 1996). As the results of the relationship, both the fungus and alga/cyanobacterium partners, which mostly thrive in relatively moist and moderate environments in free-living form, have expanded into many extreme terrestrial habitats, where they would separately be rare or non-existent (Nash 1996).

1.2.2 Chemistry and lichens

The history of chemical studies in lichens has focused much on the secondary metabolites. The chemistry of lichens has captured the interest of organic chemists dating back to the early 19th century, mainly because of the uniqueness of many aromatic compounds found in lichens. The study of lichen chemistry accelerated in the early 1900’s and the first extensive chemical studies on lichens described about 150 named lichen substances and classified them on the basis of their chemical properties (Zopf 1907). Due to technical limitation at that time, inaccuracies in the structural information were unavoidable.

The study of lichen chemistry was advanced substantially by the work of Asahina and Shibata, who elucidated the structure of many lichen compounds and later summarized the properties of about 200 metabolites (Asahina & Shibata 1971). An innovative microcrystallization method was also developed, mainly by Asahina and Shibata, for rapid identification of lichen compounds. Because of the simplicity, rapidity and accuracy of the test, the technique has subsequently allowed botanists to determine the major constituents of hundreds of species (Culberson 1969). In the 1970’s, the number of known compounds identified in lichens had risen to over 400, including some lichen-specific saccharides and oligopeptides (Culberson 1969, 1970; Culberson et al. 1977), partly due to improvements in spectroscopic and separation techniques.

An interest in chemical variations, among closely related and sometimes morphologically indistinguishable lichens, spurred on the determination of the constituents of many species. Lichen thalline spot tests using standardized reagents,
such as potassium hydroxide and calcium hypochlorite, were developed to allow rapid localization of the secondary products on the thalli (medullary layer or upper cortex) and also give clues as to the chemical nature of the substances (Orange et al. 2001). For example, spotting calcium hypochlorite on the thalli will produce a red colour in the presence of \( m \)-dihydroxyphenols, while it turns green in the presence of dihydroxy dibenzofurans. Although the identities of the compounds could not be confirmed, simple thalline spot tests that differentiated between morphologically similar lichen species were incorporated in many lichen identification keys to assist botanists in identification. Quick and simple methods like paper chromatography, later replaced by thin layer chromatography (TLC), have also assisted many botanists in the identification of lichens through their chemistry. Today, TLC with standardized solvent systems, as described in Orange et al. (2001), is the most rapid and convenient method used by botanists. Improved understanding of major biosynthetic pathways along with some radioactive labelling studies in lichens has also made possible the classification of lichen compounds according to various metabolic pathways (Yamazaki et al. 1965).

A good review of the progress in the study of lichen chemistry from 1907-1999 was provided by Culberson & Culberson (2001). The review showed that the major advance in our understanding of the lichen products was largely due to the development of chromatographic methods (radial chromatography, gas chromatography and high performance liquid chromatography) for their isolation and detection, of powerful instrumental methods (infrared spectroscopy, nuclear magnetic resonance spectroscopy, mass spectrometry) for structural determination and of new synthetic techniques to confirm these structures. Huneck and Leukert in Germany and Elix in Australia, along with a group of organic chemists in Uppsala University in Sweden were some of the prominent contributors to these advances in lichen chemistry, especially in the second half of the 20\(^{th}\) century (Elix 1999). The most recent book on lichen substances summarized about 850 lichen compounds of known structure (Huneck & Yoshimura, 1996), and the number has increased again in the most recent review by Huneck (2001). Today, lichens, particularly the ascolichens, remain the group of organisms with the most extensively studied secondary chemistry (Culberson & Culberson 2001). Interest in discovering new compounds is continuously fuelled by the discovery of biologically active lichen substances.
Understanding the genes coding for biosynthetic pathways to lichen metabolites, however, remained a major challenge that required strong collaboration between lichen chemists, systematists and molecular biologists.

1.2.3 Lichen secondary metabolites

Most of the complex organic compounds isolated and identified from lichens are of fungal origin (Elix 1996). Although free-living green algae and cyanobacteria (such as Nostoc spp.) do produce various secondary metabolites, their lichenization seems to have kept them in relatively simple and undifferentiated forms. The photobiont resources are mainly diverted to photosynthesis, presumably under the control of the mycobiont. In addition, the proportion of the photobiont is generally quite small in a lichen (about 3.1% w/w) (Bednar 1963). The development of axenic culturing methods for lichen mycobionts has also proven that the isolated fungal component can produce the typical lichen metabolites when appropriate nutrients and conditions are provided (Brunauer et al. 2007; Crittenden & Porter 1991; Culberson & Armaleo 1992; Stocker-Worgotter & Elix 2002; Yamamoto et al. 1994).

Most of the secondary metabolites present in lichens are derived from the polyketide pathway (Elix 1996). The key enzymes, polyketide synthases (PKSs), involved in the polyketide biosynthesis, will be discussed in more detail in the next Section (1.2.3). Other secondary metabolic pathways are the shikimic acid and mevalonic acid pathways (Figure 1.2). Based on current understanding of pathways and structures, the lichen secondary metabolites can be further classified into various major classes. (Asahina & Shibata 1971) were the first to propose the classification of lichen secondary metabolites based on the structures and biosynthetic pathways, with some modifications later as the knowledge in the field advanced (Culberson & Elix 1989; Elix 1996) (Table 1.1). Polyketides are built primarily from combinations of acetate (acetyl-CoA) and malonate (malonyl-CoA) units. The shikimic acid pathway provides an alternative route to aromatic compounds, particularly the aromatic amino acids L-phenylalanine, L-tyrosine and L-tryptophan. Mevalonic acid pathways, on the other hand, produce mainly terpenoids, which are derived from C$_5$ isoprene units (Dewick 2002).
Figure 1.2. Probable pathways leading to the major groups of lichen metabolites. Adapted from Elix (1996).
Table 1.1. The major classes of lichen secondary metabolites based on biosynthetic pathways. Classified according to (Elix 1996).

<table>
<thead>
<tr>
<th>Pathways and major classes</th>
<th>Representative compounds</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Polyketide pathway</strong></td>
<td></td>
</tr>
<tr>
<td>1.1 Secondary aliphatic acids, esters and related derivatives</td>
<td>1.1 Protolichesterinic acid</td>
</tr>
<tr>
<td>1.2 Polyketide derived aromatic compounds</td>
<td>1.1 Roccellic acid</td>
</tr>
<tr>
<td>1.2.1 Mononuclear phenolic compounds</td>
<td>1.2.1 Orsellinic acid</td>
</tr>
<tr>
<td>1.2.2 Mononuclear phenolic compounds</td>
<td>1.2.2 β-Orsellinic acid</td>
</tr>
<tr>
<td>1.2.2a Depsides, tridepsides &amp; benzyl esters</td>
<td>1.2.2a Lecanoric acid</td>
</tr>
<tr>
<td>1.2.2b Depsidones &amp; diphenyl ethers</td>
<td>1.2.2a Gyophoric acid</td>
</tr>
<tr>
<td>1.2.2c Depsiones</td>
<td>1.2.2a Sekikaic acid</td>
</tr>
<tr>
<td>1.2.2d Dibenzofurans, usnic acids &amp; derivatives</td>
<td>1.2.2b Salazinic acid</td>
</tr>
<tr>
<td>1.2.3 Anthraquinones and related xanthones</td>
<td>1.2.2d Contortin (biphenyl)</td>
</tr>
<tr>
<td></td>
<td>1.2.2d Didymic acid (dibenzofurans)</td>
</tr>
<tr>
<td></td>
<td>1.2.2d Usnic acid</td>
</tr>
<tr>
<td></td>
<td>1.2.3 Norsolorinic acid (anthraquinone)</td>
</tr>
<tr>
<td></td>
<td>1.2.3 Haemaventosin (naphthaquinone)</td>
</tr>
<tr>
<td></td>
<td>1.2.3 Lichexanthone (xanthone)</td>
</tr>
</tbody>
</table>
### 2. Mevalonic acid pathway

#### 2.1 Di-, sester-, and triterpenes
- 2.1 16α-hydroxykaurane (diterpene)
- 2.1 Retigeric acid (sesterterpene)
- 2.1 Zeorin (triterpene)

#### 2.2 Steroids
- 2.2 Ergosterol

#### 2.3 Carotenoids
- 2.3 Zeaxanthin

### 3. Shikimic acid pathway

#### 3.1 Terphenylquinones
- 3.1 Polyporic acid
- 3.1 Thelephoric acid
- 3.2 Vulpinic acid

#### 3.2 Pulvinic acid derivatives
- 3.2 Calycin

#### 3.3 Epidithiodioxopiperazine (new)
- 3.3 Methylscabrosin
1.2.3.1 The biological significance of lichen metabolites

Production of secondary metabolites is costly to the organisms in terms of nutrient and energy, so one would expect that the plethora of metabolites produced by lichens would have biological significance to the organisms. Recent field and laboratory studies have shown that many of these compounds are indeed involved in important ecological roles (Bjerke et al. 2005; Lawrey 1995; Rikkinen 1995). Some of the possible biological functions of lichen metabolites, as summarized by (Huneck & Yoshimura 1996) are as below:

- Antibiotic activities – provide protection against microorganisms.
- Photoprotective activities – aromatic substances absorb UV light to protect algae (photobionts) against intensive irradiation.
- Promote symbiotic equilibrium by affecting the cell wall permeability of photobionts.
- Chelating agents – capture and supply important minerals from the substrate.
- Antifeedant/ antiherbivory activities – protect the lichens from insect and animal feedings.
- Hydrophobic properties – prevent saturation of the medulla with water and allow continuous gas exchange.
- Stress metabolites – metabolites secreted under extreme conditions.

1.2.3.2 Chemotaxonomy of lichens

The secondary metabolites play an important role in the taxonomy of lichens, especially at the species level. The major metabolites detected in the thallus are often incorporated into monograph describing a lichen taxon (e.g. Calvelo et al. 2005; Louwhoff & Elix 2000; Lumbush et al. 1995). The main reason is that some lichens are difficult to differentiate morphologically, but can be distinguished by their secondary metabolites, which are often present in significant quantities (Huneck & Yoshimura 1996). Many secondary compounds in lichens are also relatively stable, so that even very old herbarium samples may still be used for chemotaxonomic studies (Elix et al. 1984b).
Most morphologically defined species have constant chemistry and normally contain a cortical substance and one or more medullary substances. Three common patterns of chemical variation can be observed in lichens (Huneck & Yoshimura 1996): 1) replacement-type compounds, 2) chemosyndromic variation, and 3) accessory-type compounds. Replacement-type compounds arise in congeneric chemotypes where there is a simple replacement of one or more substances (Elix et al. 1984b). For instance, *Pseudevernia furfuracea* occurs in Europe in two races, containing either oliveroric or physodic acids, while the race from the North America produces lecanoric acid (Huneck & Yoshimura 1996). A chemosyndrome refers to a group of taxonomically related lichens characterised by producing a main compound that is often accompanied by a few biosynthetically closely related minor substances (Huneck & Yoshimura 1996). For example, four species of the *Lecanora subfusca* group (*L. elixii, L. epibryon, L. mayrhoferi* and *L. parmelinoides*) belong to a chemosyndrome with 2,5,7-trichloro-3-O-methylnorlichenxanthone as the major compound (Lumbsch et al. 1994). By contrast, accessory metabolites in lichens occur sporadically and have little or no taxonomic significance (Elix et al. 1984b).

1.2.4 Prospects and applications of lichen natural products

The current applications of lichen natural products are limited to folk medicine and perfume industry in Europe, although numerous studies have showed that many lichen compounds exhibit promising bioactivities with potential applications in medicine, agriculture and cosmetics industry. Their applications on an industrial scale are often hampered by the limited sources due to the slow growth of lichens. This section will briefly discuss the current applications and prospects of lichen natural products/secondary metabolites to underline their importance and potential commercial values.

1.2.4.1 Pharmacology and medicine

For centuries, lichens have been used as folk medicine and their use persists to present day in some parts of the world (Richardson 1988; Schultes & Von Reis 1995). The use of lichens is especially evident and well-documented in traditional Chinese medicine. (Chun 1982) reported 71 species from 17 genera (9 families)
being used for medicinal purposes in China. Lichens from the family Parmeliaceae (17 species from 4 genera), Usneaceae (13 species from 3 genera), and Cladionaceae (12 species from Cladonia) are the most commonly used. In other parts of the world, *Usnea* species are the most commonly utilized, including in India and by the Seminole tribe in Florida. A recent study of the commercial and ethnic uses of lichens in India (Upreti et al. 2003) showed that 38 different species were sold commercially. Most of the lichens are collected from the Western Himalayas and Central and Western Ghats. *Cetraria islandica* (common name: Island moss, but a lichen) has also been used widely to treat various lung diseases and catarrh, and is still sold in various parts of Europe (Elix 1996).

The bioactivities and pharmaceutical potential of lichen metabolites have been reviewed extensively (Boustie & Grube 2007; Huneck 1999; Muller 2001; Yamamoto et al. 2000) and therefore will not be discussed in detail here. In general, the bioactivities detected in lichen metabolites include antibiotic, antimycobacterial, antiviral, antifungal, anti-inflammatory, analgesic, antipyretic, antiproliferative, antitumour and cytotoxic effects. The literature on bioactivities of lichens can be traced back to the 1950s. A review of the earlier work is found in Vartia (1973). The bioactivities of a number of lichen compounds tested in some more recent studies are summarized in Table 1.2-1.5. Besides those small molecule compounds traditionally categorized as secondary metabolites or natural products, the lichen polysaccharides have also shown various bioactivities, including antitumour, immunomodulator, antiviral, and memory enhancement. The studies on the various biological effects of lichen polysaccharides have been covered in a recent review (Olafsdottir & Ingolfsdottir 2001).
<table>
<thead>
<tr>
<th>Compounds</th>
<th>Organisms</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usnic acids and derivatives</td>
<td>Gram +ve bacteria</td>
<td>(Ghione et al. 1988)</td>
</tr>
<tr>
<td></td>
<td><em>Bacteroides</em> spp.</td>
<td>(Lauterwein et al. 1995)</td>
</tr>
<tr>
<td></td>
<td><em>Clostridium perfringens</em></td>
<td>(Conover et al. 1992)</td>
</tr>
<tr>
<td></td>
<td><em>Bacillus subtilis</em></td>
<td>(Fournet et al. 1997)</td>
</tr>
<tr>
<td></td>
<td><em>Staphylococcus aureus</em></td>
<td>(Lawrey 1989)</td>
</tr>
<tr>
<td></td>
<td><em>Staphylococcus</em> spp.</td>
<td>(Proksa et al. 1996)</td>
</tr>
<tr>
<td></td>
<td><em>Enterococcus</em> spp.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Mycobacterium aurum</em></td>
<td>(Ingólfsdóttir et al. 1998)</td>
</tr>
<tr>
<td>Protolichesterinic acid</td>
<td><em>Helicobacter pylori</em></td>
<td>(Ingólfsdóttir et al. 1997)</td>
</tr>
<tr>
<td>Methylorsellinate</td>
<td><em>Epidermophyton flocosum</em></td>
<td>(Fujikawa et al. 1970)</td>
</tr>
<tr>
<td>Ethylorsellinate</td>
<td><em>Microsporum canis</em></td>
<td>(Ingólfsdóttir et al. 1985)</td>
</tr>
<tr>
<td>Methyl β-orsellinate</td>
<td><em>M. gypseum</em></td>
<td>(Hickey et al. 1990)</td>
</tr>
<tr>
<td>Methylhematommate</td>
<td><em>Trichophyton rubrum</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>T. mentagrophytes</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Verticillium achliae</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Bacillus subtilis</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Staphylococcus aureus</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Pseudomonas aeruginosa</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Escherichia coli</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Candida albicans</em></td>
<td></td>
</tr>
<tr>
<td>Alectosarmentin</td>
<td><em>Staphylococcus aureus</em></td>
<td>(Gollapudi et al. 1994)</td>
</tr>
<tr>
<td></td>
<td><em>Mycobacterium smegmatitis</em></td>
<td></td>
</tr>
<tr>
<td>1´-Chloropannarin</td>
<td><em>Leishmania</em> spp.</td>
<td>(Fournet et al. 1997)</td>
</tr>
<tr>
<td>Pannarin</td>
<td><em>Bacillus brevis</em></td>
<td>(Anke et al. 1980a; Anke et al. 1980b)</td>
</tr>
<tr>
<td>Emodin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physcion</td>
<td><em>Drechslera rostrata</em></td>
<td>(Lauterwein et al. 1995)</td>
</tr>
<tr>
<td></td>
<td><em>Alternaria alternata</em></td>
<td>(Raju &amp; Rao 1986a, b)</td>
</tr>
<tr>
<td></td>
<td><em>Aerobic and anaerobic bacteria</em></td>
<td>(Raju et al. 1985)</td>
</tr>
<tr>
<td></td>
<td><em>Pseudomonas</em></td>
<td>(Rao et al. 1989)</td>
</tr>
<tr>
<td>Pulvinic acid and derivatives</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Gram +ve and -ve bacteria</em></td>
<td>(Raju et al. 1985)</td>
</tr>
</tbody>
</table>
### Table 1.3. Antiviral activities of some lichen compounds.

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Viruses or viral enzymes</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depsidones</td>
<td>Human immunodeficiency virus (HIV) integrase</td>
<td>(Neamati et al. 1997)</td>
</tr>
<tr>
<td>Virenic acid and derivatives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Butyrolactone</td>
<td>HIV reverse transcriptase</td>
<td>(Pengsuparp et al. 1995)</td>
</tr>
<tr>
<td>Protolichesterinic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(+) - usnic acid and four other</td>
<td>Epstein-Barr virus (EBV)</td>
<td>(Yamamoto et al. 1995)</td>
</tr>
<tr>
<td>depsides</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emodin, 7-Chloroemodin 7-Chloro-1-O-methylemodin 5,7-Dichl</td>
<td>HIV, cytomegalovirus and other viruses</td>
<td>(Wood et al. 1990)</td>
</tr>
<tr>
<td>Hypericin</td>
<td></td>
<td>(Cohen et al. 1996)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Lavie et al. 1989)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Schinazi et al. 1990)</td>
</tr>
</tbody>
</table>

### Table 1.4. Antitumour and antimutagenic activities of some lichen compounds.

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Activities/cell types</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-) - Usnic acids</td>
<td>Antitumoral effect against Lewis Lung carcinoma P388 leukaemia</td>
<td>(Kupchan &amp; Kopperman 1975)</td>
</tr>
<tr>
<td></td>
<td>Mitotic inhibition</td>
<td>(Takai et al. 1979)</td>
</tr>
<tr>
<td></td>
<td>Apoptotic induction</td>
<td>(Huovinen &amp; Lampero 1989)</td>
</tr>
<tr>
<td></td>
<td>Antiproliferative effect against human HaCaT keratinocytes</td>
<td>(Cardarelli et al. 1997)</td>
</tr>
<tr>
<td>Protolichesterinic acid</td>
<td>Antiproliferative effect against leukaemia cell K-562 and Ehrlich solid tumour</td>
<td>(Bezivin et al. 2004)</td>
</tr>
<tr>
<td>Pannarin, 1’-Chloropannarin Sphaerophorin</td>
<td>Cytotoxic effect against cell cultures of lymphocytes</td>
<td>(Correché et al. 2002)</td>
</tr>
<tr>
<td>Naphthazarin</td>
<td>Cytotoxic effect against human epidermal carcinoma cells</td>
<td>(Paull et al. 1976)</td>
</tr>
<tr>
<td></td>
<td>Antiproliferative effects against human keratinocyte cell line</td>
<td>(Müller et al. 1997)</td>
</tr>
<tr>
<td>Scabrosin esters and derivatives</td>
<td>Cytotoxic effect against murine P815 mastocytoma and other cell lines</td>
<td>(Chai et al. 2004; Ernst-Russell et al. 1999a)</td>
</tr>
<tr>
<td>Euplectin Hydrocarpone</td>
<td></td>
<td>(Ernst-Russell et al. 1999b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Ernst-Russell et al. 2000)</td>
</tr>
<tr>
<td>Salazinic acid</td>
<td>Apoptotic effect against primary culture of rat hepatocytes</td>
<td>(Correche et al. 2004)</td>
</tr>
<tr>
<td>Stitic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psoromic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chrysophanol and emodin</td>
<td>Antiproliferative effect against leukemia cells</td>
<td>(Koyama et al. 1989)</td>
</tr>
<tr>
<td>derivatives</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1.5. Enzyme inhibitory activities

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Enzymes</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atranorin</td>
<td>Arginase</td>
<td>(Legaz &amp; Vicente 1983; Legaz et al. 2001)</td>
</tr>
<tr>
<td>Evernic acid</td>
<td></td>
<td>(Planelles &amp; Estrella Legaz 1987)</td>
</tr>
<tr>
<td>Norsolorinic acid</td>
<td></td>
<td>(Okuyama et al. 1991)</td>
</tr>
<tr>
<td>Physodic acid</td>
<td></td>
<td>(Matsubara et al. 1997)</td>
</tr>
<tr>
<td>Usnic acid</td>
<td></td>
<td>(Kinoshita et al. 2002)</td>
</tr>
<tr>
<td>Divarinol</td>
<td>Tyrosinase</td>
<td>(Matsubara et al. 1997)</td>
</tr>
<tr>
<td>Diphenylmethane derivatives</td>
<td></td>
<td>(Matsubara et al. 1998)</td>
</tr>
<tr>
<td>Lichen whole extracts and compounds</td>
<td>Lipoxygenase</td>
<td>(Ingólfsdóttir et al. 1985)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Ingólfsdóttir et al. 1994)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Ingólfsdóttir et al. 2002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Ogmundsdottir et al. 1998)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Sankawa et al. 1982)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Kumar &amp; Müller 1999a)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Behera et al. 2004)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Higuchi et al. 1993)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Behera &amp; Makhija 2002)</td>
</tr>
</tbody>
</table>

1.2.4.2 Potential applications in agriculture

The success of modern agriculture is, in part, built on the advances in the chemical control of pests. The ability to control weeds, and to prevent insect and microbial damage to crops by chemical means, has played a critical role in the agricultural revolution. However, increasing public concern over the effects of pesticides on the environment and human health has prompted the exploration of natural products as alternatives to the synthetic compounds currently used (Dayan et al. 1999). Lichens interact closely with photosynthetic organisms, i.e. the photobionts (algae) and the substrate to which they attached. They use their arsenal of metabolites to control (promote or inhibit) the growth of the photobionts and to protect themselves against overgrowth by lower and higher plants. Among the lower plants, (-)-usnic acid inhibits the growth of the green alga *Chlamydomonas reinhardtii* (Schimmer & Lehner 1973), while the lichen *Porpidia albocaerulescens* inhibits the growth of the mosses *Hedwigia ciliata* and *Anomodon attenuatum*, and the liverwort *Porella playtyphylla* (Heilman & Sharp 1963). By contrast, sodium rocellate (rocellic acid) promotes the growth of *Ulva lactuca* at $10^{-7}$ M (Garcia et al. 1982).
Many lichen compounds and extracts are also effective growth inhibitors of higher plants. A total of 33 lichen compounds (including aliphatic acids, depsides, depsidones, dibenzofurans, pulvinic acid derivatives and xanthones) was tested for their growth regulatory effects on five higher plants — *Lepidium sativum, Pisum sativum, Avena sativa, Cucurbita pepo and Phaseolus coccineus* (Huneck & Schreiber 1972). In another study, Barbatic acid, 4-O-demethylbarbatic acid, diffirataic acid, evernic acid, lecanoric acid and orsellinic acid inhibited the growth of lettuce seedlings (Nishitoba et al. 1987). Barbatic acid, lecanoric acid and the tridepside gyrophoric acid inhibited photosynthetic system II, like some herbicides such as atrazine and diuron, by interrupting photosynthetic electron transport in isolated chloroplasts (Dayan & Romagni 2001).

The insecticidal and antiherbivory activities of some lichen compounds has also suggested that these secondary metabolites are an important defence mechanism for lichens. For example, caperatic acid and extracts of the lichens *Flavoparmelia baltimorensis* and *Xanthoparmelia cumberlandia* have antiherbivory activities against the snail *Pallifera varia* (Lawrey 1983, 1989), while methyl-β-orsellinate, ethyl hematommate and 5-chlorohematommate show nematocidc activity on larvae of *Toxocara canis* (Ahad et al. 1991). The various bioactivities exhibited by the lichen secondary metabolites suggest that the compounds could potentially be used as an alternative to synthetic herbicides, insecticides and preservatives. More studies on the effect of lichen metabolites on (lower and higher) plants, insects and animals are reviewed in Huneck (1999). A more recent review has highlighted the potential use of lichen metabolites for pest management (Dayan & Romagni 2001).

1.2.4.3 Cosmetics and perfume industry

According to Huneck (1999), two lichen species are being harvested and processed in the perfume industry (mainly at Grasse, France): 1900 tons per year of *Pseudevernia furfuracea* (“tree moss”; 1997 level) and about 700 tons per year of *Evernia prunastri* (“oak moss”, “mousse de chene”; 1999 level). The aromatic esters in the lichens are the components that contribute to the characteristic “green” scent and act as fixatives in the perfume. Since harvesting lichens in large quantity is not sustainable due to their slow growth, alternative methods of producing the metabolites should be considered.
In the cosmetics industry, usnic acid is used as preservative in cosmetic creams (Seifert & Bertram 1995) Many lichen compounds are also photoprotectants that absorb UV light (BeGora & Fahselt 2000) making them suitable for use as UV-protectant in sunscreens. Atranorin, pannarin, gyrophoric acid and usnic acid are also applied in suntan preparations (Fernandez et al. 1996). Atranorin was also proven to be an effective elastase inhibitor (Proksa et al. 1994). Skin elastase is responsible for the sagging and wrinkling of skin and so atranorin might be useful as an anti-wrinkle agent in cosmetic products. Divarinol and bis-(2,4-dihydroxy-6-\(n\)-propyl-phenyl)-methane are effective tyrosinase inhibitors that could potentially be used as skin-whitening agents in creams (Kinoshista et al. 1994a; Kinoshista et al. 1994b; Matsubara et al. 1997).

1.2.5 The polyketide factories

Polyketides are a major class of secondary metabolites from bacteria, fungi, plants, and animals. The term “polyketide” was first coined by Collie in 1907 (Bentley & Bennett 1999). As the name implies, it generally means compounds having a series of keto (C=O) groups on the carbon backbone. Polyketide synthases (PKSs) are enzymes analogous to fatty acid synthases (FASs) that catalyse the consecutive condensation of acyl-CoA units, forming a chain of polyketomethylene. However, unlike FASs, PKSs leave most, if not all, of the keto group unreduced (Figure 1.3) (Hopwood 1997). The partially reduced keto groups formed either a hydroxyl or an enoyl group. In a broader sense, polyketides could include all secondary metabolites having a precursor of polyketide origin or involving a PKS at any point of their biosynthesis. Bently & Bennet (1999) provide a comprehensive review of polyketide research history including the origin of the term “polyketide” and how the meaning evolved to fit the new discoveries.
PKSs are usually divided into type I and type II, analogous to the FAS classification. There is also a type III PKS (chalcone synthase) that is mostly found in plants. A review by Shen (2003) provides some good insights into the current type I, II and III classification of PKSs and how new discoveries have challenged the existing paradigm. However, the classification of PKSs into type I, II and III is still useful in most cases and will be discussed briefly.

Like the type I FAS, type I PKS are encountered as multifunctional enzymes in bacteria and fungi. The enzymes are arranged in modules and catalyze condensation reactions non-iteratively according to the module orientation (Shen 2003). Hence, the length of the polyketide carbon chain and the sequence steps by which it is built are determined by the number and order of modules in the polypeptides constituting the PKS (Figure 1.4). Such enzymes are also called modular PKSs and control the biosynthesis of well-known antibiotics such as erythromycin A (Cortes et al. 1990).
and rifamycin (August et al. 1998). Each type I PKS module consists of several domains with defined functions, separated by short spacer regions. Some of those common domains are as below:

- Acyl-transferase (AT)
- Acyl-carrier-protein (ACP) with an SH group on the cofactor, a serine-attached 4'PP (4'-phospho-pantethein)
- Ketosynthase (KS)
- Keto-reductase (KR)
- Dehydratase (DH)
- Enoyl-reductase (ER)
- Thio-esterase (TE)

By contrast, type II PKSs, resemble the type II FASs in forming multi-enzyme subunits (Figure 1.4). Each PKS contains a ‘minimal’ set of core subunits (the two ß-ketoacyl:ACP synthase subunits KSα and KSβ, an acyl carrier protein (ACP), and possibly a malonyl-CoA:ACP transacylase) that is required for in vivo polyketide biosynthesis (Moore & Piel 2000). This family of PKSs produces mainly polycyclic aromatic polyketides, such as tetracenomycin C (Gramajo et al. 1991) and actinorhodin (Rudd & Hopwood 1979).

Type III PKSs are also called chalcone synthases (CHSs). They also include the stilbene synthases from plants. They are homodimeric enzymes acting iteratively like type II PKSs but independent of ACP (acting directly on acyl CoA substrates) (Figure 1.4). Although CHS is by definition a PKS because it catalyzes the linking of acyl CoA subunits by repetitive decarboxylative condensations, there are significant differences in the biochemical processes (Hopwood 1997). Sequence comparisons also suggest that CHS is phylogenetically distinct from all other groups of PKSs and all known FASs (Hopwood 1997). Type III PKSs were previously thought to be only in plants or specific classes of bacteria, but have also been found recently in fungi (Funa et al. 2007; Seshime et al. 2005).
Most PKSs found in fungi are the large multifunctional type I PKSs. By contrast with the bacterial modular (non-iterative) type I PKS, the fungal type I PKS functions iteratively as a multidomain homodimeric enzyme. The iterative fungal type I PKS are more similar to the vertebrate FASs structurally and functionally (Hopwood 1997).
One of the simple and most widely studied iterative fungal type I PKS is 6-methylsalicylic acid synthase (MSAS), which is commonly found in *Aspergillus* and *Penicillium* (Beck et al. 1990; Fujii et al. 1996). It has the domain architecture of KS–AT-DH-KR-ACP. The four domains identified resemble the corresponding sites in rat FAS, and are collinear in the two synthases (Bentley & Bennett 1999). As exemplified in the biosynthesis of 6-methylsalicylic acid (Figure 1.5), the KS and AT domains of the MSAS are used iteratively for three rounds of condensations, while a keto group is reduced to hydroxyl group by the KR domain after the second round of condensation. Unlike in vertebrate FASs where all the domains (KS, AT, DH, ER, KR, and ACP) are used in almost every round of chain extension (Figure 1.3), the programming of the functional domains of MSAS is evident (2 x condensation > keto reduction > 1 x condensation) and with precise control of chain length (Campuzano & Shoolingin-Jordan 1998). The catalytic sequence of the functional domains in other iterative type I PKSs also appears to be programmed and will be illustrated in the following examples.

In general, the fungal iterative type I PKSs can be grouped into aromatic or non-reducing (NR-PKS), partially-reducing (PR-PKS) and highly-reducing PKS (HR-PKS) according to the presence or absence of additional β-keto processing domains, as
well as the extent of reduction of the polyketide product (Figure 1.6). Phylogenetic studies based on the KS amino acid sequences showed that the grouping of the fungal PKSs into major clades correlated with the grouping of NR-, PR- and HR-PKSs (Bingle et al. 1999; Kroken et al. 2003; Nicholson et al. 2001). These studies suggested a strong relationship between KS phylogeny and PKS function.

As shown in the example given in Figure 1.6A (*Colletotrichum lagenarium* PKS1), NR-PKSs lack the reducing functional domains and therefore all the keto groups in the chain remain unreduced. The unreduced polyketide chain often cyclizes to form aromatic ring(s) by aldol and/or Claisen condensation (Shen 2000). For example, the *C. lagenarium* PKS1 catalyzes four rounds of extensions followed by a non-enzymatic aldol cyclization and an enzymatic Claisen cyclization involving the thioesterase (TE) or Claisen cyclase domain (CYC) (Fujii et al. 2000). The only PR-PKS known so far is the MSAS mentioned above, where the KR domain reduces the keto group at C-5 to a hydroxyl group (Figure 1.5 and 1.6B). The DH domain in MSAS could be involved in dehydration of the hydroxyl group to enoyl immediately after keto reduction by the KR domain or following the aldol cyclization (Figure 1.6B) (Bedford et al. 1995).

For HR-PKSs, an enoyl reductase (ER) domain is present alongside with the KR and DH domains. Their domain architecture and orientation is parallel to vertebrate FASs, except that a C-methyltransferase (CMeT) domain is present in some HR-PKSs between the DH and ER domains (Figure 1.6C). The *Aspergillus terreus* lovastatin diketide synthase (LDKS) is one of the simplest HR-PKSs characterized. It catalyzes only one extension (with the KS domain), and then reduces the keto group at C-3 to a hydroxyl group (with KR) followed by dehydration to enoyl (with DH), and further reduction to methylene (with ER). The transfer of the methyl group to C-2 by CMeT domain may occur immediately after chain extension or after the complete reduction by ER domain. Like many methyltransferases, the methyl group transferred by the CMeT domain originates from S-adenosyl methionine (SAM). The CMeT domain is also found in some NR-PKSs, e.g. *M. purpureus* PKSCT and *A. strictum* AsPKS1 (Bailey et al. 2007; Shimizu et al. 2005).
A) Non-reducing polyketide synthase (NR-PKS) - *C lagenarium* PKS1 (2187 a.a.)

B) Partially-reducing polyketide synthase (PR-PKS) – *P. patulum* MSAS (3038 a.a.)

C) Highly-reducing polyketide synthase (HR-PKS) – *A. terreus* LDKS (3038 a.a.)

**Figure 1.6.** Examples of typical domain architectures of A) non-reducing, B) partially reducing, and C) highly reducing PKSs and their corresponding products. KS, ketosynthase; AT, acyltransferase, DH, dehydratase; KR, ketoreductase, ACP, acyl carrier protein; TE, thioesterase; CYC, Claisen cyclase; CMeT, C-methyltransferase. Related references: *Colletotrichum lagenarium* PKS1 (Fujii et al. 2000); *Penicillium patulum* MSAS (Beck et al. 1990); *Aspergillus terreus* LDKS (Hendrickson et al. 1999; Kennedy et al. 1999).
Fungal Type I PKSs sometimes utilize a starter unit other than the common acetyl-coA. For example, the *A. nidulans* PKSST utilizes a hexanoate starter supplied by a specialised dimeric fatty acid synthase to produce norsolorinic acid (the polyketide intermediate for sterigmatocystin) (Brown et al. 1996a; Yu & Leonard 1995). By contrast, the *C. lagenarium* PKS1 incorporates malonyl-coA as both starter unit and extender units to produce tetrahydoxynaphthalene (THN) (Fujii et al. 2000) (Figure 1.6A). Yet, with the few exceptions known so far, most of the PKSs utilize acetyl-coA as the starter unit and malonyl-coA as the sole extender units.

1.2.6 The unique lichen compounds and the ubiquitous PKSs

Since lichen metabolites are mostly of fungal origin, it seems rather surprising that, of all the lichen metabolites identified, only a small minority occur in other fungi (Elix 1996). Most of those compounds unique to lichens are polyketides. Recent phylogenetic analysis of lichen PKS genes inferred from KS domain sequences suggests, however, that the PKS genes in lichens are not so much different to other fungi after all (Grube & Blaha 2003; Schmitt et al. 2005). Biosynthetic routes of only a few lichen polyketides have been investigated so far, and hypothetical pathways are often proposed on the basis of what is known for the biosynthesis of analogous fungal products (Mosbach 1973; Turner & Aldridge 1983) as well as inference from biomimetic organic synthesis (Elix et al. 1984b). The possible biosynthetic routes to the lichen polyketides will be discussed here with emphasis on the characteristic lichen compounds. The discussion will also refer to the recent improved understanding of the genetics and biochemistry of fungal polyketide biosynthesis. Most of the structures of lichen metabolites described here are based on Huneck & Yoshimura (1996) and Huneck (2001).

1.2.6.1 Lichen polyketides common to other fungi

Among the lichen polyketides, those from classes 1.2.1 and 1.2.3 (Table 1.1) are found commonly in many fungi. Class 1.2.1 consists of monoaromatic compounds that are formed mostly from tetraketide chains, while compounds in class 1.2.3 are polyaromatic compounds, which are formed from longer polyketide chains.
A total of 32 lichen monoaromatic compounds was described in Huneck & Yoshimura (1996) and a further 26 new ones were described in Huneck (2001). The most common monaromatic compounds found in lichens are orsellinic acid and its derivatives. Orsellinic acid is the simplest polyketide compound and is also found in *Penicillium* spp. as well as in some bacteria. Only three condensations of malonyl-CoA are required for the formation of the tetraketide chain, followed by a spontaneous non-stereospecific (non-enzymatic) proton removal that leads to aldol cyclization at C-3 and C-5 (Figure 1.7) (Woo et al. 1989). Orsellinic acid synthase (OAS) was isolated from *Penicillium madriti* as early as 1968 (Gaucher & Shepherd 1968), but the PKS gene responsible for its production remains an enigma. (Fujii et al. 1996) attempted to isolate the OAS gene from *P. cyclopium* (which produces penicillic acid via orsellinic acid) by using *P. patulum* MSAS synthase gene as a probe. Despite the structural similarity of 6-methylsalicylic acid and orsellinic acid, no homologous band was detected in the genomic DNA of *P. cyclopium*. Instead, a *P. patulum* MSAS gene homolog, *atx*, was isolated from *Aspergillus terreus* and was shown to produce 6-methylsalicylic acid as well. The biosynthesis of 6-methylsalicylic acid was discussed earlier (Section 1.2.4, Figure 1.5). Recently, the OAS genes isolated from some bacteria were shown to have high similarity to the fungal MSAS gene (Ahlert et al. 2002; Gaisser et al. 1997). The fact that an OAS gene similar to MSAS was not found in fungi suggests that the fungal OAS is probably structurally different from its bacterial counterpart.

\[
\text{Acetyl-CoA} + 3 \times \text{malonyl-CoA} \rightarrow \text{orsellinic acid}
\]

*Figure 1.7.* Biosynthesis of orsellinic acid.

The occurrence of 6-methylsalicylic acid in lichens has not been reported, but various orsellinic acid and phloroacetophenone derivatives have been found in lichens. Some of the orsellinic acid derivatives contain a long alkyl chain at C-6 of the aromatic ring. As many of these monoaromatics are involved in the biosynthesis of the di- and tri-aryl derivatives (class 1.2.2, Table 1.1), which are common in lichens,
the biosynthesis of these monoaromatic compounds will be discussed in more detail later. Some of the monoaromatic compounds isolated from lichens were methanolysis or cleavage products of these di- and tri-aryl derivatives (Huneck & Yoshimura 1996).

Polycyclic aromatic compounds such as anthraquinones and xanthones are found widely in fungi and higher plants (Dewick 2002; Peres et al. 2000; Thomas 2001). Although some anthraquinones and xanthones are peculiar to lichens, the biosynthetic route should be similar to the non-lichen fungi or higher plants (Culberson 1969). For example, norsolorinic acid (intermediate of aflatoxins and sterigmatocystins), which is common in Aspergillus spp., is also found in the lichens along with its derivatives, such as averthrin and averantin (e.g. in Solorina crocea) (Huneck & Yoshimura 1996).

Up to now, at least five fungal PKS genes that produce polyaromatic compounds have been characterised – the 1,3,6,8-tetrahydroxynaphthalene (THN) pentaketide synthase (Fujii et al. 1999), YWA1 heptaketide synthase (Watanabe et al. 1999), norsoloronic acid synthase (Feng & Leonard 1995), bikaverin synthase (Linnemannstons et al. 2002), and cercosporin synthase (Choquer et al. 2005). These non-reducing fungal PKS genes will be discussed in more detail later in Chapter 3. The well-studied norsolorinic acid, YWA1 naphthopyrone and bikaverin biosynthesis will be used as examples to illustrate the common traits with some lichen polyaromatic metabolites (Figure 1.8). Due to their structural similarity, compounds in Figure 1.8B and other related lichen metabolites are thought to be biosynthesized via similar pathways to those illustrated in Figure 1.8A. The three fungal PKSs involved in the biosynthesis of YWA1 naphthopyrone, norsolorinic acid, and bikaverin have similar PKS domain architecture and are also phylogenetically related to one another (see Chapter 3). Therefore, similar PKSs could be involved in the biosynthesis of those structurally similar polyaromatic lichen metabolites. The hypothetical pathways and polyketide chain folding pattern of some common lichen polyaromatic compounds, including chromones, xanthones, and anthraquinones, were proposed by Mosbach (1969) on the basis of the biosynthesis of related fungal metabolites understood at that time. Alternative folding patterns of the polyketide chains to those proposed by Mosbach (1969) is also possible, to form the final
products. Further radioactive labelling and biochemistry studies will be required to elucidate the pathways and polyketide folding pattern of these compounds.

1.2.6.2 Polyketide compounds common in lichens

The unique feature of lichen secondary metabolites is the ubiquitous compounds belonging to class 1.2.2 (Table 1.1). Most are formed by intermolecular coupling of two (occasionally three and four) monoaromatic units through ester, ether and/or carbon-carbon linkages. As phenol moieties are evident in all these compounds, they will be collectively referred to as coupled phenolics. Each aromatic unit in these compounds, including those joined by carbon-carbon linkages, is thought to originate from separate polyketide chains. The coupled phenolics are divided into depsides, diphenyl ethers, depsidones, depsones, biphenyls and dibenzofurans/usnic acids on the basis of the different bonds that join the two aromatic units (Figure 1.9).
Figure 1.8. Examples of fungal metabolite biosynthesis related to some lichen metabolites. A, biosynthesis and folding pattern of some well-characterized fungal polyketides; B, some related lichen polyaromatic metabolites.

* Using malonyl-CoA as a starter unit as described in (Ma et al. 2007) would have resulted in an additional carboxyl-group at the end of the polyketide chain (at C-18) in bikaverin biosynthesis. Since C-18 is not involved in a Claisen condensation as in THN biosynthesis by C. lagenarium PKS1 (Figure 1.5), the carboxyl group might be removed by decarboxylation before or after aldol condensation of C-17 and C-12.
Among all the classes of coupled phenolics, only depsones are unique to lichens. Coupled phenolics of other classes, such as depsides and depsidones, are occasionally found in the non-lichenized fungi but not as frequently in lichens (see Section 1.2.5.2.4 and 1.2.5.2.5). The frequency of occurrence of each class of coupled phenolics can be estimated from the number of compounds described so far (Huneck 2001; Huneck & Yoshimura 1996) (Figure 1.9). Although usnic acids and related compounds are technically not dibenzofurans (the aromacity is lost in the second ring), they are normally grouped together due to their structural similarity.

**Figure 1.9.** Various groups of coupled phenolics present in lichens where monoaromatic units were bonded by ester, ether and/or carbon-carbon linkages. The approximate number of compounds known [based on Huneck & Yoshimura (1996)] is indicated in parentheses. New compounds have been discovered more recently and were reviewed in Huneck (2001).
1.2.6.2.1 Basic monoaromatic units that form lichen coupled phenolics

There are several types of basic monoaromatic units that are bonded to form the various coupled phenolic compounds, mainly orsellinic acid [1], β-orsellinic acid [2] and methylphloroacetophenone [3] (Figure 1.10). There is also a fourth type of monoaromatic units – alkylresorcylic acid, where an alkyl chain of various lengths is attached to the orsellinic acid moiety (Figure 1.11). Together they form the basis of almost all coupled phenolics present in lichens.

Orsellinic acid [1], β-orsellinic acid [2] and methylphloroacetophenone [3] are all formed by a tetraketide chain via the polyketide pathway. The biosynthesis of orsellinic acid from a simple tetraketide chain was described earlier in Section 1.2.5.1 (see Figure 1.7). The tetraketide chain that forms β-orsellinic acid and methylphloroacetophenone has a methyl group present at C-4, which is likely to be added by a C-methyltransferase (CMeT) domain present in the respective PKSs. β-orsellinic acid is formed by a non-enzymatic (C-2 and C-7) aldol condensation, while methylphloroacetophenone is formed by a Claisen condensation of C-1 and C-6. This hypothesis for the formation of orcinol and phloroglucinol moiety via alternative cyclization of a tetraketide chain was originally proposed by Birch & Donovan (1953) and later used by Mosbach (1969) to infer the biosynthesis of lichen metabolites. The thioesterase (TE) domain found in many NR-PKSs also functions as a Claisen cyclase (CYC) domain. For example, the corresponding domain in A. nidulans WA catalyses the Claisen cyclization involved in forming the second ring of YWA1 naphthopyrone (Fujii et al. 2001). It is possible that the cyclization of the C-methylated tetraketide to methylphloroacetophenone [3] is mediated by such a CYC domain (Figure 1.10).
Figure 1.10. Proposed biosynthetic routes to orsellinic acid [1], β-orsellinic acid [2], and methylphloroacetophenone [3] via the polyketide pathway. PKS domains involved in the biosynthesis are KS, ketosynthase; AT, acyltransferase; ACP, acyl carrier protein, and CYC, Claisen cyclase. The three monoaromatic units formed the basis of various coupled phenolics present in lichens.
Figure 1.11. Proposed biosynthesis of tetraketide [4] and pentaketide [5] alkylresorcylic acids. Adapted from (Funa et al. 2007).

The biosynthesis of alkylresorcylic monoaromatic units is likely to be initiated by an acyl-CoA starter unit of various chain lengths instead of acetyl-CoA. The ability of some PKSs to incorporate medium-long chain acyl-CoA as a starter unit is evident in the biosynthesis of norsolorinic acid (the precursor to sterigmatocystin and aflatoxins) (Townsend et al. 1984; Yabe & Nakajima 2004). Here, a dedicated fatty acid synthase is involved in the biosynthesis of hexanoyl-CoA as the starter unit for the corresponding PKS (Aspergillus parasiticus PKSA, A. nidulans STCA). A similar mechanism might be involved in the biosynthesis of the alkylresorcylic acid units in lichens. The starter units may also be produced by the degradation of common longer chain fatty acids, e.g. C18 acids, such as oleic or linoleic acid etc.

Two different kinds of alkylresorcylic acid moieties are commonly observed among the coupled phenolics present in lichens (Figure 1.11) – alkylresorcylic acids [4] with a saturated alkyl chain [4] and the 2'-oxoalkylresorcylic acids [5] with a keto group at the C-2’ of the alkyl chain. Examples of coupled phenolics that incorporated a 2'-oxoalkylresorcylic acid moiety are haemophaein (dibenzo-furan), β-alectoronic
acid (diphenyl ether), microphyllinic acid (depside), etc. Often these compounds exist in tautomers in lactol and open chain form, as in haemophaien and β-alectoronic acid (see Table 1.9 and 1.10). The carbonyl at the C-2’ of the 2’-oxoalkylresorcylic acid [5] is likely to originate from a malonyl-CoA as a result of four rounds of condensation by a pentaketide synthase, while the alkylresorcylic acid [4] probably arises from a tetraketide chain. Hence, they will be referred as pentaketide [5] and tetraketide [4] alkylresorcylic acids respectively (Figure 1.11). Recently, the intriguing discovery of a type III PKS from *Neurospora crassa* that catalyzes the synthesis of 2’-oxoalkylresorcylic acids suggests that a similar type III PKS could be involved in the biosynthesis of some coupled phenolics with alkylresorcinol moieties (Funa et al. 2007).

Occasionally, coupled phenolics with 3’- and 1’-oxoalkylresorcylic acid moieties are also found in lichens. For example, the depside miriquidic acid has a carbonyl group at the C-3’ of the alkyl chain on the A-ring, while a carbonyl group is located at the C-1’ of the alkyl chain on the A-ring of depsidone loxodin (Figure 1.12A). The keto group of these o xoalkylresorcylic acids is located at an odd numbered carbon (e.g. C-1’, C-3’, C-5’, etc.) of the alkyl chain and is not likely to originate from the polyketide pathway. The keto group is probably the result of a separate oxidation catalysed by an enzyme such as a monooxygenase. A mechanism similar to one of the steps in the biosynthesis of aflatoxins could be involved (Figure 1.12B). After reduction of norsoloronic acid to averatrin, the intermediate could be converted to 5’-hydroxyaveratrin by averatrin monooxygenase, whereby a hydroxyl group is added to C-5’ of the hexanoyl chain. This process is followed by oxidation of the 5’-hydroxyl group by 5’-hydroxyaveratrin reductase to give 5’-oxoaveratrin (Yabe & Nakajima 2004). This does not, however, exclude the possibility that the keto group on the alkyl chain could be added before its incorporation to form the alkylresorcylic acid.

Besides the more common monoaromatic units mentioned above, a few coupled phenolic compounds in lichens incorporate 3,5-dimethylorsellinic acid, such as the depsides brialmontin (see Table 1.7) and phenarctin. The additional methyl group is likely to be added by the CMeT domain of a PKS via a similar mechanism to that proposed for β-orsellinic acid [2] (Figure 1.10).
Figure 1.12. Examples of lichen depside/depsidones with oxoalkylresorcylic acid moieties, in which the keto group in the alkyl chain has not originated from the polyketide pathway. A) Depside (miriquidic acid) with a 3'-oxoresorcylic acid moiety and depsidone (loxodin) with a 1'-oxoresorcylic acid moiety. B) A biosynthetic step in aflatoxin biosynthesis involving a monooxygenase and a reductase according to Yabe & Nakajima (2004).

