INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI[®]
800