1.2.6.2.2 The formation of various linkages in lichen coupled phenolics

Among the coupled phenolics present in lichens, there are three basic bonding modes joining the two monoaromatic units, the ester bond (depsides and benzyl esters), the ether bond (diphenyl ether) and the carbon-carbon bond (biphenyls) (see Figure 1.9). Combinations of these three basic linkages leads to the formation of depsidones, dibenzofurans and depsones. The possible biosynthetic steps that lead
to the formation of the three basic linkages will be discussed here before proceeding to the individual groups of coupled phenolics.

The formation of the ester bond in depsides via esterification is likely to be effected by a mechanism analogous to that commonly observed in lipid metabolism. The esterification between a benzoate and a phenol moiety is presumably catalyzed by a specific ester synthase/esterase that drives the reaction towards ester (depside) synthesis (Figure 1.13).

![Figure 1.13. Proposed biosynthesis steps for ester bond formation in depsides.](image)

The ether and carbon-carbon linkages in coupled phenolics in lichens is likely to be formed by oxidative phenolic coupling (Mosbach 1969) (Figure 1.14). Such a mechanism was first illustrated in the famous experiment on usnic acid synthesis via oxidative coupling of methylphloroacetophenone in the presence of potassium ferricyanide (Barton et al. 1956) and was later used to explain the coupling of phenolic units observed in the biosynthesis of various natural products (Barton & Cohen 1957). Enzymatic oxidative coupling to form usnic acid using horseradish peroxidase *in vitro* has also been demonstrated (Penttila & Fales 1966). This confirmed that the coupling of phenolics via generation of phenoxy radicals is common in biological systems. The coupling of phenolic units is initiated by one-electron oxidation by a peroxidase in the presence of hydrogen peroxide or organic peroxides. The phenoxy radical is mesomeric, with the unpaired electron effectively on the oxygen as well as carbon in the *ortho* and *para* positions. Thus, the phenoxy radical (A, B and C) may couple to produce various types of biphenyls and diphenyl ethers (Figure 1.14). The A-C type carbon-carbon bond is common among lichen dibenzofurans and usnic acids, while the B-C type ether bond is common among depsidones.
Figure 1.14. Coupling of phenols via one-electron oxidation to form various types of biophenyls and diphenyl ethers. Box: formation of phenoxy radical (B) and its mesomeric forms (A and C) via one-electron oxidation. Adapted from Dewick (2002).

The ether bond in usnic acids probably arises via an alternative mechanism. Here nucleophilic attack of a phenolic hydroxyl group on an enone system leads to the ether linkage, but the details will be discussed later (Section 1.2.5.2.8).

1.2.6.2.3 Derivatization of the basic monoaromatic units

The enormous diversity of lichen coupled phenolics is mainly generated via three routes. The first two routes were discussed in the previous two sections: 1) by incorporating different types of basic monoaromatic units originating from separate polyketide pathways, and 2) by coupling of the monoaromatic units via various linkages. The third type of variation, which amplifies the structural diversity of lichen coupled phenolics, is via post-PKS modification of the basic monoaromatic units. The modifications include O-methylation, oxidation, halogenation, decarboxylation and esterification with short-medium chain fatty acids. Some of these post-PKS modification reactions may occur after polyketide cyclization, while some may take place following the coupling of the aromatic units.
A good example of such derivatization of β-orsellinic acid was given by (Culberson 1969), where the methyl groups at C-3 and C-6 of the β-orsellinic acid unit is subjected to increasing degrees of oxidation (Figure 1.15). The couplings of these β-orsellinic acid derivatives form the depsides and depsidones in Table 1.6. Interestingly, some of the theoretically possible β-orsellinic acid derivatives formed by similar oxidative mechanisms are seldom or have yet to be found among the lichen depsides and depsidones (Figure 1.15, in grey), though a few have been identified more recently, e.g. chalybaeizanic acid (Elix & Wardlaw 1999).

Figure 1.15. Common β-orsellinic acid derivatives incorporated in depsides and depsidones with the methyl groups at C-3 and C-6 subjected to increasing level of oxidation. In grey, theoretically possible β-orsellinic acid derivatives but seldom or yet to be found in lichen depsides and depsidones. Adapted from Culberson (1969).
Table 1.6. The distribution of β-orsellinic acid derivatives from Figure 1.15 into various depsides and depsidones. The assignment of A and B rings is according to Figure 1.9. Decarboxylated units are indicated by (-CO2) after the number of the unit. Table adapted from Culberson (1969).

<table>
<thead>
<tr>
<th>Monoaromaric units (Figure 1.15) A-B</th>
<th>para-depsides</th>
<th>depsidones</th>
<th>meta-depsides</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1 (-CO₂)</td>
<td>barbatic acid, diffractaic acid, 4-O-demethyl-barbatic acid</td>
<td>hypoprotocetraric acid</td>
<td></td>
</tr>
<tr>
<td>1-1</td>
<td>vicanicin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-1</td>
<td>atranorin, chloroatranorin, baemomycesic acid</td>
<td>virensic acid</td>
<td></td>
</tr>
<tr>
<td>4-1</td>
<td>squamatic acid</td>
<td></td>
<td>hypothamnolic acid</td>
</tr>
<tr>
<td>3-2</td>
<td>fumarprotocetraric acid, protocetraric acid, physodalic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-3</td>
<td></td>
<td>haemathamnolic acid</td>
<td></td>
</tr>
<tr>
<td>3-6</td>
<td>norstictic acid, stictic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-3</td>
<td></td>
<td>thamnolic acid</td>
<td></td>
</tr>
<tr>
<td>3-7</td>
<td>salazinic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-8 (-CO₂)</td>
<td></td>
<td>psoromic acid</td>
<td></td>
</tr>
</tbody>
</table>

In a biosynthetic study using the calcium alginate-immobilized lichen *Evernia prunastri*, it was demonstrated that sodium azide (a common metallo-oxidase inhibitor) and pyrazole (an alcohol dehydrogenase inhibitor) inhibited the biosynthesis of atranorin (Vicente et al. 2003). By contrast, the addition of NADH co-factor and oxygen increased its production. This suggested that an oxidase was involved in the oxidation of the methyl group at C-3 of the β-orsellinic acid moiety to a hydroxymethyl group, while an alcohol dehydrogenase was involved in converting the hydroxymethyl group to aldehyde, thus forming the haematommic acid moiety of atranorin (Figure 1.16). The oxidation of the β-orsellinic acid moiety might occur before or after the formation of depside (ester) bond.
Figure 1.16. Biosynthesis of haematommic acid and its incorporation into atranorin. Adapted from Vicente et al. (2003).

Other examples of post-PKS modifications of the aromatic moieties, including chlorination and O-methylation, are illustrated for chloroatranorin in Figure 1.17A. Oxidation of the alkyl chain to form oxoalkylresorcylic acids as described earlier could also be considered as a type of diversification of the basic monoaromatic units (see Figure 1.12). In addition, esterification by short-medium fatty acids is also common in lichen depsidones (e.g. the formation of a fumaryl ester in fumarprotocetraric acid, Figure 1.17B). Such modification is likely to occur after the coupling of the monoaromatic units to depsidones. A biosynthetic study using alginate-immobilized cells of the lichen Cladonia verticillata demonstrated that the fumaryl ester at the C-3 of the β-orsellinic acid unit at ring-B of fumarprotocetraric acid is added via a reducing, flavin-dependent, coupling reaction that uses a pool of succinyl-CoA from the mycobiont (Fontaniella et al. 2000). The esterification by the fumaric acid is likely to occur after oxidation of the methyl group at C-3 to a hydroxymethyl group.
Figure 1.17. Post-PKS modifications of the aromatic units of chloroatranorin (A) and fumarprotocetraric acid (B).

1.2.6.2.4 Depsides

Traditionally, the lichen depsides are divided into orcinol and β-orcinol depsides based on the presence of the respective moieties in the depsides. The classification does not, however, reflect the biosynthetic origin of those monoaromatic monomers. The β-orcinol depsides refer to depsides that incorporate almost solely β-orsellinic acid units but they also include depsides such as brialmontin1 and 2 formed by 3,5-dimethylorsellinic acids. On the other hand, the orcinol depsides collectively include depsides that incorporate monoaromatic units from different polyketide pathways – orsellinic acid, tetraketide resorcylic acids and pentaketide resorcylic acids. The various types of lichen depsides formed from different monoaromatic units are summarized in Table 1.7.
Table 1.7. Examples of lichen depsides from coupling of various aromatic units via ester bond. (continue next page)

<table>
<thead>
<tr>
<th>Position of ester bond</th>
<th>Monomeric aromatic intermediates</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>para-depsides</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ester bond form with the carboxylic acid group at C-1 and hydroxyl group at C-4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The two basic moieties found in lichen depsides:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>orcinol</td>
<td></td>
<td></td>
</tr>
<tr>
<td>β-orcinol</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* the B-ring is decarboxylated.
Table 1.7 (continued). Examples of lichen depsides from coupling of various aromatic units via an ester bond.

<table>
<thead>
<tr>
<th>meta-depsides</th>
<th>ortho-depsides</th>
<th>tri- and tetra-depsides</th>
<th>benzyl esters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydroxylolation at C-5 followed by esterification</td>
<td>Ester bond form with the hydroxyl group at C-2</td>
<td>tri- and tetra-aryl esters with three and four monoaromatic units.</td>
<td>Hydroxylolation of the methyl group at C-6 to hydroxymethyl followed by esterification</td>
</tr>
<tr>
<td><img src="image" alt="meta-depsides" /></td>
<td><img src="image" alt="ortho-depsides" /></td>
<td><img src="image" alt="tri-depsides" /></td>
<td><img src="image" alt="benzyl esters" /></td>
</tr>
<tr>
<td><img src="image" alt="meta-depsides" /></td>
<td><img src="image" alt="ortho-depsides" /></td>
<td><img src="image" alt="tri-depsides" /></td>
<td><img src="image" alt="benzyl esters" /></td>
</tr>
</tbody>
</table>

Depsidies can be subdivided into *para-* and *meta-*depsides (Table 1.7) based on the position where the ester is bonded to the B-ring. Ortho-depsides also occur in lichens but so far only isolecanoric acid (Table 1.7) is known. Tri- and tetra-depsides are also occasionally found in lichens but most of them incorporate orsellinic acid as the monomeric units. The majority of aromatic units in tri- and tetra-depsides is joined with ester bonds at the para-positions and the simplest is gyrophoric acid (comprising three orsellinic acid units). Crustinic acid is an exception, as it has the ester bonded at the meta-position in the third ring (Table 1.7). The occurrence of benzyl esters in lichens is relatively rare. These compounds are likely to result from esterification of a hydroxymethyl group at C-6’ in the second β-orsellinic acid unit. Only four such lichen benzyl esters were described in Huneck & Yoshimura (1996) while another three benzyl esters (echinocarpic acid, caonechinocarpic acid and hypoallectoraric acid) were identified later (Elix et al. 1995; Elix & Wardlaw 1996). Besides the homodimeric depsides illustrated in Table 1.7, mixed depsides with orcinol and β-orcinol moieties.
in either the A- or B-ring are also common in lichens. No known lichen depside is produced from methylphloroacetophenone, perhaps because it lacks the carboxyl group, but theoretically methylphloroacetophenone could constitute the B-ring of a depside.

The biosynthesis of some depsides has been studied using radioactively labelled precursors. The biosynthesis of gyrophoric acid in *Umbilicaria papulosa* was studied by “feeding” the lichen thallus in agitated culture with $^{14}$C-labelled malonyl-CoA (Mosbach 1964). The incorporation of $^{14}$C into the C-2, C-4 and carboxyl group of each orsellinic acid moiety supported the hypothesis that the monoaromatic unit originated from acetyl-CoA and three malonyl-CoAs via a polyketide pathway (Figure 1.18A). Subsequent experiments with the β-orcinol depside atranorin using $^{14}$C-labelled formic acid resulted in specific incorporation of $^{14}$C in the β-methyl and aldehyde group of atranorin (Figure 1.18B) (Yamazaki & Shibata 1966). The result was in agreement with the proposed β-orsellinic acid [2] pathway illustrated in Figure 1.10, where the carbon atom is probably sequestered from formic acid and incorporated into the tetraketide chain via S-adenosylmethionine.

![Figure 1.18. Labelling pattern in A) gyrophoric acid and B) atranorin. 14C from malonyl-CoA (●) and formic acid (△). (Mosbach 1964; Yamazaki & Shibata 1966).](image)

No direct evidence is available on the biosynthesis of *meta*-depsides. Their biosynthesis is likely to involve an additional hydroxylation step at the C-3’ or C-6’ of the B-ring (e.g. haemathamnolic acid, Table 1.7), due to the lack of a hydroxyl group at the *meta*-position in the orsellinic/β-orsellinic acid moieties. The co-occurrence and interconversion of the depsides *meta*- and *para*-scrobiculin (which exist in dynamic equilibrium) in *Lobaria scrobiculata* led to the hypothesis that *meta*-depsides are converted from *para*-depsides via intramolecular rearrangement (acyl migration) after
hydroxylation of the carbon at the *meta*-position of the B-ring (Figure 1.19A) (Elix & Gaul 1986). However, the absence of the intermediate methyl nordivaricatate (non-hydroxylated form of *para*-scrobiculin) in *Lobaria scrobiculata* (Elix 2001) suggests that hydroxylation of the mononuclear B-ring before the ester formation is equally plausible (Figure 1.19B). Here, *para*- or *meta*-scrobiculin may form by esterification of methyl divarate or methyl 3-hydroxydivarate (Figure 1.19B). Intramolecular rearrangement by acyl migration could then follow to form the observed equilibrium mixture of the isomorphic forms of scrobiculin. Regardless of which theory is true, enzymatic hydroxylation of the B-ring in the *meta*-position (C-3’ or C-6’) is likely to be the key step in the biosynthesis of *meta*-depsides.

---

**Figure 1.19.** Proposed biosynthetic routes to *meta*- and *para*-scrobiculin. Route A), hydroxylation of nordivaricatic acid at C-3’ leads to *para*-scrobiculin; route B) hydroxylation of divaric acid followed by intermolecular esterification to form scrobiculin. Acyl migration results in interconversion of *para*- and *meta*-scrobiculin.
The enzymatic hydroxylation of aromatic compounds is commonly observed in microorganisms, plants and animals. At least three types of such enzyme are effective – dioxygenase, monooxygenase (including P450s) and peroxidase (Ullrich & Hofrichter 2007) (Figure 1.20). Hydroxylation by both monooxygenase and peroxidase involves an epoxide intermediate (4) (Figure 1.20). A dioxetane intermediate (7) has been proposed in the hydroxylation by dioxygenase, but this mechanism has been criticized because of the high endothermicity of dioxetane formation from a ground state oxygen molecule (Bugg & Winfield 1998; Sawaki & Foote 1983).

**Figure 1.20.** Basic routes of enzymatic hydroxylation of aromatic compounds via dioxygenase (DO), monooxygenase (MO), and peroxidase (PO). DH, dehydrogenase; Re, rearrangement. (1) aromatic substrate, (2) cis-dihydrodiol, (3) catecholic product, (4) epoxide intermediate, (5) phenolic product, (6) cis,trans-dihydrodiol, (7) hypothetical dioxetane intermediate. Adapted from Ullrich and Hofrichter (2007).
When the aromatic compound is phenolic, other enzymes may be involved in C-hydroxylation. Thus, tyrosinase (phenol oxidase) may catalyze hydroxylation at the \textit{ortho}-position of a phenol (Figure 1.21A), while an indirect hydroxylation of phenol by laccase or phenol-oxidizing peroxidases may result in hydroxylation at the \textit{para}-position (Figure 1.21B) (Ullrich & Hofrichter 2007). The process is initiated by one-electron oxidation of the phenol to a phenoxy radical analogous to the mechanism for phenolic oxidative coupling as illustrated earlier in Figure 1.14.

![Figure 1.21. Hydroxylation of a phenolic substrate (8) by A) tyrosinase (Tyr) and by B) a laccase (Lacc) or phenol-oxidizing peroxidases (PO) indirectly. (3) catechol product, (9) o-benzoquinone product, (10) two mesomeric forms of the phenoxy radical that can disproportionate to a cyclodienone cation (11) and a phenolic molecule (8), (12) unstable hydroxycyclohexadienone tautomer rearranging to a hydroquinone (13) or p-benzoquinone (14) product. Adapted from Ullrich and Hofrichter (2007).](image)

The formation of depsides from two monoaromatic polyketides is particularly common in lichens (Culberson 1969), but rare in other organisms (e.g. the common lichen depside, lecanoric acid, has also been isolated from the fungus \textit{Pyricularia} sp.,
which is a plant pathogen) (Umezawa et al. 1974). Two structurally similar depsides were also isolated from the plants Papaver rhoes (poppy), Myrciaria cauliflora (jaboticaba) and Vaccinium marcocarpon (cranberry) (Hillenbrand et al. 2004; Reynertson et al. 2006), but the structures of the monoaromatic units incorporated suggest that they are derived via the shikimic acid pathway (Figure 1.22).

\[
\begin{align*}
R &= \text{H, CH}_3 \\
\end{align*}
\]

**Figure 1.22.** Two structurally related depsides isolated from plants – Papaver rhoes, Myrciaria cauliflora and Vaccinium marcocarpon.

1.2.6.2.5 Depsidones

Like depsides, lichen depsidones are divided into orcinol and β-orcinol depsidones based on the presence of the respective mononuclear moieties. There are also mixed depsidones that have both orcinol and β-orcinol moieties. The same basic monoaromatic units involved in depsides are commonly involved in the formation of lichen depsidones (Table 1.8). The occurrence of depsides and depsidones with the same monoaromatic units, such as divaricatic acid (para-depside) and divaronic acid (depsidone) (Table 1.7 and 1.8), has led to the hypothesis that the depsides are precursors of depsidones. However, the corresponding depsides and depsidones with the same monoaromatic units do not always occur together in the same lichens and not all depsidones discovered have a corresponding depside. The ester linkage in most depsidones is bonded at the para-position of the B-ring, just as in para-depside. An exception known is a rare depsidone, cyclographin (Table 1.8), which is related to the meta-depside haemathamnolic acid (Table 1.7).
Table 1.8. Lichen depsidones. Coupling of various aromatic units via an ester bond and an ether bond.

<table>
<thead>
<tr>
<th>Position of ester bond</th>
<th>Monomeric aromatic intermediates</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>**Depsidones with ester bond corresponding to **&lt;br&gt;<strong>para-depsides</strong></td>
<td><img src="image1" alt="Structure" /> orsellinic acid</td>
<td><img src="image2" alt="Structure" /> Gangaleoidin</td>
</tr>
</tbody>
</table>
| | ![Structure](image3) tetraketide resorcylic acids  
R= C_{n}H_{2n+1} | ![Structure](image4) Divaronic acid |
| | ![Structure](image5) pentaketide resorcylic acids  
R= C_{n}H_{2n+1} | ![Structure](image6) Glomellonic acid |
| | ![Structure](image7) β-orsellinic acid | ![Structure](image8) Protocetraric acid |
| **Depsidones with ester bond corresponding to **<br>**meta-depsides** | ![Structure](image9) β-orsellinic acid | ![Structure](image10) Cyclographin |

The initial theory for the biosynthesis of depsidones envisaged oxidative coupling of a depside intermediate (Culberson 1969; Elix et al. 1984b). The corresponding phenoxy radicals are generated via one-electron oxidation of a phenol group via the mechanism described previously (see Figure 1.14). This theory was supported by the co-occurrence of the depside olivetoric acid and the corresponding depsidone physodic acid in the lichen *Cetraria ciliaris* (Culberson 1964) (Figure 1.23). Further evidence was provided by the laboratory synthesis of diploicin by oxidation of a
corresponding depside with manganese dioxide (Brown et al. 1961) (Figure 1.24).
The depside used for synthesis of diploicin does not occur naturally in lichens, but is
structurally similar to a lichen depside – tumidulin (methyl 3,5-dichlorolecanorate)
(Figure 1.24, Box).

![Diagram of chemical structures]

**Figure 1.23.** Proposed phenol oxidative coupling of lichen depside (olivetoric acid) to
form depsidone (physodic acid). Based on Culberson (1964).

![Diagram of chemical structures]

**Figure 1.24.** Oxidation of a synthetic depside to diploicin with manganese dioxide.
Box, tumidulin is a depside structurally related to diploicin. Based on Brown et al.
(1961).

Two alternative theories for the biosynthesis of depsidones involve 1) the
oxidative coupling via a benzophenone intermediate, or 2) a meta-depside
intermediate undergoing Smiles rearrangement to diphenyl ether. The first theory is
supported by chemical synthesis in which oxidative coupling of corresponding
benzophenones readily yielded grisadienediones, which rearranged to depsidones
under mild acidic or basic conditions (Sala & Sargent 1981) (Figure 1.25). This theory
envisaged that depsidones could be formed from a single long polyketide chain
instead of from two separate aromatic polyketides. Benzophenones are structurally similar to and could be intermediates in xanthone biosynthesis in fungi. However, the proposed benzophenone intermediate in Figure 1.25 and an intermediate with a carboxyl group (required for the biosynthesis of depsidones with a carboxyl group in the B-ring) would not originate from a continuous polyketide chain, as they do not match any polyketide folding pattern (Figure 1.26). This mechanism has also recently been discredited by as yet unpublished labelling studies (in litt. to J.A. Elix).

![Diagram showing proposed depsidone (diploicin) biosynthesis via a benzophenone intermediate. Adapted from Elix et al. (1984b).](image)

**Figure 1.25.** Proposed depsidone (diploicin) biosynthesis via a benzophenone intermediate. Adapted from Elix et al. (1984b).

![Diagram showing different polyketide folding patterns that would produce a benzophenone moiety but not the benzophenone that is crossed. Box, a carboxyl group at the benzophenone moiety that would be required for production of most depsidone carboxylic acids via the mechanism illustrated in Figure 1.25.](image)

**Figure 1.26.** Different polyketide folding patterns that would produce a benzophenone moiety but not the benzophenone that is crossed. Box, a carboxyl group at the benzophenone moiety that would be required for production of most depsidone carboxylic acids via the mechanism illustrated in Figure 1.25.
The second theory, for biosynthesis of lichen depsidones via a *meta*-depside intermediate, is supported by the chemical synthesis where two depsidones, divaronic acid and stenosporonic acid, were prepared by a biomimetic-type approach that involved a Smiles rearrangement of a precursor *meta*-depside to a diphenyl ether in the key step (Elix et al. 1987). A biosynthetic route was proposed for divaronic acid and other depsidones (Figure 1.27), by combining the possible pathway for *meta*-depside biosynthesis via acyl migration (see Figure 1.19). A similar mechanism could be envisaged for cyclographin (see Table 1.8 for formula), where a *para*-depside intermediate is directly converted to the corresponding diphenyl ether followed by intramolecular esterification to form this depsidone. For both mechanisms, hydroxylation of the B-ring at the *meta*-position is required. Enzymes involved in such hydroxylation have been discussed in the previous section.

![Figure 1.27. Proposed biosynthesis of divaronic acid (depsidone) via a meta-depside intermediate. Based on Elix et al. (1987).](image)
More recently, preliminary experiments in D. Amaleo’s laboratory indicated that a dioxygenase was involved in converting a \textit{para}-depside precursor to the corresponding depsidone (Figure 1.28) (in litt. to J.A. Elix). An unpublished labelling study in C.F. Culberson’s lab (in litt. to J.A. Elix) has also demonstrated that \textit{para}-depsides are the immediate precursors of depsidones via a two-step oxidation-elimination process, which further supports the dioxygenase mechanism. According to J.A. Elix, this is the most likely mechanism for depsidone biosynthesis.

![Figure 1.28. Proposed mechanism for biosynthesis of depsidones from a \textit{para}-depside precursor mediated by a dioxygenase. NADH is thought to act as a hydrogen donor (in litt. to J.A. Elix).](image)

Several depsidones have been found in fungi other than lichens. For example, nidulin and other structurally related depsidones were isolated from \textit{Aspergillus nidulans} (Dean et al. 1954; Dean et al. 1953), while unguinol, 2-chlorounguinol and folipastatin (also structurally related to nidulin) were isolated from \textit{A. unguis} (Hamano et al. 1992). Eight mutagenic depsidones (mollicellins A, B, C, D, E, F, G, and H) were also isolated from \textit{Chaetomium mollicellum} (Stark et al. 1978). More recently, a novel depsidone, excelsione, has been isolated from an endophytic fungus of a native plant in New Zealand (Lang et al. 2007). The A-ring of excelsione is structurally similar to cryptostictic acid, while the B-ring is similar to variolaric acid (Figure 1.29). The structural similarity of excelsione to lichen depsidone strongly suggests that excelsione is produced via a mechanism analogous to that of lichen depsidones.
1.2.6.2.6 Diphenyl ethers

Diphenyl ethers are relatively rare in lichens when compared to depsides and depsidones (see Figure 1.9). Except for micareic acid and canesolide, most of the diphenyl ethers known have the ether bond joined at the meta-position of the B-ring (e.g. β-alectoronic acid) and are often related to depsidones (containing the same or similar monoaromatic units) (Table 1.9). Buellolide and canesolide are two diphenyl ethers that are related to diploicin, and all three have a decarboxylated B-ring. To date, leprolomin is the only diphenyl ether that incorporates a methylphloroacetophenone moiety.

Figure 1.29. Excelsione and the structurally related lichen depsidones – variolaric acid and cryptostistic acid.
Table 1.9. Lichen diphenyl ethers. Coupling of aromatic units via ether bond.

<table>
<thead>
<tr>
<th>Monomeric aromatic intermediates</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Orsellinic acid" /></td>
<td><img src="image2" alt="Buellolide" /></td>
</tr>
<tr>
<td><img src="image3" alt="Tetraketide resorcylic acids" /></td>
<td><img src="image4" alt="Micareic acid" /></td>
</tr>
<tr>
<td><img src="image5" alt="Pentaketide resorcylic acids" /></td>
<td><img src="image6" alt="β-Alectoronic acid" /></td>
</tr>
<tr>
<td><img src="image7" alt="Methylphloroacetophenone" /></td>
<td><img src="image8" alt="Leprolomin" /></td>
</tr>
</tbody>
</table>

The co-occurrence of several diphenyl ethers with structurally related depsidones has led to the hypothesis that they are simply hydrolysis products of depsidones. The diphenyl ethers were sometimes referred as ‘pseudodepsidones’ due to their apparent biosynthetic relationship (Comber et al. 1989; Huneck 2001). An example of such co-occurrence of structurally related diphenyl ethers and depsidones are loxodin [1], loxodinol [2] norlobaridone [3], isonorlobaridone [4] and norlobariol methylpseudoester [5] in *Xanthoparmelia scabrosa* and *Xanthoparmelia amplexula* (Figure 1.30). The co-occurrence of these compounds in the same lichen suggests a biosynthetic relationship between the depsidones and diphenyl ethers (McCaffery 1997).
Figure 1.30. Proposed biosynthetic relationship between the depsidones – loxodin [1], norlobaridone [3], and the diphenyl ethers – loxodinol [2], isonorlobaridone [4] and norlobariol methylpseudoester [5]. [5] could be obtained directly from norlobaridone by methanolysis.

Not all diphenyl ethers, however, co-occur with or have a structurally related depsidone in the same or phylogenetically related lichens. At least the micareic acid and leprolomin require an alternative mechanism to account for their formation. For
micareic acid, a mechanism via Smiles rearrangement of a \textit{para}-depside, prasinic acid (superlatolic acid), is possible (Figure 1.31). The co-occurrence of the two compounds in the lichen \textit{Micarea prasina} (Huneck & Yoshimura 1996) further supports this theory.

\begin{center}
\includegraphics[width=\textwidth]{figure1.31.png}
\end{center}

\textbf{Figure 1.31.} Proposed Smiles rearrangement of prasinic acid to micareic acid.

Since a carboxyl group is absent in the methylphloroacetophenone moiety of leprolomin, the formation of an ester bond to produce depsidone is not possible. The most likely biosynthetic route to leprolomin is via phenol oxidative coupling (B-C type coupling, Figure 1.14). Therefore, leprolomin could be seen as biosynthetically related to usnic acid and contortin (a biphenyl) as they could all result from coupling of different mesomeric forms of methylphloroacetophenonoxy radicals (Figure 1.32). Since these three compounds do not co-occur in the same lichens, it suggests that the enzyme (peroxidase) that catalyzes the phenol oxidation could incorporate the specific mesomeric form of methylphloroacetophenonoxy radicals to form respectively leprolomin, usnic acid and contortin.
Figure 1.32. Proposed phenol oxidative coupling of methylphloroacetophenone to form leprolomin (diphenyl ether), usnic acid (dibenzofuran-like) and contortin (biphenyl).
1.2.6.2.7 Depsones

Depsones are rare when compared to depsides and depsidones. So far only eight have been described (Huneck 2001; Huneck & Yoshimura 1996). All eight depsones are formed solely by tetraketide resorcylic acids of various alkyl chain lengths. The biosynthesis of depsones is thought to proceed via phenol oxidative coupling, as in depsidones.

The best-known depsone is picrolichenic acid. It was the first depsone to be discovered and was found in Pertusaria amara (Wachtmeister 1958). Subsequently, most depsones are named after it (e.g. subpicrolichenic acid, megapicrolichenic acid, hyperpicrolichenic acid, etc.). The unusual structure of picrolichenic acid was central to the early development of the theory of phenolic coupling as a possible mechanism for the biosynthesis of depsides and diphenyl ethers (Culberson & Dibben 1972). Picrolichenic acid is an isomer to the depsidone colensoic acid and is related to the depsides 2-O-methylnanziac acid and 2-O-methylperlatic acid, and the diphenyl ether epiphorellic acid I. All four compounds incorporate two pentenylresorcylic acids, albeit bonded via alternative linkages (Figure 1.33). The discovery of 2-O-methylperlatic acid and 2'-O-methylperlatic acid, which resemble the proposed depside precursor of picrolichenic acid in six species of Pertusaria, supports the proposed biosynthetic relationship between depsides and depsones (Culberson & Dibben 1972).

More recently, two unusual trinuclear hybrid compounds, friesiic acid and 2'-O-methylfriesiate, with both depside and depsone moieties, have been isolated from the lichen Hypocenomyce friesii (Elix et al. 2004) (Figure 1.34).
Figure 1.33. Proposed biosynthetic relationship between picrolichenic acid (depsone), colensoic acid (depsidone), epiphorellic acid I and perlatolic acid (depside). 2-O-methylanziaic acid is the likely precursor of picrolichenic acid, while anziaic acid is a possible precursor of colensoic acid and epiphorellic acid I.
1.2.6.2.8 Dibenzofurans and usnic acid homologs

Dibenzofurans are the third most abundant group of coupled phenolics in lichens after depsides and depsidones. They mostly consist of orcinol-type monoaromatic units, which include orsellinic acid and the resorcylic acids (Table 1.10). Although the formation of dibenzofuran with $\beta$-orsellinic acid and its derivatives is theoretically possible, such a compound has never been found in lichens. Usnic acid and related compounds (e.g. placodiolic acid) are formed by coupling of methylphloroacetophenone. They are technically not dibenzofurans due to the lost of aromaticity in the second ring but their biosynthesis is likely to involve similar mechanisms.

The chemistry and biosynthesis of usnic acid is well-studied due to its interesting structure, wide distribution in lichens and various biological activities (Ingolfsdottir 2002). After the demonstration of usnic acid synthesis via oxidative coupling of methylphloroacetophenone (Barton et al. 1956), the biosynthetic route was further confirmed by isotope incorporation experiments (Taguchi et al. 1966). The incorporation study by Taguchi et al. (1996) showed that the radioactively-labelled methylphloroacetophenone, and not phloroacetophenone, was incorporated into usnic acid, suggesting that the methylation occurred at an earlier stage in the biosynthesis. This agrees with the proposed mechanism for methylphloroacetophenone biosynthesis in Figure 1.10 (Section 1.2.5.2.1).
**Table 1.10.** Dibenzofurans and usnic acid homologs. Coupling of aromatic units via carbon-carbon and ether linkages.

<table>
<thead>
<tr>
<th>Compound type</th>
<th>Monomeric aromatic intermediates</th>
<th>Examples</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dibenzofuran</strong></td>
<td><img src="image1" alt="Dibenzofuran" /></td>
<td><img src="image2" alt="Porphyric acid" /></td>
<td>The A-ring of subdidymic acid is decarboxylated before coupling.</td>
</tr>
<tr>
<td></td>
<td>orsellinic acid</td>
<td><strong>Didymic acid</strong></td>
<td>The A-ring of haemophaein is decarboxylated before coupling (Mosbach 1969).</td>
</tr>
<tr>
<td></td>
<td>tetraketide resorcylic acids $R= C_{2n+1}H_{2n+1}$</td>
<td><strong>Haemophaein</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pentaketide resorcylic acids $R= C_{2n+1}H_{2n+1}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Usnic acid and placodiolic acid derivatives</strong></td>
<td><img src="image3" alt="Usnic acid and placodiolic acid derivatives" /></td>
<td><strong>Usnic acid</strong></td>
<td>The B-ring of both usnic acid and placodiolic acid lose their aromaticity (Dewick 2002).</td>
</tr>
<tr>
<td></td>
<td><img src="image4" alt="Methylphloroacetophenone" /></td>
<td><strong>Placodiolic acid</strong></td>
<td></td>
</tr>
</tbody>
</table>

A revised mechanism for usnic acid biosynthesis based on Dewick (2002) is illustrated in Figure 1.35. After the radical coupling, the ether bond is formed by nucleophilic attack of the phenol on the enone system. Only the left-hand ring is restored to aromaticity by keto-enol tautomerism, because radical coupling occurs on the methyl-containing position *para* to the original phenol on the right-hand ring (Dewick 2002). Since two equivalent phenol groups are present as potential nucleophiles, and two equivalent enone systems are also available, four different products can theoretically be formed (Taguchi et al. 1966) (Figure 1.36). Dewick
(2002) reasoned that the reaction must be enzyme-mediated since most of the time only one product, usnic acid, is formed. The alternative nucleophilic attack by a different phenol group on the same enone system [route (b)] would lead to isousnic acid (B, Figure 1.36), which has been found along with usnic acid in Cladonia metis at a ratio of 1:10 (isousnic: usnic acid) (Taguchi & Shibata 1967). The other two isoforms of usnic acid (C and D, Figure 1.36) have not been found so far. Other compounds closely related to usnic acid in lichens are placodiolic acid (Table 1.10) and isoplagodiolic acid.

Figure 1.35. Biosynthesis of usnic acid via phenol oxidative coupling follows by a nucleophilic attack of a phenol group on to an enone system (Dewick 2002).
Although usnic acid has only been identified in lichens, closely related compounds have been found in filamentous fungi, e.g. the mycousnine and isomycousnine in *Mycosphaerella nawae* (Sassa & Igarashi 1990) and cercosporamide and usnic acid amide in *Cercosporidium henningsii* (Conover et al. 1992). Mycousnine and isomycousnine are identical to isoplacodiolic and isopseudoplacodiolic acid from the lichens *Haematoma flexuosum* and *H. matogrossense* (Huneck 2001).

There has been no direct biosynthetic study of dibenzofurans but their pathway is usually inferred from the established usnic acid biosynthetic mechanism (Mosbach 1969). Two types of radical coupling are commonly observed among the dibenzofurans. Poryphyrilic acid and strepsilin are used as examples to illustrate the case (Figure 1.37). The radical coupling of monoaromatic units to polyphyrilic acid is similar to usnic acid (Figure 1.37A); since there is no methyl group at both sites where the radical coupling occurs (unlike in usnic acid), and the aromaticity of both rings in the dibenzofurans can be restored. The biosynthesis of strepsilin involves decarboxylation of the left-hand ring (Figure 1.37B); the decarboxylation could occur after the radical coupling but before the heterocyclic ring formation. Unlike usnic acids and its homologs (including placodiolic acid and isoplacodiolic acids), which are
mainly heterodimers of methylpholoroacetophenone units, dibenzofurans often incorporate different monoaromatic units to form a heterodimer (e.g. didymic acid, Table 1.10).

![Proposed biosynthesis for A) poryphyrilic acid and B) strepsilin. Adapted from Mosbach (1969).](image)

**Figure 1.37.** Proposed biosynthesis for A) poryphyrilic acid and B) strepsilin. Adapted from Mosbach (1969).

### 1.2.6.2.9 Biphenyls

As present, contortin (Figure 1.38), isolated from *Psoroma contortum* (Elix et al. 1984a), is the only biphenyl known from lichens. The most direct route for its biosynthesis would be via phenolic coupling of two methylpholoroacetophenone units (as described in Section 1.2.5.2.6, Figure 1.32), followed by O-methylation of the hydroxyl groups (alternatively it may arise from phenolic coupling of two molecules of 2,4-di-O-methyl-3-methylphloroacetophenone). Since the mechanism for carbon-carbon bond formation between two monoaromatic units is common in lichens, as exemplified in the formation of the usnic acids, dibenzofurans and depsones, it is interesting that contortin is the only biphenyl known. Another compound also formed from two monoaromatic units bound by a carbon-carbon linkage is the diphenylmethane, bis-(2,4-dihydroxy-6-n-propylphenyl)methane, found in *Protousnea*
sp. (Kinoshista et al. 1994a). The mechanism involved in the biosynthesis of diphenyl methane is unknown.

![Contortin and bis-(2,4-dihydroxy-6-n-propylphenyl)-methane](image)

**Figure 1.38.** Lichen biphenyl (contortin) and diphenylmethane [bis-(2,4-dihydroxy-6-n-propylphenyl)-methane].

1.2.6.3 Lichens as combinatorial chemists

The most intriguing feature observed among the lichen coupled phenolic compounds is their ability to generate a huge diversity of compounds by coupling a limited set of monoaromatic units via different linkages and subsequently modify the side chains or substituents on the aromatic rings (secondary derivatization). Although mostly homodimeric compounds are shown in section 1.2.5.2.4-8 for simplicity, many of the coupled compounds are derived from different monoaromatic units. This approach, which involves mix-and-match of different basic building blocks, greatly resembles the combinatorial synthesis approach used by a chemist to generate libraries of small molecular weight compounds for drug screening.

As exemplified in Figure 1.15, sequential oxidation of the C-3 and C-5 substituents of β-orsellinic acid can lead to 15 possible derivatives (16, if unmodified β-orsellinic acid is included). Of the 16 monoaromatic units, not all are known to be incorporated into lichen depsides and depsidones and together, they form at least 20 different compounds (see Table 1.6). Assuming that β-orsellinic acid and its 15 derivatives can be bonded to form para-depsides, meta-depsides or depsidones, there would be at least 768 (16x16x3) possible combinations. Further modifications by decarboxylation, O-methylation, halogenation and/or esterification with different chain length fatty acids at different sites will further amplify the diversity. One can imagine that, if other basic monoaromatic units (orsellinic acid, tetraketide and pentaketide resorcylic acids with various alkyl chain lengths) and different linkages
(dibenzofurans, depsones, diphenyl ethers, etc.) were to be included, the possibilities are almost unlimited (Figure 1.39). This raises the interesting question – how do lichens do combinatorial synthesis?

**Figure 1.39.** The hierarchy of variations introduced at each stage of the biosynthesis of lichen coupled phenolic compounds. The variations introduced at each level increase the diversity by several folds.

The genes encoded for expression of secondary metabolites in fungi are normally clustered together, although a few exceptions do exist (Hoffmeister & Keller 2007; Keller & Hohn 1997). Sequencing of the biosynthetic gene clusters that produce homologous compounds have allowed synteny comparisons of the gene clusters and have provided clues as to genetic differences that account for the variation in the chemical structure of the products. So far, only a few homologous
gene clusters have been sequenced, e.g. the lovastatin and compactin biosynthetic gene clusters, and the sterigmatocystin and aflatoxins biosynthetic clusters. For example, aflatoxin biosynthesis in *A. parasiticus* involves a cluster of 25 genes in a contiguous 70 kb region (Yu et al. 2004). Similarly, a 60 kb cluster with 25 genes in *A. nidulans* is involved in biosynthesis of the aflatoxin precursor, sterigmatocystin (Brown et al. 1996b). More recently, biosynthetic genes responsible for production of dothistromin, a compound similar to the aflatoxin precursor versicolorin B, have been cloned from *Dothistroma pini* (*Mycosphaerella pini*, a pine tree pathogen) (Bradshaw et al. 2006; Bradshaw & Zhang 2006; Zhang et al. 2007). Interestingly, the dothistromin biosynthetic genes are fragmented into three mini-clusters. Comparison of the three related biosynthetic gene clusters has shown that many homologous genes are involved in the production of aflatoxins, sterigmatocystin and dothistromin, although the gene orientation is not fully conserved among them (Figure 1.40). For instance, the biosynthesis of norsolorinic acid, the common precursor for sterigmatocystin, aflatoxin B₁, and dothistromin, involves a PKS homolog (norsolorinic acid synthase) present in the three gene clusters (Figure 1.41). It would be expected that the biosynthetic genes in lichens are also clustered and homologous genes are involve in the biosynthesis of structurally similar compounds.

As discussed in section 1.2.5.2, there are still many questions and hypotheses regarding the biosynthesis of lichen coupled phenolics that remain to be answered or tested. For example, the proposed PKSs responsible for production of the basic monoaromatic units, the hypothetical peroxidases involved in phenol oxidative coupling, enzymes involved in the depsidone formation, etc. are all yet to be determined. Although there are disputes about which mechanisms or enzymes might be involved in the coupling of various phenolic units, a PKS (or more than one) is undeniably involved in the biosynthesis of these compounds.
Figure 1.40. Comparison of the sterigmatocystin (ST), dothistromin (DOT) and aflatoxin (AF) gene clusters and their corresponding metabolic products. Genes predicted to be homologous and have similar functions are in the same colour. Putative AflR binding sites in promoter regions of the genes are indicated by stemmed balls. Figure reproduced from Zhang (2000).
Figure 1.41. Biosynthesis of sterigmatocystin, aflatoxin B₁, and dothistromin via the common precursor – norsolorinic acid. The homologous PKS genes involved in the biosynthesis of norsolorinic acid are stcA (sterigmatocystin), pksA (dothistromin), and aflC (aflatoxin B₁).
Assuming the biosynthetic genes are also clustered together in lichens, identification of the PKS gene involved in the biosynthesis of a specific compound would also locate the related genes in the same pathway. Studying the biochemistry of individual enzymes coded in the gene cluster would provide valuable insight into the various coupling mechanisms and post-PKS modifications. It is reasonable to predict that the enzymes and genes involved in the same coupling mechanism would be similar among the lichens. Although little is known about the genetics of lichen secondary metabolism, it is possible that lichens generate the diverse array of coupled phenolic compounds by varying the combinations of genes in the cluster/pathway at a few key points (i.e. at the specific level of variation in Figure 1.39). For example, a biosynthetic pathway/gene cluster produces a particular depside by combining two different monoaromatic polyketide units, arbitrarily termed A and B. A genetic recombination event that replaces the “depside synthase” with a “dibenzofuran synthase” in the pathway/gene cluster will produce a different compound (a dibenzofuran instead of depside) but with the same A-B monoaromatic subunits. Likewise, changing the PKS gene in the pathway/gene cluster would result in a new depside with a different combination of monoaromatic subunits (e.g. A-C), while changing the post-PKS tailoring enzymes would result in a depside with A-B’ monoaromatic subunits. Studying the different biosynthetic gene clusters responsible for the production of different but related compounds by comparing the sequences and gene orientations would provide clues to the genetic basis of biosynthesis in lichens as well as an understanding of the evolution of secondary metabolic pathways.

The understanding of the genetics and biosynthesis of lichen coupled phenolics would open up an interesting aspect of research in future, in which new combinations of genes or gene fragments might be artificially introduced to produce new coupled phenolic compounds. This concept is not new and has been termed “combinatorial biosynthesis”, i.e. the systematic modification and interchange of biosynthetic genes (or parts of genes encoded for functional domains) with the consequential production of “unnatural” or “hybrid products” (Rix et al. 2002). It has been a subject of intensive research in bacteria, especially among the actinomycetes. Due to the rapid increase in literature and numerous advances, the subject has also been extensively reviewed
at frequent intervals (e.g. Baltz et al. 2006; Floss 2006; Wilkinson & Micklefield 2007). Due to their modular nature, PKS and non-ribosomal peptide synthase (NRPS) genes have been central to these studies. Most studies have focused upon early steps catalyzed by PKSs and NRPSs, but there have also been significant developments in combinatorial biosynthetic approaches in post-PKS tailoring steps (Rix et al. 2002).

By contrast with bacterial systems, combinatorial biosynthesis in fungi is still relatively undeveloped. Recently, a novel hexaketide naphthalene was synthesized by a chimeric polyketide synthase that was derived from a combination of gene fragments of pentaketide (*Colletotrichum lagenarium* PKS1) and heptaketide (*Aspergillus nidulans* WA) synthases (Watanabe & Ebizuka 2002). The potential to combine different fragments of fungal PKS genes to generate metabolic diversity is more challenging than bacterial modular PKS due to the iterative nature of fungal PKSs. It is, however, a potentially useful way of studying the intricate mechanisms of fungal PKSs. The combinatorial approaches reflected in the biosynthesis of lichen coupled phenolics is somewhat different; instead of varying the PKS genes to increase the metabolite diversity, the lichens use only a small number of monoaromatic polyketide units as basic building blocks, but modify them with post-PKS tailoring enzymes and combine them with other polyketide units to produce the diverse array of coupled phenolics. The post-PKS tailoring enzymes play a more important role here in generating the observed diversity. Investigating the genetics and biosynthesis of lichen coupled phenolics is likely to reveal novel biosynthetic genes that could be applied in combinatorial biosynthesis.
1.3 Research scope and objectives

The review in this chapter demonstrates that there are still many unanswered questions regarding the biosynthesis of lichen polyketides, in particular the coupled phenolics, and that identifying the PKSs involved in the biosynthesis is the key to understanding these problems. Thus, the major goal of this thesis was to investigate the potential PKS genes involved in biosynthesis of some characteristic lichen polyketides. The overall research comprised the following specific aims:

1. Screening for lichen PKS genes potentially involved in biosynthesis of coupled phenolic compounds (Chapter 3). This was accomplished by development of new degenerate primers that targeted the KS domains of potential groups of PKS genes and analysis of the KS domain sequences using phylogenetic methods.

2. Development of a method for locating and cloning of PKS genes from environmental lichen DNA (Chapter 4). The selected KS domain PCR products obtained from Chapter 2 were used as probes to recover the full-length PKS genes from the environmental lichen DNA. This method bypassed the cumbersome lichen mycobiont culturing stage.

3. Attempts to characterize the lichen PKS gene by heterologous expression and identification of the polyketide product (Chapter 5). *Aspergillus nidulans* and *A. oryzae* were tested as potential heterologous hosts.
CHAPTER 2  COMMON MATERIALS AND METHODS

2.1 Introduction ................................................................................................................... 77

2.2 Molecular biology techniques ...................................................................................... 77

2.2.1 Small-scale lichen DNA extraction ..................................................................... 77

2.2.2 Large-scale lichen DNA extraction .................................................................... 77

2.2.3 Polymerase chain reaction (PCR) ........................................................................ 78

2.2.4 Gel electrophoresis (GE) .................................................................................... 80

2.2.5 DNA purification and gel extraction ................................................................. 81

2.2.6 DNA sequencing ................................................................................................. 81

2.2.7 Plasmid miniprep ................................................................................................. 81

2.2.8 Transformation of *Escherichia coli* by electroporation ...................................... 82

2.2.9 Colony-direct PCR ............................................................................................... 82

2.2.10 Cloning of PCR products .................................................................................... 82

2.2.11 Restriction digestion and ligation of DNA ......................................................... 83

2.2.12 Southern blotting ............................................................................................... 83

2.3 Bioinformatic analysis and software ........................................................................... 83

2.3.1 Raw sequence editing ......................................................................................... 83

2.3.2 Multiple sequence alignment and phylogenetic analysis ................................... 84

2.3.3 Gene annotation and plasmid map drawing ......................................................... 86

2.3.4 Drawing of chemical structures ........................................................................... 86
2.1 Introduction

This chapter describes the common materials and methods, primarily in molecular biology, that were used repetitively throughout the thesis. Unless otherwise stated, standard molecular techniques were adopted (Sambrook & Russell 2001) and all molecular biological kit reagents were used as supplied by the manufacturers following manufacturers’ instructions. Specific materials and methods for a particular chapter will be described separately in the corresponding chapter.

2.2 Molecular biology techniques

2.2.1 Small-scale lichen DNA extraction

Lichen DNA from small scale extractions was used for all routine PCR amplifications. Lichen thalli were brushed to remove field debris and ground into a powder under liquid nitrogen with a mortar and pestle for DNA extraction. The DNA was extracted using a DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany). DNA extracts (5 μL each) were visually examined by gel electrophoresis (Section 2.2.4) and compared with standard DNA marker (GeneRuler 1kb DNA Ladders; Fermentas, Lithuania) to estimate the concentration.

2.2.2 Large-scale lichen DNA extraction

For Southern blotting and cloning purposes, large amount of DNA of high integrity and purity was required and thus the DNA was extracted with a different method at larger scale. Lichen thalli were washed to remove field debris, dried in vacuum and frozen at -80 °C. Dried samples were ground into a powder under liquid nitrogen with a mortar and pestle for DNA extraction. DNA was extracted using the anion-exchange chromatography method described for plants (Gauch et al. 1998) with minor modification – 4 g (instead of 2 g) of ground thalli were used for the extraction and purified through a Genomic-tip 500/G anion exchange column (QIAGEN). The column was washed with a double volume of washing buffer before elution.


2.2.3 Polymerase chain reaction (PCR)

All PCR thermal cycles were performed in a PCR Express Thermal Cycler (ThermoHybaid, Middlesex, UK) or GeneAmp 2400 (PerkinElmer, Waltham, MA). DNA amplifications were carried out using Taq polymerase (Invitrogen, Carlsbad, CA) with the buffer supplied or alternatively, GoTaq Green Master Mix (Promega, Madison, WI). For amplification of PCR products longer than 3 kb, an enzyme cocktail of Taq DNA polymerase and HotStar HiFidelity DNA polymerase (QIAGEN) was used following the protocol for long PCR products as described by the manufacturer. Nuclease-free water (Promega) was used in all PCR reactions.

Custom oligonucleotide primers (Table 2.1) were ordered from commercial oligonucleotide synthesis services from either GeneWorks (Thebarton, SA, Australia) or Micromon Services (Clayton, VIC, Australia). The primer stocks were made up to a concentration of 25 μM for non-degenerate primers and 50 μM for degenerate primers with nuclease-free water (Promega). For all PCR amplification, 0.5 μL of each forward and reverse primer was used in a 25 μL PCR reaction (final concentration at 0.5 μM for non-degenerate primers, and 1 μM for degenerate primers). For degenerate primers, equal concentrations of forward and reverse primers were used even when both primers varied in the degree of degeneracy. A negative control (without template DNA) was incorporated in all thermal cycles. Some oligonucleotide primers were not custom-ordered but supplied by manufacturers along with the products. NZ-f and NZ-r primers were supplied along with the BigEasy Linear cloning system (Lucigen, Middleton, WI), while 3’CDS A, UPM and NUP primers were supplied along with the SMART RACE cDNA amplification kit (Clontech Laboratories, Mountain View, CA).
**Table 2.1.** The list of primers used in this study, with their corresponding sequences, in the order of appearance in this thesis.

<table>
<thead>
<tr>
<th>PCR primer</th>
<th>Oligonucleotide sequence 5´ to 3´</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITS1F (Gardes &amp; Bruns 1993)</td>
<td>CTTGGTGCATTTAGAGGAAGTAA</td>
</tr>
<tr>
<td>ITS4 (White et al. 1990)</td>
<td>TCCTCCGCTTAATTGATATGC</td>
</tr>
<tr>
<td>LC1 (Bingle et al. 1999)</td>
<td>GATCIAAGTTTTTTAATATG</td>
</tr>
<tr>
<td>LC2c (Bingle et al. 1999)</td>
<td>GTCGCGTCCGGTCCATTTTC</td>
</tr>
<tr>
<td>LC3 (Bingle et al. 1999)</td>
<td>GCCGACARATGGAYCCICA</td>
</tr>
<tr>
<td>LC5c (Bingle et al. 1999)</td>
<td>IGAGTIGCRTGIGCYTIC</td>
</tr>
<tr>
<td>NRKS-F</td>
<td>CAGCGGCCATGGCCCTGACARCKTAYGARGC</td>
</tr>
<tr>
<td>NRKS-R</td>
<td>GCCGCTGCCGGTCCRTGCATYTC</td>
</tr>
<tr>
<td>PRKS-F</td>
<td>GAGGCCGAAGCAGATGGAYCCYCCNARCA</td>
</tr>
<tr>
<td>PRKS-R</td>
<td>AGCGGGGTGGAGTGSCRTGNGYCCYTC</td>
</tr>
<tr>
<td>NR3KS-F</td>
<td>TACCAGCCTCGACCAYMNNTYTT</td>
</tr>
<tr>
<td>NR3KS-R</td>
<td>GGGGTGGAGGGTGCCRTGNGYCCYTC</td>
</tr>
<tr>
<td>M13/pUC forward</td>
<td>ACCACCGGTTCCTTCAAGCGTAGC</td>
</tr>
<tr>
<td>M13/pUC reverse (-46)</td>
<td>GCCAGGGTTTTCCCAAGTCAAGC</td>
</tr>
<tr>
<td>XsNR3KS2-F</td>
<td>AGTCCTCAGAATAGGGCAGTCA</td>
</tr>
<tr>
<td>XsNR3KS2-R</td>
<td>AGAAGCTGCAATATGCAAGCA</td>
</tr>
<tr>
<td>NRMcT-F</td>
<td>CAGCATCCTGGAGATGGGNGCNGG</td>
</tr>
<tr>
<td>NRMcT-R</td>
<td>TCAGTCGGACAACAGCACCACCCYTC</td>
</tr>
<tr>
<td>XsNR3KS2-iF</td>
<td>AGACGGCTACTGCGAGGGTGAAG</td>
</tr>
<tr>
<td>CMeT011a-F</td>
<td>GATGTTCTGCTTCTGGCTTGTC</td>
</tr>
<tr>
<td>CMeT011a-R2</td>
<td>CACCAGCCTTCTACAGGCGCA</td>
</tr>
<tr>
<td>RT52KS-F</td>
<td>ACGATGTCGAAACCTTGGGT</td>
</tr>
<tr>
<td>RT52KS-R</td>
<td>TCGAAAGCTGCTTGGCTG</td>
</tr>
<tr>
<td>GAP1-F</td>
<td>TGCGAAGTATCTTGTCTGCGT</td>
</tr>
<tr>
<td>NZ-F2</td>
<td>CAGTGTGATTACATTGTCG</td>
</tr>
<tr>
<td>NZ-R4</td>
<td>TT ATTGACAGCGCAGCC</td>
</tr>
<tr>
<td>attB1Cse-N</td>
<td>GGGGACAAAGTTTTGTAACAAAAAAGCAGGTTACTGCCACCATCGGCGACATTCAACCTTTCTCAAG</td>
</tr>
<tr>
<td>attB2Cse-C</td>
<td>GGGGACACATTGTAACAAAAAGCTGGGTATTCAAGTCAGTGTACCTTGGTCAC</td>
</tr>
<tr>
<td>3´-CDS A</td>
<td>AAGCAGTGCGTATCAACGCGAGTAC(T)30VN</td>
</tr>
<tr>
<td>UPM (Universal Primer A Mix)</td>
<td>Long (0.4 µM): CTAATACGACTCACTATAGGGAAGCAGCAGTTGATCAACG</td>
</tr>
<tr>
<td></td>
<td>Short (2 µM): CTAATACGACTCACTATAGGGAAGCAGCAGTTGATCAACG</td>
</tr>
<tr>
<td>NUP (Nested Universal Primer A)</td>
<td>AAGCAGTGCGTATCAACGCGAGTAC</td>
</tr>
<tr>
<td>GAP3-F</td>
<td>TGACGGCTTTTGATGAGCCTCG</td>
</tr>
<tr>
<td>3`RACEpksMT</td>
<td>ACCGAGCCCTTCTGGGTGCGA</td>
</tr>
</tbody>
</table>
Standard touchdown PCR cycle conditions were employed for all PCR amplifications to minimize time required for optimizing individual PCR (Figure 5.21). The optimal annealing temperature ($T_a$) range for the touchdown cycles for each primer pair was calculated using the “Primer Test” function embedded in FastPCR 3.6.6 by Kalendar (www.biocenter.helsinki.fi/bi/programs/fastpcr.htm). The formula used for the calculation of $T_a$ in FastPCR 3.66 was modified from (Rychlik et al. 1990) and applied to degenerate primers as well. Touchdown PCR cycles started at the highest annealing temperature based on the calculated $T_a$ range and decreased 1°C per cycle for 6 cycles, followed by an additional 29-34 cycles at the lowest annealing temperature of the calculated $T_a$ range. Initial denaturation was at 95°C for 3 mins and denaturation steps during the cycle were at 94°C for 30 seconds each. Extension steps during the cycles were at 72°C for 1 minute in most standard cycles, with a final extension step for 7-10 minutes. For longer products, the extension time was prolonged by 1 minute with every additional 1 kb product expected. To improve PCR specificity, a simple hot start approach was used for all PCR cycles, in which the thermal cycler was pre-warmed to above 80°C before the tubes were placed on the heating block.

![Figure 2.1. Standard touchdown PCR cycle conditions. $T_a$, optimal annealing temperature.](image)

### 2.2.4 Gel electrophoresis (GE)

DNA was electrophoresed with a Sub Cell GT gel electrophoresis apparatus (BioRad, Hercules, CA) using agarose gel (AppliChem, Darmstadt, Germany). For resolving PCR products and plasmid DNA, 1.0 - 1.4% agarose was used at 6 V/cm in...
TBE buffer. For Southern Blotting purposes, DNA was resolved on 0.7% agarose gel at 4V/cm in TAE buffer. TAE and TBE buffer were prepared from chemicals and reagents from Sigma (St. Louis, MO) following the recipes as described elsewhere (Sambrook & Russell 2001). The gel was stained in ethidium bromide solution (6 μg/mL) for 10 min, destained in running water for 30 min, and photographed using a gel documentation system (GelDoc, Bio-Rad). Standard DNA markers (GeneRuler 1 kb or 100 bp Plus DNA Ladders; Fermentas) was used to estimate the size and concentration of the DNA when required.

2.2.5 DNA purification and gel extraction

Qiaquick PCR Purification Kit (QIAGEN) was used for general purification of PCR products and DNA from other enzymatic reactions. For extraction of DNA smaller than 10 kb from agarose gels, the corresponding region was excised and extracted using a Qiaquick Gel Purification Kit (QIAGEN). DNA larger than 10 kb was extracted from agarose gels using a QIAEX II Gel Extraction Kit (QIAGEN).

2.2.6 DNA sequencing

DNA templates for sequencing were column-purified using methods as described in Section 2.2.5 before proceeding with cycle sequencing. Cycle sequencing chemistry was performed with the PRISM BigDye Terminator Mix version 3.1 (Applied Biosystems, Foster City, CA) in a GeneAmp 2400 (PerkinElmer). Capillary electrophoreses and detection was carried out at the Micromon DNA Sequencing Facilities, Monash University (Clayton, VIC, Australia) using a 3730S Genetic Analyser (Applied Biosystems).

2.2.7 Plasmid miniprep

Plasmid was prepared from 1-2 mL overnight bacterial culture by the standard alkaline-SDS lysis method (minipreparation) as described in Sambrook & Russell (2001) or using QIAquick miniprep kit (Qiagen) when plasmid of higher purity was required.
2.2.8 Transformation of *Escherichia coli* by electroporation

Bacterial strains used for general cloning were DH5α (Invitrogen) or SURE cells (Stratagene). The electrocompetent cells were prepared according to Sambrook & Russell (2001). Electrocompetent cells (50 μL) were mixed with plasmid DNA (0.1-1 μg) and the mixture was chilled on ice for 1 min. If the competent cells had been stored at -70 °C, they were first thawed on ice for 10 min before mixing with DNA. The entire mixture was transferred into a chilled electroporation cuvette (2 mm gap) (Molecular Bioproducts, San Diego, CA). The Gene Pulser apparatus (Bio-Rad) was set at 25 mF capacitance, 2.0 kV, and 200 Ω resistance. The mixture was pulsed once at these settings and the cuvette was immediately filled with 1 mL of Luria-Bertani (LB) broth medium (Sambrook & Russell 2001). The resultant mixture in LB medium was transferred to a culture tube and incubated at 37 °C for one hour before transferring to LB-agar plates with the appropriate antibiotic for selection.

2.2.9 Colony-direct PCR

For blue-white screening, the white colonies were picked using autoclaved/sterilized toothpicks and suspended in 10 μL of PCR grade water (Fisher Biotec). To lyse the cells, the cell suspensions were heated to 100 °C for 5 min and cooled to room-temperature. After a brief centrifugation, the DNA containing supernatant (1 μL) was transferred into a PCR reaction mixture and the PCR amplification proceeded according to Section 2.2.3.

2.2.10 Cloning of PCR products

The final extension step of the PCR cycles (Section 2.2.3) was run for 10 minutes at 72 °C to ensure proper A-tailing. PCR products were column-purified before ligation with pGEM-T Easy vector using the pGEM-T Easy Cloning System (Promega). Electrocompetent cells were transformed with 1-2 μL of the ligation mixture (as described in Section 2.2.9). After growing the transformants on LB-agar-ampicillin (100 μg/mL) plates overnight, 5-8 colonies from each PCR reaction were screened with colony-direct PCR (Section 2.2.10).
2.2.11 Restriction digestion and ligation of DNA

All restriction enzymes used were either from Promega or New England Biolabs (Ipswich, MA). They were chosen based on their buffer compatibility during double-digestions. All the enzymes were used with the buffer supplied. For common digestion of plasmids and PCR products, 1-3 U of enzyme was used per µg DNA and incubated for 2-4 hours. For genomic DNA digestions for Southern blot purposes, approximately 5 U of enzyme per µg DNA was used and the digestion was incubated overnight (16 hours). Plasmids and DNA fragments were ligated using T4 DNA ligase (Promega) and incubated at 16°C for 4 hours or 4°C overnight.

2.2.12 Southern blotting

Digoxygenin (DIG)-labelled DNA probes were generated from PCR products or DNA fragments using a DIG-High Prime kit (Roche Applied Sciences, Mannheim, Germany). Approximately 10-20 µg of genomic DNA was digested with either a single or combination of two restriction enzymes and incubated overnight at 37°C to ensure complete digestion (Section 2.2.11). Digested genomic DNA was electrophoresed on a 0.7% (w/v) agarose gel (Section 2.2.4) and was transferred onto positively charged nylon membranes (Roche Applied Sciences). The membranes were hybridized with the DIG-labelled probe and detected with the DIG Luminescent Detection System (Roche Applied Sciences) and Kodak BioMax Light-1 Film (Eastman Kodak, NY). Kodak GBX developer and fixer were used to develop the film exposed to the chemiluminescent blot for 25-60 min.

2.3 Bioinformatic analysis and software

2.3.1 Raw sequence editing

All sequence electrophoregrams were visually examined and the raw sequences were edited using BioEdit v7.0.52 (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) (Hall 1999) and/or Sequence Scanner v1.0 (Applied Biosystems). Ambiguous base calls at both ends of sequences were removed and in some cases, forward and reverse
complement sequences were aligned pairwise on BioEdit v7.0.52 to check for accuracy.

### 2.3.2 Multiple sequence alignment and phylogenetic analysis

Multiple sequences were aligned with the ClustalW (Thompson et al. 1994) function embedded in Bioedit (Hall 1999), MEGA 4.0 (Tamura et al. 2007) and/or Vector NTI (Invitrogen) (Lu & Moriyama 2004). For DNA sequence alignment, the default parameters in ClustalW were used. For protein sequence alignment, the parameters recommended by Hall (2004) were followed: pairwise alignment parameters (gap opening penalty = 35.00, gap extension penalty = 0.75); multiple alignment parameters (gap opening penalty = 15.00, gap extension penalty = 0.30).

Phylogenetic analysis were performed using MEGA 4.0 (Tamura et al. 2007). The specific phylogenetic method/algorithm and parameters used for constructing the phylogenetic tree will be described in the Materials and Methods of the corresponding chapter. The list of the reference sequences used in this thesis and their corresponding GenBank accession number are listed in Table 2.2.

<table>
<thead>
<tr>
<th>Genus</th>
<th>Species</th>
<th>Gene/protein</th>
<th>GenBank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acremonium</td>
<td>strictum</td>
<td>pks1/ PKS1/ AsPKS1</td>
<td>CA87161</td>
</tr>
<tr>
<td>Alternaria</td>
<td>solani</td>
<td>pksN/ PKSN</td>
<td>BAD83684</td>
</tr>
<tr>
<td>Aspergillus</td>
<td>nidulans</td>
<td>wA/ WA</td>
<td>Q03149/CAA46695</td>
</tr>
<tr>
<td>Aspergillus</td>
<td>fumigatus</td>
<td>alb1/ ALB1/ pksP/ PKSP</td>
<td>EAL94057/ XP_756095</td>
</tr>
<tr>
<td>Aspergillus</td>
<td>nidulans</td>
<td>stcA/STCA/ pksST/ PKSST</td>
<td>EAA61613/ XP_681094/ Q12397</td>
</tr>
<tr>
<td>Aspergillus</td>
<td>parasiticus</td>
<td>pksL1/ PKSL1</td>
<td>AAS66004/ AAC41675/ Q12053</td>
</tr>
<tr>
<td>Aspergillus</td>
<td>terreus</td>
<td>at4/ AT4</td>
<td>BAB88689</td>
</tr>
<tr>
<td>Aspergillus</td>
<td>terreus</td>
<td>at1/ AT1</td>
<td>BAB88688</td>
</tr>
<tr>
<td>Aspergillus</td>
<td>terreus</td>
<td>at5/ AT5</td>
<td>BAB88752</td>
</tr>
<tr>
<td>Aspergillus</td>
<td>terreus</td>
<td>ATEG_10080</td>
<td>XP_001209382/ EAU29529</td>
</tr>
<tr>
<td>Aspergillus</td>
<td>parasiticus</td>
<td>pksL2/ PKSL2</td>
<td>AAC23536</td>
</tr>
<tr>
<td>Aspergillus</td>
<td>terreus</td>
<td>lovB/ LovB/ LNKS</td>
<td>AAD39830/Q9Y8A5/Q0C8M3</td>
</tr>
<tr>
<td>Aspergillus</td>
<td>terreus</td>
<td>lovF/ LovF/ LDKS</td>
<td>AAD34559</td>
</tr>
<tr>
<td>Aspergillus</td>
<td>ochraceus</td>
<td>LC35-12</td>
<td>AAT92023</td>
</tr>
<tr>
<td>Bipolaris</td>
<td>oryzae</td>
<td>pks1/ PKS1</td>
<td>BAD22832</td>
</tr>
<tr>
<td>Bombyx</td>
<td>mori</td>
<td>FAS p270</td>
<td>AAB53258</td>
</tr>
<tr>
<td>Botryotinia</td>
<td>fuckeliana</td>
<td>pks14/ PKS14/ BfPKS14</td>
<td>AAR90250</td>
</tr>
<tr>
<td>Botryotinia</td>
<td>fuckeliana</td>
<td>pks15/ PKS15/ BfPKS15</td>
<td>AAR90251</td>
</tr>
<tr>
<td>Organism</td>
<td>Protein Name</td>
<td>Accession</td>
<td></td>
</tr>
<tr>
<td>----------------------------------</td>
<td>--------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>Botryotinia fuckeliana</td>
<td>pks16/ PKS16</td>
<td>BfPKS16</td>
<td>AAR90252</td>
</tr>
<tr>
<td>Botryotinia fuckeliana</td>
<td>pks18/ PKS18</td>
<td>BfPKS18</td>
<td>AAR90254</td>
</tr>
<tr>
<td>Botryotinia fuckeliana</td>
<td>pks19/ PKS19</td>
<td>BfPKS19</td>
<td>AAR90255</td>
</tr>
<tr>
<td>Botryotinia fuckeliana</td>
<td>pks20/ PKS20</td>
<td>BfPKS20</td>
<td>AAR90256</td>
</tr>
<tr>
<td>Botryotinia fuckeliana</td>
<td>pks17/ PKS17</td>
<td>BfPKS17</td>
<td>AAR90253</td>
</tr>
<tr>
<td>Caenorhabditis elegans</td>
<td>FAS</td>
<td></td>
<td>NP_492417</td>
</tr>
<tr>
<td>Ceratocystis resiniifera</td>
<td>pks1/ PKS1</td>
<td></td>
<td>AAO60166</td>
</tr>
<tr>
<td>Cercospora nicotianae</td>
<td>ctb1/ CTB1</td>
<td></td>
<td>AAT69682</td>
</tr>
<tr>
<td>Coccomyces pocillarium</td>
<td>&quot;0202F&quot;</td>
<td></td>
<td>AY918721</td>
</tr>
<tr>
<td>Coccomyces maritimum</td>
<td>&quot;1471A&quot;</td>
<td></td>
<td>AY918717</td>
</tr>
<tr>
<td>Coccomyces coccophorum</td>
<td>&quot;0299A&quot;</td>
<td></td>
<td>AY918715</td>
</tr>
<tr>
<td>Cochliobolus heterostrophus</td>
<td>pks19/ PKS19</td>
<td>ChPKS19</td>
<td>AAR90273</td>
</tr>
<tr>
<td>Cochliobolus heterostrophus</td>
<td>pks20/ PKS20</td>
<td>ChPKS20</td>
<td>AAR90274</td>
</tr>
<tr>
<td>Cochliobolus heterostrophus</td>
<td>pks18/ PKS18</td>
<td>BfPKS18</td>
<td>AAR90272</td>
</tr>
<tr>
<td>Cochliobolus heterostrophus</td>
<td>pks23/ PKS23</td>
<td>ChPKS23</td>
<td>AAR90277</td>
</tr>
<tr>
<td>Cochliobolus heterostrophus</td>
<td>pks22/ PKS22</td>
<td>ChPKS22</td>
<td>AAR90276</td>
</tr>
<tr>
<td>Cochliobolus heterostrophus</td>
<td>pks21/ PKS21</td>
<td>ChPKS21</td>
<td>AAR90275</td>
</tr>
<tr>
<td>Cochliobolus heterostrophus</td>
<td>pks11/ PKS11</td>
<td>ChPKS11/ FUM1</td>
<td>AAR90266</td>
</tr>
<tr>
<td>Cochliobolus heterostrophus</td>
<td>pks2/ PKS2</td>
<td>ChPKS2</td>
<td>AAR90257/ ABB76806</td>
</tr>
<tr>
<td>Cochliobolus heterostrophus</td>
<td>pks5/ PKS5</td>
<td>ChPKS5</td>
<td>AAR90260</td>
</tr>
<tr>
<td>Cochliobolus heterostrophus</td>
<td>pks6/ PKS6</td>
<td>ChPKS6</td>
<td>AAR90261</td>
</tr>
<tr>
<td>Cochliobolus heterostrophus</td>
<td>pks10/ PKS10</td>
<td>ChPKS10</td>
<td>AAR90262</td>
</tr>
<tr>
<td>Colletotrichum lagenarium</td>
<td>pks1/ PKS1</td>
<td></td>
<td>BAA18956</td>
</tr>
<tr>
<td>Dothistroma/ Mycosphaerella</td>
<td>septosporum/ pini</td>
<td>pksAPKSA</td>
<td>AAZ95017</td>
</tr>
<tr>
<td>Exophiala lecanii-corni</td>
<td>pks1/ PKS1/ EIPKS1</td>
<td></td>
<td>AAN75188</td>
</tr>
<tr>
<td>Gallus</td>
<td>gallus</td>
<td>FAS</td>
<td>P12276</td>
</tr>
<tr>
<td>Gibberella zeae</td>
<td>pks12/ PKS12</td>
<td></td>
<td>AAS57296/ AAU10633</td>
</tr>
<tr>
<td>Gibberella fujikuroi</td>
<td>pks4/ PKS4</td>
<td></td>
<td>CAS92399</td>
</tr>
<tr>
<td>Gibberella zeae</td>
<td>pks13/ PKS13</td>
<td>GzPKS13</td>
<td>ABB90282</td>
</tr>
<tr>
<td>Gibberella zeae</td>
<td>pks4/ PKS4</td>
<td></td>
<td>AAS57289/ ABB90283</td>
</tr>
<tr>
<td>Gibberella/ Fusarium</td>
<td>pks3/ PKS3/ GzPKS3</td>
<td></td>
<td>AAS57288</td>
</tr>
<tr>
<td>Gibberella/ Fusarium</td>
<td>pks11/ PKS11</td>
<td>GmPKS11/ FUM1</td>
<td>AAR92218/ ABD72181/ AAD43562</td>
</tr>
<tr>
<td>Glarea lozoyensis</td>
<td>pks1/ PKS1</td>
<td></td>
<td>BAA18956</td>
</tr>
<tr>
<td>Glarea lozoyensis</td>
<td>pks2/ PKS2</td>
<td></td>
<td>AAX35547</td>
</tr>
<tr>
<td>Homo sapiens</td>
<td>FAS</td>
<td></td>
<td>AAC50259</td>
</tr>
<tr>
<td>Micromonospora echinospora</td>
<td>calO5/ CalO5</td>
<td></td>
<td>AAM70355</td>
</tr>
<tr>
<td>Monascus purpureus</td>
<td>pks1/ PKS1</td>
<td></td>
<td>CAC94008</td>
</tr>
<tr>
<td>Monascus purpureus</td>
<td>pksCT/ PKSCT</td>
<td></td>
<td>BAD44749</td>
</tr>
<tr>
<td>Nectria haematococca</td>
<td>pgl1/ PGL1/ FvPGL1</td>
<td>FVEG_03695 (Broad Inst. MIT)</td>
<td></td>
</tr>
<tr>
<td>Nectria haematococca</td>
<td>pgl1/ PGL1/ NhPGL1</td>
<td>108347 (JGI database)</td>
<td></td>
</tr>
<tr>
<td>Nectria haematococca</td>
<td>pksN/ PKSN</td>
<td></td>
<td>AAS48892</td>
</tr>
<tr>
<td>Nodulisporium sp.</td>
<td>melanin PKS</td>
<td></td>
<td>AAD38786</td>
</tr>
<tr>
<td>Ochrolechia pallescens</td>
<td>&quot;1023B&quot;</td>
<td></td>
<td>AY918729</td>
</tr>
<tr>
<td>Ochrolechia androgyna</td>
<td>&quot;1368E&quot;</td>
<td></td>
<td>AY918725</td>
</tr>
<tr>
<td>Penicillium nordicum</td>
<td>otapksPN/ OTAPKSPN</td>
<td>AAP33839</td>
<td></td>
</tr>
<tr>
<td>Penicillium citrinum</td>
<td>mlcA/MlcA</td>
<td></td>
<td>BAC20564</td>
</tr>
<tr>
<td>Penicillium citrinum</td>
<td>mlcB/MlcB</td>
<td></td>
<td>BAC20566</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genus</th>
<th>Species</th>
<th>Accession</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pertusaria</td>
<td>hemisphaerica</td>
<td>&quot;1367A&quot;</td>
<td>AY918760</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>flavida</td>
<td>&quot;1371J&quot;</td>
<td>AY918753</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>pertusa</td>
<td>&quot;1369E&quot;</td>
<td>AY918785</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>scaberula</td>
<td>&quot;1448C&quot;</td>
<td>AY918789</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>panyrga</td>
<td>&quot;5718&quot;</td>
<td>AY918783</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>coronata</td>
<td>&quot;1033B&quot;</td>
<td>AY918748</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>scaberula</td>
<td>&quot;0327C&quot;</td>
<td>AY918787</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>gibberosa</td>
<td>&quot;1061A&quot;</td>
<td>AY918756</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>xanthoplaica</td>
<td>&quot;1079A&quot;</td>
<td>AY918797</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>subverrucosa</td>
<td>&quot;1067A&quot;</td>
<td>AY918793</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>amara</td>
<td>&quot;1473A&quot;</td>
<td>AY918735</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>amara</td>
<td>&quot;1066A&quot;</td>
<td>AY918739</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>mesotropa</td>
<td>&quot;1064A&quot;</td>
<td>AY918777</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>flavida</td>
<td>&quot;1371L&quot;</td>
<td>AY918754</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>alpescens</td>
<td>&quot;1017B&quot;</td>
<td>AY918734</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>oculata</td>
<td>&quot;0457C&quot;</td>
<td>AY918778</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>pertusa</td>
<td>&quot;1369A&quot;</td>
<td>AY918784</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>oculata</td>
<td>&quot;0457D&quot;</td>
<td>AY918779</td>
</tr>
<tr>
<td>Pertusaria</td>
<td>amara</td>
<td>&quot;1066F&quot;</td>
<td>AY918740</td>
</tr>
<tr>
<td>Streptomyces</td>
<td>viridochromogens</td>
<td>aviM/ AviM</td>
<td>AAK83194</td>
</tr>
<tr>
<td>Varicellaria</td>
<td>rhodocarpa</td>
<td>&quot;0528A&quot;</td>
<td>AY918798</td>
</tr>
<tr>
<td>Wangiella/Exophiala</td>
<td>dermatitidis</td>
<td>pks1/ PKS1/ WdPKS1</td>
<td>AAD31436</td>
</tr>
<tr>
<td>Xanthoria</td>
<td>elegans</td>
<td>pks1/ PKSI</td>
<td>ABG91136</td>
</tr>
</tbody>
</table>

2.3.3 Gene annotation and plasmid map drawing

NCBI BLAST tools (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) (including Blastn, Blastp, PSI-Blast, Blastx) was used to search for sequence similarity and infer homology. The searches of functional domains were assisted by NCBI conserved domain search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and EBI InterProScan (http://www.ebi.ac.uk/InterProScan/). The assignments of open-reading frames (ORFs) were assisted by Frameplot ver. 2.3.2 by Ishikawa & Hotta (http://www.nih.go.jp/~jun/cgi-bin/frameplot.pl) and by NCBI ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The annotation of genes and plasmid/restriction map drawing was performed with Vector NTI (Invitrogen).

2.3.4 Drawing of chemical structures

All structural formulae and chemical drawings were illustrated with ACD/ChemSketch Freeware Version 8.17 (Advanced Chemistry Development, Toronto, Ontario).
CHAPTER 3  DIVERSITY OF POLYKETIDE SYNTHASE GENES IN LICHENS

3.1  Introduction ........................................................................................................ 89
3.1.1  Exploring the genetic diversity of fungal PKSs ........................................ 89
3.1.2  Correlating the structure of polyketides and PKS genes with phylogenetic analysis ................................................................. 93
3.1.3  Research goals ................................................................................................. 96

3.2  Lichen identification and chemical analysis ............................................. 97
3.2.1  Materials and methods .................................................................................. 97
3.2.1.1  Lichen collection and identification ....................................................... 97
3.2.1.2  Thin layer chromatography (TLC) ......................................................... 98
3.2.1.3  High performance liquid chromatography (HPLC) ............................... 98
3.2.1.4  PCR amplification of internal transcribed spacer regions .............. 99
3.2.2  Results ............................................................................................................. 99
3.2.2.1  Lichen specimens and identification ...................................................... 99
3.2.2.2  Chemical analysis of lichens ................................................................. 105
3.2.2.3  Internal transcribed spacer (ITS) regions ......................................... 111
3.2.3  Discussion ....................................................................................................... 112

3.3  NR- and PR-type KS domain-specific degenerate primers .............. 113
3.3.1  Materials and methods .................................................................................. 114
3.3.1.1  Design of degenerate primers ................................................................. 114
3.3.1.2  PCR amplification of KS domain ........................................................... 114
3.3.1.3  KS domain sequence analysis .............................................................. 114
3.3.2  Results ............................................................................................................. 115
3.3.2.1  NRKS degenerate primers ................................................................. 115
3.3.2.2  PRKS degenerate primers ................................................................. 115
3.3.2.3  PCR results and KS domain sequence analyses ........................... 118
3.3.3  Discussion ....................................................................................................... 120
3.3.3.1  NRKS amplified KS domains ............................................................. 121
3.3.3.2  PRKS amplified KS domain ............................................................... 121

3.4  Clade III NR-type KS domain-specific degenerate primers ............. 124
3.4.1  Materials and methods .................................................................................. 125
3.4.1.1  Design of degenerate primers ................................................................. 125
3.4.1.2  PCR amplification of KS domain and sequence analysis .................. 125
3.4.2  Results ............................................................................................................. 125
3.4.2.1  NR3KS degenerate primers ................................................................. 125
3.4.2.2  PCR results and KS domain sequence analyses ........................... 127
3.4.3  Discussion ....................................................................................................... 128

3.5  Phylogenetic analysis of lichen KS domains .............................................. 130
3.5.1  Materials and methods .................................................................................. 131
3.5.1.1 Multiple sequence alignment .................................................... 131
3.5.1.2 Construction of phylogenetic tree ............................................. 131

3.5.2 Results..........................................................................................132

3.5.3 Discussion....................................................................................134
  3.5.3.1 Non-reducing PKS Clade I.................................................... 134
  3.5.3.2 Non-reducing PKS Clade II................................................... 139
  3.5.3.3 Non-reducing PKS basal to Clades I and II.............................. 142
  3.5.3.4 Non-reducing PKS Clade III................................................. 143
  3.5.3.5 Partial-reducing PKSs........................................................... 147
  3.5.3.6 Highly-reducing PKSs........................................................... 148

3.6 Conclusions and overview ............................................................... 149
3.1 Introduction

When this project was first proposed in August 2004, there was virtually no publication on characterized PKS gene from lichenized-fungi; neither was there any full-length sequence in the public domain. Although the number of research papers on fungal polyketide synthases has been increasing exponentially, the only available sequences at the inception of this work were the handful of partial β-ketoacyl synthase (KS) domain sequences (total of 15) from the genus *Lecanora*, plus three from other genera (Grube & Blaha 2003). The main objective of the research in this chapter was, therefore, to extend the diversity of fungal PKS genes that could be detected in lichens. The second objective of this study was to investigate if there was a correlation between the polyketide compounds and the PKS genes detected in the individual lichens with the intention of detecting PKS genes potentially involved in biosynthesis of characteristic lichen compounds (coupled phenolics). A brief review on the sampling of fungal PKS diversity and PKS phylogenetic analysis is given below.

3.1.1 Exploring the genetic diversity of fungal PKSs

The most common approach used for surveying the PKS genetic diversity of an organism is by PCR amplification using degenerate primers based on conserved regions of known PKSs. This approach allows the amplification of PCR products that are fragments of new PKS genes and has been successfully employed in both bacteria (e.g. (Ayuso-Sacido & Genilloud 2005; Ehrenreich et al. 2005; Liu et al. 2003; Mofit & Neilan 2001; Ostash et al. 2005; Savic & Vasiljevic 2006; Seow et al. 1997; Snyder et al. 2003) and fungi (see below). The approach has also been used for screening PKS genes present in metagenomes from soil, marine sponges and stromatolites (Burns et al. 2005; Ji et al. 2006; Schirmer et al. 2005; Seow et al. 1997; Wawrik et al. 2005). The main advantage of this approach is the ability to amplify new PKS gene fragments, which can be sequenced and analyzed for desired
properties. Homologous probes could then be generated from the selected PKS gene fragments to locate and clone the candidate PKS genes and associated gene clusters from a genomic or metagenomic library (See also Chapter 4 Section 4.1 for discussion of the use of degenerate primers to facilitate cloning).

The degenerate primer strategy was employed in this study to explore the diversity of PKS genes in environmental lichen samples. The DNA extracted from environmental lichen samples is, in a sense, metagenomic in nature, as it consists of DNA from both symbiotic components (mycobiont and photobiont) and possibly DNA from other environmental sources. Since the major secondary metabolites in lichens are produced by the mycobionts (see Chapter 1 Section 1.2.3), the fungal PKS genes should be responsible for the production of most polyketide compounds detected in lichens. Therefore, degenerate primers that amplify specifically the PKS genes from the fungal component in lichens are preferred.

The prevalent paradigm for polyketide biosynthesis is that the process is mediated by the iterative type I PKSs in fungi, while bacteria use the modular type I and the iterative type II PKSs. Besides these two, there are also the type III PKSs (also known as chalcone synthases) that are predominantly found in plants (see Chapter 1 Section 1.2.4 for an introduction to the various PKS types). Although new discoveries of iterative type I PKSs in bacteria (Shen 2003) as well as type III PKSs in fungi and bacteria (Funa et al. 1999; Seshime et al. 2005) have blurred the boundaries, the paradigm still holds for most cases so far. With a few exceptions (such as the 6-methylsalicylic acid synthases), most fungal iterative type I PKSs known to date are sufficiently different from their bacterial counterparts to allow the development of fungal-specific degenerate primers, which should be biased to the PKS genes originating from the mycobiont in lichens. Since targeting the type III PKSs might amplify the genes from the photobionts or bacteria, the study conducted here focused mainly on the fungal iterative type I PKSs. For the purposes of the following content, the term “PKS” from this point will refer solely to the fungal iterative type I PKSs unless otherwise specified.

The β-ketoacylsynthase (KS) domain is the most conserved domain among the minimal PKS domains, which also include an acyltransferase (AT), and an acyl-carrier protein (ACP) domain (Aparicio et al. 1996; Bingle et al. 1999; Donadio & Katz
The conserved regions in the KS domain are useful for designing degenerate primers that amplify fragments of new fungal PKS genes. KS domain primers (KS1 and KS2) based on the two conserved amino acid regions in the KS domains of type I fungal and bacterial PKSs were first used to detect mycotoxin-producing fungal strains (Keller 1995). The primers successfully amplified the KS region of fumonosin PKS gene *fum5* from cDNA of *Gibberella fujikuroi* (Proctor et al. 1999).

Bingle et al. (1999) later designed two pairs of KS primers (LC1/2c and LC3/5c) that targeted specifically the WA- (non-reducing) and MSAS-type (partially-reducing) fungal PKSs. In a subsequent study, another pair of KS primers (KS3/4c) selective for the HR-PKSs was also developed by the group (Nicholson et al. 2001). Together all three pairs of primers (LC1/2c, LC3/5c and KS3/4c) could distinguish between NR-, PR- and HR-PKSs (Nicholson et al. 2001) (see Chapter 1 Section 1.2.5 for classification of fungal PKSs). The development of these KS primers exploits the slight amino acid sequence differences at the conserved regions of KS domains among different classes of PKSs, which in turn are reflected in the corresponding DNA sequences. These primer sets remain one of the most powerful tools when the aim is to locate and clone a PKS gene responsible for the production of a particular polyketide compound. As the group (HR-, PR- or NR-PKSs) to which the PKS gene belongs can often be predicted from the structure (the extent of reduction) of the corresponding polyketide product, these KS primer sets could narrow down the search for the target PKS gene.

A slightly different approach was used in another study in which the aim was to detect a wide range of PKS genes using KS primers (FPKSKSU-2 and FPKSKSD-1) with very high degeneracy (128 and 512 respectively) (Sauer et al. 2002). After cloning of the amplified products, the identities of the PKS gene fragments were confirmed by colony-hybridization with mixed probes consisting of ten fungal KS domain fragments. Using this approach, 12 new KS fragments were obtained from 11 fungal endophytes isolated from wild cranberry (*Vaccinium macrocarpon*). The cloned KS fragments could be segregated into three clusters, distributed among the NR-PKSs. One cluster did not group with any known fungal PKS at that time. Other sets of KS primers were also developed to explore the PKS genes in other fungi,
such as a group of insect- and nematode-associated fungi (Lee et al. 2001) and the wood-decay fungus *Xylaria* sp. BCC1067 (Amnuaykanjanasin et al. 2005).

As the compounds unique to lichens are mainly coupled phenolics that consist of two to three aromatic rings, the effort in this chapter was focused on the detection of fungal PKS genes responsible for the production of aromatic compounds in lichens. Furthermore, compounds used to assist in lichen identification are also mainly aromatic compounds. Since both NR-PKSs and PR-PKSs produce aromatic polyketides in fungi (see Chapter 1 Section 1.2.4), these two classes of PKSs were targeted specifically. For the same reasons, Grube & Blaha (2003) used the LC1/2c and LC3/5c primers developed by Bingle et al. (1999), which target NR-PKSs and PR-PKSs respectively, to detect the potential genes for aromatic polyketides in lichens. No PR-PKS genes were detected with the LC3/5c primers but the LC1/2c primers amplified 15 NR-type KS domains from the genus *Lecanora* and another three from other genera. The limited number of KS sequences detected was partly because direct sequencing of some gel-purified PCR products yielded mixed sequences. This is due to multiple PKS gene homologs are often present in a single fungus and the degenerate primers could amplify more than one product of the same size, which appeared as a single band when electrophoresed on agarose gel.

A subsequent study by Schmitt et al. (2005) also used the LC1/2c primers, but found more fungal NR-PKS genes in the order Pertusariales that were not available at the inception of this project. To avoid mixed sequences, Schmitt et al. (2005) subcloned all the PCR products into a TA-cloning vector prior to sequencing. As a result, more KS domain sequences were obtained. But the diversity and number of lichen PKS genes detected in the study is still likely to be an underestimate, based on the number of PKS genes found in other fungi from whole genome sequencing projects (e.g. up to 25 PKS genes in some species) (Kroken et al. 2003; Varga et al. 2003). More PKS gene sequences have become available since the commonly used KS primers (LC1/2c and LC3/5c) were developed by Bingle et al. (1999). With the intention of detecting a wider range of PKS genes in lichens, new sets of KS-specific degenerate primers were developed in this study, with the recently available PKS genes taken into consideration.
3.1.2 Correlating the structure of polyketides and PKS genes with phylogenetic analysis

So far, only a limited number of fungal PKSs has been linked directly to the structure of the polyketide products (Cox 2007). The variations in PKS domain architecture and their inherent programming encoded in the protein are the qualities that account for the structural diversity of polyketides. The multi-domain polyketide synthases express selectivity for the starter unit, use and control the different accessory domains for sequential modification of the polyketide chain at specific site(s), and at the same time regulate the chain length and timely release of the final product (Cox 2007). Although the details of the control and programmed use of these functional domains in a PKS is still poorly understood, it was known that the presence/absence of the functional domains affected the overall structure of the polyketide products.

Assuming that radical recombination of a PKS gene that results in domain gain/loss/shuffling does not happen too regularly, it is logical to expect that the evolutionary relationships of a KS domain (as the most conserved domain present in all PKSs) would reflect the domain architecture of a particular PKS. Earlier cladistic analyses showed that the segregation of the KS domain sequences is consistent with the NR-, PR- and HR-PKS groupings, which correspond with the extent of reduction of the polyketides being produced and the functional domains present in the PKSs (Bingle et al. 1999; Nicholson et al. 2001).

The availability of a large number of PKS gene sequences from recently completed fungal whole genome sequencing (WGS) projects has allowed the extensive phylogenetic analysis of fungal PKSs across several fungal genomes (Kroken et al. 2003). Although the functions of most PKSs from WGS projects are uncharacterized, the full-length PKS gene sequences allow the assignment of functional domains present in the corresponding proteins. Phylogenetic analysis based on KS protein sequences revealed a significant correlation between the PKS genealogy and domain architecture (Figure 3.1). Additional to the same three major classes of PKSs observed in previous analyses, the NR-PKSs (fungal non-reducing PKSs) with no reductive domains and HR-PKSs (fungal reducing PKSs) with additional reductive domains were each further segregated into four subclades.
On the other hand, the PR-PKSs were nested within a large clade that comprised bacterial type I PKSs (bacterial PKSs, Figure 3.1). The PKSs that grouped into each subclade shared certain common traits:

1. the extent of reduction of the polyketide product based on the presence or absence of ketoreductase (KR), enoyl reductase (ER) and dehydratase (DH) domain in the PKS,

2. the occurrence of additional accessory domains such as, C-methyltransferase (CMeT), thioesterase (TE) and other domains related to nonribosomal peptide synthases (NRPSs), and

3. the chain length of the polyketide product.

For example, PKSs that produced long reduced polyketides (e.g. lovastatin and citrinin/compactin nonaketides) are grouped together in reducing PKS (HR-PKS) clade I, while PKSs that produced short reduced polyketides (e.g. lovastatin and citrinin/compactin diketides) are in reducing PKS clade II (Figure 3.1). On the other hand, the non-reducing clade I consisted solely of PKSs that produced long unreduced polyketides, which cyclized into polyaromatic compounds (e.g. sterigmatocystin, aflatoxin, etc.).

The analyses by Kroken et al. (2003) and related studies (Bingle et al. 1999; Nicholson et al. 2001; Varga et al. 2003) suggest that the KS domain phylogeny could potentially be used to infer the domain architecture of new PKSs and predict the type of polyketide products (non-reduced or reduced; methylated or non-methylated) they might produce. Thus, it was hypothesized that if a particular PKS homolog is frequently detected among lichens producing a specific group of biosynthetically-related compounds and that the predicted domain architecture for this PKS homolog also matches the structure/biosynthetic routes of these compounds, a correlation might be established between them. Testing this hypothesis on several phylogenetically-related lichen samples that produced compounds of similar biosynthetic origins would determine if such prediction was supported. The second objective of this chapter was therefore to test this hypothesis with lichens in which this information was not known.
Figure 3.1. Genealogy of type I PKSs, inferred by maximum parsimony analysis of the KS domain. Taken from Kroken et al. (2003). KS, ketosynthase; AT, acyl transferase; DH, dehydratase; ER, enoyl reductase; KR, ketoreductase; ME, C-methyltransferase; PP, phosphopanteinylation site (acyl carrier protein).
3.1.3 Research goals

Following the two objectives underlined above, the experiments in the first part of this chapter (Section 3.2) were focused on the identification of the lichen specimens collected and the detection of the polyketide compounds in the specimens. Section 3.3 and 3.4 focused on expanding the diversity of lichen PKS genes using newly designed degenerate primers and testing them on the collected lichen specimens. The last experimental section of this chapter (Section 3.5) concentrates on the phylogenetic analysis of the new lichen KS domains obtained and comparison with other known PKSs to infer possible relationships between the PKS genes and the lichen compounds detected earlier in Section 3.2. The use of the newly acquired KS domains in locating and cloning of new PKS genes from lichens will be demonstrated in Chapter 4.
3.2 Lichen identification and chemical analysis

Lichens were collected from a variety of habitats in Victoria, Australia. Most of the collections were aimed at the lichens from the family Parmeliaceae, while some specimens were collected at random. The Parmeliaceae is one of the most widespread families of lichens and is known for its metabolic complexity. Many of the compounds produced by lichens in the Parmeliaceae are also unique to lichens; some are biologically active, such as usnic acids, depsides and depsidones (See Chapter 1 section 1.3.1). The lichens were identified using a combination of morphological observation and chemical analysis. They were largely guided and assisted by S. H. Louwhoff (Royal Botanical Gardens, Victoria) and J. A. Elix (Australian National University) as experience and skill is required to identify most species accurately. J. A. Elix also assisted with the HPLC analysis and identification of lichen compounds by matching the chromatography peaks with an extensive HPLC-UV lichen compound library that he had built over decades of research. For molecular identification, the internal transcribed spacer regions 1 and 2 of the lichens were also amplified by PCR and compared with the DNA sequences in GenBank.

3.2.1 Materials and methods

3.2.1.1 Lichen collection and identification

Lichen specimens were collected from various parts of Victoria, Australia from three main locations, i.e. the Mount Waverley residential area (suburban, around Toiram Road), Wilsons Promontory National Park (coastal, around Tidal River), Wyperfeld National Park (mallee, around Casuarina campground) and Melbourne Central Business District (CBD) (Figure 3.2). Vouchers with details of collection sites were deposited in the Herbarium at the RBG, Victoria (MEL). Their identities were confirmed by observing their morphological characteristics at macroscopic and microscopic level and guided by keys described in Flora of Australia (Archer 1992; Elix 1994). Chemical spot tests as described by Orange et al. (2001) were also performed on the lichen cortex and/or medulla as required in the identification keys. Occasionally, further chemical analysis by thin layer chromatography (TLC) was required to resolve the identity of some lichens to the species level.
3.2.1.2 Thin layer chromatography (TLC)

About 200 mg of dried lichen thalli were broken into small pieces and extracted with 0.5 mL of acetone. The extracts was spotted on a TLC plate and resolved with standardized solvent systems as described elsewhere (Orange et al. 2001). The compounds were identified by comparing their retention factor ($R_f$), colour and UV fluorescence to authentic compounds or to extracts from lichens with known chemical constituents.

3.2.1.3 High performance liquid chromatography (HPLC)

HPLC was performed on lichen extracts when the compounds did not resolve adequately on the TLC or if they exhibited a more complex metabolic profile that required more detailed analysis. The lichen metabolites were identified by HPLC, by comparison with retention index (RI) values calculated from benzoic acid and solorinic acid controls as described previously by Feige et al. (1993) and Elix et al.
(2003). Essentially, the lichen fragments were brushed clean of obvious organic substrate material and extracted with warm methanol prior to analysis using a HP1050 Series System (Hewlett Packard, Australia) with a Hypersil 5 µ C18 column (250 mm by 4.6 mm) (Phenomenex, Australia) and a UV spectrometric detector operating at 254 nm with a flow rate of 1 ml/min. Two solvent systems were used: 1% aqueous orthophosphoric acid and methanol in the ratio 7:3 (A) and methanol (B). The run started with 100% solvent A and was raised to 58% B within 15 min, then to 100% solvent B within a further 15 min, followed by isocratic elution in 100% Solvent B for a further 10 min. The HPLC was coupled to a photodiode array detector for ultraviolet spectroscopic comparisons. By this means, the ultraviolet spectra observed for the various components eluting in the HPLC chromatogram were computer-matched against a library of ultraviolet spectra recorded for authentic compounds under identical conditions. For each substance to be identified, the correlation of ultraviolet spectra to the spectra of authentic compounds in the library had to be and was greater than 99.9%.

3.2.1.4 PCR amplification of internal transcribed spacer regions

The identities of 12 lichens were confirmed by molecular methods due to ambiguity in morphology- and chemistry-based identification. The internal transcribed spacer (ITS) regions 1 and 2 in the nuclear ribosomal repeat unit of the mycobiont were determined by direct sequencing of the PCR product (gel purified, Section 2.2.5) amplified with ITS1-F and ITS4 primers as described previously (Gardes & Bruns 1993; White et al. 1990). The lichen DNA used for PCR was extracted according to Section 2.2.1 and PCR conditions were according to Section 2.2.3. The resultant sequences were queried against the GenBank non-redundant nucleotide database.

3.2.2 Results

3.2.2.1 Lichen specimens and identification

A total of 27 lichen specimens was collected and identified (Figure 3.3). All of the specimens collected were ascomyceteous lichens belonging to the Pezizomycotina lineage (Lecanoraceae), with the majority from the Parmeliaceae and a few from the
Teloschistaceae, Cladoniaceae, Ramalinaceae and Physciaceae. Not all specimens were identified to the species level. This is due to the collected specimens 1) being too small, 2) lacking the reproductive parts (apothecia etc.), 3) being morphologically diverse, 4) being in a mixture of more than one lichen and mosses, or a combination of the above reasons.

**Figure 3.3.** The lichens collected from Mount Waverley (MW), Wilsons Promontory National Park (WP), Wyperfeld National Park (WF), and Melbourne central business district (CBD). Figure continues next page.
Ramalinaceae: *Ramalina* sp. (MW)

Physciaceae: *Physcia jackii* (MW)

Parmeliaceae: *Flavoparmelia soredians* (MW)

Parmeliaceae: *Flavoparmelia caperatulla* (MW)

Teloschistaceae: *Teloschistes* sp. (MW)

Parmeliaceae: *Xanthoparmelia semiviridis* (WF)

Figure 3.3 (continued).
Parmeliaceae: Parmotrema chinense (WP)

Parmeliaceae: Parmotrema reticulatum (WP)

Parmeliaceae: Parmelia cunninghamii (WP)

Teloschistaceae: Teloschistes chrysophthalmus (WP)

Figure 3.3 (continued).
Figure 3.3 (continued).
Figure 3.3 (continued).
3.2.2.2 Chemical analysis of lichens

The compounds detected in all lichen specimens collected were mainly polyketides (Table 3.1). They can be divided into four groups based on their biosynthetic origin (the putative precursors produced by PKSs). The orcinol, β-orcinol, and methylphloroacetophenone precursors are all monoaromatic units. The fourth group consisted of compounds from long and unreduced polyketide chains and the only member detected was parietin, which is an anthraquinone. Among the orcinol and β-orcinol derivatives, there were compounds formed by mono-, di- and tri-aryls of the monoaromatic units. The structural formulae of all the compounds detected are shown in Figure 3.4.

A few distinct chemotypes could be observed among the lichen specimens collected. For example, parietin was predominantly found in lichens from the family Telochistaceae (resulting in the orange appearance of the lichen thalli), while coupled phenolic compounds were not detected. Three lichens (Punctelia borreri, Hypogymnia subphysodes and Carnoparmelia sp.) produced only atranorin and lecanoric acid-related compounds. Protocetraric acid-related depsidones were detected along with usnic acid in Flavoparmelia spp., Xanthoparmelia semiviridis and Usnea oncodeiodes, but the common lichen depside, atranorin, was noticeably absent. The menegaziaic acid and stictic acid-related depsidones were found mostly in the genera Menegazzia and Parmotrema, but atranorin was only present in the genus Parmotrema and not in Menegazzia.
<table>
<thead>
<tr>
<th>Collection number</th>
<th>Genus</th>
<th>Species</th>
<th>Orcinol-derivatives</th>
<th>β-orcinol derivatives</th>
<th>Methylphloro - acetophenone derivatives</th>
<th>Polyaromatic/ long polyketides</th>
</tr>
</thead>
<tbody>
<tr>
<td>001*</td>
<td>Flavoparmelia</td>
<td>caperatulla</td>
<td></td>
<td>atranorin, protocetraric acid, conprotocetraric acid</td>
<td>usnic acid</td>
<td></td>
</tr>
<tr>
<td>002*</td>
<td>Thysanothecium</td>
<td>scutellatum</td>
<td>divaric acid, methyl divarate, <strong>divaricatic acid</strong>, sekikaic acid, stenosporic acid</td>
<td>usnic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>003</td>
<td>Xanthoria</td>
<td>parietina</td>
<td></td>
<td>protocetraric acid</td>
<td></td>
<td>parietin</td>
</tr>
<tr>
<td>004</td>
<td>Flavoparmelia</td>
<td>rutidota</td>
<td></td>
<td>protocetraric acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>005</td>
<td>Ramalina</td>
<td>sp.</td>
<td></td>
<td>salazinic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>006*</td>
<td>Physcia</td>
<td>jackii</td>
<td></td>
<td>methyl b-orsellinate, atranorin, haematomomic acid, methyl isohaematommate, methyl haematomatone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>007*</td>
<td>Flavoparmelia</td>
<td>soredians</td>
<td>consalazinic acid, salazinic acid, galbinic acid, protocetraric acid</td>
<td>usnic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>008</td>
<td>Flavoparmelia</td>
<td>caperatulla</td>
<td></td>
<td>protocetraric acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>010</td>
<td>Teloschites</td>
<td>sp.</td>
<td></td>
<td>protocetraric acid, usnic acid</td>
<td></td>
<td>parietin</td>
</tr>
<tr>
<td>011*</td>
<td>Xanthoparmelia</td>
<td>semiviridis</td>
<td>succinprotocetraric acid, fumarprotocetraric acid, conprotocetraric acid, consuccinprotocetraric acid, confumarprotocetraric acid, protocetraric acid, virensic acid</td>
<td>usnic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>012</td>
<td>Parmotrema</td>
<td>chinense</td>
<td>atranorin, consictic acid, acetic acid, cryptosictic acid, menegaziaic acid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>013</td>
<td>Flavoparmelia</td>
<td>caperatulla</td>
<td>protocetraric acid</td>
<td>usnic acid</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 3.1.* Compounds detected in the lichens collected by TLC and/or HPLC (*). Colour indicates the compound classes: monoaromatics, depsides, depsidones, dibenzofurans, anthraquinones. Major compounds detected in HPLC are in bold.
Table 3.1 (continued). Compounds detected in the lichens collected by TLC and/or HPLC (*).

<table>
<thead>
<tr>
<th>Collection number</th>
<th>Genus</th>
<th>Species</th>
<th>Orcinol-derivatives</th>
<th>β-orcinol derivatives</th>
<th>Methylphloro-acetophenone derivatives</th>
<th>Multiaromatic rings/ long polyketides</th>
</tr>
</thead>
<tbody>
<tr>
<td>014*</td>
<td>Parmotrema</td>
<td>reticulum</td>
<td></td>
<td>atranorin, chloroatranorin, constictic acid, peristictic acid, cryptostictic acid, stictic acid, menegazziaic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>015*</td>
<td>Punctelia</td>
<td>borreri</td>
<td>orsellinic acid, methyl orsellinate, lecanonic acid, orcinyl lecanorate, gyrophoric acid</td>
<td>atranorin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>016*</td>
<td>Parmelia</td>
<td>cunninghamii</td>
<td></td>
<td>atranorin, chloroatranorin, consalazinic acid, salazinic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>017</td>
<td>Teloschites</td>
<td>chrysophthalmus</td>
<td></td>
<td>consalazinic acid, salazinic acid</td>
<td></td>
<td>parietin</td>
</tr>
<tr>
<td>018</td>
<td>Menegazzia</td>
<td>caesiopruinosa</td>
<td></td>
<td>consalazinic acid, salazinic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>019</td>
<td>Menegazzia</td>
<td>subpertusa</td>
<td></td>
<td>consalazinic acid, salazinic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>020</td>
<td>Menegazzia</td>
<td>platytrema</td>
<td></td>
<td>consalazinic acid, salazinic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>021</td>
<td>Hypogymnia</td>
<td>subhysodes</td>
<td></td>
<td>atranorin, physodalic acid, physodic acid, oxyphysodic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>023</td>
<td>Flavoparmelia</td>
<td>caperatulia</td>
<td></td>
<td>protocetraric acid, usnic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>026</td>
<td>Canoparmelia</td>
<td>sp.</td>
<td>lecanoric acid</td>
<td>atranorin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>027*</td>
<td>Usnea</td>
<td>oncodeoides</td>
<td>consalazinic acid, salazinic acid, galbinic acid, protocetraric acid</td>
<td>usnic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>028</td>
<td>Usnea</td>
<td>oncodeoides</td>
<td></td>
<td>salazinic acid</td>
<td>usnic acid</td>
<td></td>
</tr>
<tr>
<td>029</td>
<td>Teloschistes</td>
<td>spinosus</td>
<td></td>
<td></td>
<td></td>
<td>parietin</td>
</tr>
<tr>
<td>030</td>
<td>Cladonia</td>
<td>sp.</td>
<td>merochlorophic acid, homosekikaic acid</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Orcinol derivatives
(including orsellinic acid derivatives and alkyl resorcylic acid derivatives)

**Monoaromatics**

- orsellinic acid (R=H)
- methyl orsellinate (R=CH₃)
- divaric acid (R = H)
- methyl divarate (R=CH₃)

**Depsides**

- divaricatic acid
- sekikaic acid
- stenosporic acid
- homosekikaic acid (R₁=H; R₂=CH₃)
- merochlorophaic acid (R₁= CH₃; R₂=H)
- lecanoric acid
- gyophoric acid
- orcinyl lecanorate

**Figure 3.4.** The chemical structures of compounds detected in the lichen specimens (Table 3.1).
**β-Orcinol derivatives**

**Monoaromatics**

- Haematommic acid (R=H)
- Methyl haematommate (R=CH$_3$)

**Depsides**

- Atranorin
- Chloroatranorin

**Depsidones**

- Protocetraric acid
- Conprotocetraric acid
- Succinprotocetraric acid
- Succinprotocetraric acid

**Figure 3.4 (continued).** The chemical structures of compounds detected in the lichen specimens (Table 3.1).
β-Orcinol derivatives

*Depsidones* (continued)

Figure 3.4 (continued). The chemical structures of compounds detected in the lichen specimens (Table 3.1).
3.2.2.3 Internal transcribed spacer (ITS) regions

The sizes of the ITS PCR products were 650-900 bp. The closest Blastn matches on the GenBank database are shown in Table 3.2. Sequences that share over 99% identity were morphologically and chemically the same species, while sequences that shared 96-98% identity were mostly of the same genus (due to the absence of the corresponding species in GenBank database), except for *Thysanothecium scutellatum* (002) (96%) and *Hypogymnia subphysodes* (021) (98%). Only one lichen specimen (*Physcia jackii*) had <96% matching to a sequence in GenBank (87%). No corresponding sequence for this species was in the database.

**Table 3.2.** Closest BLASTn matches of the internal transcribed spacer (ITS) 1, 5.8s ribosomal RNA gene, and ITS 2 of selected lichens. *The expected (E) value for all the closest Blast hit = 0.0

<table>
<thead>
<tr>
<th>No.</th>
<th>Species</th>
<th>Closest Blastn match*</th>
<th>GenBank accession</th>
<th>Identity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td><em>Flavoparmelia caperatulla</em></td>
<td><em>Flavoparmelia rutidota</em></td>
<td>DQ299906</td>
<td>501/509 (98%)</td>
</tr>
<tr>
<td>002</td>
<td><em>Thysanothecium scutellatum</em></td>
<td><em>Ramalea cochleata</em></td>
<td>AF453267</td>
<td>531/552 (96%)</td>
</tr>
<tr>
<td>003</td>
<td><em>Xanthoria parietina</em></td>
<td><em>Xanthoria parietina</em></td>
<td>DQ888716</td>
<td>562/564 (99%)</td>
</tr>
<tr>
<td>006</td>
<td><em>Physcia jackii</em></td>
<td><em>Physcia tribacia</em></td>
<td>AY498691</td>
<td>672/772 (87%)</td>
</tr>
<tr>
<td>007</td>
<td><em>Flavoparmelia soredians</em></td>
<td><em>Flavoparmelia soredians</em></td>
<td>AY586562</td>
<td>500/503 (99%)</td>
</tr>
<tr>
<td>008</td>
<td><em>Flavoparmelia caperatulla</em></td>
<td><em>Flavoparmelia rutidota</em></td>
<td>DQ299906</td>
<td>501/509 (98%)</td>
</tr>
<tr>
<td>011</td>
<td><em>Xanthoparmelia semiviridis</em></td>
<td><em>Xanthoparmelia semiviridis</em></td>
<td>AF451746</td>
<td>511/514 (99%)</td>
</tr>
<tr>
<td>015</td>
<td><em>Punctelia borrisi</em></td>
<td><em>Punctelia borrisi</em></td>
<td>DQ394373</td>
<td>533/541 (98%)</td>
</tr>
<tr>
<td>016</td>
<td><em>Parmelia cunninghamii</em></td>
<td><em>Parmelia pseudotenuirima</em></td>
<td>AY037002</td>
<td>480/492 (97%)</td>
</tr>
<tr>
<td>021</td>
<td><em>Hypogymnia subphysodes</em></td>
<td><em>Hypotrachyna revoluta</em></td>
<td>DQ279523</td>
<td>499/505 (98%)</td>
</tr>
<tr>
<td>027</td>
<td><em>Usnea oncodeoides</em></td>
<td><em>Usnea rubrotincta</em></td>
<td>DQ232664</td>
<td>533/552 (96%)</td>
</tr>
<tr>
<td>028</td>
<td><em>Usnea oncodeoides</em></td>
<td><em>Usnea rubrotincta</em></td>
<td>DQ232664</td>
<td>533/552 (96%)</td>
</tr>
</tbody>
</table>
3.2.3 Discussion

Except for Cladonia spp. (030), the compounds detected in the individual lichen specimens were mostly as expected and corresponded with the chemistry described in those identification keys (Archer 1992; Elix 1994) and phytochemical register of lichens (Walker & Lintott 1997). Based on the metabolites detected (Table 3.1) and re-examination of the specimens according to Archer (1992), the Cladonia spp. (030) is possibly a mixture of C. ramulosa and C. merochlorophaea grown along with mosses on the same substrate (dead wood). Minor compounds related to the major metabolites were also detected in these specimens analysed by HPLC (Table 3.1). The lichen specimens that produced coupled phenolic compounds were chosen as the main subjects for experiments in the following sections.

BLASTn searches of the ITS region sequences were mostly consistent with the identification, except for Thysanothecium scutellatum (002) and Hypogymnia subphysodes (021). Further examination revealed that 021 was a mixed specimen that consisted of H. subphysodes and Hypotrachyna revoluta (the secondary metabolites detected in 021 in Table 3.1 corresponded to H. subphysodes). Since the ITS PCR products from 021 did not show mixed sequences, it was concluded that thallus fragments of mostly H. revoluta had been included in DNA extraction (extracts consisted mostly of DNA from H. revoluta) while thallus fragments of H. subphysodes was used in chemical analysis. Based on Walker & Lintott (1997), H. subphysodes produces atranorin and gyrophoric acid as the major secondary metabolites. Other lichens with ITS region sequences that share less than 99% identity with the closest BLASTn match are due to the corresponding sequence of the particular species not being available in the database. This was also the case for T. scutellatum (002). No significant sequence mixture was observed in the electropherograms of the amplified ITS regions listed in Table 3.2; this suggests that fungal DNA from other sources was absent or negligible in the DNA extracts (sequence mixture would be observed if fungal DNA other than the mycobiont was present in the DNA extracts, as ITS1F/4 primers bind non-selectively to all fungi). PCR amplification with the primers ITS-1F and ITS-4 failed in some lichen DNA extracts (e.g. 023 and 030); this is probably due to the impurities that were not totally removed during the purification.
3.3 NR- and PR-type KS domain-specific degenerate primers

To survey the diversity of the polyketide synthase genes in lichens, two pairs of new degenerate primers analogous to LC1/2c and LC3/5c (Bingle et al. 1999) were designed, with the intention of covering a larger range of PKSs than the original primers. Two new primer pairs (NRKS and PRKS) were developed to target the KS domains of NR-PKSs and PR-PKSs respectively, as these produce aromatic polyketides. Additional KS domain sequences to those employed by Bingle et al. (1999) were considered when designing the new degenerate primers.

The LC1/2c primers that were specific to the KS domain of WA-type PKS genes were designed based on the only two NR-PKS sequences that were available at that time, i.e. *Aspergillus nidulans* WA (NR-PKS clade I) and *Colletotrichum lagenarium* PKS1 (NR-PKS clade II). Besides *A. nidulans* WA and *C. lagenarium* PKS1, other well-characterised NR-PKSs are the *A. nidulans* STCA and *A. parasiticus* PKSL2 that produce norsolorinic acid (precursor for both sterigmatocystin and aflatoxin). Preliminary multiple alignment analysis suggests that the LC1/2c is probably not compatible with STCA-type PKS genes. To incorporate all NR-PKSs including the STCA-type PKS genes, particular attention was paid to regions conserved among them when the new pair of degenerate primers (NRKS-F/R) was designed.

The target group for the second primer pair, PRKS, were the PR-PKSs. Only one fungal PKS homolog has been identified in this group so far – the 6-methylsalicylic acid synthase (6MSAS), which has been found in fungi such as *Penicillium patulum*, *A. terreus* (ATX), *A. parasiticus* (PKSL2), and *Glarea lozoyensis* (PKS2) (Beck et al. 1990; Feng & Leonard 1998; Fujii et al. 1996; Lu et al. 2005). All the PKSs produce the same compound – 6-methylsalicylic acid. The degenerate primers, LC3/5c (Bingle et al. 1999), were based on the 6MSASs from *A. terreus* (ATX) and *A. parasiticus* (PKSL2). Although 6MSASs are the only known members of PR-PKSs in fungi, they are highly similar to the bacterial orsellinic acid synthases (OAS) found in actinomycetes, i.e. *Streptomyces viridochromogenes* AviM (Gaisser et al. 1997) and *Micromonospora echinospora* CalO5 (Ahlert et al. 2002), which were both unique bacterial iterative Type I PKSs. Orsellinic acid is the intermediate to many compounds in lichens such as lecanoric acid and gyrophoric acid (see Figure 3.4 and Chapter 1 Section 1.2.5.2.1). On the assumption that the lichen OASs could be
similar to bacterial OASs and fungal MSAS, KS sequences of the two known bacterial OASs were also considered when designing the primers.

### 3.3.1 Materials and methods

#### 3.3.1.1 Design of degenerate primers

Both NRKS and PRKS degenerate primers were designed using the Consensus-Degenerate Hybrid Oligonucleotide Primers (CODEHOP) program (http://bioinformatics.weizmann.ac.il/blocks/codehop.html) developed by (Rose et al. 1998). Multiple sequence alignment of KS domain protein sequences of NR-PKSs and PR-PKSs respectively, were converted to Block format and provided as input into the CODEHOP program using the default parameters. The output was a list of forward and reverse primers corresponding to the various conserved blocks. Primers predicted to bind to regions conserved within the group (NR-PKSs or PR-PKSs), but not PKSs of other classes, were chosen.

#### 3.3.1.2 PCR amplification of KS domain

The lichen DNA templates used for the PCR were from Section 2.2.1. PCR was performed using NRKS-F/R and PRKS-F/R primers (Section 2.2.3 and Table 2.1) with 5-10 ng of template DNA; 5 µL of the PCR products were examined by gel electrophoresis. Subsequently, these PCR products (20 µL) were gel purified (Section 2.2.5) for sequencing (Section 2.2.6). LC1/2c and LC3/5c primers were also tested on selected lichen specimens, but most of those PCR products were not sequenced.

#### 3.3.1.3 KS domain sequence analysis

The GenBank non-redundant (nr) protein database was queried with the edited sequences using BlastX in NCBI (http://www.ncbi.nlm.nih.gov/BLAST/) to search for similar protein sequences. The Conserved Domain Database (CDD – 23523 PSSMs) was also queried with the translated sequences using the Conserved Domain Search (cds) (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) tool in NCBI to check for the presence of conserved domains and their function.
3.3.2 Results

3.3.2.1 NRKS degenerate primers

From the multiple sequence alignment (Figure 3.5), there were two amino acids in the STCA and PKSL1 that were different from the “DPRFFNM” motif used to design the LC1 forward primers (the corresponding regions in STCA and PKSL1 were “DPRFFSI” and “DPRFFGI” respectively). Since the last two amino acids of the regions correspond to the 3’ end of the forward primer, it is very important that they match to allow amplification. To incorporate the STCA-type PKSs in the primer design, a different conserved region was chosen for the NRKS-F forward primer – “QR(M/L)AL(T/L/M) (T/S)(T/A)YEA”. As the CODEHOP program yielded primers consisting of a 3´ degenerate core and a 5´ non-degenerate/consensus clamp, the “(T/A)YEA” is the core degenerate region that should prime the NR-PKS genes included in the CODEHOP input (Figure 3.5). Although this stretch of sequence is also similar to the LNKS and other HR-PKSs, the bias in the reverse primer sequence to NR-PKS was likely to favour the amplification of NR-PKSs. For the NRKS-R reverse primer, the same region, “EMHGTGT”, corresponding to the LC2c primer was adopted (Figure 3.5), and contrasted with the “EAHATST” in PR-PKSs and “EAHGTHT” in HR-PKSs.

3.3.2.2 PRKS degenerate primers

For PRKS primers specific to the PR-PKSs, the Block alignment of fungal MSASs and bacterial OASs was used as the input. The multiple sequence alignment showed that the KS domains of OASs were surprisingly similar to the fungal MSASs (Figure 3.6). Primers corresponding to the two conserved regions used for the design of LC3/5c primers (“AEQMDPQ” and “EAHATST”) were chosen, since both regions are conserved within the fungal MSASs and bacterial OASs (Figure 3.6).
Figure 3.5. Multiple sequence alignment of NR-PKSs (in bold) and the conserved regions corresponding to NRKS-F/R and LC1/2c primers. Col, *Colletotrichum lagenarium*; Pep, *Penicillium patulum*; Asn, *Aspergillus nidulans*; Asf, *A. fumigatus*; Asp, *A. parasiticus*; Ast, *A. terreus*. 6MSAS (a PR-PKS) and LNKS (a HR-PKS) are outliers for comparison.
**Figure 3.6.** Multiple sequence alignment of PR-PKSs (in bold) and the conserved regions corresponding to PRKS-F/R and LC3/5c primers. Stv, *Streptomyces viridochromogenes*; Mie, *Micromonospora echinospora*. WA (a NR-PKS) and LNKS (a HR-PKS) are outgroups for comparison.
3.3.2.3 PCR results and KS domain sequence analyses

The product sizes obtained by NRKS and PRKS primers were 650-900 bp (Figure 3.7). PCR amplification of the KS domain using the primers was successful for most lichen DNA samples except a few (Table 3.3). From the 16 lichen specimens, a total of 13 KS domain sequences was obtained (nine NR-, three HR- and one PR-type KS domains) using the NRKS and PRKS degenerate primers, along with some using the LC-series primers of Bingle et al. (1999). Several PCR products produced mixed sequences when sequenced directly (Table 3.3). All 19 sequences showed high similarity (>70%) to other fungal KS amino acid sequences. Based on the PKS domain architecture of the closest BLASTx matches, the lichen KS domain sequences were assigned as NRKS, PRKS or HRKS (Table 3.3). Not all of the 28 lichens collected were included in the KS domain PCR as some had similar chemistry or compounds not of interest in this study.

For the NR-type KS domains, a total of 9 KS domain sequences was obtained with the NRKS primers (Table 3.3). X. parietina failed to produce any products using either LC1/2c or NRKS primers, while Parmotrema chinense did not produce any products with NRKS primers as well. Only one KS sequence was obtained from LC1/2c primers – NRKS1 of T. scutellatum. A total of 6 KS domain sequences was obtained with the PRKS primers (Table 3.3), but only one sequence (PcuPRKS1) matched closely with known PR-PKSs (MSAS). All KS domain PCR products amplified by the LC3/5c primers yielded mixed sequences or did not yield any readable sequences.
Figure 3.7. Gel electrophoreses of the PCR products amplified using A) NRKS-F/R primers and B) PRKS-F/R primers. The numbering of the wells corresponds to the collection number for the lichens specimens in Figure 3.3 and Table 3.3. Lane 1= *F. caperatulla*, 2 = *T. scutellatum*, 6 = *P. jackii*, 7 = *F. soredians*, 8 = *F. caperatulla*, 11 = *X. semiviridis*, 12 = *P. chinense*, 14 = *P. reticulatum*, 15 = *P. borreri*, 16 = *P. cunninghamii*, 21 = *H. revoluta*, 23 = *F. caperatulla*, 27=28 = *U. oncodeoides*, 30 = *Cladonia* sp..
Table 3.3. PCR products obtained from the lichen DNA using LC1/2c, LC3/5c, NRKS-F/R and PRKS-F/R primers. PCR products that yielded clean sequences are shaded; those that yielded mixed sequences or failed the sequencing are unshaded. (-) indicates no PCR products were obtained; blank indicates no PCR reaction was carried out.

<table>
<thead>
<tr>
<th>no.</th>
<th>Species</th>
<th>LC1/2c</th>
<th>LC3/5c</th>
<th>NRKS</th>
<th>PRKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>Flavoparmelia caperatulla (Fca)</td>
<td>820bp</td>
<td>800bp</td>
<td>680bp (NRKS1)</td>
<td>700bp</td>
</tr>
<tr>
<td>002</td>
<td>Thysanothecium scullatum (Tsc)</td>
<td>700bp (NRKS1)</td>
<td>550+800bp</td>
<td>680bp (NRKS2)</td>
<td>700+800bp</td>
</tr>
<tr>
<td>003</td>
<td>Xanthoria parietina (Xpa)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>006</td>
<td>Physcia jackii (Pja)</td>
<td>750bp</td>
<td>-</td>
<td>680+800bp</td>
<td>650bp (NRKS1)</td>
</tr>
<tr>
<td>007</td>
<td>Flavoparmelia soredians (Fso)</td>
<td>700+820bp</td>
<td>700+800bp</td>
<td>750bp (NRKS1)</td>
<td>700bp (HRKS1)</td>
</tr>
<tr>
<td>008</td>
<td>Flavoparmelia caperatulla (Fca)</td>
<td>820bp</td>
<td>700+800+1000bp</td>
<td>680bp [same as 001]</td>
<td>700bp</td>
</tr>
<tr>
<td>011</td>
<td>Xanthoparmelia semiviridis (Xse)</td>
<td>700+820bp</td>
<td>-</td>
<td>680bp (NRKS1)</td>
<td>700bp</td>
</tr>
<tr>
<td>012</td>
<td>Parmotrema chinense (Pch)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>014</td>
<td>Parmotrema reticulatum (Pre)</td>
<td>680+700bp</td>
<td>700bp (HRKS1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>015</td>
<td>Punctelia borreri (Pbo)</td>
<td>680bp (NRKS1)</td>
<td>700bp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>016</td>
<td>Parmelia cunninghamii (Pcu)</td>
<td>680+750bp</td>
<td>700bp (PRKS1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>021</td>
<td>Hypotrachyna revoluta (Hre)</td>
<td>680+700bp</td>
<td>400bp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>023</td>
<td>Flavoparmelia caperatulla (Fca)</td>
<td>680bp [same as 001]</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>027</td>
<td>Usnea oncodeoides (Uon)</td>
<td>-</td>
<td>700bp (HRKS1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>028</td>
<td>Usnea oncodeoides (Uon)</td>
<td>680bp (NRKS1)</td>
<td>700bp [same as 027]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>030</td>
<td>Cladonia spp. (Csp)</td>
<td>680bp (NRKS1)</td>
<td>700bp</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.3.3 Discussion

The variation in product size observed in Figure 3.7 and Table 3.3 is probably due to the presence of an intron commonly found upstream of the KS domain catalytic core. Some specimens were particularly difficult for obtaining PCR products (including the PCR with ITS1-F/4 primers), e.g. X. parietina (030), which might be due to the presence of inhibitors in the lichen DNA extracts. Only one KS domain PCR product was successfully sequenced using the LC-series primers (TscNRKS1), while most produced mixed sequences or did not yield any sequences. This is probably due to the lack of a non-degenerate “ adaptor” at the 5’ end of these
primers, which led to priming at multiple sites during the sequencing cycle reaction. Bingle et al. (1999) added a short stretch of “adaptor” sequence to the LC-series primers to facilitate PCR-cloning and sequencing, but the adaptor sequence was not attached to the LC-series primers used in this study.

A non-degenerate “clamp” was present in primers designed using the CODEHOP program, which should have improved the specific binding of these primers at the 5’ and 3’ end of the PCR products during the sequencing cycle. However, mixed sequences were still obtained for some NRKS-F/R and PRKS-F/R PCR products. This could be due to the presence of multiple products of the same size or cross-contamination of PCR products during gel excision of bands for column purification (for samples that yielded multiple bands). Further sequencing of the mixed PCR products by PCR-cloning was not carried out.

3.3.3.1 NRKS amplified KS domains

By comparing the number of bands on the agarose gel, the LC1/2c primers seems to produce more diverse PCR products, especially for *F. soredians* (007) and *X. semiviridis* (011) (Table 3.3). However, an extra band was obtained for *P. jackii* with NRKS primers compared to the LC1/2c primers. The number of bands of different sizes obtained do not reflect the number of PCR products, as often different KS domains yielded the same-size PCR products. This is particularly true for single band products that gave mixed sequences. Nevertheless, the results showed that the two pairs of primers were inclined to pick up different NR-type KS domains. For example, the TscNRKS2 amplified using the NRKS-F/R primers is significantly different from TscNRKS1 (59.4% identity), which was amplified using the LC1/2c primers. In retrospect, the overlooked conserved region “WDLLYKGLDV” upstream of the LC1 region (Fig 3.2) is probably a better target for the NR-specific degenerate KS forward primer, as it was highly conserved within the NR-PKS group.

3.3.3.2 PRKS amplified KS domain

From the PRKS primers, no PR-type KS domains were obtained except for a 6MSAS homolog in *P. cunninghamii*. Although the 6MSAS is reported as being ubiquitous among ascomycete fungi (Beck et al. 1990; Fujii et al. 1996; Kroken et al.
2003; Lu et al. 2005; Nicholson et al. 2001), no 6-methylsalicylic acid or its derivatives is known in lichens. A previous attempt by Grube & Blaha (2003) also failed to obtain any PCR products with the LC3/5c primers specific to PR-type KS domains. Hence, it was a surprise that a KS domain with such a high amino acid identity to those of Aspergillus parasiticus PKSL2 (74%) and Penicillium patulum MSAS (71%) was amplified from the lichen P. cunninghamii. Direct sequencing of the ITS1F/4 PCR product of P. cunninghamii indicated that the KS domain was unlikely to be from other fungal DNA from the environmental source (but it could be of bacterial origin). Hence, this might be an indication that the P. cunninghamii PRKS1 is a cryptic MSAS gene, as was the case for the ATX gene in Aspergillus terreus (Fujii et al. 1996). [Note: MSAS-type or PR-PKS genes were recently reported in lichens of the genus Pertusaria (Schmitt et al. 2008)]

Unexpectedly, the PRKS primers also obtained several KS domain fragments from F. caperatulla, P. reticulatum and U. oncodeiodes that showed high similarity to HR-PKSs, such as LNKS and LDKS (Table 3.3). This is the first report of the presence of HR-type KS domains from lichens. Some of the HR-PKSs produce pharmaceutically important compounds such as the cholesterol-lowering statins (e.g. lovastatin, compactin/ mevastatin and squalestatin), and therefore the lichen HR-PKSs should be investigated in future.

Another unexpected result from the PRKS primers was the discovery of a NR-type KS domain in P. jackii that showed high similarity to a subgroup of NR-PKSs with an additional C-methyltransferase (CMeT) domain – the clade III NR-PKSs (Kroken et al. 2001). The PjaNRKS1 showed 62% identity to BfPKS18 and 55% identity to BfPKS16. The only NR-PKS with a CMeT domain that has been characterized is from Monascus purpureus and is involved in citrinin biosynthesis (Shimizu et al. 2005). The CMeT domain is predicted to be responsible for the methylation of the polyketide backbone of citrinin. [Note: More recently, another clade III NR-PKS, AsPKS1 (methylorcinaldehyde synthase), with a CMeT domain has been discovered in Acremonium strictum (Bailey et al. 2007)]. For P. jackii, the dominating secondary metabolite is atranorin, which is made up of a methyl β-orsellinate and haematomatic acid linked with an ester bond (see Figure 3.4 for chemical formula). Noticeably, other β-orsellinic acid derivatives, i.e. methyl β-orsellinate, haematomomic
acid, methyl isoahaematommate, and methyl haematomate, were also detected in \textit{P. jackii} by HPLC. Since a methylation at the aromatic ring via a similar mechanism in citrinin biosynthesis is required for the formation of $\beta$-orsellinic acid (see Chapter 1 Figure 1.10), the PjaNRKS1 that was predicted to possess a CMeT domain may be part of a PKS gene involved in $\beta$-orsellinic acid synthesis.

In lichens, a large group of depsides and depsidones ($\beta$-orcinol depsides/depsidones) are derived from $\beta$-orsellinic acid (see Chapter 1 Section 1.2.5). If the clade III NR-PKSs are indeed involved in biosynthesis of $\beta$-orsellinic acid, the clade III NR-type KS domains should occur more frequently than the present result suggests, since 12 out of the 14 lichen species tested with NRKS and PRKS primers produce $\beta$-orsellinic acid derived compounds (Table 3.1). Surprisingly, lichen KS phylogenetic studies by Grube & Blaha (2003) and Schmitt et al. (2005) did not reveal any clade III NR-type KS domain. Both studies also utilized the LC1/2c to amplify the KS domains. Therefore, it is likely that both the LC1/2c and NRKS primers that were supposed to amplify NR-type KS domains were sub-optimal for clade III NR-type KS domains. A closer examination of the sequence alignment that included the clade III NR-type KS domains (PKSCT and BfPKS18) showed that this was the case (see Section 3.4). On the other hand, the ability of PRKS primers to amplify the clade III NR-type KS domain from \textit{P. jackii} is probably a coincidence. Since a touch-down PCR approach with high to low annealing temperature was adopted, the amplification will tolerate a certain degree of mismatch between the primers and the DNA template. The detection of a clade III NR-type KS domain in \textit{P. jackii} may be ascribed to the lack of PR- and HR-PKS genes in the lichen. Thus, it was hypothesized that a new pair of degenerate primers designed to target specifically the clade III NR-type KS domains might detect the presence of clade III NR-PKSs in other lichens.
3.4 Clade III NR-type KS domain-specific degenerate primers

Clade III NR-PKSs is the only group of non-reducing PKSs that possess a CMeT domain. Since the formation of β-orsellinic acid inevitably needs a C-methyltransferase for the methylation at the aromatic ring, it is logical to hypothesize that a clade III NR-PKS is involved in the biosynthesis of β-orsellinic acid – a precursor to many compounds in lichens, such as depsides and depsidones (see Section 1.2.5.2.1, Figure 1.10). The detection of a clade III NR-type KS domain in *P. jackii*, which produces several β-orsellinic acid derivatives, supports this possibility (see Section 3.3.3.2). The C-methylation of β-orsellinic acid, a monoaromatic polyketide (tetraketide), is similar to another simple aromatic polyketide precursor – methylphloroacetophenone, which is the precursor to some unique lichen metabolites such as usnic acid, placodiolic acid and lepromin (see Section 1.2.5.2.1, Figure 1.10). Based on the alternative cyclization hypothesis proposed by Birch & Donovan (1953) and later Mosbach (1973), both β-orsellinic acid and methylphloroacetophenone could be derived from a C-4 methylated tetraketide chain by internal aldol condensation and Claisen condensation respectively. Since cyclisation by Claisen condensation involving the CYC domain is evident in other NR-PKSs and also present in some of the clade III NR-PKSs (e.g. *C. heterostrophus* PKS23), clade III NR-PKS could also be involved in the biosynthesis of usnic acid and other methylphloroacetophenone derivatives (e.g. placodiolic acid, leprolomin, contortin etc.)

Many of the lichen specimens collected in this study contained β-orsellinic acid-derived depsides/ depsidones and the methylphloroacetophenone-derived usnic acid (Section 3.2.2.2, Table 3.1). To investigate the presence of clade III NR-PKSs in lichens and their possible correlation with the compounds above, a third pair of degenerate primers, NR3KS-F/R, was designed and tested on these lichens.
3.4.1 Materials and methods

3.4.1.1 Design of degenerate primers

The method used to design the NR3KS-F/R degenerate primers was essentially same as described in Section 3.3.1.1. The KS domain protein sequences of *M. purpureus* PKSCT (the only characterized clade III NR-PKS at that time) and other uncharacterized clade III NR-PKSs described in Krogen et al. (2003) were included into the Block alignment as the input for the CODEHOP program. The output of the CODEHOP program was examined closely by comparing KS domains of other PKS classes to choose the most selective primers to clade III NR-type KS domains.

3.4.1.2 PCR amplification of KS domain and sequence analysis

PCR protocols were as described in Chapter 2 Section 3.2.1.4. The newly designed NR3KS-F/R primers (Figure 3.8, Table 2.1) were used in the PCR with 5-10 ng of template DNA and 5 µL of the PCR products were examined by gel electrophoresis (section 2.2.4). Successful amplifications were column- or gel-purified and sequenced as described previously (Section 2.2.5 & 2.2.6). Selected PCR products that yielded mixed sequences were cloned using the method described in Section 2.2.9 and the positive clones were sequenced. The resultant sequences were analysed as described previously (Section 3.3.1.3).

3.4.2 Results

3.4.2.1 NR3KS degenerate primers

The “FDHKFF” region upstream of the KS domain, which were conserved within the clade III NR-PKSs but not other PKSs, was identified (Figure 3.8). A primer from the CODEHOP output list, which corresponds to this region, was chosen as the forward primer (NR3KS-F). The same region that corresponds to the NRKS-R and PRKS-R primers was chosen for the NR3KS-R primer. Although the conserved sequence “EAHGTGT” was identical to LNKS and other HR-PKSs, the forward primer (NR3KS-F) was deliberately biased to clade III NR-type KS domains (Figure 3.8).
**Figure 3.8.** Multiple sequence alignment of the clade III NR-type KS domains and the corresponding sites of NRKS-F/R primers and other KS primers (NRKS-F/R, PRKS-F/R, LC1). Representatives of the HR-, PR- and non-clade III NR-PKSs are shown as outgroup to highlight the differences in the sequences of the KS degenerate primers.
NR3KS-F (degen=64)
5´- C TAC GAC GCC TTC GAC cay mrn tty tt -3´
Y D A F D H K F F

NR3KS-R (degen=16)
3´- cty cgn gtr ccG TGG CCG TGG GG -5´
E A H G T G T P

Figure 3.8 (continued). The sequences of NR3KS-F/R primers and the corresponding amino acids.

3.4.2.2 PCR results and KS domain sequence analyses

A total of 14 PCR products was obtained but only three were successfully sequenced (Table 3.4). The rest either failed the sequencing or showed mixed sequences in the electrophoregrams. Two of the mixed PCR products (from X. semiviridis and P. chinense) were cloned and sequenced, which resulted in four additional KS domain sequences (two from each lichen). For low quality mixed sequences (as appeared in the electrophoregrams), the raw sequence were used as the direct input for a BLASTx search to check for any similarity to KS domains. All the KS sequences (including the mixed sequences) showed the greatest similarity to clade III type PKSs with a CMeT domain in BLAST searches. The closest homologs among the PKSs analysed in Kroken et al. (2003) are included in Table 3.4 for comparison, and all are clade III NR-PKSs (see Figure 3.1).
Table 3.4. PCR products obtained using the NR3KS-F/R primers and the closest homolog among the PKSs analysed in Kroken et al. (2003).

<table>
<thead>
<tr>
<th>No.</th>
<th>Species</th>
<th>PCR product size</th>
<th>Designation</th>
<th>Closest homolog among the PKSs analysed in Kroken et al. (2003)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td><em>F. caperatulla</em></td>
<td>750 bp</td>
<td>NR3KS1</td>
<td><em>B. fuckeliana</em> BiPKS17</td>
</tr>
<tr>
<td>003</td>
<td><em>X. parietina</em></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>006</td>
<td><em>P. jackii</em></td>
<td>700 bp</td>
<td>(same as PRKS PC product)</td>
<td><em>B. fuckeliana</em> BiPKS18</td>
</tr>
<tr>
<td>007</td>
<td><em>F. soredians</em></td>
<td>1100 bp (750 bp (faint))</td>
<td>mixed sequence</td>
<td><em>B. fuckeliana</em> BiPKS2</td>
</tr>
<tr>
<td>008</td>
<td><em>F. caperatulla</em></td>
<td>750 bp</td>
<td>Same as 001</td>
<td></td>
</tr>
<tr>
<td>011</td>
<td><em>X. semiviridis</em></td>
<td>(750 bp (cloned)) NR3KS1</td>
<td>NR3KS2</td>
<td><em>B. fuckeliana</em> BiPKS19</td>
</tr>
<tr>
<td>012</td>
<td><em>P. chinense</em></td>
<td>(750 bp (cloned)) NR3KS1</td>
<td>NR3KS2</td>
<td><em>B. fuckeliana</em> BiPKS16</td>
</tr>
<tr>
<td>014</td>
<td><em>P. reticulatum</em></td>
<td>750 bp</td>
<td>NR3KS1</td>
<td><em>B. fuckeliana</em> BiPKS16</td>
</tr>
<tr>
<td>015</td>
<td><em>P. borreri</em></td>
<td>1100 bp (750 bp (faint) 400 bp (faint))</td>
<td>Not KS domain N/A N/A</td>
<td>aspartyl-tRNA synthetase</td>
</tr>
<tr>
<td>016</td>
<td><em>P. cunninghamii</em></td>
<td>750 bp</td>
<td>mixed sequence</td>
<td><em>B. fuckeliana</em> BiPKS18</td>
</tr>
<tr>
<td>021</td>
<td><em>H. revoluta</em></td>
<td>1100 bp (750 bp)</td>
<td>mixed sequence</td>
<td><em>B. fuckeliana</em> BiPKS17</td>
</tr>
<tr>
<td>023</td>
<td><em>F. caperatulla</em></td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>027</td>
<td><em>U. oncodeoides</em></td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>028</td>
<td><em>U. oncodeoides</em></td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>030</td>
<td><em>Cladonia</em> (mixed)</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

N/A – Not available (not successfully sequenced)

*For good quality sequences, one of the longest ORFs of the six-frame translated sequences was queried against the GenBank database using BLASTp search. For low quality mixed sequences (as appeared in the electrophoregrams), the raw sequence were used as the direct input for a BLASTx search to check for any similarity to KS domains. All sequences showed >50% amino acid identity to clade III NR-PKSs, except for PCR products from *F. soredians* (007, which showed only 30% identity to *B. fuckeliana* BiPKS2) and *P. borreri* (015, not KS domain).

3.4.3 Discussion

The results showed that the clade III NR-type PKSs indeed occurred relatively frequently, as predicted. This is the first report of the widespread occurrence of clade III NR-PKSs in lichens. The clade III NR-type KS domains have not been reported from previous phylogenetic studies on lichen PKSs (Grube & Blaha 2003; Opanowicz et al. 2006; Schmitt et al. 2005), including the most recent one (Schmitt et al. 2008). BLASTx searches using the raw mixed sequences amplified with NR3KS-F/R primers showed that most sequences are closest to clade III NR-PKS, except for a *P. borreri* PCR product (1.1 kbp). This suggests that these mixed PCR products might consist of multiple clade III NR-type KS domains, as demonstrated for the NR3KS-F/R PCR products from *X. semiviridis* and *P. chinense* (PCR cloning and sequencing showed
that two different clade III NR-type KS domains were present in their NR3KS-F/R PCR products). All of the lichens in which the clade III NR-type KS domains were amplified, produced $\beta$-orsellinic acid-derived compounds, and/or usnic acids. This is in agreement with the hypothesis that clade III NR-type KS domains are involved in biosynthesis of $\beta$-orsellinic acid and methylphloroacetophenone as precursors for coupled phenolic compounds in lichens.

Overall, the results demonstrated that the newly developed NR3KS-F/R primers could specifically amplify the clade III NR-type KS domains. The following section (3.5) analysed the phylogenetic relationship between the 19 lichen KS domain sequences amplified with the three newly designed primers (NRKS-F/R, PRKS-F/R and NR3KS-F/R) and other fungal KS domains.
3.5 Phylogenetic analysis of lichen KS domains

To infer the function of the lichen PKS genes based on the KS domain sequences obtained from the previous sections (3.3 and 3.4), a KS domain phylogenetic tree based on the study of Kroken et al. (2003) was constructed. The 19 lichen KS domain sequences were aligned with fungal PKS sequences retrieved from GenBank, including most of the KS sequences used in Kroken et al. (2003). The bacterial PKSs were not included in the analysis, except for *S. viridochromogenes avLM* and *M. echinospora calO5*, which both produced orsellinic acid (a compound of interest in this study). Since the emphasis was on the NR-PKSs, more KS sequences from this class were included, whereas only the KS sequences from the characterized HR-PKSs and PR-PKSs were included in the phylogenetic analysis. To obtain a better picture of the diversity of PKS genes that are present in lichens, some lichen KS sequences from other studies (Grube & Blaha 2003; Opanowicz et al. 2006; Schmitt et al. 2005) were also included in the phylogenetic analysis. The KS sequences of Schmitt et al. (2005) are especially useful in inferring the relationship between KS phylogeny and compound structure, as the HPLC chemical data of the lichens are provided along with the phylogenetic analysis. Since many of the KS sequences are highly similar (with some sharing 98-100% identity when translated into amino acids), some KS sequences were represented by only one of them when the group as a whole showed >95% identity in amino acids, to avoid inflation of the phylogenetic tree due to over-representation of a particular homolog. Although PKSs with >95% amino acid identity at the KS domains cannot be assumed to have identical function, orthologous PKSs with such high similarity often produce identical polyketide products [e.g. the *A. nidulans WA* and *A. fumigatus Alb1p* (the KS domains share 88% identity), the *P. patulum MSAS* and *A. parasiticus PKSL2* (the KS domains share 68% identity)]. Additional PKS sequences characterized more recently were also included in the multiple sequence alignment and phylogenetic analysis.
3.5.1 Materials and methods

3.5.1.1 Multiple sequence alignment

Multiple sequence alignment was performed as described in Section 2.3.2. Selected multiple-aligned amino acid sequences of KS domains available in the online supporting information of Kroken et al. (2003) were used as the basis to add more sequences into the alignment (profile alignment mode in ClustalX). GenBank accession numbers of the protein sequences used in the multiple alignment and phylogenetic analyses are listed in Table 2.2.

3.5.1.2 Construction of phylogenetic tree

The evolutionary history of KS domains was inferred using the Minimum Evolution (ME) method (Rzhetsky & Nei 1992). The bootstrap consensus tree inferred from 1000 replicates (Felsenstein 1985) was taken to represent the evolutionary history of the taxa analyzed (Felsenstein 1985). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) were shown next to the branches (Felsenstein 1985). The tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl & Pauling 1965) and are in the units of the number of amino acid substitutions per site. The ME tree was searched using the Close-Neighbor-Interchange (CNI) algorithm (Nei & Kumar 2000) at a search level of 1. The Neighbor-joining (NJ) algorithm (Saitou & Nei 1987) was used to generate the initial tree. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There was a total of 128 positions in the final dataset. The resulting phylogenetic tree was artificially rooted with animal fatty acid synthases according to the phylogenetic tree in Kroken et al. (2003). Other phylogenetic methods such as ME with interior branch test, NJ and maximum parsimony (MP) were also tested with this same dataset for comparison with the bootstrapped ME tree. All the phylogenetic analyses were conducted in MEGA ver. 4.0 (Tamura et al. 2007).
3.5.2 Results

The resulting tree from 108 KS domains (Figure 3.9) showed a distinct major clade of NR-PKSs along with the HR- and PR-PKSs clades. The animal fatty acid synthases formed the root of the tree (rooted artificially). The phylogenetic relationship of the PKSs inferred by the KS domains was consistent with Kroken et al. (2003). The NR-PKSs further segregated into two major clades. Two subclades, corresponding to clade I and clade II NR-PKSs of Kroken et al. (2003), and several NR-PKSs basal to clade I and II, formed one of the major NR-PKS clades. The other NR-PKS major clade consisted of solely clade III NR-PKSs. The NR-PKSs that formed the base to clades I and II included the KS domains corresponding to clade VII and XI of (Schmitt et al. 2005).

As expected, PKSs (with complete sequence available) of similar domain architecture were grouped together in the same clade. Some minor changes of the common domain architecture assigned to each clade in Kroken et al. (2003) were observed, with the addition of newly characterized PKS sequences. Out of the 19 lichen KS domains obtained in this study, three were found in the NR-PKS clade I, one in the NR-PKS clades II, eight in the NR-PKS clade III, and three were basal to NR-PKS clades I and II. The eight KS domains in NR-PKS clade III were further segregated into two subclades along with other PKSs. Only one lichen KS domain (P. cunninghamii NRKS1) was grouped with PR-PKSs, while three other KS domains grouped with the HR-PKSs. The resulting bootstraped consensus minimum evolution (ME) tree in Figure 3.9 was not significantly different from the ME tree tested with the interior branch test (Nei & Kumar 2000) and was also similar to the tree constructed with the Neighbour-joining (NJ) method. Due to the limited computer processing power, the maximum parsimony (MP) tree was not finalized (the branch lengths were not calculated or only included limited sites); however, the initial MP trees showed similar topology to the ME tree in Figure 3.9.
Figure 3.9. KS domain phylogeny inferred from Maximum Evolution (ME) method (bootstrap = 1000 replicates). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test were shown next to the branches. The evolutionary distances are in the units of the number of amino acid substitutions per site. Classification of PKSs are based on Kroken et al. (2003).
3.5.3 Discussion

The KS domain ME tree in Figure 3.9 suggests that clades I and II NR-PKSs may have a closer phylogenetic relationship to each other than to NR-PKSs in clade III; this is in agreement with the observation that these NR-PKSs (clade I and II) have almost identical domain architecture. Based on the observations in Kroken et al. (2003), the corresponding PKSs of the lichen KS domains obtained in this study can be reasonably assumed to have similar domain architecture and function with other PKSs grouped together in the same clade. On this assumption, the following sections examined the potential role of the lichen KS domains in the individual clade by comparing them to the known PKSs of the same clade. Some new hypotheses were formed based on these speculations and require investigation in future.

3.5.3.1 Non-reducing PKS Clade I

The PKSs of NR-PKS clade I are known for producing long, unreduced polyketide chains that are precursors to sterigmatocystin, aflatoxin, heptaketide pyrone YWA1 and bikaverin (Table 3.5). Dothistromin and cercosporin are the two polyaromatic polyketide products that were linked to the clade I NR-PKSs more recently (Table 3.5). Three of the lichen KS domains amplified in this study fall into this clade, i.e. *F. soredians* NRKS1, *T. scutellatum* NRKS2 and *Cladonia* spp. (030) NRKS1 (Figure 3.9). However, compounds that are likely to originate from long unreduced polyketides, such as the polyaromatic anthraquinones, were not detected in these three lichens (Table 3.6).

Collectively, the NR-PKSs in clade I have the common domain architecture of KS-AT-ACP-(ACP)-(ACP)-CYC/R, which shows more variability than that originally proposed in Kroken et al. (2003), i.e. KS-AT-ACP-(ACP)-CYC [Note: domains that are sometimes absent are in parentheses]. All PKSs from the NR-PKS clade I, except three PKSs (*Fusarium verticillioides* PGL1, *Gibberella zeae* PKS3 and *Nectria haematococca* PGL1) possess a thioesterase (TE) domain. The TE domain is also referred as a Claisen-cyclization (CYC) domain due to the cyclization reaction it catalyses (Fujii et al. 2001). The recently discovered *F. verticillioides* PGL1, *G. zeae* PKS3 and *N. haematococca* PGL1 fall into this clade but they do not possess a CYC domain. Instead, a reductase (R) domain is present at the C-terminal of the PKSs
The R-domain may be involved in a reductive release of the polyketide chain (Cox 2007) (see Section 3.5.3.4). The F. verticillioides PGL1 and G. zeae PKS3 are involved in dark perithecial pigment production, but the structure of the polyketide product has not been resolved (Proctor et al. 2007). Duplicate copies of acyl-carrier protein (ACP) domains are also present in some PKSs in this clade, with the recently characterized Dothistroma septosporium PKSA having up to three copies of the ACP domain.

TscNRKS2 and CspNRKS1 formed a small subclade with Cercospora nicotianae CTB1, between the YWA1 heptaketide synthases (A. nidulans WA and A. fumigatus ALB1) and the norsolorinic acid octaketide synthases (A. nidulans PKSST and A parasiticus PKSL1) (Figure 3.9). The YWA1 heptaketide synthases produce a heptaketide chain intermediate (YWA1), which is shortened to a pentaketide chain by a protein (Ayg1P) and cyclized into tetrahydroxynaphthalene (a precursor for DHN-melanin and the blue-green conidial pigment of A. fumigatus) (Table 3.5, 1), whereas the C. nicotianae CTB1 is involved in the biosynthesis of cercosporin (a red pigment) (Table 3.5, 4). Chen et al. (2007) proposed a pathway to the polyketomethylene intermediate (monomer) of cercosporin via a pentaketide chain but its similarity to the clade I NR-PKSs suggests that a chain length shortening mechanism analogous to Ayg1p could be involved (Table 3.5, 4).
<table>
<thead>
<tr>
<th>Characterised PKS genes</th>
<th>Pathway</th>
<th>Polyketide intermediate/ final product</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) YWA1 heptaketide synthases</td>
<td></td>
<td></td>
<td>(Fujii et al. 2001; Fujii et al. 2004; Tsai et al. 2001; Watanabe et al. 1999)</td>
</tr>
<tr>
<td>- <em>A. nidulans</em> WA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[KS-AT-ACP-ACP-TE/CYC]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- <em>A. fumigatus</em> PKSP (alb1p)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[KS-AT-ACP-ACP-TE/CYC]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1) Norsolorinic acid (octaketide) synthases</td>
<td></td>
<td></td>
<td>(Bradshaw et al. 2006; Bradshaw &amp; Zhang 2006; Chang et al. 1995; Crawford et al. 2006; Feng &amp; Leonard 1995; Minto &amp; Townsend 1997; Yu et al. 2002; Yu &amp; Leonard 1995)</td>
</tr>
<tr>
<td>- <em>A. nidulans</em> PKSST (stcA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[KS-AT-ACP-ACP-TE/CYC]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- <em>A. parasiticus</em> PKSL1(PKSA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[KS-AT-ACP-TE/CYC]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- <em>Mycosphaerella pini</em> PKSA (Dothistroma septosporum)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[KS-AT-ACP-ACP-ACP-TE/CYC]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3) Bikaverin (nonaketide) synthase</td>
<td></td>
<td></td>
<td>(Linnemannstons et al. 2002; Ma et al. 2007; McInnes et al. 1976)</td>
</tr>
<tr>
<td>- <em>Fusarium fujikuroi</em> PKS4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Gibberella moniliformis)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[KS-AT-ACP-TE/CYC]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4) Penta/heptaketide? synthase
- *Cercospora nicotianae* CTB1
  [KS-AT-ACP-ACP-TE/CYC]

<table>
<thead>
<tr>
<th>Cercosporin</th>
</tr>
</thead>
</table>

Common PKS architecture: KS-AT-ACP-(ACP)-(ACP)-CYC/R [Note: domains that are variable in presence are in parentheses]

Table 3.6. Lichen KS domains that grouped with clade I NR-PKSs and compounds detected in the corresponding lichens.

<table>
<thead>
<tr>
<th>Lichen KS domains in NR-PKS clade I</th>
<th>Compounds detected in the lichens</th>
<th>Compounds related to clade I NR-PKSs</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. scutellatum (Tsc) NRKS2</td>
<td>alkyl resorcylic acid derivatives – divaric acid, methyl divarate, divaricatic acid, sekikaic acid and stenosporic acid, methylphloroacetophenone derivatives – usnic acid</td>
<td>Unknown</td>
</tr>
<tr>
<td>Cladonia spp. (Csp) NRKS1</td>
<td>alkyl resorcylic acid derivatives – merochlorophaeic acid and homosekikaic acid, β-orcinol derivatives – fumarprotocetraric acid</td>
<td>Unknown</td>
</tr>
<tr>
<td>F. soredians (Fso) NRKS1</td>
<td>β-orcinol derivatives – consalazinic acid, salazinic acid, galbinic acid, protocetraric acid, methylphloroacetophenone derivatives – usnic acid</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

Class of lichen compounds predicted to be produced by clade I NR-PKSs: Anthraquinones and related xanthones Class 1.2.3 (Table 1.1)
Similar long chain polyketide compounds were not detected in both *T. scutellatum* and *Cladonia* spp. (030). Instead, several alkylresorcylic acid-derived compounds and usnic acid were detected in *T. scutellatum*, while two orcinol depsides and a β-orcinol depsidone were detected in *Cladonia* spp. (030) (Table 3.6). Since all of these compounds are derivatives of monoaromatic tetraketides (see Section 1.2.6.2), it was concluded that both TscNRKS2 and CspNRKS1 are unlikely to be involved in their biosynthesis. TscNRKS2 and CspNRKS1 also grouped closely with *F. verticillioides* PGL1 and *G. zeae* PKS3, which are involved in dark perithecial pigment production. It is not clear if the clade I NR-PKSs found in *T. scutellatum* (with pale brown pigmentation at the ascocarp) and *Cladonia* spp. 030 (brown pigmentation in the apothecia of *C. ramulosa*) could be involved in melanin or pigment production, as polymerized melanin could be difficult to extract and might not be detected in HPLC analysis, but it is equally possible that the two lichen clade I NR-PKSs are cryptic genes only activated under certain conditions.

In the case of FsoNRKS1, the KS domain grouped with norsolorinic acid synthases (NSAS) – *A. nidulans* PKSST (87.7%), *A. parasiticus* PKSL1 (86.9%), and *D. septosporum* (85.5%). Such high similarity suggests that FsoNRKS1 is part of a PKS that has similar, if not identical, function to its counterparts. However, no nonaketide or norsolorinic acid-like compounds were detected in the *F. soredians* specimens. *F. soredians* produces several β-orcinol depsidones and usnic acid, which are all derived from methylated tetraketides (Table 3.6), and therefore the involvement of FsoNRKS1 in biosynthesis of these compounds is unlikely. Interestingly, two naphthapyrones (euplectin and coneuplectin) and an anthraquinone (skyrin) that are all biosynthetically related to norsolorinic acid-derived compounds were found in a lichen of the same genus – *Flavoparmelia euplecta* (Ernst-Russell et al. 2000).

The ability of NRKS-F/R to amplify a NSAS homolog from *F. soredians* is expected. Based on the multiple alignment analysis in Figure 3.5, the LC1/2c primers would not be able to amplify the NAS homologs, and this corresponds to the lack of detection of a NSAS-like KS domain in Schmitt et al. (2005), who employed the LC1/2c primers. In another ME phylogenetic analysis (data not shown), the clade I NR-type *Xanthoria parietina* KS1 obtained by Grube & Blaha (2003) with LC1/2c
primers grouped more closely with the *F. verticillioides* PGL1, *G. zeae* PKS3 and *N. haematococca* PGL1 and not the norsolorinic acid synthases. The NSAS-type PKS genes may be present in the lichens investigated in the two studies (Grube & Blaha 2003; Schmitt et al. 2005), but could not be amplified by the LC1/2c primers due to the mismatch shown earlier.

### 3.5.3.2 Non-reducing PKS Clade II

Clade II NR-PKSs share similar domain architecture with NR-PKS clade I, except that they are thought to produce shorter-chain (pentaketide) products (Table 3.7). All the members of the clade known so far possess a CYC domain. The only polyketide product that has been linked to clade II NR-PKSs is tetrahydroxynaphthalene (THN), which is the precursor for dihydroxynaphthalene (DHN)-melanin. DHN-melanin is associated with spore pigmentation (Tsai et al. 1998). The best-characterised PKS in this clade is *C. lagenarium* PKS1, a pentaketide synthase known to utilize malonyl-coA instead of acetyl-coA as the starter unit to produce THN (Fujii et al. 2000). The PKS homologs of this pentaketide synthase are widespread among many fungi, including *Glarea* sp., *Nodulisporium* sp., *Wangiella dermatitidis*, *Exophiala lecanii-corni*, *Bipolaris oryzae*, *Ceratocystis resinifera*, *C. heterostrophus* etc. and are all involved in melanin biosynthesis (Kroken et al. 2003) (Table 3.7).

Only two lichen KS domains in this study were found to group with this clade – *T. scutellatum* NRKS1 and *P. borreri* NRKS1. They had no pentaketide compound, but alkyl resorcylic acid-derivatives (Tsc), usnic acid (Tsc), orsellinic acid derivatives (Pbo) and β-orcinol depside atranorin (Pbo), were detected in these two lichens. All these compounds are formed from tetraketide monoaromatic precursors.

The two lichen KS domains may be part of a PKS that produces the precursor (THN) for melanin and pigments, since *P. borreri* has a greenish-grey upper surface and *T. scutellatum* shows pale brown pigmentation at the ascocarps. The TscNRKS1 also grouped closely with the KS domains in “clade V” according to the groupings in Schmitt et al. (2005), which included KS domains from *Coccotrema pocillarium*, *Pertusaria occulata*, and a group of KS sequences from *Lecanora* spp. and *Tephromela atra* obtained by Grube & Blaha (2003). Both *T. atra* and *P. occulata* were characterized by deeply blackish ascomata, and the *Lecanora* spp. also
developed dark pigments locally in their thalli (Schmitt et al. 2005; Grube & Blaha 2003). So, it is possible that the “clade V” lichen PKSs, including the TscNRKS1, are involved in melanin production.

The clade II NR-PKSs could also be involved in biosynthesis of orsellinic acid derivatives detected in the lichens, since tetraketide orsellinic acid was found as a byproduct in A. oryzae transformants carrying the C. lagenarium PKS1 gene (Fujii et al. 1999). This suggests that a PKS1 homolog with minor variations could be involved in the production of the orcinol moiety in lichens. The PboNRKS1 is especially diverged from the melanin-producing PKSs and formed a small subclade by itself with an uncharacterised PKS gene – A. terreus at5 (Figure 3.9). The detection of several orsellinic acid derivatives in P. borerr (Table 3.8) is in agreement with the hypothesis that the PboNRKS1 might be an orsellinic acid synthase. In T. scutellatum, however, the orcinol-type compounds detected were all derivatives of propyl resorcylic acid, and therefore the participation of TscNRKS1 in their biosynthesis is unlikely.
### Table 3.7. Characterised polyketide synthases in the non-reducing (NR) clade II and their corresponding products.

<table>
<thead>
<tr>
<th>Characterised PKS genes</th>
<th>Pathway</th>
<th>Polyketide intermediate/ final product</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) THN pentaketide synthases</td>
<td></td>
<td></td>
<td>(Cheng et al. 2004; Feng et al. 2001; Fujii et al. 1999, 2000; Fulton et al. 1999; Moriwaki et al. 2007; Paolo et al. 2006; Tanguay et al. 2006; Zhang et al. 2003)</td>
</tr>
<tr>
<td>[KS-AT-ACP-TE/CYC]</td>
<td></td>
<td>DHN-Melanin</td>
<td></td>
</tr>
<tr>
<td>- C. lagenarium PKS1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- G. lozoyensis PKS1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Exophiala lecanii-corni ElPKS1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Wangiella dermatitidis WdPKS1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Nodulisporium sp. PKS1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Bipolaris oryzae PKS1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Ceratocystis resinifera PKS1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Common PKS architecture: KS-AT-ACP-(ACP) -TE/CYC

### Table 3.8. Lichen KS domains that grouped with clade II NR-PKSs and compounds detected in the corresponding lichens.

<table>
<thead>
<tr>
<th>Lichen KS in NR-PKS clade II</th>
<th>Compounds detected in the lichens</th>
<th>Compounds related to this NR-PKS clade II</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. scutellatum (Tsc) NRKS1</td>
<td>alkyl resorcylic acid derivatives - divaric acid, methyl divarate, divaricatic acid, sekikaic acid and stenosporic acid methylphloroacetophenone derivatives - usnic acid</td>
<td>DHN-melanin?</td>
</tr>
<tr>
<td>P. subrudecta (Pbo) NRKS1</td>
<td>orsellinic acid derivatives - orsellinic acid, methyl orsellinate, lecanoric acid, orcinyl lecanorate, gyophoric acid β-orcinol derivatives - atranorin</td>
<td>orsellinic acid and related mono-, di-, tri-aryl derivatives?</td>
</tr>
</tbody>
</table>

Class of lichen compounds predicted to be produced by clade II NR-PKSs: DHN-melanin?
3.5.3.3 Non-reducing PKS basal to Clades I and II

According to Kroken et al. (2003), basal to clades I and II are several uncharacterized NR-PKSs that have similar domain architecture to clade I and clade II NR-PKSs, but did not form a significant clade. The presence of a TE/CYC domain in these PKSs is variable. The phylogenetic analysis of Kroken et al. (2003) showed this group consisted of at least two small subclades and an isolated *Gibberella zeae* PKS14 (see Figure 3.1). According to the phylogenetic tree in Figure 3.9, the clades VII and XI, based on Schmitt et al. (2005), also fall into this category.

The three new lichen KS domains in this study that could be included in this group of PKS are *X. semiviridis* NRKS1, *F. caperatulla* NRKS1 and *U. oncodeoides* NRKS1. All three KS domains grouped together with *B. fuckeliana* PKS15 (62-73% amino acid identity) and were also similar to *G. zeae* PKS13 (42-49%). Both *B. fuckeliana* PKS15 and *G. zeae* PKS13 have the same domain architecture – KS-AT-ACP-TE/CYC. Recently, *G. zeae* PKS13 and PKS4 (a HR-PKS) was shown to involved in the biosynthesis of the mycotoxin zearalenone (Kim et al. 2005). The structure of zearalenone [6-(10-hydroxy-6-oxotrans-1-undecenyl)-β-resorcylic acid lactone] consists of a polyketide with four unreduced ketone groups and a highly-reduced chain (Figure 3.10). Thus, it was proposed that the *G. zeae* PKS4 produces the highly-reduced polyketide chain, which is then incorporated by *G. zeae* PKS13 as a starter unit, followed by three malonyl-CoA condensations and aldol cyclization, to yield the zearalenone structure (Cox 2007; Kim et al. 2005) (Figure 3.10). In this case, the *G. zeae* PKS13 acts as a tetraketide synthase. The aldol cyclization of the aromatic moiety in zearalenone is analogous to orsellinic acid (see Figure 1.10), and so a *G. zeae* PKS13 homolog that incorporates acetyl-CoA as a starter unit, instead of the highly-reduced polyketide, should produce orsellinic acid.

Nonetheless, no orsellinic acid derivatives or zearalenone-like compounds were detected in *X. semiviridis*, *F. caperatulla* and *U. oncodeoides*. All three lichens produce β-orcinol depsidones (β-orsellinic acid derivatives) and usnic acid (methylphloroacetophenone derivative) as major secondary metabolites. Both of the key precursors, β-orsellinic acid and methylphloroacetophenone, required a CMeT domain for methylation at the aromatic ring (see Figure 1.10). The absence of a CMeT domain in *B. fuckeliana* PKS15 and *G. zeae* PKS13 suggests that XsNRKS1,
FcaNRKS1 and UonNRKS1 could not be involved in the biosynthesis of these major metabolites, although they are potentially tetraketide synthases.

![Proposed biosynthesis of zearalenone by *G. zeae* PKS13 (a NR-PKS) and PKS4 (a HR-PKS). The highly-reduced polyketide chain produced by PKS4 is in grey. Box: aldol cyclization of the tetraketide moiety that is analogous to orsellinic acid biosynthesis.](image)

**Figure 3.10.** Proposed biosynthesis of zearalenone by *Gibberella zeae* PKS13 (a NR-PKS) and PKS4 (a HR-PKS). The highly-reduced polyketide chain produced by PKS4 is in grey. Box: aldol cyclization of the tetraketide moiety that is analogous to orsellinic acid biosynthesis.

3.5.3.4 Non-reducing PKS Clade III

Unlike the ambiguous NR-PKSs basal to clade I and II, clade III NR-PKSs formed a distinct clade. This clade was further segregated into two subclades with four new lichen KS domains in each subclade (Figure 3.9). The NR-PKSs in clade III are characterised by the presence of a C-methyltransferase (CMeT) domain not found in other NR-PKSs, but the present of a CYC domain is variable. The presence of a CMeT domain leads to the hypothesis that the clade III NR-PKSs are involved in biosynthesis of the many β-orsellinic acid and methylphloroacetophenone derivatives in lichens (see Section 3.4.3).

Only two clade III NR-PKSs has been characterized so far. The *M. purpureus* PKSCT involved in citrinin biosynthesis is the first clade III NR-PKS to be characterized (Shimizu et al. 2005) (Table 3.9, 1). It has a domain architecture of KS-AT-ACP-CMeT-R. The C-MeT domain was proposed to be involved in the methylation of the polyketide intermediate of citrinin (Shimizu et al. 2005). An unprecedented reductase (R) domain instead of the common CYC domain is present at the C-terminal of PKSCT, but its role in biosynthesis of citrinin is unknown. The
function of the R-domain has only become more apparent recently with the discovery and characterization of a methylorcinaldehyde synthase (MOS) from *Acremonium strictum* (AsPKS1), which may be involved in xenovulene A biosynthesis (Bailey et al. 2007) (Table 3.9, 2). The domain architecture of MOS is identical to PKSCT and also possesses an R-domain, which is likely to catalyze a reductive release of the tetraketide chain, resulting in an aldehyde terminal instead of the usual carboxyl terminal in orsellinic acid. The high protein similarity of PKSCT to MOS suggests that PKSCT is probably a tetraketide synthase like MOS and its R-domain should perform a similar function. On this basis, the PKSCT was proposed to synthesize the citrinin intermediate using a methylated diketide starter unit involving a reductive release (Bailey et al. 2007; Cox 2007) (Table 3.9, 1).

Four of the lichen KS domains – *P. chinense* NR3KS1 and NR3KS2, *P. reticulatum* NR3KS1 and *H. pulv... domain analysis (Figure 3.9). The CYC domain is present in *C. heterostrophus* PKS23, but not in *C. heterostrophus* PKS22 and *B. fuckeliana* PKS16 (Figure 3.9). These three lichens (*P. chinense*, *P. reticulatum* and *H. revoluta*) produce β-orsellinic acid derivatives but not methylphloroacetophenone derivatives (Table 3.10); these four lichen KS domains may be involved in biosynthesis of these β-orsellinic acid-derived compounds.

The *P. jackii* NRKS1 is the first lichen KS domain known to group with the clade III NR-PKSs (Section 3.3.3.2). It grouped closely with the uncharacterized *B. fuckeliana* PKS18 and *A. terreus* ATEG_10080. Both *B. fuckeliana* PKS18 and *A. terreus* ATEG_10080 possess a KS-AT-ACP-CMeT domain architecture and a domain of undetermined function at the C-terminal. A NCBI conserved domain search indicated that the C-terminal domain is homologous to alpha/beta hydrolase (pfam07859) and predicted esterase (COG0400) domains. EBI Interproscan also detected a conserved motif with a serine active site (PROSITE accession no. PS00120), which is common in lipases, at the C-terminal region of both PKSs. This domain is different from the previously known TE/CYC domain found at C-terminal of most NR-PKSs, but it may also be involved in polyketide chain release. It is possible that PjaNRKS1 is involved in biosynthesis of β-orsellinic acid derivatives in *P. jackii*. 

144
Table 3.9. Characterised polyketide synthases in the non-reducing (NR) clade III and their corresponding products.

<table>
<thead>
<tr>
<th>Characterised PKS genes</th>
<th>Pathway</th>
<th>Polyketide intermediate/ final product</th>
<th>References</th>
</tr>
</thead>
</table>
| 1) Citrinin (tetraketide) PKS  
  - *Monascus purpureus* PKSCT  
| 2) Methyl orcinaldehyde (tetraketide) synthase (MOS)  
  - *Acremonium strictum* AsPKS1  
  [KS-AT-ACP-CMeT-R] | Xenovulene | ![Xenovulene pathway](image) | (Bailey et al. 2007; Raggatt et al. 1997) |

Common PKS architecture: KS-AT-ACP-CMeT-(CYC/R)
<table>
<thead>
<tr>
<th>Lichen KS in NR-PKS clade III</th>
<th>Compounds detected in the lichens</th>
<th>Compounds related to this NR-PKS clade III</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. chinense</em> (Pch) NR3KS1</td>
<td><em>β</em>-orcinal derivatives – atranorin, constictic acid, stictic acid, cryptostictic acid, menegazziaic acid</td>
<td><em>β</em>-orcinal derivatives</td>
</tr>
<tr>
<td><em>NR3KS2</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>P. reticulatum</em> (Pre) NR3KS1</td>
<td><em>β</em>-orcinal derivatives – atranorin, chloroatranorin, constictic acid, peristictic acid, stictic acid, cryptostictic acid, menegazziaic acid</td>
<td><em>β</em>-orcinal derivatives</td>
</tr>
<tr>
<td><em>H. revoluta</em> (Hre) NR3KS1</td>
<td><em>β</em>-orcinal derivatives – atranorin*</td>
<td><em>β</em>-orcinal derivatives</td>
</tr>
<tr>
<td></td>
<td><em>orcinol derivatives – gyrophoric acid</em></td>
<td><em>β</em>-orcinal derivatives</td>
</tr>
<tr>
<td></td>
<td><em>based on Walker &amp; Lintott (1997) and Elix (1994)</em></td>
<td></td>
</tr>
<tr>
<td><em>P. jackii</em> (Pja) NRKS1</td>
<td><em>β</em>-orcinal derivatives – methyl β-orsellinate, atranorin, haematommiac acid, methyl isohaematommate, methyl haematommate</td>
<td><em>β</em>-orcinal derivatives</td>
</tr>
<tr>
<td><em>X. semiviridis</em> (Xse) NR3KS1</td>
<td><em>β</em>-orcinal derivatives – fumarprotocetraric acid, succinprotocetraric acid, conprotocetraric acid, consuccinprotocetraric acid, protocetraric acid, virensic acid methylphloroacetophenone derivatives – usnic acid</td>
<td><em>β</em>-orcinal derivatives &amp; usnic acid</td>
</tr>
<tr>
<td><em>NR3KS2</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>F. caperatulla</em> (Fca) NR3KS1</td>
<td><em>β</em>-orcinal derivatives – protocetraric acid, conprotocetraric acid, atranorin methylphloroacetophenone derivatives – usnic acid</td>
<td><em>β</em>-orcinal derivatives &amp; usnic acid</td>
</tr>
</tbody>
</table>

Class of lichen compounds predicted to related to NR-PKS clade III: *β*-orcinal derivatives in Class 1.2.1 and 1.2.2 (see Table 1.1)
The *X. semiviridis* NR3KS1 is most similar to *B. fuckeliana* PKS19. However, the *B. fuckeliana* PKS19 sequence in GenBank is incomplete and the domain architecture of this PKS is unknown. It is likely that both *B. fuckeliana* PKS19 and the PKS corresponding to XseNR3KS1 have a KS-AT-ACP-CMeT domain architecture, but the presence of a C-terminal releasing domain could not be predicted. Since *X. semiviridis* produces both β-orsellinic acid derivatives and usnic acid (Table 3.1); XseNR3KS1 may be involved in biosynthesis of either monoaromatic precursor (β-orsellinic acid or methylphloroacetophenone) that lead to these compounds.

Two lichen KS domains, *F. caperatulla* NR3KS1 and *X. semiviridis* NR3KS2, formed a small subclade with *M. purpureus* PKSCT and *Acremonium strictum* AsPKS1, along with the uncharacterized *B. fuckeliana* PKS17 and *C. heterostrophus* PKS21 (Figure 3.9). The high similarity of *F. caperatulla* NR3KS1 and *X. semiviridis* NR3KS2 to PKSCT and AsPKS1 (MOS) indicates that both of the corresponding lichen PKSs might possess a R-domain at the C-terminal. Both *F. caperatulla* and *X. semiviridis* produced usnic acid and a variety of β-orcinol depsidones that are biosynthetically related to protocetraric acid (Table 3.10). The likely absence of a CYC domain, which catalyzes Claisen cyclisation crucial to the formation of methylphloroacetophenone, suggests that the two PKSs (FcaNR3KS1 and XseNR3KS2) are unlikely to biosynthesize usnic acid, but possibly the β-orsellinic acid derivatives found in these two lichens.

3.5.3.5 Partial-reducing PKSs

In this study, the only PR-PKS detected was from *P. cunninghamii* PRKS1, which nested among other fungal MSAS (6-methylsalicylic acid synthase), i.e. *P. patulum* 6MSAS, *A. parasiticus* PKSL2, *G. lozoyensis* PKS2 (Figure 3.9). The high similarity shared between PcuPRKS1 and other fungal MSASs suggests they probably have similar, if not the same, functions. It was previously hypothesized that the PR-PKSs could produce orsellinic acid, as do the two bacterial orsellinic acid synthases (OASs), *M. echinospora* CalO5 and *S. viridochromogens* AviM, which are similar to fungal PR-PKSs/6MSASs. However, no 6-methylsalicylic acid or orsellinic acid derivatives were detected in *P. cunninghamii*, and so PcuPRKS1 is possibly part of a cryptic PKS gene. In a recent study, the 6MSAS homologs were also detected in
several lichens of the genus *Pertusaria* (Schmitt et al. 2008). Similarly, no 6-methylsalicylic acid was detected in these lichens and the function of these PKSs is unknown. Schmitt et al. (2008) proposed that these lichen PR-PKSs may be involved in biosynthesis of typical lichen metabolites (coupled phenolic compounds).

3.5.3.6 Highly-reducing PKSs

Three lichen KS domains were found in this clade. Two of them, *F. soredians* HRKS1 and *P. reticulatum* HRKS1, are sister to the *A. terreus* lovB and *P. citrinum* mlcA of clade II HR-PKS, which produce the nonaketide moiety of lovastatin and compactin (citrinin) respectively. Another HR-type lichen KS domain, *U. oncodeoides* HRKS1, grouped with *A. terreus* (lovF) and *P. citrinum* (mlcB) of clade I HR-PKSs, which produce the diketide moieties of lovastatin and compactin (citrinin) respectively. The recently characterised squalestatin triketide synthase from *Phoma* sp. (Cox et al. 2004) is also another member of this clade (GenBank accession no. AY217789, sequence not included in the phylogenetic analysis).

The detection of HR-PKSs in the three lichens was unexpected and was not the focus of this study. No highly-reduced polyketides were detected in all three lichens above using the HPLC and TLC methods described previously (Section 3.2.2.2, Table 3.1). Due to the lack of a chromophore (or a conjugated system) in most highly-reduced polyketides, their presence in these lichens may not be detected by the UV-spectrophotometer coupled to the HPLC instrument or the TLC plates viewed under the UV lamp. Several macrolides and related aliphatic acids, which are likely to be produced by HR-PKSs, were previously isolated from lichens, e.g. aspicilin from *Aspicilia caesiocinerea* and *Lecanora* spp. (Huneck et al. 1973; Quinkert et al. 1985), lepranthin from *Arthonia impolita* (Polborn et al. 1995), dasypogalactone from *Usnea filipendula* (formerly *U. dasypoga*) (Suwarso et al. 1999), and bourgeanic acid from *Ramalina* spp. (Bodo et al. 1973; Follmann & Huneck 1968) (Figure 3.11). Since HR-PKSs seem to be a rich source of pharmaceutically relevant compounds (e.g. lovastatin, mevastatin, squalestatin are all important cholesterol lowering agents) a more detailed exploration of lichen HR-PKSs in future would be worthwhile.
Figure 3.11. Examples of lichen compounds formed from highly-reduced polyketides. * A homodimeric compound with two identical polyketide chains; ** A homotrimeric compound with three identical polyketide chains.

3.6 Conclusions and overview

The main objective of this study was to develop new degenerate primers that would allow the detection of a larger variety of PKS. The NRKS-F/R primers amplified NR-type KS domains that are different from those amplified from LC1/2c. By contrast, KS domains ranging from PR-, HR-to clade III NR-PKSs have been obtained with the PRKS-F/R primers. More significantly, the NR3KS-F/R primers targeted specifically the clade III NR-PKSs, which were characterised by the presence of a CMeT domain. Although only a total of 19 KS domains was amplified and sequenced with the three sets of primers, KS domains from a broad range of PKSs were sampled. The low number of KS domains obtained could be partly attributed to the mixed sequences obtained from many KS domain PCR products. These multiple PCR products of the
same size could be isolated by PCR-cloning, as demonstrated for *X. semiviridis* and *P. chinense* (XseNR3KS1, XseNR3KS2, PchNR3KS1 and PchNR3KS2).

Almost all of the metabolic patterns detected in this study are consistent with those described in the identification keys and phytochemical register (Archer 1992; Elix 1994; Walker & Lintott 1997), with a few lichens having additional compounds detected at lower concentration but being biosynthetically related to those already known to present in the corresponding lichens. Consequently, it is rather surprising that most KS domains amplified from the lichens pointed to PKS genes with domain architecture not matching the compounds detected. These PKS genes may be silent in the absence of a particular trigger, or the compounds may be produced at concentrations below the detectable limit. A closer examination of the lichen KS sequences analysed by Schmitt et al. (2005) also observed similar trends, whereby many PKSs could not be matched with the compounds present in the organisms. This suggests that the metabolic diversity of lichens is potentially underestimated.

Nevertheless, the sampling of PKS genes from lichens (and fungi in general) via a degenerate primer PCR approach is likely to be incomplete. Fungal genome sequencing has showed that there are numerous PKS genes in fungi. For example, 7-25 PKS genes were detected in each genome of the saprobic ascomycetes studied in Kroken et al. (2003), while up to 30 PKS genes were found in some *Aspergillus* genomes (Varga et al. 2003). Most studies have also obtained fewer KS domains than expected (Bingle et al. 1999; Grube & Blaha 2003; Nicholson et al. 2001; Sauer et al. 2002; Schmitt et al. 2005). Therefore, it is expected that there are more PKS genes in the lichen genomes not detected in this and previous studies.

The phylogenomic comparison of PKS genes acquired from fungal genome sequencing projects (Kroken et al. 2003) showed that the common ancestor of the four classes of Pezizomycotina (Sordariomycetes, Eurotiomycetes, Dothiomycetes and Leotiomycetes) had at least eight types of PKS genes (four subclades of HR-PKSs and four subclades of NR-PKSs, see Fig 3.1). Gene duplications lead to some fungi having more than one copy of a PKS gene homolog, while gene losses account for the absence of a particular PKS subclade in some fungi. The diversity of PKSs from the ascolichens (mainly the Lecanoromycetes) sampled so far, shows that they could, in general, fit into this “eight ancestral fungal PKSs” model (although only two
HR-PKSs subclades from lichens have been identified so far. Therefore, it is likely that the Lecanoromycetes (also in the order of Pezizomycotina) have also inherited the set of PKS genes from the common Pezizomycotina ancestor. Regardless of whether or not this hypothesis reflects the true fungal PKS phylogeny, this study indicates that the KS domain phylogeny provides a good guide for the exploration of different types of PKSs in ascomycetous lichens, as it allowed the prediction of the PKS domain architecture and type of polyketide products. For example, to obtain fragments of NR-PKS genes with a CMeT domain, degenerate primers were designed to target the clade III NR-PKSs based on the study by Kroken et al. (2003). Using this approach, eight lichen KS domains were amplified and the phylogenetic analysis showed that these KS domains grouped with the other clade III NR-type KS domains, as expected.

Another objective of this chapter was to find out if the metabolic profile of lichens could be matched to the corresponding KS domains on the basis of the relationship between KS domain phylogeny and PKS domain architecture, by comparing the KS domains from lichens with similar metabolic profiles. The results from this study have shown that this approach was not particularly successful; this could be due to a few reasons:

1) Unlike the genome-wide PKS sampling from whole-genome sequencing projects as shown in Kroken et al. (2003), the sampling of PKS genes in an organism using a degenerate primer approach is probably incomplete. Not KS domains of all PKS genes could be amplified. The problem could be overcome using more sets of degenerate primers that target smaller sets (clades) of PKS genes.

2) The metabolic profile detected by HPLC did not match the inherent metabolic potential of lichens; therefore the number of PKS genes probably exceeds the number of polyketide compounds detected. Some PKS genes may be silent in the organism or only expressed under specific conditions.

3) Convergent evolution would further complicate the prediction of the product of a PKS based on its similarity to other characterized PKSs. A well-known example is the similarity in the product from both the clade I NR-type Aspergillus spp. WA/ALB1 (heptaketide synthases) and the clade II NR-type C. lagenarium PKS1 (a pentaketide...
synthase), which both produce THN (tetrahydroxynaphthalene). In the case of the WA/ALB1, another protein, Ayg1P, is involved in shortening the heptaketide chain to pentaketide. Recent mutations in specific accessory domains that resulted in functional changes of the PKS could also cloud the prediction of PKS function based on KS domain phylogeny. For example, the A. terreus LovB possesses a non-functional ER domain (Hendrickson et al. 1999; Kennedy et al. 1999) but the KS domain phylogeny would predict that LovB should have an ER domain.

Despite the limitations, this study demonstrated that the degenerate primer approach remained a powerful tool for rapid screening of PKS genes in lichens at a relatively low cost. The range of PKSs detected could be extended by combining multiple sets of degenerate primers that target specific clade of PKS genes. With the aid of phylogenetics, the KS domain sequences could be used to obtain a crude prediction of the domain architecture of the corresponding PKSs, and to infer the type of polyketide they would produce.

Based on the proposed biosynthetic routes for β-orsellinic acid and methylphloroacetophenone, the clade III NR-PKSs are likely to be the key to many characteristic lichen compounds derived from the two monoaromatic units. The KS domain PCR sampling and phylogenetic analysis performed in this chapter support this hypothesis. The following chapter focused on locating and cloning of lichen PKS genes of this clade with the intention of characterizing them by heterologous expression.
CHAPTER 4 LOCATING AND CLONING LICHEN PKS GENES

4.1 Introduction ........................................................................................................ 155
  4.1.1 Rationale and research goal ........................................................................ 160

4.2 Locating and cloning of partial PKS genes in Flavoparmelia caperatulla ...................... 161
  4.2.1 Materials and methods .............................................................................. 162
    4.2.1.1 Southern blotting .............................................................................. 162
    4.2.1.2 Construction of partial genomic library ............................................. 162
    4.2.1.3 Screening of partial genomic library ................................................... 163
  4.2.2 Results .......................................................................................................... 163
    4.2.2.1 Southern blotting .............................................................................. 163
    4.2.2.2 Construction and screening of partial genomic library ...................... 165
  4.2.3 Discussion .................................................................................................... 167
    4.2.3.1 High molecular weight lichen DNA for Southern and cloning ......... 167
    4.2.3.2 Southern-guided library construction and PCR screening ............. 168

4.3 Locating and cloning of partial PKS genes in Xanthoparmelia semiviridis .................... 170
  4.3.1 Materials and methods .............................................................................. 170
    4.3.1.1 Southern blotting .............................................................................. 171
    4.3.1.2 Construction of X. semiviridis partial genomic library ................... 171
    4.3.1.3 Screening of partial genomic library ................................................... 171
  4.3.2 Results .......................................................................................................... 172
    4.3.2.1 Southern blotting .............................................................................. 172
    4.3.2.2 Construction and screening of partial genomic library ...................... 172
    4.3.2.3 Bioinformatic analysis ....................................................................... 174
  4.3.3 Discussion .................................................................................................... 177
    4.3.3.1 Southern-guided PKS gene cloning .................................................. 177
    4.3.3.2 Bioinformatic analysis ....................................................................... 178

4.4 Detection of CMeT domain in NR-PKS genes ....................................................... 180
  4.4.1 Materials and methods .............................................................................. 181
    4.4.1.1 Design of NR-type CMeT domain-specific primers ......................... 181
    4.4.1.2 PCR amplification of CMeT domain .................................................. 181
    4.4.1.3 Southern blotting with CMeT domain probe ..................................... 181
    4.4.1.4 KS-CMeT domain-hopping PCR ....................................................... 182
  4.4.2 Results .......................................................................................................... 182
    4.4.2.1 NRMMeT degenerate primers .......................................................... 182
    4.4.2.2 PCR amplification of CMeT domains ............................................... 183
    4.4.2.3 Southern blotting with CMeT domain probe ..................................... 185
    4.4.2.4 Extending the PKS gene sequence by domain-hopping PCR .......... 185
  4.4.3 Discussion .................................................................................................... 186
    4.4.3.1 Clade III NR-type CMeT domain ...................................................... 186
    4.4.3.2 X. semiviridis clade III NR-PKS gene (xsepks1) .............................. 189
4.5  Cloning of a full-length PKS gene from *X. semiviridis* ..... 190

4.5.1  Materials and methods.................................................................191
  4.5.1.1  Southern blotting with the xsepks1 KS-CMeT probe ................. 191
  4.5.1.2  Construction and screening of partial genomic library with large
           inserts 191

4.5.2  Results.....................................................................................192
  4.5.2.1  Southern blotting with the xsepks1 KS-CMeT probe ................. 192
  4.5.2.2  Cloning of the full-length PKS gene........................................ 193
  4.5.2.3  Sequence and PKS domain analysis....................................... 194

4.5.3  Discussion...............................................................................197
  4.5.3.1  A clade III NR-PKS from *X. semiviridis (xsepks1)*................. 197
  4.5.3.2  Regions surrounding the xsepks1.......................................... 203
4.1 Introduction

Locating and cloning of a PKS gene responsible for production of a polyketide compound could be an enormous task, especially before the wide availability of fungal PKS gene sequences. Schumann & Hertweck (2006) have recently reviewed the technological aspects of fungal PKS research; the review also covered the various approaches and advances used previously for isolating PKS genes. The purpose of this introduction is to provide an overview of the relative success of the various approaches and to consider their applicability for lichens before proposing a method to be used in this thesis. Recent and past literature organized into groups that employed similar strategies are summarized in Table 4.1. With the extensive literature on fungal PKS now available and the fast pace of the development in the area, it is likely that the list is now incomplete.

Due to the significant differences between fungal iterative type I PKSs and bacterial modular type I PKSs, heterologous probing of fungal PKS genes with bacterial PKS genes was unsuccessful (Schumann & Hertweck 2006). Several methods were developed to circumvent this problem. One of the first PKS genes known, 6-methylsalicylic acid synthase (MSAS), was isolated by screening the DNA expression library of *Penicillium patulum* using antibody raised against the purified MSAS (Beck et al. 1990) (Table 4.1, approach 2). At about the same time, the *w*A gene responsible for green conidial pigmentation was isolated from *Aspergillus nidulans* using a mutant complementation technique with a cosmid that could be rescued from the mutants (Mayorga & Timberlake 1990) (Table 4.1, approach 3). For screening of the cDNA/DNA library with polyclonal antibody, cross-reactivity with other proteins besides the candidate PKS could occur, such as the ubiquitous ubiquitin homologs observed by Hendrickson et al. (1999). In the case of mutant complementation, it is a disadvantage if the restored phenotype is not easily observable. For example, in the search for the lovastatin PKS, 6000 mutants were screened for restored lovastatin production by HPLC/TLC after transformation of *A. terreus* with a cosmid library (Hendrickson et al. 1999).
Table 4.1. The various approaches used previously in locating and cloning of fungal PKS genes.

<table>
<thead>
<tr>
<th>Approaches</th>
<th>PKS genes (product/phenotype)</th>
<th>References</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Hybridization with bacterial heterologous PKS probes</td>
<td>-</td>
<td>(Schumann &amp; Hertweck 2006)</td>
<td>Fungal PKS is too different from the modular bacterial PKS – unsuccessful.</td>
</tr>
<tr>
<td>2) Antibody screening of DNA/cDNA expression library</td>
<td>- <em>Penicillium patulum 6msas</em> [6-methylsalicylic acid (6-MSA)] - <em>Aspergillus terreus lovB/LNKS</em> (lovastatin nonaketide)</td>
<td>(Beck et al. 1990) (Hendrickson et al. 1999)</td>
<td>Polyclonal antibody is raised against the purified PKS protein. Cross reactions were observed e.g. ubiquitin (Hendrickson et al. 1999).</td>
</tr>
<tr>
<td>3) Mutant complementation with cosmid library</td>
<td>- <em>A. nidulans wa</em> (melanin) - <em>A. fumigatus pksP</em> (melanin) - <em>A. terreus lovB/LNKS</em> - <em>Nectria haematococca pksN</em> (red pigment)</td>
<td>(Mayorga &amp; Timberlake 1990) (Langfelder et al. 1998) (Hendrickson et al. 1999) (Graziani et al. 2004)</td>
<td>Useful for colour compounds (pigments) or other obvious phenotypes. In locating the LNKS gene, 6000 clones were assayed by TLC/HPLC.</td>
</tr>
<tr>
<td>4) Hybridization with heterologous fungal PKS probe</td>
<td>- <em>A. terreus atx</em> (6-MSA) - <em>A. parasiticus pksP1</em> (6-MSA)</td>
<td>(Fujii et al. 1996) (Feng &amp; Leonard 1998a)</td>
<td>Only 6msas homologs were isolated when 6msas was used as probe.</td>
</tr>
<tr>
<td>5) Identification of gene locus via restriction enzyme-mediated integration (REMI) mutagenesis.</td>
<td>- <em>Cochliobolus heterostrophus pks1</em> (T-toxin) - <em>Cercospora nicotianae ctb1</em> (cercosporin)</td>
<td>(Lu et al. 1994; Yang et al. 1996) (Choquer et al. 2005; Chung et al. 2003)</td>
<td>Non-producing mutants were selected by loss of antimicrobial activity (T-toxin) or red pigmentation (cercosporin). In both cases, plasmid including the flanking regions was rescued by digestion, religation and transferring into <em>E. coli</em>.</td>
</tr>
<tr>
<td>6) Differential display Reverse Transcriptase-PCR (DDRT-PCR) of organism under production and non-production conditions</td>
<td>- <em>Giberella fujikuroi pks4</em> (bikaverin) - <em>Penicillium nordicum otapksPN</em> (ochratoxin A)</td>
<td>(Linnemannstons et al. 2002) (Farber &amp; Geisen 2004; Karolewicz &amp; Geisen 2005)</td>
<td>Differential polyketide production was achieved by varying the nitrogen and/or carbon source in the growth media. Differentially expressed gene fragments were visualized and extracted from gel.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td><strong>7)</strong> Suppression Subtractive Hybridization-PCR (SSH-PCR) of organism under production and non-production conditions</td>
<td><em>A. ochraceus pks</em> (ochratoxin A)</td>
<td>(O’Callaghan et al. 2003)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subtractive hybridization removed the genes expressed in both production and non-production conditions and hence enriched the amplification of differentially expressed genes.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| **8)** Hybridization with homologous probe generated from degenerated primers based on conserved motifs | - *Gibberella fujikuroi fum5* - PKS genes in *Phoma* sp.  
- *Monascus purpureus pksCT* - *Phoma* sp. *pks1* (SQTKS) - PKS genes in insect- and nematode-associated fungi - PKS genes of *Vaccinium macrocarpon* endophytes and *Glarea lozoyensis pks1*. - *P. citrinum mlcA* | (Keller et al. 1995)  
(Bingle et al. 1999; Nicholson et al. 2001)  
(Shimizu et al. 2005)  
(Cox et al. 2004)  
(Lee et al. 2001)  
(Sauer et al. 2002; Zhang et al. 2003)  
(Wagschal et al. 1996) |
| After locating and cloning the genes, functional analysis was normally followed later by either targeted gene disruption and/or heterologous expression. |   |
| **9)** Whole genome sequencing (WGS) | - PKS genes of pathogenic and saprobic ascomycetes  
- *Gibberella* (*Fusarium*) spp. and zearalenone-related PKSs  
- *C. heterostrophus* T-toxin PKSs  
- *A. nidulans* hybrid PKS-NRPSs | (Kroken et al. 2003)  
(Gaffoor et al. 2005; Gaffoor & Trail 2006; Kim et al. 2005; Lysoe et al. 2006; Malz et al. 2005; Voigt et al. 2007)  
(Baker et al. 2006)  
(Bergmann et al. 2007) |
| PKS genes were retrieved from fungal WGS database using highly conserved regions in PKSs and FASs, (e.g. KS domain) as a query for BLAST searches. Identified ORFs of high similarity to PKS genes could be amplified directly from the genome for subsequent functional analyses. |   |
A useful technique for locating a PKS associated with a particular phenotype is via tagged mutagenesis using restriction enzyme-mediated integration (REMI) pioneered by Lu et al. (1994) in locating T-toxin PKS (Table 4.1, approach 5). In this method, the fungal strain was transformed with a linearized, nonhomologous plasmid carrying a selective marker along with an excess of the restriction enzyme used to linearize the plasmid. Fragments of the corresponding PKS gene could be recovered from those mutants with lost phenotype by using the plasmid as homologous probe. The other approaches took advantage of the differential expression of genes under production and non-production conditions (Table 4.1, approaches 6 & 7). Like the approaches (2) and (3), these strategies are very useful especially when no prior knowledge is available about the PKS gene, but they all rely on easily observed phenotypes (e.g. pigmentation) or a simple assay that can be performed on large number of mutants simultaneously (e.g. antimicrobial assay used for detecting T-toxin production).

Alternatively, probing a genomic library with fungal PKS genes cloned by other means should theoretically able to recover novel PKS genes by lowering the hybridization stringency (Table 4.1, approach 4). The P. patulum 6MSAS gene has been attempted as a heterologous probe to recover PKS genes from various fungi. However, the probe could only hybridize to 6MSAS homologs (Feng & Leonard 1998b; Fujii et al. 1996), indicating that other types of PKS genes are quite different from 6MSAS. The need to find useful probes that can localize various PKS genes in fungi has led to the development of degenerate primers based on conserved motifs (Table 4.1, approach 7). The earliest employment of degenerate primers in locating fungal PKS was in 1995 (Keller et al. 1995). The degenerate primers (KS1 and KS2) were based on the KS domain of fungal and bacterial type I PKS and were used to locate the aflatoxin gene cluster (Keller et al. 1995); they later facilitated the cloning of the fumonisin PKS gene (fum5) (Proctor et al. 1999).

The subsequent primer sets developed to target KS domains of a specific class of PKSs are especially useful (Bingle et al. 1999; Nicholson et al. 2001). The degenerate primer approach has been used as an effective means to estimate the genetic diversity of polyketide biosynthesis in an organism (as reviewed in Chapter 3, section 4.1) and also facilitated the cloning of several PKS genes (Table 4.1,
approach 7). The advantage of this approach is that the short KS domain sequences obtained can be analyzed for replication, and for discrimination or selection of desired traits (based on clustering or phylogenetic analysis) before proceeding with cloning the full-length gene or the associated gene cluster. This was one of the main objectives of the experiments in Chapter 3. However, as discussed in Chapter 3 Section 3.6, PKS cloning based on the degenerate primer approach has its own shortcomings. PKS genes responsible for production of a particular compound are searched based on a preconceived paradigm (or the limited knowledge on fungal PKSs). When an unexpected gene is involved in the biosynthesis of the target compound, it can be missed. This limitation to the degenerate primer PCR approach and other aspects discussed in the previous chapter (Section 3.6) may be resolved when more PKS genes are characterized. By contrast, other approaches (Table 4.1 approach 3, 5, 6, 7) that are blind to the identity of DNA samples but rely on detecting the genetic basis of phenotypical differences are less biased, but more time-consuming and relatively inefficient.

As sequencing technology becomes cheaper and more efficient, fungal whole genome sequencing (WGS) projects are increasing at a fast pace, with many medically, pharmaceutically, industrially and economically important fungal strains now being sequenced or in the process. Studies such as these of Kroken et al. (2003) and Varga et al. (2003) that performed genome-wide annotations of PKS genes from various fungi have revealed numerous fungal PKS gene sequences, and have allowed simultaneous phylogenomic comparisons of several fungi, which have in turn provided insights to the evolution of PKS genes. The fungal WGS projects also accelerated many fungal PKS research programs in which the candidate full-length PKS gene could be located and amplified from the organism directly, based on the sequences available from the WGS database, for further characterization, bypassing the time-consuming library screening and cloning (Table 4.1 approach 8). Nonetheless, the application is still limited only to several fungal genomes where sequencing and annotations have been completed (or are at the draft assembly stage). For non-standard/ non-model organisms, degenerate primers still provide one of the most convenient and less time-/ cost-demanding approaches for locating PKS genes.
For unculturable or difficult-to-culture and slow-growing fungi (such as those in lichens), mutant complementation and other approaches relying on mutagenesis or differential expression are not feasible. The slow growth and lack of understanding of the culture conditions are the main limitation. For the same reason, development of transformation systems for these fungi is relatively cumbersome. Therefore, generation of short homologous probes by degenerate primer PCR seems to be the best approach to locate the desired PKS gene.

4.1.1 Rationale and research goal

The previous chapter demonstrated the successful development of new degenerate primers and expanded the diversity of PKS genes that could be detected in lichens. The study showed that uncharacterized PKS genes from the WGS projects could be used to design clade-specific degenerate primers and to predict the domain architecture, based on the phylogenetic relationship of KS domains. In this chapter, the study will focus on locating and cloning of candidate PKS genes from *Flavoparmelia caperata* and *Xanthoparmelia semiviridis* using the KS domain fragments obtained from previous chapter as homologous probes.

Isolation of whole PKS genes from environmental lichen samples possesses additional challenges, as the DNA is metagenomic in nature. The genome size of lichen-fungi and its algal component is unknown but the average ascomycetous fungal genome contains 25-35 Mbp, while the unicellular algae genomes range from 12.6 Mbp (*Ostreococcus tauri*) to 121 Mbp (*Chlamydomonas reinhardtii*). To have a genomic library that include both organisms and have 99% coverage, a very large number of colonies will have to be screened. There would also be DNA from bacteria potentially associated with the environmental lichen thalli. Although the library would consist of mostly DNA from the mycobiont, the number of colonies that have to be screened in order to obtain a clone carrying the candidate gene is likely to be larger than screening a genomic library from pure fungal culture. To circumvent this problem, a Southern blotting-guided approach was used to create a subgenomic library or a “genomic file” based on the band size hybridized with selected KS domain probes.
As discussed in Chapter 3 Section 3.5.3.4, the lichen KS domains that grouped together with clade III NR-PKSs in the phylogenetic analysis are of particular interest. The presence of CMeT domain at the C-terminal of the NR-PKSs in this clade indicates that they are most likely to be responsible for the biosynthesis of β-orsellinic acid and methylpholoroacetophenone – the two common precursors for many unique bioactive lichen compounds, including the β-orcinol depsides/ depsidones and usnic acids. This makes the lichen PKSs of this clade an attractive target for cloning and further characterization. For this reason, the subsequent cloning experiments will focus on the clade III NR-PKSs in *F. caperatulla* and *X. semiviridis*.

### 4.2 Locating and cloning of partial PKS genes in *Flavoparmelia caperatulla*

*F. caperatulla* was chosen mainly because of its wide availability. The lichen is an endemic species throughout Victoria. It is also relatively tolerant to urban pollution and is found on trees in many older neighbourhoods. A tree with a few large patches of *F. caperatulla* (diameter > 20 cm) growing on its bark was located opposite Mount Waverley primary school. The homogeneity of the specimen (samples can be collected from a single large patch) and the location (ease of repeated collections) of the specimens were additional advantages. *F. caperatulla* is from the family Parmeliaceae – one of the largest families of lichens and known to produce diverse metabolites. The type of KS domain PCR products obtained from this lichen (Section 3.5.2) and its metabolic profile (Section 3.2.2, Table 3.1) were similar to another two interesting lichens (*U. oncodeiodes* and *X. semiviridis*) of the same family. All three lichens contain usnic acid, as well as several biosynthetically related β-orcinol depsidones. This lichen was thus a good starting point for the initial experiments to test the strategy of locating and cloning lichen PKS genes.

Two *F. caperatulla* KS domain PCR products (NRKS1 and NR3KS1), as described in Chapter 3, were selected as probes for Southern blotting. The NRKS1 was chosen due to its grouping with other lichen KS domains (*U. oncodeiodes* and *X.
*semiviridis* that shared similar metabolic profiles and the possibility that this group of PKSs would produce orsellinic acid due to their similarity to *Gibberella zeae* PKS13 (the NR-PKS in zearalenone biosynthesis, see Section 3.5.3.3). NR3KS1 was chosen based on the interest in the potential presence of a methyltransferase domain in the PKS gene, as it grouped with the clade III NR-PKSs (see Section 3.4.3.4).

### 4.2.1 Materials and methods

#### 4.2.1.1 Southern blotting

Both purified PCR products (NRKS1 and NR3KS1) were digoxygenin (DIG)-labelled and used as probes for Southern blotting, using methods as described in Section 2.2.12, and the resulting probes will be referred as FcNRKS1 and FcNR3KS1 respectively. High molecular weight DNA of *F. caperatulla* was extracted with the method described in Section 2.2.2 and digested separately with three common six-base restriction endonucleases (*Bam*HI, *Eco*RI and *Xho*I), which did not cut either FcNRKS1 or FcNR3KS1. More than one band of different size should be observed if there are similar or duplicated genes that hybridize with the probes in the organism. About 20 µg of genomic DNA was added for each 100 µL digestion reaction to 20 U of enzyme and incubated overnight at 37°C. After gel electrophoresis and transfer of the DNA to a positively-charged nylon membrane, the membrane was hybridized with the FcNRKS1 probe followed by detection. The same membrane was stripped and rehybridized with the FcNR3KS1 probe.

#### 4.2.1.2 Construction of partial genomic library

*F. caperatulla* DNA digested with *Eco*RI was resolved on a 0.7% agarose gel under the same conditions as for Southern blotting and DNA with a size corresponding to the signal (5 kb) on the Southern blot was excised from the gel and purified (see Results in 4.2.2.2-3). Purified pUC19 plasmid digested with the *Eco*RI (dephosphorylated and not dephosphorylated) was ligated with the size-fractionated DNA and transferred into electrocompetent *E. coli* (SureCells, Stratagene), using the method described in Section 2.2.8, to generate a subgenomic library. Transformed clones carrying the inserts were detected as white colonies as opposed to blue on LB agar medium.
4.2.1.3 Screening of partial genomic library

The partial genomic library was screened by a systematic PCR-based assay in a 96-well format adapted from Kim et al. (2003). Essentially, 12 colonies from each row grown in a 96-well microtitre plates were pooled together for plasmid DNA extraction. The pooled plasmid miniprep DNA from each group was then screened by PCR using primers corresponding to the KS probe used in Southern blotting (NR3KS-F/R degenerate primers). Individual clones from the positive plasmid pools were rescreened to locate the positive clones. A non-degenerate primer specific to the internal sequence of the FcNR3KS1 probe (FcNR3KS1-iF) and pUC universal primers were used to screen the 12 groups of pooled plasmids. Since the orientation of the insert was unknown, both forward and reverse M13 (pUC) universal primers were used consecutively with FcNR3KS1-iF. Positive clones were sequenced using pUC universal primers and FcNR3KS1-iF to determine which of them was carrying the *F. caperatulla* NR3KS1.

### 4.2.2 Results

#### 4.2.2.1 Southern blotting

Using the DNA extraction method described in Section 2.2.2, DNA of high molecular weight >21 kb (e.g. Figure 4.1, lane 5 uncut, *F. caperatulla*) was successfully obtained in sufficient quantity (>100 µg from 4 g of lichen thalli) and purity, which allowed complete digestion and detection of discrete bands in Southern blotting.

Southern blotting with the FcNRKS1 probe yielded the same band number and sizes when hybridized at both high (65°C) and lower (60°C) stringency wash conditions (Figure 4.1). The number of bands ranged from 2-5 in each lane with the three different restriction enzymes. Two bands were observed for DNA digested with *Eco*RI, but more than two bands with *Bam*HI and *Xho*I. Conversely, Southern blotting with FcNR3KS1 probe using the same DNA-bound nylon membrane (stripped and re-hybridized) showed a single band with each enzyme (Figure 4.2).
Figure 4.1. Southern blotting using FcNRKS1 probe. Left, digested (BamHI, EcoRI, HindIII) and undigested (lane unct) genomic DNA electrophoresed on a 0.7% agarose gel along with the FcNRKS1 PCR product (lane PCR) as positive control. Right, chemiluminescent detection of the corresponding blot.

Figure 4.2. Southern blotting using FcNR3KS1 probe. Left, same as Figure 4.1 (left). Right, chemiluminescent detection of the corresponding blot (stripped and reprobed with FcNR3KS1 probe).
4.2.2.2 Construction and screening of partial genomic library

A subgenomic library was constructed based on the 5 kb band signal detected in the EcoRI-digested DNA hybridized with NR3KS1 probe, as it was of the ideal size to be cloned into the commonly used pUC plasmids. A total of 144 white colonies was screened with the NR3KS-F and -R degenerate primers in the first round. However, as shown in Figure 4.3A, multiple bands were obtained when the pooled plasmids (group IG, IH and IIB) were amplified with the NR3KS degenerate primers. Multiple bands were also observed when 12 individual clones from the group IG were screened individually (Figure 4.3A), and none of the bands were of the same size as the NR3KS PCR product of *F. caperatulla* DNA. To circumvent this problem, a non-degenerate internal primer specific to FcNR3KS1 (FcNR3KS1-iF, Figure 4.4) was used in conjunction with M13 (pUC) universal primers to screen the 12 groups of pooled plasmids. With the FcNR3KS1-iF/M13 reverse primer combination, a single discrete band was amplified from three clones – IG4, IIA4 and IIB7. However, the three clones produced bands of different sizes (1.1, 1.6 and 1.3 kb respectively). The miniprep plasmids of the three clones were digested with EcoRI to check for their insert size and all had a 5 kb insert.

None of the three clones contained inserts that matched a polyketide synthase gene on GenBank. Sequences from IIA4 (forward and reverse) did not match any sequences in the database when queried using BlastX in NCBI, except for a few fungal hypothetical protein sequences with bit scores <40. Sequencing of IG4 with pUC reverse primer resulted in a closest match with a putative histidine kinase, HHK9p, from *Cochliobolus heterostrophus* (Genbank accession AAR29888) [Identity = 87/250 (34%), Positives = 140/250 (56%)], while sequencing with FcNR3KS1-iF primer did not obtain any significant matches (bit scores <40). For IIB7, a sequence obtained with FcNR3KS1-iF matched an odorant receptor protein, 22a, from *Drosophila mauritiana* (Genbank accession ABP57130) [Identity = 94/95 (98%), Positives = 95/95 (100%)].
Figure 4.3. PCR screening of the F. caperatulla subgenomic library with NR3KS-F/R primers: A) PCR products from pooled plasmids of 12 colonies each lane (IA-IH, IIA-IIID), B) PCR products from the individual 12 colonies from group G (lane 1-12).

Figure 4.4. The approximate location of FcNR3KS1-iF primer binding site on the F. caperatulla NR3KS1 KS domain.
4.2.3 Discussion

4.2.3.1 High molecular weight lichen DNA for Southern and cloning

When using the Qiagen DNeasy plant mini extraction kit (see Section 2.2.1), lichen DNA extracts showed mild shearing and generally had low molecular weights (<25 kb). As mentioned in the manual/handbook supplied by the manufacturer (Qiagen), the integrity of the DNA extract obtained from DNeasy kit is more suitable for PCR analysis than for Southern blotting. The quantity and concentration obtained by this means was also not sufficient for Southern blotting purposes. Inhibition of enzymatic digestion was observed with some of the DNA extracts, possibly due to the high content of polysaccharides and phenolic substances in the lichen thalli that inhibit enzyme activity.

To scale up the DNA extraction and remove the impurities, several other previously published methods (besides the one described in Materials and Methods) were attempted, such as CTAB, double phenol-chloroform extraction, dialysis tubing etc. (Armaleo & Clerc 1995; Cubero et al. 1999; Sambrook & Russell 2001). All these methods either led to significantly low yield, or to incomplete removal of the impurities in the DNA extract, in which coprecipitation of the polysaccharides was observed during the final alcohol precipitation step of most protocols (formation of gel blob) or sometimes a yellow to brownish tinge (possibly the phenolic substances) was visible in the final DNA extract.

The DNA extraction method in Section 2.2.2 (Csaikl et al. 1998; Gauch et al. 1998), which used an ion-exchange resin column (Genomic-tip) from Qiagen, yielded DNA of high molecular weight (50-120 kb). The high integrity of the DNA was due to the lack of external physical force (such as centrifugation) throughout the extraction procedure; DNA was allowed to bind, wash and elute from the column by gravitational flow. Although the exact size of the DNA was not determined on pulse-field gel electrophoresis (PFGE), it migrated slower than a 21 kb DNA band on agarose gel electrophoresis (e.g. Figure 4.1, lane 5 uncut, F. caperatulla) and a large amount of DNA trapped in the well (at the wall close to the cathode) was also observed when visualizing the agarose gel under UV. This was presumably due to the large sizes of the DNA, which might be overcome by lowering the agarose concentration and electrophoresis voltage or by using PFGE. Nevertheless, through
this method, a sufficient quantity of DNA (>250 µg) was obtained from a single extraction (from 4 g of lichen thalli) with sufficiently high purity that it allowed complete digestion and detection of discrete bands in Southern blotting (Figure 4.1 and 4.2).

4.2.3.2 Southern-guided library construction and PCR screening

The Southern blot result with FcNRKS1 probe suggests that there are at least two copies of similar PKS genes in *F. caperatulla*. For the FcNR3KS1 probe, only a single copy of the PKS gene was detected. Superimposing the signals of one blots over the other showed no overlapping band, except for a 3 kb band in the *BamHI* lane. This suggests that the two probes did not hybridize to the same PKS gene.

Since hybridization with FcNR3KS1 probe only showed a single copy gene in the lichen DNA, the PKS gene corresponding to FcNR3KS1 was chosen as the initial cloning target. Screening of the pooled plasmid and individual colonies from the *F. caperatulla* subgenomic library with the degenerate primers NR3KS-F/R produced multiple bands and so it may have amplified non-specifically against a background with pUC19 plasmid and possibly bacterial genomic DNA carried over from plasmid miniprep. Therefore, degenerate primers might not be a good choice for PCR library screening. Three “positive” clones, which yielded single PCR product, were identified using the FcNR3KS1-iF primer specific to the FcNR3KS1 probe. However, sequences from all three clones did not match any PKS gene in GenBank database.

Although Southern blotting yielded a discrete single band and indicated the presence of FcNR3KS1 in the 5 kb region of *EcoRI*-digested DNA, the results suggested that the approach appeared to have failed. If more colonies were screened, perhaps the right clone could be identified. When re-examined closely, weaknesses could be found in the strategy, possibly leading to the failure to clone the 5 kb fragment containing FcNR3KS1:

1. Cloning with single-digested DNA led to low cloning/ transformation efficiency
2. Screening the library with degenerate primers resulted in non-specific amplification, hence false positives.

3. When screening the library with FcNR3KS1-iF specific primer and pUC universal primer, the amplicon size for the positive clone was unknown and therefore could not be distinguished from the false positives.

Keeping these in mind, modification of the strategy to use double-digested DNA for library construction and employment of a specific primer pair that would amplify the KS domain fragment (of known product size) for library screening might improve the success rate. The use of forward and reverse primers specific to the probe but avoiding binding sites corresponding to highly conserved regions in the KS domain should also increase the specificity for PCR screening. While the experiments for *F. caperatulla* were going on, the plan for cloning *X. semiviridis* PKS genes had started. The knowledge gained from Southern blotting and DNA library construction of *F. caperatulla* was implemented on this second lichen, as shown in the following section.
4.3 Locating and cloning of partial PKS genes in *Xanthoparmelia semiviridis*

*X. semiviridis* previously belonged to the genus *Chondropsis*, but is now in the genus *Xanthoparmelia*, as described in a recent reclassification (Blanco et al. 2004). This species is found mostly in the cooler and drier parts of Southern Australia and has also been reported in New Zealand (Galloway 1980; Rogers 1971). The unusual feature of *X. semiviridis* is that it has no attachment to its substrate. The thalli lie flat on the ground when it is moist but curl up and move along with the wind when the weather is dry. This feature makes this species particularly suitable for this study, as DNA from other sources is minimized. In addition, the lichen was found in large quantities on the ground under *Allocasuarina luehmannii* (buloke) trees at the collection site. The specimens used in this study were collected again from the same location as described previously in Wyperfeld National Park (Chapter 3 Section 3.2.1.1) to obtain a larger quantity sufficient for gene cloning.

For similar reasons to choosing *F. caperatulla*, *X. semiviridis* was chosen from the family Parmeliaceae, which represents the majority of the lichens in the collection, and it shared a similar metabolic profile and KS domain PCR products with *Usnea oncodeiodes* and *F. caperatulla*. Like the other two lichens, it contains usnic acid, as well as several biosynthetically related β-orcinol depsidones. In addition, two copies of clade III NR-type KS domains were found in this lichen. As mentioned earlier, the corresponding clade III NR-PKSs were of particular interest due to their potential to be involved in the biosynthesis of important polyketide precursors in lichens.

### 4.3.1 Materials and methods

The PKS gene cloning strategy employed for *X. semiviridis* was essentially same as that used for *F. caperatulla* but with modifications to overcome the problems faced previously (Section 4.2.3.2).
4.3.1.1 Southern blotting

Both probes (XsNR3KS1 and XsNR3KS2) selected from *X. semiviridis* for testing were from clade III NR-PKS. Due to the cloning difficulties experienced in the previous cloning experiment (Section 4.2), double-digested lichen DNA was used for Southern blotting in order to generate asymmetrical sticky-ends for cloning into pUC19 and hence prevent religation of the plasmid and inserts, eliminating the need for dephosphorylation. Genomic DNA digested with single enzymes was also loaded into the gel for Southern detection. This helped to determine if the bands detected were double- or single-digested fragments. Five of the six enzymes selected (EcoRI, HindIII, BglII, XhoI and BamHI) either do not cut the probes at all or otherwise cut near 5’ or 3’ end of the probes. PstI cut in the middle of NR3KS1. Therefore, only one band was expected in the blot if the individual probe hybridized to a single-copy gene, except for PstI-digested DNA. Due to the increased enzyme combinations, only ~10 µg was loaded into each well to conserve the limited genomic DNA.

4.3.1.2 Construction of *X. semiviridis* partial genomic library

*X. semiviridis* DNA was resolved on 0.7% agarose gel under the same conditions for Southern blotting as before and DNA fragments with size 3.8-4 kb, corresponding to the signal on the Southern blot (see Results 4.3.2.2), were excised from the gel. Purified pUC19 plasmid digested with the appropriate restriction enzymes (BamHI and HindIII) was ligated with the size-fractionated DNA and transformed with electrocompetent *E. coli* (SureCells, Stratagene) to generate a subgenomic library.

4.3.1.3 Screening of partial genomic library

PCR screening of the partial genomic library was essentially same as in 4.2.1.4, except that non-degenerate forward and reverse primers (XsNR3KS2-F/R) specific to the probe (XsNR3KS2) were used for amplification. *X. semiviridis* DNA was used as a positive control in the PCR to estimate the size of the PCR product.
4.3.2 Results

4.3.2.1 Southern blotting

Southern blotting with the XsNR3KS1 probe showed two distinct bands for EcoRI and HindIII single-digested DNA (Figure 4.5B), while signals that resembled two slightly overlapping bands were observed for BglII and Xhol single-digested DNA. In the BamHI-HindIII and BglII-Xhol double-digested DNA, the bands were smaller than 1 kb. No band appeared in the lane with EcoRI-PstI double-digested DNA. By contrast, hybridization with XsNR3KS2 probe showed a single distinct band for all lanes, except for lanes with Xhol, BglII-Xhol and EcoRI-PstI digested DNA. A 3.8 kb band was detected in the BamHI-HindIII double-digested genomic DNA. The size of the band was smaller than those for BamHI and HindIII single-digested genomic DNA, as expected for a BamHI-HindIII double-digested fragment.

4.3.2.2 Construction and screening of partial genomic library

Based on the detected 3.8 kb target BamHI-HindIII double-digested fragment, a subgenomic library was constructed. A positive clone (no. 52) was detected after screening of 60 white colonies. The positive clone, pUC52KS, was first sequenced with the M13 universal primers and XsNR3KS2-F/R specific primers. The pUC52KS insert was confirmed to contain the XsNR3KS2 fragment. Based on an EcoRI site found upstream of XsNR3KS2, the BamHI-EcoRI (~2.2 kb) fragment was subcloned into pUC19 for further sequencing using M13 universal primers (Figure 4.6). A gap in the sequence was closed by primer walking using the primer GAP1-F.
Figure 4.5. Southern blots of *X. semiviridis* DNA hybridized with XsNR3KS1 and XsNR3KS2 probes. A) Digested genomic DNA electrophoresed on a 0.7% agarose gel, B) blot probed with XsNR3KS1, C-D) blot probed with XsNR3KS2, band targeted for cloning/library constriction is circled. Ba, *Bam*HI; Bg, *Bgl*II; Ec, *Eco*RI; H, *Hind*III; P, *Pst*I; X, *Xho*I; BaH, *Bam*HI+*Hind*III, BgX, *Bgl*II+*Xho*I; EcP, *Eco*RI+*Pst*I.
Figure 4.6. The plasmid map of pUC52KS containing a partial PKS gene (xsepks1) from X. semiviridis. The small arrows indicate the approximate binding sites for the specific primers XsNR3KS2-F/R. Bent arrows indicate the putative translation start sites and transcriptional directions.

4.3.2.3 Bioinformatic analysis

In BLASTx results, a 2264 bp region of the 3.8 kb insert showed a high similarity to the 5′ region of PKS genes (Figure 4.6, 5′xsepks1). The two closest BLASTx matches were Botryotinia fuckeliana PKS17 (AAR90253) and Monascus purpureus PKSCT (BAD44749) followed by other uncharacterized non-reduced type PKS genes (mostly clade III NR-PKSs). This 2264 bp region was tentatively assigned as the 5′ fragment of a PKS gene (xsepks1) in X. semiviridis. An alignment gap corresponding to a 21 amino acid region (interrupted by a stop codon) in the 5′ xsepks1 was consistently found in the BLASTx pairwise alignments with other PKS protein sequences. Tentatively, the corresponding region (63 bp) flanked with a typical “GT...AG” motif was assigned as a putative intron (intron1) in the 5′ xsepks1. A 438 bp region upstream of the putative 5′ PKS gene showed significant amino acid homology (56%, 82/146) with oxidoreductase CtnB (BAE95339) in the M. purpureus citrinin biosynthetic gene cluster. The conserved domain search (cds) result in NCBI showed that both CtnB and the partial homolog of CtnB in pUC52KS were related to a cluster of orthologous genes designated as COG0400 – predicted esterase. The partial 5′ ORF homologous to ctnB transcribes in the opposite direction to xsepks1 (Figure 4.6) and was tentatively designated as xseB.
The region between the two partial genes (\textit{xseB} and \textit{xsepks1}) transcribing in opposite directions was analysed for potential promoter elements that co-regulate the two genes (Figure 4.7), based on the regulatory sites detected in the intergenic region of \textit{nor1-pksA}, which is involved in \textit{A. parasiticus} aflatoxin biosynthesis (Ehrlich et al. 2002). Three putative TATA boxes were found in the \textit{xseB-xsepks1} intergenic region. One of the TATA boxes was located at 89 bp upstream of the putative translational start of \textit{xsepks1}, while the other two TATA boxes were located at 69 bp and 182 bp upstream of the putative translational start of \textit{xseB} on the reverse strand. A consensus motif (5´SYGGGG-3´), typical of binding sites for CreA, a globally acting transcription factor normally involved in carbon catabolite repression (Gonzalez et al. 1997; Panozzo et al. 1997), was found in the middle of the intergenic region. Four potential AreA binding sites (5´-HGATAR-3´), which regulate genes involved in nitrogen metabolism (Burger et al. 1991; Muro-Pastor et al. 1999), were also found (two in each forward and reverse strand). The \textit{xseB-xsepks1} intergenic region was also analysed for potential binding sites for pathway-specific Zn$_2$Cys$_6$-type transcriptional activators, such as the AflR that regulates the \textit{A. nidulans/A. parasiticus} aflatoxin pathway and the yeast GAL4 transcriptional activator. Three 5´-TCGN$_x$CGA-3´ motifs corresponding to the \textit{A. nidulans} AflR binding site (Fernandes et al. 1998) and one 5´-CGGN$_x$CCG-3´ motif corresponding to the yeast GAL4 binding site (Reece & Ptashne 1993) were detected in the \textit{xseB-xsepks1} intergenic region (where N$_x$ represents a variable spacing between the inverted repeat triplets).
Figure 4.7. Potential promoter elements found in the intergenic region between xsepks1 and xseB. Numbering started from the BamHI site of the pUC52KS insert.
4.3.3 Discussion

4.3.3.1 Southern-guided PKS gene cloning

Overall, the results showed a successful approach for cloning PKS genes from *X. semiviridis*. Due to the reduced amount of DNA loaded in each well, the signals of the Southern blots (Figure 4.5) were relatively weak. Both XsNR3KS1 and XsNR3KS2 probes hybridized successfully to the genomic blot. Based on the XsNR3KS1 sequence, only one band was expected in the *Eco*RI and *Hind*III single digested genomic DNA, since there is no *Eco*RI site and *Hind*III cut at 60 bp before the 3´ end of XsNR3KS1. The detection of more than one band in each lane suggested that there are possibly two similar copies of PKS genes in *X. semiviridis*. Together with the *xsepks1* detected by the XsNR3KS2 probe, there may be at least three clade III NR-type PKS genes in *X. semiviridis*.

The number of DNA fragments cloned into the library was narrowed down by size-fractionation following the 3.8 kb band detected in Southern blots and the target fragment was obtained by screening a relatively small number of colonies (60 colonies). Using the primers specific to XsNR3KS2, a discrete band was obtained for the positive plasmid pool and the positive clone itself. This showed that the PCR screening approach was more effective than using degenerate primers or combinations of plasmid universal primers and a specific primer, as demonstrated in the previous section (Section 4.2). Alternatively, Southern or dot blotting could be used to confirm the positive clones before sequencing.

Despite the overall success of the approach, Southern blotting results with double-digested DNA often produced short positive bands, which were not favourable for cloning the full-length PKS gene. For example, hybridization with the XsNR3KS1 probe did not produce any bands larger than 2 kb (Figure 4.5B). A further problem is that the sensitivity of the DIG-chemiluminescent detection with CSPD seems to reduce proportionately with the size of the target DNA fragment. Weaker signals were observed with bands smaller than 1 kb in the DIG-labelled DNA marker and the two smallest bands (564 and 125 bp) were sometimes hardly detectable. This probably explained the missing bands in some of the lanes. (e.g. lane X, BgX and EcP in Figure 4.5 C). The missing bands in the lanes could be due to the frequent cutting sites in the region flanking XsNR3KS2, which lead to shorter DNA
fragments. This presented an obvious disadvantage for this approach, where many digestion reactions have to be attempted to find the restriction enzyme combinations that yield the ideal fragment size for cloning. This is not feasible when the amount of biomass is very limited, as a large amount of DNA is required. For cloning of multiple genes from environmental samples (or unculturable organisms), a lambda or cosmid genomic library with huge coverage would be preferred, as the corresponding library could be screened repetitively.

4.3.3.2 Bioinformatic analysis

Two partial ORFs (5´ xsepks1 and a ctnB homolog) were found in the insert of pUC52KS (Figure 4.6). The high similarity of the 5´ xsepks1 translated sequence to M. purpureus PKSCT and other clade III NR-PKSs (including BfPKS17 and ChPKS21) supported the prediction that XsePKS1 is a member of the clade III NR-PKSs and so might be expected to contain a CMeT domain near the C-terminal of the protein. The CtnB was assigned as a putative oxidoreductase by (Shimizu et al. 2007) based on its similarity (26%) to an oxidoreductase, mlcF (BAC20563), found in the Penicillium citrinum compactin biosynthetic gene cluster (Abe et al. 2002). The conserved domain search (cds) results in NCBI showed that both CtnB and the partial homolog of CtnB in pUC52KS are related to a cluster of orthologous genes designated as COG0400 - predicted esterase, while no conserved domain was detected in the mlcF oxidoreductase. This indicated that CtnB and the homolog in X. semiviridis might have a different function to the mlcF oxidoreductase, but they remain to be characterized.

The region between the two partial genes (xseB and xsepks1) transcribing in opposite directions might contain promoter elements that co-regulate the two genes. Several regulatory sites were found in the intergenic region of nor1-pksA involved in the A. parasiticus aflatoxin biosynthesis (Ehrlich et al. 2002). Similar to xsepks1 and the xseB, nor1 and pksA also transcribed divergently. Besides the TATA boxes commonly found upstream of an ORF, potential binding sites for global-acting transcriptional factors were detected in the xseB-xsepks1 intergenic region. The presence of potential binding sites for CreA and AreA transcriptional factors suggested that xsepks1 and xseB could be expressed in response to nutritional
factors in *X. semiviridis*. However, the putative recognition sites for transcription factors (BrI and BrA) involved in conidiation and other developmental processes (Adams et al. 1988; Andrianopoulos & Timberlake 1994; Chang & Timberlake 1993; Lee & Adams 1996; Sewall et al. 1990) were not found in the *xseB-xsepks1* intergenic region, implying that the products from the two genes were probably not involved in these developmental processes.

Besides the global acting regulators, pathway-specific Zn$_2$Cys$_6$-type transcriptional activator is commonly found in secondary metabolite gene clusters (Hoffmeister & Keller 2007). One of the best characterized PKS-related Zn$_2$Cys$_6$-type transcriptional activators is the AflR, which regulates the aflatoxin pathway in *A. nidulans* and *A. parasiticus* (Ehrlich et al. 1999a; Ehrlich et al. 1999b; Ehrlich et al. 2002; Fernandes et al. 1998). Transcription factors of the Zn$_2$Cys6 class (such as the well-known yeast GAL4 transcriptional factor) are positive regulators that dimerize and appear to bind to GC-rich palindromic sequences such as 5´-CGGN$_x$CGG-3´ or 5´CCGN$_x$CCG-3´ (Reece & Ptashne 1993). However, Fernandes et al. (1998) discovered that AflR binds to a slightly different palindromic sequence, 5´-TCGN$_x$CGA-3´, on the promoter regions of several genes involved in the aflatoxin pathway of *A. nidulans*. Similar AflR binding sites were also observed in the *nor1-pksA* intergenic region in *A. parasiticus* (Ehrlich et al. 2002). In the *xseB-xsepks1* intergenic region, four such motifs were found (three 5´-TCGN$_x$CGA-3´ motifs and one 5´-CGGN$_x$CCG-3´). Recently, the *M. purpureus* citrinin biosynthesis transcriptional activator CtnR/CtnA has also been characterized (Shimizu et al. 2007). The disruption of the CtnR/CtnA gene led to the reduction of *pksCT* and *orf5* (a membrane transporter) transcription, and barely detectable citrinin production. The similarity of *xsepks1* to *pksCT* and the presence of potential AflR- and GAL4-like binding sites in the promoter region suggests that the *xsepks1* transcription could also be regulated by a Zn$_2$Cys$_6$-type transcriptional activator.

Although the PKS gene fragment was extended from 750 bp to at least 2264 bp, it was far from complete. The size of known PKS genes ranges from 5322 bp (*6msas*) to 9114 bp (*lovB*) (*pksCT* is 9269 bp long). Following from this, the effort was focused on confirming the presence of the CMeT domain and extending the 5´ fragment in pUC52KS to obtain the full-length *xsepks1*. 

179
4.4 Detection of CMeT domain in NR-PKS genes

Apart from the conserved regions of the KS and AT domains, which are commonly used to design degenerate primers, the CMeT domain of HR-PKSs has also been targeted (Nicholson et al. 2001). The degenerate primers derived from the CMeT domain of \textit{A. terreus} LNKS and LDKS (CMeT1 and CMeT2c/3c) were used to probe the PKS gene involved in Fusarin C biosynthesis (Song et al. 2004) and facilitated the cloning of squalestatin tetraketide synthase (Cox et al. 2004). Clade III NR-PKSs also possess a CMeT domain, but it is normally located closer to the C-terminal after the ACP domain, as opposed to being in the middle of the PKS right after the AT domain in HR-PKSs. From preliminary multiple sequence alignment analysis, the CMeT domains of the HR-PKSs were different from those of NR-PKSs. This led to the hypothesis that it would be possible to target specifically the CMeT domains of NR-PKSs using degenerate primers.

One of the objectives of the experiments in this section was to detect the presence of a CMeT domain in \textit{X. semiviridis} and other lichens collected. Based on the KS phylogenetic analysis in Chapter 3 Section 3.5.3.4, the XsNR3KS2 (\textit{xsepks1}) and other lichen KS domains, which grouped with clade III NR-PKS, were predicted to contain a CMeT domain. The detection of CMeT domains in lichens that possess such KS domains would confirm the hypothesis. The second objective was to confirm the presence and contiguity of a CMeT domain at the C-terminal of XsePKS1 by bridging the KS and CMeT domains with PCR (domain-hopping PCR). Since the CMeT domains of NR-PKSs are found almost at the end of the gene, bridging the KS and CMeT domains would also extend the sequence greatly. In this section, a pair of NR-type CMeT specific degenerate primers was developed and tested on the lichen DNA samples used in Chapter 3. This new primer pair was then used to extend the 5’ fragment of \textit{xsepks1} obtained from the previous section (4.3) and to confirm the presence of a CMeT domain in the gene as hypothesized.
4.4.1 Materials and methods

4.4.1.1 Design of NR-type CMeT domain-specific primers

Using the same approach as described in Chapter 3 Section 3, another pair of CODEHOP primers, this time targeting the CMeT domains of NR-PKSs, was designed based on the multiple sequence alignment of the CMeT domains of known clade III NR-PKSs (M. purpureus PKSCT, and B. fuckeliana PKS17 and PKS18). To ensure the primer only amplified the CMeT domain of NR-PKSs and not HR-PKSs, the specificity of the degenerate primers was checked using the in silico PCR function in FastPCR 3.6.6 by Kalendar (www.biocenter.helsinki.fi/bi/programs/fastpcr.htm) with various fungal PKS genes (clade III NR-PKSs and HR-PKSs with CMeT domains) as templates.

4.4.1.2 PCR amplification of CMeT domain

The newly designed degenerate primers (NRMeT-F/R) were tested on X. semiviridis as well as those of other lichens – F. caperatulla (001, 008 & 023), X. parietina (003), P. jackii (006), F. soredians (007), P. chinense (012), P. reticulatum (014), P. subrudecta (015), P. cunninghamii (016), H. pulverata (021), U. oncodeoides (027 & 028) and Cladonia sp. (030), under the PCR conditions described in Section 2.2.3. PCR products from selected lichens that yielded mixed sequences were cloned and resequenced (eight clones for each PCR product).

4.4.1.3 Southern blotting with CMeT domain probe

Southern blotting was performed to check for the number of CMeT domains of NR-PKSs present in X. semiviridis and the potential use of the CMeT PCR product as a probe for locating clade III NR-PKS genes. The CMeT domain PCR products (which contained a mix of two CMeT domains, as shown in sequencing results in Section 4.4.2.2) amplified from X. semiviridis were DIG-labelled. The same membrane used in Section 4.3.1.2 was hybridized with the CMeT probe (XsNRMeT) after the previous probe (XsNR3KS2) was stripped from the membrane.
4.4.1.4 KS-CMeT domain-hopping PCR

The forward XsNR3KS2-iF primer derived from the 3′ end of the xsepkst fragment in pUC52KS and a reverse CMeT011a-R2 primer derived from a chosen CMeT domain sequence amplified from X. semividis (see results 4.4.2.2 and 4.4.2.3) were used to bridge the region between the two domains. Based on the sequences of M. purpureus PKSCT and B. fuckeliana PKS17 (which were both highly similar to the 5′ xsepkst fragment), the KS-CMeT PCR product was expected to be longer than 4 kb. Hence, a long distance PCR protocol as described in Section 2.2.3 was employed. X. semiviridis DNA with high integrity (prepared from large-scale DNA extraction, Section 4.2.1.1) was used for the long-distance PCR.

A combination of shotgun sequencing and primer walking approaches were used to sequence the KS-MT PCR product. First, the KS-MT PCR product was digested with combinations of BamHI-SalI and BamHI-PstI. Individual bands larger than 500 bp were excised from the gel, subcloned into pUC19, and sequenced with the M13/pUC universal primers. Overlapping fragments were assembled with the CAP3 sequence assembly program (http://pbil.univ-lyon1.fr/cap3.php) by Huang & Madan (1999). The small contigs were compared with M. purpureus PKSCT to estimate the order and location of the fragments. The gaps in between the small contigs were filled in by primer walking using specific primers directed at the unknown regions.

4.4.2 Results

4.4.2.1 NRMeT degenerate primers

A pair of primers (NRMeT-F and NRMeT-R) corresponding to a region conserved within the CMeT domains of NR-PKSs but biased away from the CMeT domains of reducing PKSs (Figure 4.8) was chosen from the CODEHOP output list. As shown in Figure 4.8, the selected conserved region “ILEMGAG” and “EGWWLF” were biased towards the CMeT domain of clade III NR-PKSs (in bold) and not those from HR-PKSs (i.e. LNKS, LDKS and FUM1).
NRMeT-F primer (degeneracy=16)

5’- C AGC ATC CTG GAG atg ggn gcn gg -3’
S I L E M G A G

M. purpureus PKSCT  NTGEPRIILEM:ACTGGTVKMLP
B. fuckeliana PKS17  KDGEPILILEM:ACTGGTITKMVO
B. fuckeliana PKS18  KDGEPILILEM:ACTGGTITKMVO
A. terreus LNKS     HRYQSMNILEGACTGGATKYVLA
A. terreus LDKS     HKNPRARILEIGGCTGATQLVVD
G. moniliformis FUM1 HTRPRLVEIGACTGGGAQVILE

NRMeT-F primer (degeneracy=8)

3’- cty ccc acc acC GAC AAG CTG CT -5’
E G W W L F D

M. purpureus PKSCT  WVDIFGLLEGWGLFDCCRHALQ
B. fuckeliana PKS17  WVDIFGLLEGWGLFDCCRHALQ
B. fuckeliana PKS18  WVDIFGLLEGWGLFDCCRHALQ
A. terreus LNKS     RLGFIIFLGAFAWGAGVDDGCTEP
A. terreus LDKS     DLFHFLGGLPGWWSREEFQSTP
G. moniliformis FUM1 MVNLIMGILPCGWLGAEEGRVPEP

Figure 4.8. Design and relative binding positions of the NR3MeT-F and -R primers.

4.4.2.2 PCR amplification of CMeT domains

A band of the expected size (approximately 380 bp) was amplified from the X. semiviridis DNA with the primers NRMeT-F/R (Figure 4.9 lanes 1 & 2). For the other lichen DNA samples tested, bands of the same size were obtained only in F. caperatulla (001), F. soredians (007), Punctelia subrudecta (015) and Cladonia sp. (030). No PCR product was obtained from the other lichen DNA samples tested. Direct sequencing of the gel-purified PCR product produced good quality sequences for P. subrudecta and Cladonia sp., but yielded mixed sequences for F. caperatulla, F. soredians and X. semiviridis. The preliminary BLASTx analysis of all the sequences above (including the mixed raw sequences) showed similarity to the CMeT domain of known clade III NR-PKSs.
To overcome the mixed sequences obtained, the CMeT domain PCR products from *F. caperatulla* and *X. semiviridis* were cloned and sequenced. Out of the eight colonies, only one uniform sequence was obtained for *F. caperatulla* (FcNRMeT1). FcNRMeT1 showed high similarity to *C. heterostrophus* PKS21 (62% identity), *M. purpureus* (57% identity) and *B. fuckeliana* PKS17 (56% identity). Two sequences (XsNRMeT1 and XsNRMeT2) that showed similarity to the CMeT domains of clade III NR-PKSs were identified from *X. semiviridis*. The XsNRMeT1 showed a greater amino acid similarity to *M. purpureus* PKSCT (66% identity) and *B. fuckeliana* PKS17 (68% identity) than to XsNRMeT2. The XsNRMeT2, however, showed 67% identity to a putative PKS of *Aspergillus clavatus* NRRL1 (GenBank accession no. XP_001273603), while only 51% identity to *M. purpureus* PKSCT. Hence, it was predicted that XsNRMeT1 could be contiguous and downstream of XsNR3KS2. All the putative lichen CMeT domains that were sequenced contained the conserved S-adenosylmethionine binding site in the CMeT domain (E/DXGXGXG), a glycine-rich region generally known as motif I (Fauman et al. 1999; Fujii et al. 2005; Kagan & Clarke 1995; Miller et al. 2003).
4.4.2.3 Southern blotting with CMeT domain probe

One to three bands were detected on the Southern blot (Figure 4.10A) hybridized with XsNRMeT mixed probe (contains XsNRMeT1 and XsNRMeT2). Superimposing the Southern blot signals of XsNR3KS2 on the CMeT blot showed two overlapping bands – a ~10 kb band in the BglII digested DNA, and a ~20 kb band in the EcoRI digested DNA (Figure 4.10A-B). Although a ~4 kb band was also detected in the BamHI-HindIII digested DNA with the XsNRMeT1 probe, the size was slightly larger than the band detected with XsNR3KS2 probe in Section 4.3.2.1.

![Southern Blot Diagram](image)

**Figure 4.10.** Southern blots of *X. semiviridis* hybridized with A) XsNRMeT mixed probe, and B) XsNR3KS2 probe. Grey dotted arrows indicate the bands of similar sizes detected in both blots. Ba, BamHI; Bg, BglII; Ec, EcoRI; H, HindIII; P, PstI; X, XhoI; BaH, BamHI+HindIII; BgX, BglII+XhoI; EcP, EcoRI+PstI.

4.4.2.4 Extending the PKS gene sequence by domain-hoping PCR

Forward and reverse specific non-degenerate primers (XsNR3KS2-iF and CMeT011a-R2) were designed, based on the 3´ end of the xsepks1 fragment in pUC52KS and the internal region of the XsNRMeT1 respectively. A band of the expected size (4.5 kb) was obtained and sequenced. Sequence analysis by pairwise alignment showed that the KS-MT PCR product was homologous to the *M.*
purpureus PKSCT (54 % identity). The assembled sequence (Figure 4.11) confirmed that the 5’ xsepks1 and XsNRMeT1 were contiguous. Preliminary sequence analyses using the NCBI conserved domain search (cds) and EBI Interproscan indicated that an AT and an ACP domain were present between the KS (XsNR3KS2) and CMeT (XsNRMeT1) domains of xsepks1 as expected (more detailed PKS domain analysis will be presented in the Section 4.5.2.3). Based on the sequence deduced so far, xsepks1 matched the domain architecture of clade III NR-PKSs (i.e. KS-AT-ACP-CMeT) described by Kroken et al. (2003), as predicted.

Figure 4.11. Overlapping regions of 5’xsepks1 (pUC52KS insert), KS-CMeT PCR product (4.5 kb), and XsNRMeT1 PCR product/probe.

4.4.3 Discussion

4.4.3.1 Clade III NR-type CMeT domain

This study has shown that the NRMeT primers can specifically target the CMeT domain of clade III NR-PKS genes and bias away from the CMeT domain of HR-PKS genes. Nicholson et al. (2001), Cox et al. (2004) and Song et al. (2004) have also previously demonstrated the utility of CMeT domain-specific degenerate primers to probe the HR-PKS genes. Not all lichens that produced PCR products for the clade III NR-type KS domain-specific primers (NR3KS-F/R) also yielded NR-type CMeT domain PCR products and vice versa (Table 4.2). This suggests that both the NR3KS and NRMeT primers may not be optimal for amplifying the respective domains of all clade III NR-PKS genes and that the number of lichen clade III NR-PKSs detected in Chapter 3 Section 3.4.2.2 may be an underestimation. All the
mixed sequences (multiple products) resulting from both NR3KS-F/R and NRMeT-F/R primers in Table 4.2 showed the closest matches to corresponding domains (KS and CMeT domain respectively) of clade III NR-PKSs when the raw sequences were used to query the GenBank non-redundant protein database with BLASTx. Although these mixed PCR products were not cloned and resequenced, except for the two candidate lichens, it is likely that the mixed KS and CMeT domain-like sequences represent multiple clade III NR-PKS genes in the respective lichens.

Table 4.2. Comparison of lichen KS and CMeT domains PCR results.

<table>
<thead>
<tr>
<th>Strain No.</th>
<th>Species</th>
<th>NR3KS-F/R products</th>
<th>NRMeT-F/R products</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td><em>F. caperatulla</em></td>
<td>1. FcNR3KS1</td>
<td>FcNRMeT1</td>
</tr>
<tr>
<td>006</td>
<td><em>P. jackii</em></td>
<td>1. same as PRKS</td>
<td>-</td>
</tr>
<tr>
<td>007</td>
<td><em>F. soredians</em></td>
<td>mixed sequence</td>
<td>mixed sequence</td>
</tr>
<tr>
<td>011</td>
<td><em>X. semiviridis</em></td>
<td>1. XsNR3KS1</td>
<td>1. XsNRMeT1 (CMeT011a)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. XsNR3KS2</td>
<td>2. XsNRMeT2 (CMeT011b)</td>
</tr>
<tr>
<td>012</td>
<td><em>P. chinense</em></td>
<td>1. PcNR3KS1</td>
<td>-</td>
</tr>
<tr>
<td>014</td>
<td><em>P. reticulatum</em></td>
<td>PrNR3KS1</td>
<td>-</td>
</tr>
<tr>
<td>015</td>
<td><em>P. subrudecta</em></td>
<td>Not KS domain</td>
<td>PsNRMeT1</td>
</tr>
<tr>
<td>016</td>
<td><em>P. cunninghamii</em></td>
<td>mixed sequence</td>
<td>-</td>
</tr>
<tr>
<td>021</td>
<td><em>H. pulverata</em></td>
<td>1. mixed sequence</td>
<td>-</td>
</tr>
<tr>
<td>027</td>
<td><em>U. oncodeoides</em></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>028</td>
<td><em>U. oncodeoides</em></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>030</td>
<td>Cladonia (mixed)</td>
<td>-</td>
<td>CNRMeT1</td>
</tr>
</tbody>
</table>

All the lichens that had either the clade III NR-type KS domain and/or CMeT domain of clade III NR-PKSs produced β-orsellinic acid derivatives and/or usnic acid (derived from methylphloroacetophenone). These results further strengthen the hypothesis that clade III NR-PKSs could be involved in producing these compounds (Chapter 3 Section 3.5.3.4). Based on the hypothesis, it was as expected that neither NR3KS nor NRMeT primers produced any product from *X. parietina*, as it does not produce any compound derived from β-orsellinic acid or methylphloroacetophenone. Further optimizing the degenerate primers and PCR conditions, or Southern blotting with mixed clade III NR-type KS domain probes, would be expected to improve the detection of clade III NR-PKS gene in lichens and other fungi in general.
Figure 4.12. Comparison of the KS (left) and CMeT (right) domain phylogeny of clade III NR-PKSs. The evolutionary relationships were inferred using the Minimum Evolution (ME) method described in Section 3.5.1.2. Dotted lines linked the KS and CMeT domains that are part of the same PKS protein. Scales indicate the number of amino acid substitutions per site.
To compare the relationship within the CMeT domains and their correlation to KS domain phylogeny, a simple bootstrapped phylogenetic tree was constructed with the lichen CMeT domains and other clade III NR-type CMeT domains (HR-type CMeT domains were used as outgroup) (Figure 4.12). When compared to a subtree of the clade III NR-type KS domains (from Chapter 3 Section 3.5.2, Figure 3.9), most CMeT domains showed a similar topology to the KS domains, except for *B. fuckeliana* PKS20 (BfPKS20). This showed that the contiguity of the two domains on the same gene could be inferred by comparing the KS and CMeT domain phylogeny (as demonstrated for the XsNR3KS2 and XsNRMeT1, where long distance PCR had successfully bridged the two sequences).

4.4.3.2 *X. semiviridis* clade III NR-PKS gene (*xsepks1*)

Up to three bands were detected in the Southern blot of *X. semiviridis* DNA with XsNRMeT mixed probe, which contained XsNRMeT1 and XsNRMeT2. This suggests that one of the two CMeT domain probes might hybridize with a third clade III NR-PKS in *X. semiviridis*. The results matched with previous Southern blots that hybridized with XseNR3KS1 and XsNR3KS2 probe (Section 4.3.2.1) and also implied the presence of a third clade III NR-PKS genes in *X. semiviridis* that was undetected by KS- and CMeT-domain PCR.

Nicholson et al. (2001), Cox et al. (2004) and Song et al. (2004) have previously demonstrated the utility of CMeT domain-specific degenerate primers to locate and clone the HR-PKS genes. This study has shown that the NRMeT primers that target specifically the CMeT domain of NR-PKS genes could also be employed in locating and cloning clade III NR-PKS genes. As demonstrated in *xsepks1*, the CMeT domain sequence obtained could also be useful for PKS gene walking by domain-hoping PCR. This strategy is especially useful for the clade III NR-PKSs, since their CMeT domain is normally located at the C-terminal end after the ACP domain, as opposed to HR-PKSs. After assembling the sequences from pUC52KS (5´ *xsepks1*) and the KS-CMeT fragment (Figure 4.11), it was noticed that the *xsepks1* sequence was still incomplete, as the C-terminal and the stop codon at the 3´ end were missing. Therefore, the following section focused on cloning the complete *xsepks1*. 
4.5 Cloning of a full-length PKS gene from *X. semiviridis*

The KS-CMeT domain-hoping PCR greatly extended the *xsepks1* sequence from the original 5’ fragment obtained in pUC52KS. Based on its alignment to *pksCT*, the assembled sequence probably contributed 70-80% of the total length of *xsepks1*. The remaining 3’ end of the gene could theoretically be obtained via a 3’ rapid amplification of cDNA ends (3’ RACE). With the 3’ end sequence, a long-distance PCR could then be used to obtain the whole PKS gene in a single fragment. However, poor RNA quality obtained from environmental lichen samples could be a problem for cDNA synthesis, especially if the gene is a low-copy transcript.

With the high molecular weight DNA obtained from *X. semiviridis*, it was possible to construct a library with a large insert that would contain the whole XsePKS1 gene and perhaps the associated genes in the gene cluster. The advantage of cloning the whole PKS gene in a large insert is that the DNA fragment would be cloned directly from the organism and therefore no mutation be introduced by PCR (mutation in the coding sequence introduced by PCR might result in non-functional protein). The contiguous and high-fidelity PKS gene fragment would facilitate the construction of expression plasmid for heterologous expression of the gene.

To construct a library with large insert, an equally effective but cheaper alternative to the commercial lambda and cosmid/fosmid cloning system was sought. The recently developed pJAZZ-KA linear vector marketed by Lucigen as BigEasy Linear Cloning System was used for cloning of the full-length *xsepks1*, due to its ability to accommodate a large insert without losing plasmid stability and its relative ease of use (the linear plasmid could be transformed, propagated and isolated using the same methods as for other small circular plasmids). Using a similar Southern-guided approach to that described in Section 4.3.1.2, a partial genomic library with large inserts (>10 kb) was constructed and screened for the full-length *xsepks1*. 
4.5.1 Materials and methods

4.5.1.1 Southern blotting with the xsepks1 KS-CMeT probe

The 4.5 kb KS-CMeT PCR product obtained from Section 4.4.1.3 was DIG-labelled and used as a probe for Southern blotting. To find a two-enzyme combination that would yield a larger fragment and overlapped with the KS-CMeT region, an eight-base cutter (NotI) and six-base cutter (XbaI) that were not used in previous Southern blotting were included, along with EcoRI and HindIII that were used earlier (4.3.2.1).

4.5.1.2 Construction and screening of partial genomic library with large inserts

Based on the KS-CMeT Southern blotting results, another partial genomic library of larger inserts (13-14 kb) was constructed using the BigEasy Linear cloning system (Lucigen, Middleton, WI) with the NotI-digested pJAZZ-KA vector (dephosphorylated). The partial genomic library was screened by a systematic PCR-based assay in 96-well format as described in Section 4.3.1.4 with RT52KS-F and -R (specific primers derived from region flanking the putative intron of xsepks1 upstream of XsNR3KS2, see Section 4.3.2.2 Figure 4.6). Individual clones from the groups with positive results were then rescreened to locate the positive clone(s) and further confirmed by Southern blotting with XsNR3KS2 probe. The miniprep plasmid DNA of the positive clones was reconfirmed by Southern blotting using the XsNR3KS2 probe. One of the positive clones was sequenced using the primers flanking the cloning site of pJAZZ-KA (NZ-f and NZ-r, supplied by the manufacturer Lucigen) and primers derived from the known sequence (pUC52KS and KS-CMeT fragment), followed by primer walking to complete the gap. The resulting sequence from pJpks-ID5 was analysed using bioinformatic programs and methods described in Section 2.3.3.
4.5.2 Results

4.5.2.1 Southern blotting with the xsepks1 KS-CMeT probe

This Southern blot (Figure 4.13) was originally performed with the intention of mapping the restriction sites internal to and flanking the 3’ end of the 4.5 kb KS-CMeT PCR products, to seek an overlapping fragment that would include the still unknown 3’ end of xsepks1. The information obtained from this Southern blot would also be useful for the construction of the following partial genomic library with large insert. Since the pJAZZ-KA linear cloning vector was supplied as NotI predigested arms, it would be useful if genomic DNA digested with NotI would produce a fragment large enough to contain the full-length xsepks1. Based on the 12 kb EcoRI-NotI band detected in the blot and the sequence from pUC52KS, it was deduced that a NotI single digestion would yield a ~13 kb fragment that would contain the full-length xsepks1 and its downstream region (Figure 4.14).
4.5.2.2 Cloning of the full-length PKS gene

The second *X. semiviridis* partial genomic library with larger inserts (12-14 kb) was constructed based on the deduction from Southern blot analysis. The *NotI*-digested genomic DNA of corresponding size was cloned into the pJAZZ-KA vector. Four out of 16 groups of pooled plasmid DNA (ID, IIA, IIF and IIH) from two 96-well microtitre plates were positive. Individual clones (12 for each group) from groups ID and IIH were screened with the same primers and the positive clones were determined as ID5 and IIH6; the plasmid DNA extracted from the two clones was checked on agarose gel and subjected to Southern hybridization with NR3KS2 probe for confirmation. Both clones yielded a positive band of the expected size (13 kb, *NotI* cut fragment). Sequencing of the plasmid pJpks-ID5 confirmed that the clone possessed the complete *xsepks1* (Figure 4.15).

**Figure 4.14.** Overlapping region of 5´*xsepks1* in pUC52KS and KS-CMeT probe, and a *NotI* site predicted at about 12 kb downstream of the EcoRI site (2144).

**Figure 4.15.** Plasmid pJpks-ID5 isolated from a *X. semiviridis* partial genomic library. The small black arrows represent the approximate binding sites of the primers used for sequencing the pJpks-ID5. // represents region that has not been sequenced.
4.5.2.3 Sequence and PKS domain analysis

The total length of \textit{xsepks1} determined from the assembled sequence was 6558 bp, which included a 63 bp intron. The 63 bp intron was predicted at 651 bp downstream of the putative start codon. No further intron-like sequence was found in the putative coding region and a putative poly-A signal “AAATAAA” was found at 700 bp downstream of the putative \textit{xsepks1} stop codon. The translated sequence of \textit{xsepks1} (2164 amino acids) exhibited an end-to-end amino acid homology with that of \textit{B. fuckeliana} PKS17 (52.7\% identity) and \textit{M. purpureus} PKSCT (52.4\% identity), except that XsePKS1 was predicted to be shorter and to lack a C-terminal region found in both PKSs. Comparison of XsePKS1 with characterized PKSs showed that typical active-site sequences of KS, AT and ACP domains (Fujii et al. 1996; Pazoutova, et al. 1997) were well conserved in XsePKS1 (Figure 4.16A-C). XsePKS1 also contained the highly conserved S-adenosylmethionine binding site in the CMeT domain (E/DXGXGXG), a glycine-rich region generally known as motif I (Fauman et al. 1999; Fujii et al. 2005; Kagan & Clarke 1995; Miller et al. 2003) (Figure 4.17). Other motifs typical of the CMeT domains are also found in XsePKS1, except that the proposed key aromatic amino acid, tyrosine or phenylalanine (Fujii et al. 2005), seemed to be lacking in motif III of Miller et al. (2003) (Figure 4.17). Besides that, the XsePKS1 N-terminal also contained the GXCXG motif as well as the histidine and glutamine residues conserved among the starter unit-ACP transacylase (SAT) domains, as defined by Crawford et al. (2006) (Figure 4.18). Overall, the domain architecture of the PKS protein (XsePKS1) encoded by \textit{xsepks1} is consistent with that expected from a clade III NR-PKS (KS-AT-ACP-CMeT) as described by Kroken et al. (2003) with a SAT domain at the N-terminal (Figure 4.19).
Figure 4.16. Multiple sequence alignment of the catalytic cores of the minimum PKS functional domains - β-ketoacyl synthase (KS), acyltransferase (AT), and acyl carrier protein (ACP). The proposed key amino acids in the active sites are shaded.

A. β-ketoacyl synthase (KS)

<table>
<thead>
<tr>
<th>Species</th>
<th>Protein</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>X. semiviridis</td>
<td>XsePKS1</td>
<td>KISHYFGWTGPSLTIDTACSSSVAVHSAACKAILLG</td>
</tr>
<tr>
<td>M. purpureus</td>
<td>pksCT</td>
<td>KISHHFGWTGPSLTLDTACSSSVAHQACRSILSG</td>
</tr>
<tr>
<td>A. nidulans</td>
<td>wa</td>
<td>RINYYFKPSGSVSVTDACSSLAAIHLCNSIWRN</td>
</tr>
<tr>
<td>P. patulum</td>
<td>MSAS</td>
<td>RISYHNLNMGSTAVDAAASSLVAIHGHQAIRLG</td>
</tr>
<tr>
<td>G. moniliformis</td>
<td>FUM1</td>
<td>RISYEDLKGPSMTIKACCSSSLALHEAIVAIRAG</td>
</tr>
<tr>
<td>A. terreus</td>
<td>lovF</td>
<td>RVSHFYDLRGPSSIDTACSTTLALHLAIQLRAG</td>
</tr>
</tbody>
</table>

B. Acyltransferase (AT)

<table>
<thead>
<tr>
<th>Species</th>
<th>Protein</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>X. semiviridis</td>
<td>XsePKS1</td>
<td>SCAMSWIDCGIEVAAVGHSFGELTALCISGVLSLE</td>
</tr>
<tr>
<td>M. purpureus</td>
<td>pksCT</td>
<td>SCAKAWIDSLKVASVGGHSFEGIALCVSNAVSLK</td>
</tr>
<tr>
<td>A. nidulans</td>
<td>wa</td>
<td>ALSSFWASLIGTPSLGLGLGFAAMNAGVLSTS</td>
</tr>
<tr>
<td>P. patulum</td>
<td>MSAS</td>
<td>GLSALLQQSNITPQAVHSGVGEIASSVAGALSPA</td>
</tr>
<tr>
<td>G. moniliformis</td>
<td>FUM1</td>
<td>EATADNKLYIGSVKPNLGSEASGVSSVAKVSLAL</td>
</tr>
<tr>
<td>A. terreus</td>
<td>lovF</td>
<td>ALVRLTVSWNIQPVAVTIESGGEAAAAYAIGALT</td>
</tr>
</tbody>
</table>

C. Acyl carrier protein (ACP)

<table>
<thead>
<tr>
<th>Species</th>
<th>Protein</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>X. semiviridis</td>
<td>XsePKS1</td>
<td>DADQIKDDSDLVEMGISLMGMELAHEVEVAFICTL</td>
</tr>
<tr>
<td>M. purpureus</td>
<td>pksCT</td>
<td>EPDEVKDDSDLVGLGISLMELAREVDALFKTT</td>
</tr>
<tr>
<td>A. nidulans</td>
<td>wa</td>
<td>SESMSDDDLVFADYGVSLLSTVTGKYREELNLM</td>
</tr>
<tr>
<td>P. patulum</td>
<td>MSAS</td>
<td>TAEHVDSKAALADLGVSVMVLRLRLQQLLTLKIAV</td>
</tr>
<tr>
<td>G. moniliformis</td>
<td>FUM1</td>
<td>PVEEMDPTASLTSLG潇VLVTEIIRNWDITRFGVE</td>
</tr>
<tr>
<td>A. terreus</td>
<td>lovF</td>
<td>TDSEMSATQTLAGIGVDSLVAELRWNITAKFNVD</td>
</tr>
</tbody>
</table>

Figure 4.17. Multiple sequence alignment of the CMeT domains in XsePKS1 and other known PKSs showing the conserved motifs as defined by (Kagan & Clarke 1994; Miller et al. 2003). Proposed key amino acids are indicated by (*).
Figure 4.18. Multiple sequence alignment of the SAT domains in XsePKS1 and other known PKSs showing the catalytic triads (*) as proposed by Crawford et al. (2006).

Figure 4.19. Comparison of the PKS domain architecture of XsePKS1 and other PKSs. The first four PKSs belong to the clade III NR-type. PKS domains: SAT, starter unit-ACP transacylase; KS, β-ketoacyl synthase; AT, acyltransferase; ACP, acyl carrier protein; CMeT, C-methyltransferase; R, reductase; CYC, Claisen cyclase/thioesterase; Est?, esterase/lipase (COG0657); DH, dehydratase; KR, ketoreductase; ER, enoyl reductase; C, condensation.
A small gap (700 bp estimated size) of sequence information located at 1.8 kb downstream of \textit{xsepks1} remained unknown due to repeated sequencing failure. Nevertheless, no significant ORF was found in the region directly downstream of \textit{xsepks1}. Sequencing of 2418 bp from the 3′ end \textit{NotI} site of pJpks-ID5 insert found a region of 1452 bp (about 2.5 kb downstream of \textit{xsepks1}) that showed homology to fungal non-LTR (long terminal repeat) retrotransposon with reverse transcriptase (RT) activity (Figure 4.20, RT-like). The translated sequence of the RT-like region (484 amino acids) showed a 31.7% protein identity with the ORF2 of Tad1-1, an active long interspersed nuclear element (LINE)-like element of \textit{Neurospora crassa} (Cambareri et al. 1994). Also found downstream of the RT-like ORF was a short 225 bp region (Figure 4.20), \textit{ctnR-like} that showed low similarity (29%, 18/62 amino acids) to \textit{ctnR} – the citrinin biosynthesis transcriptional activator (Shimizu et al. 2007).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Figure_4.20.png}
\caption{Assembled nucleotide sequence combining pUC52KS insert and pJpks-ID5. ORFs that showed homology to protein sequences in the GenBank database are indicated as blocks. Ambiguous or incomplete ORFs are shaded in grey. A gap of 700 bp estimated size remained unsequenced. RT, reverse transcriptase.}
\end{figure}

\subsection*{4.5.3 Discussion}

\subsubsection*{4.5.3.1 A clade III NR-PKS from \textit{X. semiviridis} (\textit{xsepks1})}

This study demonstrated a successful strategy, for the first time, for targeting and cloning an intact lichen PKS gene encoding a protein of desired domain architecture from field-collected lichen (\textit{X. semiviridis}). The amino acid sequence conservation in KS, AT, ACP and CMeT domains and the transcriptional activity observed in \textit{A. nidulans} strongly suggest that \textit{xsepks1} encodes a functional PKS and is a member of the clade III NR-PKSs. The only lichen PKS gene available in the public domain so far is a NR-PKS cDNA from \textit{Xanthoria elegans} (GenBank accession ABG91136) with domain architecture and amino acid sequences similar to those of \textit{A. nidulans} WA (KS-AT-ACP-ACP-CYC), for which the experimental results are yet to be published.
Besides the \textit{xsepks1}, there is at least one other PKS gene that is similar to the clade III NR-PKSs in \textit{X. semiviridis}, as indicated by \textit{XsNR3KS1} and \textit{XsNRMeT2}, with the possibility of a third one. However, the contiguity of \textit{XsNR3KS1} and \textit{XsNRMeT2} was not confirmed. A BLASTx search with \textit{xsepks1} found many uncharacterized PKS genes with similar KS-AT-ACP-CMeT domain architecture in various fungal genomes. Until Shimizu et al. (2005) recently cloned and characterized the citrinin PKS gene (\textit{pksCT}), no other clade III NR-PKSs had been linked to the polyketide products. More recently, another clade III NR-PKS – \textit{A. strictum} AsPKS1, which encoded for a methylorcinaldehyde synthase, has been characterized (Bailey et al. 2007). Interestingly, no similar genes have so far been detected in previous studies on the phylogenetics of lichen PKS genes (Grube and Blaha 2003; Schmitt et al. 2005; Opanowicz et al. 2006). This is probably due to the limitations of the degenerate primers employed in their studies, since KS domains and CMeT domains belonging to clade III NR-PKS were successfully amplified with \textit{NR3KS} and \textit{NRMeT} primers from other lichens (see Section 4.4.1.2 and Chapter 3 Section3.4.2.2).

A closer examination of the CMeT domains amplified with \textit{NRMeT} primers indicates that CMeT domain of \textit{XsePKS1} possesses the I-IV consensus motifs as described by Miller et al. (2003). By contrast, the CMeT domain of \textit{XsePKS1} (\textit{CMeT011b}) seems to lack the typical motif III and instead shows a conserved (LXXSQHI) motif with \textit{M. purpureus} PKSCT, \textit{B. fuckeliana} PKS17 and other similar clade III PKS genes (Figure 4.17). Fujii et al. (2005) compared the CMeT domain sequences of reducing PKS genes and showed that I-IV consensus motifs as defined by Miller et al. (2003) are present in PKSs with a catalytically competent CMeT domain. However, the proposed citrinin biosynthetic pathway indicated that a methyltransferase is essential for the formation of a methylated polyketide precursor (Barber et al. 1981; Sankawa et al. 1983; Hajjaj et al. 1999), suggesting that the CMeT domain of PKSCT is likely to be functional. In a recent review on PKS domain analysis, Cox (2007) suggested that the methylation of aromatic polyketides could possibly occur after aromatization, in contrast to the linear reduced polyketides. It is unclear if the differences in the conserved motifs account for the differences in the methyl-transfer mechanism or substrate specificity. \textit{In vitro} feeding experiments of heterologously expressed individual catalytic domain, as recently demonstrated by
Ma et al. (2006) and Crawford et al. (2006), is likely to provide insights to the mechanism of CMeT domains.

The XsePKS1 was predicted to be slightly shorter than PKSCT (Fig 4.19) and AsPKS1. A putative reductase (R) domain typical of non-ribosomal peptide synthase, which was not found in the XsePKS1, was detected at the C-terminal of PKSCT and AsPKS1. A region of about 300 bp after the putative stop codon of xsepks1 showed protein homology to the R-domain of PKSCT, but several frame shifts seem to have occurred in the relatively short region and the catalytic core seems to be missing (Figure 4.21). The function of the R-domain was unknown until the recent discovery of A. strictum AsPKS1 that produces methylocinaldehyde. A. strictum AsPKS1 possesses identical domain architecture to PKSCT, i.e. KS-AT-ACP-CMeT-R. The R-domain is thought to be involved in the reductive releases, resulting in an aldehyde group, as opposed to the more commonly observed carboxyl group, at the terminal of the polyketide chain (see Table 3.9). In the case of the citrinin biosynthesis in M. purpureus, the presence of a R-domain in PKSCT was used to explain the results of an isotopic feeding experiment (Hajjaj et al. 1999); it was proposed that a methylated diketide starter unit may be involved in the formation of citrinin in Monascus spp. (Bailey et al. 2007; Cox 2007).
Figure 4.21. The three-frames (positive) translational map of the 3' end region of \textit{xsepkS1}; the shaded amino acids indicate regions that showed homology to the reductase (R) domain of PKSCT at different frames (based on BLASTx result). Position 1 in the map corresponds to 6220 bp downstream from the putative start codon of \textit{xsepkS1}. The putative stop codon (underlined) is located at position 337 in this map.

XsePKS1 also appears to possess the starter unit-ACP transacylase (SAT) domain and the product template (PT) domain. Cox (2007) provided a comprehensive overview on the functions of these two recently discovered domains. Watanabe & Ebizuka (2002) first noted the presence of two such domains of undetermined function in \textit{Colletotrichum lagenarium} PKS1 and \textit{A. nidulans} WA. Udwary \textit{et al.} (2002) later also predicted the location of SAT and PT domains in \textit{A. parasiticus} norsolorinic acid (NA) synthase using the Udwary-Merski algorithm (UMA). The function of the SAT domain of NA synthase was recently confirmed by Crawford et al. (2006), by showing that the SAT domain catalyzes the transfer of hexanoate selectively from CoA onto the ACP in vitro. It is now known that the presence of a SAT domain is widespread among NR-PKSs. Cox (2007) extended this idea of the incorporation of an advanced starter unit by the SAT domain to other known PKS systems, such as the biosynthesis of zearalenone in \textit{G. fujikuroi} and citrinin in \textit{M. purpureus}. An advanced starter unit may also be required for XsePKS1,
since the SAT domain of XsePKS1 showed considerable similarity to PKSCT (40% protein identity).

A conserved region similar to PKSCT (47% identity) and other NR-PKSs was identified between the AT and ACP domain in XsePKS1, where the product template (PT) domain is thought to be located. The function of the PT domain is still undetermined but it is thought to be involved in the chain length control of the polyketide extension in PKSs (Cox 2007; Udwary et al. 2002). The phylogenetic analysis of PT domains in Cox (2007) showed that the NR-PKSs grouped into clades that correspond to product chain-length (Figure 4.22). According to this analysis, the citrinin PKSCT grouped together with other tetraketide synthases (A. strictum AsPKS1 and G. zearalenone PKSB). This further supports the hypothesis that PKSCT used an advance starter unit (methylated diketide acyl unit) in biosynthesis of citrinin, as only three rounds of condensations were required according to this proposed pathway. Using the region corresponding to the position 1281-1600 in XsePKS1 protein (1682-1603 in PKSCT), a maximum parsimony phylogenetic analysis of the predicted PT domain region was performed to predict the chain length of XsePKS1 (Figure 4.23). The analysis showed that XsePKS1 formed a clade with the tetraketide synthases.

**Figure 4.22.** Phylogenetic analysis of individual PT domains. NSAS, norsolorinic acid synthase; STS, sterigmatocystin synthase; THNS, tetrahydroxynaphthalene synthase; WAS, YWA1 naphthopyrone synthase. Unmodified figure taken from Cox (2007).
Figure 4.23. The PT domain phylogeny inferred using the Maximum Parsimony method (Eck & Dayhoff 1966). The bootstrap consensus tree (mid-point rooted) inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed (Felsenstein 1985). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (Felsenstein 1985). The MP tree was obtained using the Close-Neighbor-Interchange algorithm (Nei & Kumar 2000) (pg. 128) with search level 3 in which the initial trees were obtained with the random addition of sequences (10 replicates). The tree is drawn to scale, with branch lengths calculated using the average pathway method (Nei & Kumar 2000) (pg. 132) and are in the units of the number of changes over the whole sequence. All positions containing gaps and missing data were eliminated from the dataset (Complete Deletion option). There was a total of 276 positions in the final dataset, out of which 260 were parsimony-informative. Phylogenetic analyses were conducted in MEGA4 (Tamura et al. 2007).

Several β-orcinol depsidones (fumarprotocetraric acid, succinprotocetraric acid etc.) and a β-orcinol depside (atranorin) were detected by the HPLC system in the X. semiviridis methanol extract (Section 3.2.2.2, Table 3.1). All of these compounds were possibly derived from a β-orsellinic acid intermediate, potentially originating from a clade III NR-PKS. Besides these β-orsellinic acid derivatives, X. semiviridis produced usnic acid, which would require a methylpholoacetophenone precursor that is also likely to be produced by a clade III NR-PKS. The functional domains present in XsePKS1 and its product chain length inferred by PT domain phylogenetic analysis.
fulfill the requirement for biosynthesis of \(\beta\)-orsellinic acid or methylphololo-acetophenone, which are both tetraketides. However, it is unknown at this stage if the XsePKS1 or the other clade III NR-PKSs is responsible for the formation of the major, minor or perhaps undetected polyketide products in \textit{X. semiviridis}. To characterize the polyketide product, heterologous expression of the whole xsepks1 is needed and is attempted in the next chapter.

4.5.3.2 Regions surrounding the xsepks1

Based on the combined sequence of pUC52KS and pJpks-ID5, the only protein coding region within proximity and with potential post-PKS tailoring activity is the partial ORF, \textit{xseB}, located upstream of the \textit{xsepks1} (as described in Section 4.3.2.3). The XseB is homologous to CtnB, which is thought to possess oxidoreductase activity. It is possible that other biosynthetic genes are present further upstream and that \textit{xsepks1} is part of a metabolic gene cluster. The clustering of \textit{xsepks1} and \textit{xseB} and their high similarity to \textit{pksCT} and \textit{ctnB} respectively was parallel to the \textit{M. purpureus} citrinin gene cluster, although the \textit{xseB} transcribes in the opposite direction from \textit{ctnB} (Figure 4.24). Another region similar to the ORFs found in the citrinin gene cluster is a 225 bp region that showed homology to \textit{ctnR} - citrinin biosynthesis transcriptional activator (Shimizu et al. 2007). However, the 225 bp ORF found in pJpks-ID5 is too short to be a functional transcriptional activator. This raised doubts as to the regulation and expression of \textit{xsepks1} in \textit{X. semiviridis}. The similarity of \textit{xseB}, \textit{xsepks1} and the \textit{ctnR}-like region to the citrinin gene cluster (Figure 4.24) indicated that the two gene clusters could have originated from a common ancestor. Genome walking and sequencing of the upstream region might identify more PKS-related biosynthetic genes of the same cluster.

A RT-like partial ORF was also found at about 2.5 kb downstream of \textit{xsepks1} (Figure 4.24); this showed similarity to \textit{N. crassa} Tad1-1 ORF2 and other fungal hypothetical proteins with an RT-like domain. \textit{N. crassa} Tad1-1 is an active LINE-like element belonging to the fungal Class I non-LTR retrotransposon (Cambareri et al. 1994). This class of retrotransposons consists of two ORFs: 1) a GAG-like protein and 2) a reverse transcriptase-RNaseH protein (Daboussi & Capy 2003; Favaro et al. 2005). However, it was unknown if an equivalent of Tad1-1 ORF1 was present.
upstream of the *X. semiviridis* RT-like coding region, since a ~700 bp gap in the sequence remained undetermined. Insertion of transposable elements, such as a Class I retrotransposon, into and near a gene could block transcription totally or alter the transcription pattern (Daboussi & Capy 2003). The presence of the non-LTR retrotransposon at the C-terminal end may represent a recent transposition event that altered the 3´end of *xsepks1*, resulting in disruption of the R-domain of XsePKS1 and the *ctnR*-like transcriptional activator downstream of *xsepks1*.

**Figure 4.24.** Citrinin biosynthetic gene cluster in *Monascus purpureus* (AB243687) (Shimizu et al. 2007) and its parallel to *xsepks1* and associated genes.
# CHAPTER 5 HETEROLOGOUS EXPRESSION OF LICHEN PKS GENE IN *ASPERGILLUS*

## 5.1 Introduction ................................................................. 206

## 5.2 Transformation and transcription analysis of partial xsepks1 in *Aspergillus nidulans* ................................................. 210

### 5.2.1 Materials and methods .............................................. 210
- **5.2.1.1 Strain and plasmids** ........................................... 210
- **5.2.1.2 A. nidulans** transformation .................................. 210
- **5.2.1.3 Selection and PCR screening of transformants** .... 211
- **5.2.1.4 Transcription detection by RT-PCR** ...................... 211

### 5.2.2 Results ...................................................................... 212
- **5.2.2.1 A. nidulans** transformation and PCR screening ...... 212
- **5.2.2.2 Transcriptional analysis of the partial PKS gene in A. nidulans** .................................................. 213

### 5.2.3 Discussion ............................................................... 214

## 5.3 Heterologous expression of full-length xsepks1 in *Aspergillus oryzae* ............................................................... 215

### 5.3.1 Materials and methods .............................................. 218
- **5.3.1.1 Construction of entry clone by BP reaction** ........ 218
- **5.3.1.2 Construction of entry clone by LR reaction** ....... 219
- **5.3.1.3 A. oryzae** transformation ................................... 219
- **5.3.1.4 PCR screening of A. oryzae** transformants .......... 220
- **5.3.1.5 Transcription analysis of A. oryzae** transformants ... 220
- **5.3.1.6 Fermentation and organic extraction** ............... 222
- **5.3.1.7 HPLC analysis** .................................................. 222

### 5.3.2 Results ...................................................................... 223
- **5.3.2.1 A. oryzae** transformation and screening .......... 223
- **5.3.2.2 A. oryzae** transformants transcription analysis ... 223
- **5.3.2.3 Chemical analysis of A. oryzae** transformants ...... 228

### 5.3.3 Discussion ............................................................... 237
- **5.3.3.1 Transformation and expression of xsepks1 in A. oryzae** .......................................................... 237
- **5.3.3.2 Functionality of XsePKS1 protein in A. oryzae** .... 241
- **5.3.3.3 Possible functions of clade III NR-PKSs in lichens** 244
5.1 Introduction

The next step after cloning the PKS gene from the lichen X. semiviridis is to characterize its function biologically (the role it plays in the organism) and/or chemically (identify the polyketide product produced by the PKS). An overview of the technological aspect of functional analysis and heterologous expression of fungal PKS genes was put forward by Schumann and Hertweck (2006) while an insightful review on the functional analysis of individual PKS domain was given recently by Cox (2007). The various strategies as well as their limitations and suitability for application on lichens will be briefly discussed here.

Regardless of the approach adopted, genetic manipulation that involves introduction of a DNA fragment into the organism is inevitable. This includes the conventional gene disruption and complementation of a fungus, and heterologous expression of a PKS gene in a surrogate host. The two earliest developed and best-studied transformation systems are *Neurospora crassa* and *Aspergillus nidulans* (Ballance et al. 1983; Case et al. 1979; Tilburn et al. 1983; Yelton et al. 1984). The most commonly used fungal transformation method requires the generation of protoplasts with cell wall lytic enzymes followed by permeabilization of cell membranes with polyethylene glycol (PEG) and calcium chloride (Andrianopoulos & Hynes 1988; Brakhage & Langfelder 2002; Gomi et al. 1987). Other alternatives available are electroporation of geminating conidia, *Agrobacterium*-mediated transformation, and particle bombardment (Schumann & Hertweck 2006).

The common markers used for selection of transformants are auxotrophic markers or antibiotic resistance markers against hygromycin B, phelomycin, glufosiniate (BASTA), sulfonylurea and benomyl (Talbot 2001). The advantage of auxotrophic markers is the low cost involved, since transformants can be easily selected on a specific nutrient-deficient medium. However, it is only useful for a handful of fungi where auxotrophic mutant strains are available. In other cases, dominant antibiotic selection markers are more versatile so long as the targeted fungus does not have a natural resistance to the antibiotic used.

The function of PKS genes could theoretically be characterized by gene disruption or complementation of mutant strains, followed by observing the changes
in phenotype or polyketide production. In many ways, the approaches 3, 5, 6 and 7 as described in Chapter 4 (Table 4.1) to locate a particular PKS gene have simultaneously linked the gene to its phenotype or biological function. However, approaches that rely on differential expression of genes under varied conditions (such as 6 & 7 in Table 4.1) need further experimental confirmation, normally via targeted gene disruption. Similarly, PKS gene fragments obtained using degenerate primers (approach 8, Table 4.1) can be targeted for disruption to link the gene to the lost phenotype observed (Shimizu et al. 2005; Zhang et al. 2003). Targeted gene disruption is commonly achieved by double homologous integration (double cross-over), whereby a selective marker (auxotrophic or antibiotic marker) flanked by two homologous fragments of the targeted gene is introduced into the fungus, resulting in gene replacement.

Although gene disruption and mutant complementation are quick ways to link a PKS gene to its phenotype, for reasons discussed in Chapter 4 (Introduction 4.1), they are yet to be feasible for characterizing lichen PKSs. Since these approaches inevitably need culturing and genetic manipulation of the organism, the slow growth rate of lichen mycobionts in laboratory conditions is a bottleneck to this approach. Therefore, the alternative approach of characterizing the lichen PKS gene by heterologous expression was examined.

Heterologous expression, by definition, is the expression of an isolated gene in a surrogate host. The expression can be regulated by either a host promoter or by the native promoter of the gene introduced into the host. The disadvantage of characterizing the function of a PKS gene by heterologous expression is that the biological significance of the gene in the original organism remains unknown. For instance, if a PKS gene is responsible for production of a compound that is important for a developmental function (e.g. sporulation) in the original lichen, the function will remain unknown unless characterized by gene disruption/complementation or transcriptional analysis, although it is sometimes possible to infer the function from the compound(s) produced in a heterologous host. Heterologous expression, however, could theoretically allow the function of individual PKS genes to be analyzed, whereby the direct product of the foreign PKS protein unmodified by the original organism could be identified in the heterologous host, especially when the polyketide product is an intermediate to a series of
following enzymatic steps that leads to the final product. Gene disruption and complementation of the respective PKS gene would observe just the lost and restored production of the final product. Although in theory host enzymes could also sometimes modify the heterologously expressed polyketide, the unmodified product could still be isolated in most cases, since the expression of the modifying host enzymes would not be co-regulated with the heterologously expressed PKS gene, unlike in the original organism. Furthermore, heterologous expression of a PKS gene would allow high production of a desired polyketide by over-expression and serve as a basis for engineering the PKS protein in future.

Heterologous expression of fungal PKS genes is still in its infancy, as it possesses its own unique challenges (Schumann & Hertweck 2006). Firstly, the presence of introns in the fungal PKS genes may require the use of cDNA, which can be difficult to obtain sometimes (especially from environmental samples). Although there are also other means of removing the introns, such as rapid assembly of multiple exons by PCR (An et al. 2007) or ligating fragments of DNA excluding the intron, the methods are relatively cumbersome and the prediction of intron sites can be inaccurate.

Secondly, the ACP domains of fungal PKSs require posttranslational 4-phosphopanteinlyation catalyzed by a phosphopanteinyl transferase (PPTase). To date, filamentous fungi, such as *A. nidulans* and *A. oryzae*, have been used as the heterologous hosts to express the greater number of fungal PKS genes from different classes (Cox et al. 2004; Fujii et al. 1999; Fujii et al. 1996; Fujii et al. 2005; Kasahara et al. 2006; Lu et al. 2005; Watanabe et al. 2000; Watanabe et al. 1998). Other heterologous hosts, such as *Streptomyces coelicor* CH999 (Bedford et al. 1995), *Escherichia coli* (Kealey et al. 1998), *Sacharomyces cerevisiae* (Kealey et al. 1998) and even tobacco (Yalpani et al. 2001) have been used successfully for fungal PKS gene expression, but this has been limited to the 6-methylsalicylic synthase (6-MSAS) gene. Expression of fungal PKS genes in non-polyketide producing hosts such as *E. coli* and yeast required the co-expression of a PPTase gene (such as the *sfp* from *Bacillus subtilis* or the PPTase gene from *S. coelicor*) to convert the fungal PKS into the holo-form (Bedford et al. 1995). Since the phosphopanteinylation of fungal PKS by bacterial PPTases has so far only been demonstrated in 6MSAS, it is not clear if bacterial PPTases can be applied to other
fungal PKSs as well, or is limited to the 6-MSAS, which is known to resemble bacterial iterative type I PKS (Kroken et al. 2003). [Note: The *Gibberella fujikuroi* PKS4 gene, which produce bikaverin, was expressed in *E. coli* and the enzyme activity was reconstituted *in vitro* recently. The ACP domain of the purified PKS was verified to be in the holo form, likely to be phosphopantethienylated by the *E. coli* holo-ACP synthase *in vivo* (Ma et al. 2007)].

Miao et al. (2001) proposed the exploitation of the lichen biosynthetic potential by heterologous expression in a well-characterized, fast-growing fungal host, such as *A. nidulans* or *N. crassa*. *A. nidulans* was preferred due to its closer phylogenetic relationship with ascomycetous lichens. Prior to the perspective provided by Miao et al. (2001), Sinnemann et al. (2000) had demonstrated the feasibility of using *A. nidulans* as a heterologous host for studying lichen genes, by complementing a lichen (*Solorina crocea*) pyrG gene in an *A. nidulans pyrG* (uridine-dependent) mutant strain. The complemented *A. nidulans* strain transcribed and spliced the introns of *S. crocea* pyrG correctly, and grew on uridine-deficient medium. Although only mentioned as unpublished data in Miao et al. (2001), D. Armaleo (pers. comm.) also reported the successful splicing and transcription of a *Cladonia grayii* PKS gene inserted into *A. nidulans* under the regulation of the native lichen promoter. Since *A. nidulans* and *A. oryzae* are 1) native polyketide producers, 2) could potentially splice the introns in lichen PKS genes and 3) have a proven track record of successful heterologous expression of a variety of fungal PKS genes, they were chosen as prospective heterologous hosts for expressing the *X. semiviridis* xsepks1.
5.2 Transformation and transcription analysis of partial xsepks1 in Aspergillus nidulans

This experiment involved transformation of *A. nidulans* with pUC52KS (containing the 5´ *xsepks1*) and was carried out to check the ability of *A. nidulans* host to utilize the putative *X. semiviridis* promoter and splice the intron predicted as present in the 5´ *xsepks1*. The well-characterized *A. nidulans* transformation system in M.A. Davis’s laboratory at the Genetics Department in University of Melbourne, Australia was used. Transformation of *A. nidulans* was performed in M.A. Davis’s laboratory and was assisted by D. Clarke (research assistant in Davis’s lab) who was, at that time, investigating nitrogen metabolism in *A. nidulans* and required to do some transformation experiments in parallel. The *A. nidulans* strain with multiple mutation/recombination loci and the co-transformed plasmid (pMT1612) containing a glufosinate-resistance marker corresponded to what he was using at that time (see Section 5.2.1.1 below). Following selection on glufosinate-containing medium, the transformants were transferred to the plant microbiology lab at RMIT University for molecular and transcriptional analysis.

5.2.1 Materials and methods

5.2.1.1 Strain and plasmids

The *A. nidulans* strain used for the transformation was MH19635 [*biA1 pyrG89 pabaB22 gpd(p)areAHA crmA<sup>T525C</sup> pyrG<sup>+</sup> (at crmA)*]. This strain was originally used for the study of genetic regulation of nitrogen metabolism in *A. nidulans* as described by Todd et al. (2005). The media and growth conditions for *A. nidulans* MH19635 were as described by Cove (1966). pUC52KS was as described in Section 4.3.2.2 Figure 4.6 and pMT1612 was as described by Monahan et al. (2006) and contained the glufosinate resistance (*bar*) gene as a selectable marker.

5.2.1.2 *A. nidulans* transformation

The MH19635 strain was cotransformed with pUC52KS and pMT1612 using the protoplast-PEG method described previously (Andrianopoulos & Hynes 1988). Cotransformants were selected on medium containing 25 µL/mL glufosinate.
5.2.1.3 Selection and PCR screening of transformants

Genomic DNA of *A. nidulans* transformants for PCR screening was prepared by the miniprep method described by Prebble and Anderson ([http://www.aspergillus.org.uk/secure/laboratory_protocols/dnarna/dnaprep.html](http://www.aspergillus.org.uk/secure/laboratory_protocols/dnarna/dnaprep.html)). For other purposes, genomic DNA of *A. nidulans* was extracted with a DNeasy Plant Extraction Kit (QIAGEN).

Cotransformants that were able to grow on glufosinate selective medium were further screened by PCR using the RT52KS-F/R primers to detect the presence of the 5' *xsepk1*. The contiguity of the 5' *xsepk1* with the putative promoter region in the positive transformants was determined by PCR using the GAP-1F primer (previously used for sequencing of the pUC52KS) corresponding to the putative promoter region and the universal M13 (pUC) forward (-20) primer with RT52KS-R as reverse primer (Figure 5.1).

**Figure 5.1.** The plasmid map of pUC52KS and the approximate location of the binding sites for M13 forward, Gap1-F, RT52KS-F and RT52KS-R primers.

5.2.1.4 Transcription detection by RT-PCR

RT52KS-F/R specific primers flanking the putative intron were used for RT-PCR. Total RNA from the *A. nidulans* transformants grown in malt extract (20 g/L) (Oxoid, Hampshire, United Kingdom) liquid medium was extracted with a RNeasy Plant Mini Extraction Kit (QIAGEN). The first step of RT-PCR from the total RNA was carried out with MLV-reverse transcriptase (Promega) using the RT52KS-R primer. The reaction was column-purified and subjected to PCR with *Taq* polymerase (Invitrogen) using RT52KS-F and R primers. For the positive control, PCR under the same conditions was set up with genomic DNA extracted from the *A. nidulans* transformants. For the
negative control, total RNA was used directly for PCR amplification with RT52KS-F/R primers, omitting the first step (reverse transcriptase reaction).

5.2.2 Results

5.2.2.1 A. nidulans transformation and PCR screening

About 50 cotransformants were obtained from the glufosinate-containing selective medium. The transformants were subcultured on to new selective agar plates in a grid format, where a few of the colonies did not grow (Figure 5.2). Two (a10 and a22) out of 25 cotransformants showed positive amplification for PCR screening with RT52KS-F/R primers. PCR analysis of the transformants using the universal M13 (pUC) forward (-20) primer or GAP-1F with RT52KS-R as reverse primer (Figure 5.1) produced products of the expected size (3.1 and 2.3 kbp respectively) confirming the presence of at least one copy of contiguous putative promoter region and 5’ xsepks1 fragment in both transformants (Figure 5.3).

Figure 5.2. A. nidulans transformants growing on glufosinate-containing selective medium. Two transformants (a10 and a22) in the first plate carried the 5’ xsepks1.
5.2.2.2 Transcriptional analysis of the partial PKS gene in *A. nidulans*

RT-PCR of a10 and a22 transformants with RT52KS-F/R primers flanking the putative intron (Figure 5.1) yielded a product shorter than that amplified from the genomic DNA (Figure 5.4), as expected if the *A. nidulans* host was able to process the intron. Sequence comparison of the RT-PCR product with the sequence from genomic DNA confirmed that the deletion from the RT-PCR transcript corresponded to the 63 bp intron.

**Figure 5.4.** Transcription analysis of a10 and a22 transformants with RT52KS-F/R primers flanking the putative intron. Lanes 1-3, transformant a10; lanes 4-6, transformant a22; lanes 1 & 4, cDNA; lanes 2 & 5, genomic DNA; lanes 3 & 6, total RNA as negative control.
5.2.3 Discussion

The transcription of 5′ xsepks1 could be controlled by either a host or the native lichen promoter due to the random integration of pUC52KS, but the contiguity of the putative native promoter with 5′ xsepks1 as detected by PCR suggests that the latter is likely to be true. Sinnemann et al. (2000) have previously demonstrated the expression of a pyrG gene from the lichen S. crocea (Peltigeraeaceae) in A. nidulans. Transformation of A. nidulans with S. crocea pyrG, controlled by either its native promoter or the A. nidulans trpC promoter, resulted in uridine-independent strains. Miao et al. (2001) also reported as a personnel communication with other researchers (D. Armaleo, C-S. Kua and Y. Zhang) that RT-PCR analysis of transformants of A. nidulans carrying a PKS gene from the C. grayi mycobiont (cultured) showed that the lichen gene was transcriptionally active under its native promoter and that the introns were spliced correctly. This study reverified the potential use of A. nidulans as a heterologous host for expression of lichen PKS genes. A transformable host that is able to process the lichen PKS gene intron is advantageous, especially when it is difficult to obtain the full length cDNA from environmental samples. The ability to utilize lichen gene promoters also enables the possible integration of a complete biosynthetic gene cluster in the host without reconstructing the whole pathway. The next chapter focused on the effort to express the full-length xsepks1 with the intention of characterizing the PKS gene function.
5.3 Heterologous expression of full-length xsepks1 in *Aspergillus oryzae*

Following Section 5.1, which confirmed the ability of *Aspergillus nidulans* to transcribe and splice the introns of 5’ *xsepks*, the function of the PKS was explored by expressing the full-length gene followed by polyketide product detection. Although many secondary metabolites of *Aspergillus* spp. have been well-characterized, fungi in the genus produce numerous metabolites that could result in the heterologous polyketide being buried in the high biosynthetic background. As the timing and conditions for activation of the lichen promoter is unknown, it could be difficult to predict when the lichen PKS gene is being expressed and would require time for optimization. To prevent the unpredictability and reduce the biosynthetic background, an inducible and strong promoter is desired to highlight the production of the heterologous polyketide under the induced conditions.

Since secondary metabolites in fungi are typically produced after the log growth phase, promoters involved in primary metabolism, such as those expressing genes involved in the breakdown of carbohydrates, are especially useful, as they initiate expression of the heterologous PKS gene before the secondary metabolism has started. Hence, if the fermentation process is halted before the start of secondary growth phase, the secondary biosynthetic background could, in principle, be reduced greatly, while a high quantity of the heterologous polyketide could be obtained. Some examples of such inducible promoters that are available for fungal hosts are the TAKA-amylase *amyB* promoter (Tada et al. 1991), xylanase *xylP* promoter (Zadra et al. 2000), alcohol dehydrogenase *alcA* promoter (Kulmburg et al. 1992), etc. The *A. oryzae* TAKA-amylase controlled by its promoter is expressed in a relatively large percentage of the total protein in *Aspergillus* when starch or maltose is present, but the expression is repressed when glucose is available (Tada et al. 1991). The promoter of the amylase gene has been used successfully to express several fungal PKS genes in *A. nidulans* and *A. oryzae* (Cox et al. 2004; Fujii et al. 1999; Fujii et al. 1996; Fujii et al. 2005; Kasahara et al. 2006; Lu et al. 2005; Watanabe et al. 2000; Watanabe et al. 1998).
Another challenge in expression of fungal PKS genes is the construction of an expression plasmid. Since PKS genes are relatively large, ranging from 5.5 kb to over 10 kb, the use of the conventional restriction-ligation cloning method is restricted greatly by the limited restriction sites. A restriction site that would otherwise be used in a cloning scheme might also be found in an inconvenient location within the gene or vector of interest. This problem becomes greater as the sizes of the genes and vectors increase. The direction and insertion site of the PKS gene at an appropriate distance after the promoter is also critical for correct expression of the gene. To circumvent these technical problems, a Gateway expression system (pTAex3R) based on the pTAex3 vector was developed in I. Fujii’s lab. The proprietary Gateway technology (marketed by Invitrogen) is a cloning method based on the site-specific recombination properties of bacteriophage λ (Landy 1989; Hartley et al., 2000). A recent review by Hartley (2006) discussed the various alternative cloning technologies available, with specific mention of the two site-specific recombination-based Gateway and Creator technologies (marketed by Clontech) and their advantages over the conventional cloning method.

The pTAex3 vector was originally described in (Fujii et al. 1995), and contains an A. oryzae α-amylase promoter (PamyB) for starch-inducible gene expression and an Aspergillus nidulans argB gene for selection of complemented Aspergillus auxotrophic mutants. It was used successfully for the expression of dihydrogeodin oxidase of A. terreus in A. nidulans (Huang et al. 1995) and A. terreus atx gene that produced 6-methylsalicylic acid in A. nidulans (Fujii et al. 1996). Even though atx (msas) is among the smallest of fungal PKS genes, the cloning procedure was still relatively cumbersome, as illustrated in Fujii et al. (1996). Other PKS gene expression studies have also exploited the pTAex3 system (Cox et al. 2004; Fujii et al. 1999; Fujii et al. 2001; Kasahara et al. 2006; Tsai et al. 2001; Watanabe et al. 2000; Watanabe et al. 1998). The pTAex3 vector was converted into a GATEWAY destination/expression vector by inserting a GATEWAY cassette Frame A into the SmaI (blunt end) site of the pTAex3 plasmid to become pTAex3R (Figure 5.5). With the suicide ccdB gene inserted and flanking attR sites, a gene or DNA fragment flanking attL-sites in an entry clone could be swapped into the pTAex3R plasmid via a λ enzyme-mediated site-specific recombination to form the expression plasmid. Following the construction of the pTAex3R system, the GATEWAY destination vector
was successfully employed in expression of \textit{pksN} and \textit{pksF} from \textit{Alternaria solani} in \textit{A. oryzae} (Fujii et al. 2005; Kasahara et al. 2006).

\textbf{Figure 5.5.} Conversion of pTAex3 vector into a GATEWAY destination vector, pTAex3R.
The convenience of this expression system was exploited for expression of xsepks1 from X. semiviridis in A. oryzae. The construction of the xsepks1 expression plasmid and transformation of A. oryzae were carried out during a short research visit to Y. Ebizuka and I. Fujii’s lab in University of Tokyo. Part of the work (specifically in Section 5.3.1.1) was carried out by I. Fujii before my arrival in University of Tokyo due to the limited time period of the visit. Some fermentations and preliminary HPLC analysis of the A. oryzae transformants was performed in University of Tokyo. At the end of the research visit, the A. oryzae transformants and organic extracts were transferred to the lab at RMIT University for further analysis. Molecular and transcriptional analysis was performed at RMIT along with repeated fermentation experiments. The organic extracts were analysed in J.A. Elix’s lab at the Australian National University to match against a library of 1075 lichen compounds.

5.3.1 Materials and methods

5.3.1.1 Construction of entry clone by BP reaction

Using the primers with attB site attachments (attB1Cse-N and attB2Cse-C), the ORF of xsepks1 was amplified from the pJpks-ID5 plasmid with Phusion High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA).

| attB1Cse-N: 5' - GGGGACAAGTTTGTACAAAAAAGCAGGCT TACT GCCACCATGGGCGACATTCAAACCTTCTCAAGC - 3' |
| attB2Cse-C: 5' - GGGGACCACTTTGTACAAGAAAGCTGGGT TA TCAAGTCACGTGATCTTGTTCAC - 3' |

The bases in bold are the attB1 and attB2 sites respectively. The consensus Kozak sequence is underlined.

The consensus Kozak sequence “ACCATGG” was added at the start codon in the forward attB1Cse-N primer for efficient eukaryotic translation (Kozak 1984). The reverse primer attB2Cse-N binding site was located further downstream of the proposed stop codon as there was an ambiguous region that showed some homology to the N-terminal (reductase domain) of PKSCT (details were discussed in Chapter 4 Section 4.3). Hence, if there was an unpredicted intron at the N-terminal beyond the proposed stop codon, it might be removed by the A. oryzae host splicing mechanism and extend the ORF further.
To construct the entry clone, the amplified \textit{xsepks1} ORF flanking attB sites (2 \(\mu L\)) were cloned into pDONR221 entry vector (Invitrogen, Carlsbad, CA) by site-specific recombination using the BP clonase II mix (Invitrogen, Carlsbad, CA). After incubation for 1 h, the BP reaction was treated with proteinase K for 10 minutes at 37\(^\circ\)C and the reaction (10 \(\mu L\)) was transformed with 120 \(\mu L\) DH5\(\alpha\) Z-competent cells (prepared using the Z-Competent \textit{E. coli} transformation buffer set - Zymo Research, Orange, CA) and grown on LB-kanamycin plates. Plasmids were isolated from eight of the colonies grown overnight in 3 mL culture using GFX Micro Plasmid Prep Kit (GE Healthcare, Shinjuku-ku, Tokyo, Japan). The plasmids were sequenced to check for PCR amplification fidelity. One clone, pDONR-Cse3, which was most consistent with the sequence from pJpks-ID5, was chosen.

5.3.1.2 Construction of entry clone by LR reaction

The \textit{xsepks1} ORF, now flanked by the attL sites in the entry clone pDONR-Cse3, was moved into the pTAex3R destination vector using the LR Clonase II mix (Invitrogen, Carlsbad, CA) to form the expression plasmid, pTA-Cse3 (Figure 5.6). After 1 h incubation, 2.5 \(\mu L\) of the reaction was transformed with 120 \(\mu L\) DH5\(\alpha\) Z-competent cells and spread directly on LB-carbenicillin plates. Four of the colonies were picked and grown overnight in LB-carbenicillin broth. The resulting plasmids isolated from the overnight culture were checked for correct band size by digestion with \textit{BamH}I and electrophoresed on agarose gel.

5.3.1.3 \textit{A. oryzae} transformation

The auxotrophic \textit{A. oryzae} M-2-3 (argB\(^{-}\)), as described previously (Gomi et al. 1987), was used as the host for \textit{xsepks1} expression. The \textit{A. oryzae} overnight culture was protoplasted with 1.0% Yatalase (Takara BIO, Shiga, Japan) and transformed with 5-10 \(\mu g\) of DNA (pTA-Cse3) using the protoplast-PEG method as described previously (Gomi et al. 1987; Kasahara et al. 2006). The resulting transformants were selected on Czapek Dox (CD) agar (0.3\% NaNO\(_3\), 0.2\% KCl, 0.1\% KH\(_2\)PO\(_4\), 0.05\% MgSO\(_4\), 0.002\% FeSO\(_4\), 2\% glucose and 1.5\% agar). To induce the expression of \textit{xsepks1} under the amyB promoter, the transformants were grown in Czapek Dox (CD) broth medium with the glucose substituted by 2\% starch and supplemented with 1\% peptone (CD-starch).
Figure 5.6. The plasmid map of pTA-Cse3 containing the xsepks1 ORF. Small arrows indicate the binding sites of the primers used in 3’RACE.

5.3.1.4 PCR screening of A. oryzae transformants

The presence of pTA-Cse3 in the A. oryzae transformants was screened with PCR, essentially the same as described in Section 5.2.1.3. Besides the RT52KS-F/R primers, the primers that flanked the KS-CMeT interdomain region (XsNR3KS2-iF and CMeT011a-R2) (described in Chapter 4 Section 4.4.1.4 and 4.4.2.5) were also used to confirm the integrity of the xsepks1 ORF in the A. oryzae transformants.

5.3.1.5 Transcription analysis of A. oryzae transformants

To induce the xsepks1 expression, the A. oryzae transformants were grown for 2 days in 200 mL induction broth medium (CD-starch, see section 5.2.1.3) at 25°C, 150 rpm. For RNA extraction, about 200 mg of mycelium was collected from the culture.
and partially dried on filter paper. The remaining cultures were subjected to organic extraction and HPLC analysis (Section 5.3.1.6-7). Total RNA was extracted with an RNeasy Plant Mini Extraction Kit (QIAGEN). Initially, the xsepks1 transcription activity of the transformants was checked with the one-step RT-PCR method as described in Section 5.2.1.4 using the RT52KS-F/R primers flanking the intron. However, the one-step RT-PCR would yield a product for both truncated and full-length transcripts (Figure 5.7, 2). To verify the transcription of full-length xsepks1, the analysis was repeated with a two-step RT-PCR method (Figure 5.7, 3). First, cDNA was generated from mRNA with Superscript II reverse transcriptase (Invitrogen) using 3'-CDS primer A (supplied in SMART RACE cDNA Amplification Kit - Clontech Laboratories, Mountain View, CA) that binds to the poly-A tail of mRNA, therefore a truncated mRNA would not be reverse transcribed (Figure 5.7, 3). The cDNA was then used as template for PCR amplification using the RT52KS-F/R primers.

**Figure 5.7.** The differences between the one-step and two-steps RT-PCR. 1) PCR from genomic DNA: intron present in the PCR product, 2) one-step RT-PCR: both poly-A tailed mRNA and truncated mRNA were amplified, 3) two-steps RT-PCR: 3'-CDS A primer selectively reverse transcribed poly-A tailed mRNA; PCR in the second step with gene-specific primers amplified only from the xsepks1 first strand cDNA.
To check if any splicing occurred at the N-terminal or 3’ end of \textit{xsepks1}, a 3’ cDNA RACE (\textit{Rapid Amplification of cDNA Ends}) was performed with the SMART RACE cDNA Amplification Kit using the same cDNA as template and gene-specific forward primers (GAP3-F, MeT011a-F and 3´RACEpksMeT) (Figure 5.6).

5.3.1.6 Fermentation and organic extraction

\textit{A. oryzae} transformants and a control \textit{A.oryzae} strain transformed with empty pTAex3 plasmid were grown for two days in 200 mL induction medium (CD-starch) at 25°C, 150 rpm. For extraction of metabolites, the culture medium was acidified with hydrochloric acid to pH 3-4 and extracted twice with equal volumes of ethyl acetate. The organic layer was separated from the aqueous phase in a separation funnel and dried over anhydrous magnesium sulphate. The ethyl acetate was rotary evaporated and the residue was dissolved in HPLC grade methanol or acetonitrile for HPLC analysis. For extraction of metabolites in the mycelium (only for transformant 6), the mycelium was separated from the culture by suction filtration through a filter paper (Whatman No.1) and extracted with 200 mL acetone for 2 h with slow agitation. The extract was filtered to remove the mycelium and the acetone was rotary evaporated. The residue was dissolved in HPLC grade methanol for HPLC analysis.

5.3.1.7 HPLC analysis

HPLC analyses of the organic extracts from \textit{A. oryzae} transformants were either carried out at I. Fujii’s lab at the University of Tokyo or at J.A. Elix’s lab at the Australian National University. In I. Fujii’s lab, HPLC analysis was performed on a Tosoh DP 8020 pump, PD 8020 photodiode array detector, and an Inertsil ODS-3 column (4.6R150 mm) (GL Sciences, Tokyo, Japan) with a solvent system of acetonitrile containing CH$_3$COOH (1%, solvent B) and H$_2$O containing CH$_3$COOH (1%, solvent A) at a flow rate of 0.8 mL/min. Elution was started with 80:20 solvent A/solvent B for 5 min followed by a gradient from 80:20 to 0:100 over 20 min and remained isocratic for another 5 min. The UV adsorption was monitored at 240-350 nm. To identify the presence of expected compounds (\textit{β}-orsellinic acid or methylpholoacetophenone) or other lichen compounds in the extracts, the remaining extracts in the HPLC vials were dried in dry nitrogen gas and transferred to J.A. Elix’s
lab for HPLC-UV library match (the library contains 1075 compounds). In this case, the standard protocol as described in Chapter 3 Section 3.2.1.3 was used.

5.3.2 Results

5.3.2.1 A. oryzae transformation and screening

A total of 28 transformants was obtained from two rounds of transformation. Only three colonies formed during the first transformation due to low plasmid DNA concentration. The second transformation, with higher plasmid DNA concentration, produced about 22 colonies (Figure 5.8). Seven of the colonies (1, 2, 3, 4, 5, 6, 25) from both transformations were picked for further analysis. All seven colonies produced the expected band size (1 kb) when PCR-amplified with the RT52KS-F/R primers. As the pTA-Cse3 plasmid was randomly integrated into the A. oryzae chromosome, the PCR results only showed that a fragment of xsepks1 was introduced into the transformants. KS-CMeT interdomain PCR produced a 4.5 kb band, confirming the integrity of xsepks1, but only in the transformants 3, 4, 5, and 6.

![Figure 5.8. A. oryzae transformants growing on Czapek-Dox (CD) agar medium.](image)

5.3.2.2 A. oryzae transformants transcription analysis

Initial analysis using one-step RT-PCR with RT52KS-F/R primers showed a band of the expected size (0.94 kb, about 60 bp shorter than the PCR product from genomic DNA due to the intron removal) for transformants 2, 3, 4, 5, 6 and 25 (same as Section 5.2.2.2, Figure 5.4). Since the KS-CMeT interdomain PCR showed that
only transformants 3, 4, 5, and 6 maintained the integrity of xsepks1, the two-step RT-PCR with the same primers was tested for these transformants only. As opposed to one-step RT-PCT, the two-step RT-PCR should prove if the full-length gene was transcribed in the transformants, since the two-steps RT-PCR product should theoretically be amplified from a poly-A tailed mRNA (Figure 5.7). All four transformants (3, 4, 5 and 6) produced a band of expected size (0.94 kb).

A 3´ RACE was performed on transformants 5 and 6 to determine the exact translational stop of xsepks1. The 3´ end of xsepks1 was amplified using the Gap3-F gene specific primer (previously used for pJpks-ID5 sequencing in Chapter 4) and universal primer mix (UPM, supplied in the SMART RACE cDNA Amplification Kit) that binds to the adaptor sequence of 3´-CDS A primer. Based on the attB2 site in pTA-Cse3 (Figure 5.6) (the 3´ end of the insert containing xsepks1 ORF cloned from pJpks-ID5 into pTA-Cse3), the product size was expected to be about 3.1 kb if no substantial intron was removed. The 3´ RACE obtained a band slightly larger than 3 kb for both transformants but an additional weak band of ~1.6 kb was obtained for transformant 5 (Figure 5.9). The ~3.1 kb products from both transformants were gel-purified and sequenced using two internal gene-specific primers, CMeT011a-F and 3´RACEpksMeT.

![Figure 5.9. Gel electrophoresis of 3´ RACE products of transformants 5 (T5) and 6 (T6).](image)
Interestingly, the electrophoretograms returned from the sequencing showed mixed peaks for both transformants after a certain point in the sequences. Comparing the two electrophoretograms side by side indicated that the all four sequences started to mix after “TCCTAG” located at about 6495 bp downstream of the putative start codon of \textit{xsepsks1} (Figure 5.10 A-B). To verify the mixed sequences, the ~3.1 kb 3´ RACE products was subjected to a second round of nested PCR amplification using two sets of primers. The CMeT011a-F primer corresponds to 2.2 kb downstream of the GAP3-F binding site, while 3´RACEpksMeT primer is 200 bp further downstream of the CMeT011a-F primer binding site (Figure 5.6). Accordingly, using the two primers (CMeT011a-F and 3´RACEpksMeT) along with the nested universal primer (NUP, supplied in SMART RACE cDNA Amplification Kit), the expected PCR product sizes would be roughly 1.1 kb and 0.9 kb respectively. Surprisingly, each primer combination yielded an additional band slightly shorter than the expected bands for both transformants (Figure 5.11). The CMeT011a-F/ NUP PCR products (band 1.1kb and 1.05 kb) of both transformants 5 and 6 were gel extracted. The bands of the same size from both transformants were pooled together and sequenced with CMeT011a-F and 3´RACEpksMeT (internal) primers, on the assumption that the two products of the same size from both transformants should be identical (since on the electrophoretograms in Figure 5.10; the mixed peaks consisted of the same bases).

\begin{figure}
\centering
\includegraphics[width=\textwidth]{5.10.png}
\caption{Electrophoregrams of the 3.1 kb 3´ RACE products: A) transformant 5, B) transformant 6. Mixed peaks were observed in both sequences after “TCCTAG”.}
\end{figure}
The sequencing of the 1.0 kb product with both primers (CMeT011a-F and 3'RACEpksMeT) showed that a splicing of a 63 bp intron had occurred (Figure 5.12). The intron (intron2) had a “GT….AG” conserved motif and had the same size as the intron1 located near the 5´ end of xsepks1. However, sequencing of the 1.05 kb band started to get mixed peaks after the same “TCCTAG” region using both primers, possibly due to cross contamination during gel extraction. The size difference between the 1.1kb band and 1.05 kb band approximated the size of the 63 bp intron. Further analysis of the 3´ end sequence chromatogram from the 1.05 kb product (with the intron removed) showed that the poly-A tail of the mRNA started almost immediately after the attB2 site. Since the attB2Cse-C primer binding site is located at 125 bp upstream of a putative poly-A signal “AAATAAAA” found at 700 bp downstream of the putative xsepks1 stop codon in pJpks-ID5 (Section 4.5.2.3) and would not be included into the pTA-Cse3 expression vector, the polyA-tailing could be initiated by the amyB terminator (TamyB) in the pTA-Cse3 plasmid (see Figure 5.6).
In Chapter 4 Section 4.5.3.1, it was noted that a discontinuous region with low homology to a reductase (R) domain was found at the 3’ end or C-terminal of xsepks1 (Figure 4.21). Therefore, it was expected that the splicing of intron2 would result in extension of the xsepks1 ORF or restore the discontinuous R-domain at the C-terminal. Surprisingly, translated sequence of the xsepks1 cDNA showed that removal of intron2 resulted in no extension to the XsePKS1 peptide chain length, but changes in the last few amino acids and a shift of the stop codon from 2165 to 2153 (Figure 5.13). The resultant XsePKS1 peptide chain length (total 2152 amino acids) was shorter by 12 amino acids compared to the translated sequence without the removal of intron2 (2164 amino acids).

Figure 5.12. Sequencing of 3’ cDNA ends (1.05 kb 3’ RACE product, Figure 5.11) showing splicing of an intron. (Top): pairwise alignment of xsepks1 3’ cDNA end sequence to genomic sequence; (Bottom): electrophoretogram showing no mixed sequence was observed after the “TCCTAG” region.

Figure 5.13. Pairwise alignment of the XsePKS1 transcripts: 1) with and 2) without the intron2 removed. Arrows and (*) indicates the position of stop codons.
5.3.2.3 Chemical analysis of _A. oryzae_ transformants

Preliminary chemical analysis was carried out at I. Fujii’s lab in Tokyo for transformants 1, 3, 4, 5 and a control _A. oryzae_ strain (carrying empty pTAex3 plasmid) before the PCR screening and transcription analysis of the _A. oryzae_ transformants was performed at the lab in Melbourne.

Transformant 1 was obtained from the first transformation (see Section 5.3.2.1) and therefore was the first to be analysed for metabolite production. HPLC analysis of the extracts from the medium for transformant 1 showed three major products (arbitrarily named compounds A, B and C) at retention times (R_t) 10.0, 12.6 and 24.0 min not found in the control culture (Figure 5.14, Control & Transformant 1a). The \( \lambda_{\text{max}} \) values of A, B and C were 275, 308 and 280 nm respectively (Figure 5.14). The fermentation was repeated for transformant 1 along with transformants 3, 4, 5 (obtained from the second transformation) and the control strain under the same conditions. In the second batch fermentation, only compound C was observed in the extract of transformant 1. Compound C also appeared as the only compound not found in the control culture but observed in the cultures of transformants 3, 4, and 5. The chromatograms for the culture extracts of transformants 1 (2\(^{nd}\) fermentation), 3, 4, and 5 were nearly identical (Figure 5.14, Transformant 5).

In a separate fermentation experiment, orange-pigmented mycelia were observed at the neck of the flask (forming a ring) after an accidental prolonged fermentation of transformant 6 (4 days). The orange pigment was neither found previously in the control culture nor in the other transformants. Due to the time limitation in I. Fujii’s lab, the extract from transformant 6 was not analyzed until later.
Figure 5.14. HPLC analyses of medium extracts from *A. oryzae* transformants. Bottom: UV spectra of the peaks at Rt A) 10.0, B) 12.6 and C) 24.0 min monitored at 240 – 350 nm.
To identify if XsePKS1 produced any typical lichen compounds in the A. oryzae transformants, the same extracts obtained in Fujii’s lab were reanalyzed in J.A. Elix’s lab in Australian National University to match against a library of lichen compounds (Figure 5.15 – 5.17). In the extracts from transformant 1 (1st fermentation), compounds A and B were detected but compound C was not. A peak (compound D) not detected previously was now found at $R_t = 18.6$ min (Figure 5.15, Transformant 1a). For transformant 1 from the second fermentation, a peak with UV spectra corresponding to compound C ($R_t = 24$ min) remained in the extracts (Figure 5.15, Transformant 1b). Transformants 3, 4 and 5 had showed almost identical chromatograms (Figure 5.15, only Transformant 5 shown). Besides compound C, two well-resolved peaks (compounds E and F, with $R_t$ values of 7.8 and 8.9 min) not observed in the previous chromatograms were found in transformants 3, 4, and 5 with the $R_t$ at 7.6 and 8.9 min (Figure 5.15, Transformant 5).

Neither compounds A nor B produced >90% match to any of the compounds in the library, which contains 1075 lichen compounds (including $\beta$-orsellinic acid, methylacetophlorophenone and derivatives). The UV spectrum of compound C showed a 94.8% matches to methylisoplacodiolic acid, but the $R_t$ of methylisoplacodiolic acid is much longer (Figure 5.16). Compound D produced a similar spectrum to compound B and had its greatest match to haemoventosin (95.6%) (Figure 5.16).

Compound E has its highest match to ellagic acid (93.8%), whereas compound F showed an almost significant high match (98.4%) to 5-hydroxy $\beta$-orsellinic acid (Figure 5.17). However, for a target compound to be identified, it must have a ≥99.9% match to a spectrum in the reference library. Both compounds E and F also appeared in the control culture, albeit at relatively low concentrations (Figure 5.17, bottom), and so are unlikely to be the product of xsepks1. Overall, none of the expected compounds ($\beta$-orsellinic acid, methylacetophlorophenone or related derivatives) or any of the lichen compounds in the reference library was detected in the extracts of A. oryzae transformants (1, 3, 4 and 5) carrying the xsepks1.
Figure 5.15. HPLC chromatograms of extracts from transformant 1 (a, 1st fermentation; b, 2nd fermentation) and transformant 5, analyzed in J.A. Elix's lab.
Figure 5.16. UV spectra of compounds B, C and D in the extracts of *A. oryzae* transformants (Figure 5.15) and the library search results.* The structure of methylisoplacodiolic was not resolved, but a CH$_2$ is known to present in addition to the isoplacodiolic acid based on mass spectroscopy (J.A. Elix, pers. comm.).
Figure 5.17. UV spectra of compounds E and F found in transformants 3, 4 and 5. Bottom: HPLC chromatogram of the extract from the A. oryzae control strain carrying the empty pTAex3 plasmid; compounds E and F were also detected in the extract.
The medium extract of transformant 6, which was not analysed in I. Fujii’s lab, was analysed in J.A. Elix’s lab. Besides the peak for compound C, an additional peak (compound G) at $R_t = 27.3$ min not found in other transformants or the control strain was detected (Figure 5.18). Since the orange pigmentation was evident in the mycelium of transformant 6, the metabolites from the mycelium was also extracted (in I. Fujii’s lab) for HPLC analysis. The orange mycelial extract from the mycelium resulted in the HPLC chromatogram in Figure 5.18 (Transformant6, mycelium), where compound C was not detected, but a high concentration of compound G was. It is not clear if compound G was responsible for the orange pigmentation, as the HPLC spectrophotometric detector equipped with the HPLC could not operate at visible wavelength. The UV spectrum of compound D showed an almost significant match (99.0%) to subpsoromic acid (Figure 5.17, bottom), which is a colourless depsidone found in the lichen *Ocellularia praestans* (Thelotremataceae) (Elix et al. 2000). A peak corresponding to ergosterol was also detected in the mycelium extract at $R_t = 32.6$ min, as expected.

Although the orange pigment was not detected in the *A. oryzae* control culture, the pigment is unlikely to be a product of *xsepks1*. Since it was not produced until the later growth stages, it is possible that the orange pigment is a secondary metabolite of the original *A. oryzae* host. This hypothesis was supported by the observation of orange pigmentation in some transformants, after prolonged storage on CD-starch (induced) and Czapek-Dox (non-induced) agar media (Figure 5.19). The presence of the orange pigmentation and its intensity also varied from plate to plate (replicates of the same strain) under the same conditions.
Figure 5.18. HPLC analyses of extracts from the medium and mycelia of transformant 6 culture. Bottom: UV spectrum and library search results of compound G.
Figure 5.19. Orange pigmentation appeared in the mycelium of *A. oryzae* transformant 1.

Compounds A, B, C and G were the compounds that are unique to the *A. oryzae* transformants and (not found in the control strain). To obtain more of these compounds for further identification and structural elucidation, two batches of fermentations were run for transformants 1, 4, 5 and 6 in the laboratory at RMIT University. Neither attempt reproduced any of the peaks (compound A, B, C and G) found in previous cultures grown in I. Fujii’s laboratory, even though transcriptional analysis showed that *xsepks1* was actively transcribed in the fungi (Section 5.3.2.2).
5.3.3 Discussion

5.3.3.1 Transformation and expression of xsepks1 in A. oryzae

The aim of the study reported in this chapter was to characterize the function of xsepks1 by heterologous expression in A. oryzae. It was hypothesized that xsepks1 or other clade III NR-PKS genes are involved in the biosynthesis of β-orsellinic acid or methylphloroacetophenone, the two putative precursors to the major secondary metabolites in X. semiviridis, i.e. succinprotocetraric acid [1], fumarprotocetraric acid [2], and usnic acid [3] (Figure 5.20). Besides these three compounds, other β-orsellinic acid-derived depsidones related to succinprotocetraric acid and fumarpotocetraric acid (i.e. conprotocetraric acid, consuccinprotocetraric acid, confumarpotocetraric acid, protocetraric acid, and virensic acid) are also present at lower concentration in the lichen extract (Section 3.2.2.2, Table 3.1).

Based on the HPLC-UV analysis, neither the expected β-orsellinic acid nor methylphloroacetophenone was detected in all of the extracts from the A. oryzae transformants carrying xsepks1. Although compound F showed a 98.4% similar UV spectrum and retention time to 5-hydroxy-β-orsellinic acid, the compound was also present in the control culture and therefore unlikely to be a product of xsepks1. It is possible that xsepks1 produces other compounds and not β-orsellinic acid or methylphloroacetophenone. Several compounds (A, B, C, D and G) unique to the A. oryzae transformants and not the control strain were detected in the fermentation cultures performed at the lab in University of Tokyo. Among them, compound C was detected in most transformants (1, 3, 4, 5 and 6). Compounds A, B and D were only detected in transformant 1, while compound G was only detected in transformant 6 after extended fermentation. Nevertheless, the overall reproducibility of the metabolic profile of each culture batch is low and none of these compounds (A, B, C, D and G) or any new compound was detected in the subsequent rounds of fermentation carried out in the lab at RMIT University even though xsepks1 was being actively transcribed.
Figure 5.20. HPLC analyses of the methanol extract from the *X. semiviridis* thalli and the structural formula of the three major polyketide compounds detected - succinprotocetraric acid [1], fumarprotocetraric acid [2], and usnic acid [3]. The retention time (Rt) for [1], [2] and [3] were 20.6 min, 22.4 min and 28.2 min respectively. [1] and [2] are originate from β-orsellinic acid, while [3] is from methylphloroacetophenone.

The success expression of the *xsepks1* and synthesis of the corresponding polyketide product requires 1) the integration of the complete/ intact *xsepks1* gene into the *A. oryzae* chromosome, 2) transcription of the full-length *xsepks1* into mRNA, and 3) translation, folding and post-translational modification (e.g. phosphopanteinylation) of the peptide chain to yield a functional PKS protein (Figure 5.21). To explain the HPLC-UV results of the *A. oryzae* transformants observed above, each step in the process from gene to product was systematically investigated below (Figure 5.21).

1) Integration of *xsepks1* into *A. oryzae*

PCR results (with RT52KS-F/R primers) showed that the frequency of integration of the plasmid DNA (pTA-Cse3) into *A. oryzae* chromosomes was relatively high, as
fragments of xsepks1 were detected in all of the transformants analysed (Figure 5.21, table). However, KS-CMeT PCR showed that the xsepks1 ORF only retained its integrity in four out of the seven colonies analysed (transformants 3, 4, 5 and 6). Since transformant 1 might be an abortive strain with a truncated xsepks1, compounds A, B, C and D detected in transformant 1 could be an artefact.

2-3) Expression of xsepks1 in the A. oryzae transformants

One-step RT-PCR failed to detect the transcription of 5´xsepks1 in transformant 1 but did in others (Figure 5.21, table). This was expected, as transformant 1 might carry a truncated xsepks1. The two-step RT-PCR confirmed the transcription of complete xsepks1 into mRNA in transformants 3, 4, 5 and 6; this was validated further by the 3´ RACE experiments. As the transcriptional analyses were carried out in shake-flask fermentations in the laboratory at RMIT (RNA was obtained from the same induced cultures used for organic extraction), the absence of compounds A, B, C, D and G in these cultures (transformants 3, 4, 5 and 6) contradicts the premise that one of these compounds may be a product of xsepks1. This poses a possibility that the XsePKS1 protein may be translated but not be functioning in the A. oryzae transformants (3, 4, 5 and 6) and that compounds A, B, C and D observed in earlier rounds of fermentation are merely artefacts (e.g. stress-induced compounds caused by transformation). The absence of these compounds in the control strain (A. oryzae with empty pTAex3 plasmid) could be explained; as the strain was constructed previously and has been subcultured for several generations in I. Fujii’s lab, its genetic and metabolic profile is expected to be more stable compared with the newly transformed A. oryzae. The hypothesis that compounds A, B, C and D are not products of xsepks1 is also in agreement with the detection of these compounds in transformant 1, a potentially abortive transformant (not possessing a contiguous xsepks1 and the gene was not transcribed).

The translation of xsepks1 in the A. oryzae transformants could not be confirmed. Protein extraction followed by SDS-PAGE was carried out for transformants 1, 3, 4, 5 and control strain after starch-induction culture, but the gel showed no significant difference in concentration from the control strain at the expected band size (250 kD) (data not shown). According to I. Fujii (pers. comm.), this was also the case for some A. oryzae expressing a functional heterologous PKS using the pTAex3 system in previous studies.
Figure 5.21. Summary of results for expression of xsepks1 in *A. oryzae* and factors that might have affected the outcome.
5.3.3.2 Functionality of XsePKS1 protein in A. oryzae

The main difficulty in detecting the product for XsePKS1 is that the identity of the compound itself is unknown. If a compound was produced by XsePKS1 but present at low concentration, e.g. due to low productivity, or the product was broken down by the host metabolism, it would be overlooked easily. The systematic analysis in previous section and Figure 5.21 suggests that the protein (XsePKS1) might be translated but dysfunctional (not producing any compound or only at very low concentration) and that all of the unique compounds detected are unlikely to be products of XsePKS1. Even though the main catalytic domains KS, AT, ACP and CMeT of XsePKS1 are strongly conserved and possess the key amino acids in the active sites, several possible factors could result in a dysfunctional PKS protein:

1) xsepks1 may be a truncated PKS gene that has lost its function in X. semiviridis

As discussed in Chapter 4 Section 4.5.3.1, XsePKS1 lacks the reductase (R) domain found in PKSCT and AsPKS1 (MOS), although a region of low homology to the R-domain was found downstream of XsePKS1 after the putative stop codon. The splicing of intron2 at the 3’ end of xsepks1 also did not restore the R-domain after splicing. The R or TE/CYC domain is thought to be involved in polyketide chain release in the PKS. The lack of an R-domain or TE/CYC domain in XsePKS1 could result in an inefficient release of the final product and low productivity. The non-LTR retrotransposon found downstream of xsepks1 could be responsible for the disruption of the XsePKS1 C-terminal domain (see Chapter 4 Section 4.5.3.2). To test for this hypothesis, a chimeric PKS gene with the R-domain or TE/CYC domain attached to the C-terminal of XsePKS1 could be constructed and expressed in a heterologous host.

2) an additional accessory protein crucial to the biosynthesis may be required

This is the case in the lovastatin biosynthesis, whereby an additional enoyl reductase (ER)-like protein (LovC) is required for synthesis of the nonaketide moiety (dihydromonacolin L) of lovastatin due to a mutation in the ER domain in lovastatin nonaketide synthase (LNKS) (Kennedy et al. 1999). Similarly, the esterase-like xseB upstream of xsepks1 or another protein could be required for the efficient production of the XsePKS1 product due to the loss of the R-domain in xsepks1. This hypothesis
could be tested by co-expression of xsepks1 with xseB and other biosynthetic genes in proximity (perform genome walking to recover the complete xseB ORF and identify any biosynthetic genes upstream of xsepks1).

3) an advanced starter unit may be required for the biosynthesis of XsePKS1 product

As discussed in Chapter 4 Section 4.5.3.1, a starter-unit acyltransferase (SAT) is found at the N-terminal of XsePKS1. So far, the function of PKSCT (a close homolog of XsePKS1) has only been proven in citrinin biosynthesis by gene disruption (loss of citrinin production) and expression of the gene (pksCT) in M. purpureus (increased citrinin production) (Shimizu et al. 2005; Shimizu et al. 2006). Neither report identified the immediate polyketide product of PKSCT or revealed the biosynthetic steps catalyzed by PKSCT. Instead of via the citrinin pathway proposed by Sankawa et al. (1983) for A. terreus, Cox (2007) proposed that the a methylated diketide starter unit may be involved in biosynthesis of citrinin in M. purpureus. Neither citrinin nor a related compound (e.g. the precursor proposed for A. terreus) has so far been detected in A. oryzae transformants carrying M. purpureus pksCT regulated under the PamyB promoter (I. Fujii, pers. comm.); it is therefore possible that an advanced starter unit as proposed by Cox (2007) is required for the production of citrinin in M. purpureus. Further evidence that supports this hypothesis can be found in a study (O’Callaghan et al. 2003) that reported a HR-PKS gene (lc35-12) from A. ochraceus (GenBank accession AAP32477) involved in the production of ochratoxin A (a mycotoxin structurally similar to citrinin but with a phenylalanine attached to the citrinin-like core, Figure 5.22). Only a fragment of the HR-PKS gene was cloned (near the AT domain, 1506 bp), but disrupting the gene lead to the loss of ochratoxin A production in A. ochraceus (O’Callaghan et al. 2003). The LC35-12 is most similar to clade I HR-PKSs (based on BLASTp), such as C. heterostrophus PKS5 (38% protein identity at the AT domain) and A. terreus lovastatin diketide synthase (30%). This suggests that LC35-12 is possibly involved in the biosynthesis of a reduced diketide starter unit that is required for the production of the citrinin-like core in ochratoxin A.

Since the SAT domain of XsePKS1 showed similarity to the N-terminal of M. purpureus PKSCT (40% protein identity), it is possible that an advanced starter unit is required for the biosynthesis of the corresponding polyketide product. The required advanced starter unit (or analogs that could be processed by the XsePKS1 SAT
domain) might be available in the *A. oryzae* host, but its biosynthesis is not coregulated under the *amyB* promoter in the pTA-Cse3 plasmid. Therefore, if the starter unit has ever been produced, its timing of production would be expected to vary from the expression of *xsepks1*. This might explain the high variability in the metabolic profiles of the *A. oryzae* transformants from batch to batch. The requirement for an advanced starter unit could be confirmed by feeding the *A. oryzae* transformant cultures with alkylacyl-CoAs of different chain lengths (e.g. C2-C8).

![Figure 5.22. Structural formula of Ochratoxin A. Box, citrinin-like moiety; in grey, phenylalanine moiety.](image)

4) specific conditions may be required for the functioning of XsePKS1

The lack of a recognisable metabolite in *xsepks1*-transformed *A. oryzae* has parallels to a comparable unpublished study, in which a clade III NR-PKS gene with similar domain architecture (Figure 5.23) was cloned from the cultured *Cladonia grayi* mycobiont DNA library and transferred to *A. nidulans* (D. Armaleo, pers. comm.). The *C. grayi* PKS was regulated by the native lichen promoter in the *A. nidulans* host and transcription analysis showed that the introns in the transcribed PKS gene were spliced correctly (D. Armaleo, pers. comm.). However, as in the case for *xsepks1* in *A. oryzae*, no new metabolite was detected in the *A. nidulans* transformant culture transcribing the *C. grayi* PKS gene. It is possible that the *C. grayi* PKS is not functioning in the *A. nidulans* host due to the same reasons (2 and 3) as described above for *xsepks1*, but an alternative explanation is equally possible – specific conditions are required for the functioning of lichen PKS proteins in a filamentous fungal host.

Sinnemann et al. (2000) observed that the complementation of *pyrG* from the lichen *Solorina crocea* in an *A. nidulans* host is temperature-dependent. In the study, a *pyrG* gene, encoding orotidine 5'-monophosphate decarboxylase, cloned from *S. crocea*, was transferred into a uridine-dependent *A. nidulans* mutant host and
expressed under the regulation of a trpC promoter. The A. nidulans transformants could only grow at 24°C and not above (whereas the A. nidulans strain normally grows optimally at 37°C). This effect was ascribed to the source of the pyrG protein, which is a lichen that grows in subalpine and alpine environments among snow-bed communities (Sinnemann et al. 2000). It was reasoned that the metabolism and enzymes of S. crocea may be adapted to lower temperatures due to its natural habitat and the protein might not fold properly at higher temperature. Although X. semiviridis and C. grayi do not grow at such low temperature, other conditions might be required for the proper folding and functioning of the enzymes. Similar studies involving complementing an auxotrophic Aspergillus strain with a lichen gene could be carried out for X. semiviridis in future to determine if there is any specific condition required for the functioning of its proteins in a heterologous host.

Figure 5.23. A PKS gene cloned from the lichen Cladonia grayi. The small triangles indicate the intron sites (D. Armaleo, unpublished data). KS, ketosynthase; AT, acyltransferase; ACP, acyl carrier protein; CMet, methyltransferase; CYC, Claisen cyclase, P-450, cytochrome p450 protein. Instead of producing β-orcinol or methylphloroacetophenone derivatives, C. grayi produced the orcinol depside (4-O-demethylsphaerophorin) and orcinol depsidones (grayanic and 4-O-demethylgrayanic acids) in its natural thalli and mycobiont culture (Culberson & Armaleo 1992). Since a CMet domain is not required for biosynthesis of orcinol depsides and depsidones, this PKS might be responsible for the biosynthesis of other undetected metabolites.

5.3.3.3 Possible functions of clade III NR-PKSs in lichens

In Chapter 3, it was hypothesized that clade III NR-PKSs, which include xsepks1, may be involved in the biosynthesis of β-orsellinic acid or methylphloroacetophenone as precursors for β-orcinol depsides/ depsidones and usnic acid. Although the results from this study did not confirm the function of xsepks1, the results from two very recent studies supported this hypothesis. In the first study, A. strictum AsPKS1, a clade III NR-PKS, produced 3-methylorcinaldehyde when expressed in an A. oryzae
host (Bailey et al. 2007). The compound 3-methylorcinaldehyde (or $\beta$-orcinaldehyde) is structurally and biosynthetically related to $\beta$-orsellinic acid and methylphoroacetophenone as all three compounds result from cyclization of a methylated tetraketide chain, but result from different cyclization and release mechanisms (Figure 5.24). In the second unpublished study, a clade III NR-PKS gene ($\textit{pksTI}$) from $\textit{A. terreus}$ was cloned and characterized (I. Takayuki, T. Kushiro and Y. Ebizuka, pers. comm.). $\textit{A. terreus}$ PKSTI produced 3,5-dimethylorsellinic acid when expressed in an $\textit{A. oryzae}$ host; 3,5-dimethylorsellinic acid is structurally similar to $\beta$-orsellinic acid but has an extra methyl group. It is the key precursor for biosynthesis of terretonin, a polyketide-terpenoid mycotoxin in $\textit{A. terreus}$ (McIntyre & Simpson 1981). A conserved region similar to alpha/beta hydrolase (pfam07859) and esterase/lipase domain (Aes, COG0657) was detected at the C-terminal of this PKS instead of a CYC or R-domain; it may be involved in release of 3,5-dimethylorsellinic acid by hydrolytic cleavage. Based on these two studies, it is very likely that the other clade III NR-PKSs detected in $\textit{X. semiviridis}$ (see Section 4.3.3.1 & 4.4.3.2) and in other lichens (Section 3.5.3.4 and 4.4.2.2) are responsible for the biosynthesis of the methylated tetraketide monoaromatic precursors (i.e. $\beta$-orsellinic acid and methylphroroacetophenone) and their function should be investigated in future.

![Diagram](image.png)

**Figure 5.24.** Proposed biosynthetic relationship between $\beta$-orsellinic acid, methylphloroacetophenone, 3-methylorcinaldehyde and 3,5-dimethylorsellinic acid. Note: The methylation by CMeT might occur after cyclization as proposed by Cox (2007).
CHAPTER 6 CONCLUSIONS AND FUTURE DIRECTIONS

Lichens produce a diverse array of bioactive secondary metabolites, in which many are unique to the organisms. Their potential applications, however, are limited by their finite sources and the slow-growing nature of the organisms in both laboratory and environmental conditions. As the demand for novel chemical entities in medicine and other applications continues to grow, molecular approaches to natural product discovery and drug development are likely to become more attractive. This thesis set out to investigate polyketide synthase genes in lichens with the ultimate goal of providing a sustainable source of lichen natural products to support these applications.

One of the main objectives of this study was to investigate the fungal PKS genes potentially involved in biosynthesis of typical lichen metabolites, in particular the coupled phenolics. At the inception of this project, the only available PKS sequences were the handful of partial KS domain from a phylogenetic study (Grube & Blaha 2003). New degenerate primers targeting KS domains of specific clades of PKS gene have been developed and tested on various lichen samples (mainly from the family Parmeliaceae). Using these primers, KS domains from PKS clades not previously identified in lichens were amplified, including the clade III NR-PKSs, PR-PKSs and HR-PKSs. The discovery of clade III NR-PKSs in lichens was especially significant. Based on the KS domain phylogenetic analysis and compounds detected in the individual lichens, the clade III NR-PKSs was hypothesized to be involved in the biosynthesis of β-orsellinic acid and methylphloroacetophenone – the monoaromatic precursor for many lichen coupled phenolic compounds such as β-orcinol depsides/ depsidones and usnic acids. The newly developed degenerate primers could be used in future to detect and target a wider range of PKS genes in lichens and other fungi.

The study also showed the potential of using the clade III NR-type KS and CMeT domains amplified from the degenerate primers (NR3KS-F/R and NRMeT-F/R) to probe for new clade III NR-PKS genes in fungi. A method has been developed for cloning of a full-length PKS gene from environmental lichen DNA and a clade III NR-PKS gene (xsepks1) from X. semiviridis was successfully cloned. This is the first
report on the isolation of a full-length PKS gene from environmental lichen DNA. The domain architecture of \textit{xsepks1} is KS-AT-ACP-CMeT, as expected for a clade III NR-PKS, suggesting that the KS domain phylogenetic analysis is useful for prediction of functional domains in PKS. The KS domain phylogenetic analysis could potentially be used to facilitate the design of new degenerate primers that target a narrow subclade of PKSs with unique domain architecture, e.g. \textit{F. verticillioides} PGL1 homologs (KS-AT-ACP-R) and \textit{B. fuckeliana} PKS18/ \textit{A. terreus} ATEG\_10080 homologs (KS-AT-ACP-CMeT-Est).

Both \textit{A. nidulans} and \textit{A. oryzae} were tested as potential hosts for the expression of lichen PKS genes. Transcriptional analysis showed that \textit{A. nidulans} could potentially utilize the lichen PKS gene promoter and both fungal hosts could splice the introns of a lichen PKS gene. Several compounds unique to the \textit{A. oryzae} transformants carrying \textit{xsepks1} were detected, but they could not be reproduced in subsequent fermentations. Although the expected products (\(\beta\)-orsellinic acid, methylphloroacetophenone or similar monoaromatic compounds) were not detected in \textit{A. oryzae} transformants carrying the \textit{xsepks1} gene, the recent discovery of methylorcinaldehyde synthase from \textit{A. strictum} (Bailey et al. 2007) and 3,5-dimethylorsellinic acid synthase from \textit{A. terreus} (I. Takayuki, T. Kushiro and Y. Ebizuka, pers. comm.), as members of clade III NR-PKSs, supports the hypothesis that PKS genes of this clade might be responsible for the production of the two methylated tetraketide precursors (\(\beta\)-orsellinic acid and methylphloroacetophenone) in lichens. Thus, the function of other clade III NR-PKSs detected in \textit{X. semiviridis} and other lichens in this study should be investigated in future. The results of this study will help in the design of new approaches for targeting, cloning and heterologous expression of these clade III NR-PKS genes from lichens.

The recent discovery of type III PKSs as alkylresorcylic acid synthases in fungi (Funa et al. 2007) is relevant to the research on lichen polyketide biosynthesis. As discussed in Chapter 1 (Section 1.2.6.2.1), alkylresorcylic acids are a group of basic monoaromatic units, which act as the precursors for many lichen coupled phenolics, such as orcinol depsides/ depsidones/ depsones and dibenzofurans. The presence of type III PKSs in fungi and their ability to synthesize alkylresorcylic acids suggest that they might be involved in biosynthesis of orcinol-tye coupled phenolic
compounds in lichens and should also be a subject for future investigation along with the clade III NR-PKSs discovered in this study.

Overall, the work in this thesis demonstrated the prospect of using a molecular approach to access the lichen biosynthetic potential without going through the cumbersome culturing stage. However, the experimental results also revealed more questions that needed to be answered. These include the understanding of polyketide release mechanisms in PKSs, the incorporation of advanced starter units in NR-PKSs by the SAT domain, and the functioning of lichen PKSs in a heterologous system. Most PKSs also do not function independently in organisms, therefore understanding of the interactions between the PKSs and its accessory proteins will also be an important aspect of PKS research in future. Significant progress has been made recently by several studies in expression of fungal PKSs and other biosynthetic genes in bacterial systems, and in vitro chemoenzymatic synthesis using these heterologously-expressed proteins (Balibar et al. 2007; Ma et al. 2007; Schneider et al. 2007). These bacterial systems could potentially be utilized for characterizing the function of PKSs and other biosynthetic genes from lichens, as the presence of the biosynthetic background from the host, which often complicates the chemical analysis, can be avoided. Bacteria are also faster-growing than fungi and the in vitro systems allow feeding of advanced precursors without interference from the host metabolism.

Currently, the sequencing of the first lichen genome (Xanthoria parietina) is in progress at the DOE Joint Genome Institute. The genome sequence of X. parietina is likely to reveal many biosynthetic genes and gene clusters. The work presented in this thesis will provide the groundwork for investigation of these biosynthetic genes as well as isolation and expression of novel biosynthetic genes from environmental lichen samples/ unculturable fungi.
References


cleavage pathways. *Natural Product Reports* 15, 513-530.


product genes in marine and freshwater cyanobacterial cultures and genomes.  
*Applied and Environmental Microbiology* 71, 7401-7413.

Ehrlich KC, Cary JW, Montalbano BG, 1999. Characterization of the promoter for the gene encoding the aflatoxin biosynthetic pathway regulatory protein AFLR.  
*Biochimica Et Biophysica Acta-Gene Structure and Expression* 1444, 412-417.

*Gene* 230, 249-257.

*Biochimica Et Biophysica Acta-Gene Structure and Expression* 1576, 171-175.

*Australian Journal of Chemistry* 37, 1531-1538.


*Australian Journal of Chemistry* 39, 613-624.

*Australian Journal of Chemistry* 40, 1451-1464.

*Flora of Australia* 55, 1-360.


Elix JA, Wardlaw JH, 1996. Hypoalectorialic acid and conechinocarpic acid, two new benzyl esters from lichens.  

*Acta Universitatis Upsaliensis, Symbolae Botanicae Upsalienses* 32, 103-121.

Elix JA, Wardlaw JH, 1999. The structure of chalybaeizanic acid and quaesitic acid, two new lichen depsidones related to salazinic Acid.  


Hoffmeister D, Keller NP, 2007. Natural products of filamentous fungi: enzymes, genes,
and their regulation. *Natural Product Reports* 24, 393-416.


Journal 15, 299.


McCaffery LF, 1997. The total synthesis of some lichen tridepsides and a diphenyl ether, PhD. Thesis, Department of Chemistry. Australian National University, Canberra.


Polborn K, Steglich W, Connolly JD, Huneck S, 1995. Structure of the macrocyclic


Rose MS, Yun SH, Asvarak T, Lu SW, Yoder OC, Turgeon BG, 2002. A decarboxylase encoded at the Cochliobolus heterostrophus translocation-associated Tox1B locus
is required for polyketide (T-toxin) biosynthesis and high virulence on T-cytoplasm maize. *Molecular Plant-Microbe Interactions* 15, 883-893.


Todd RB, Fraser JA, Wong KH, Davis MA, Hynes MJ, 2005. Nuclear accumulation of the GATA factor AreA in response to complete nitrogen starvation by regulation of


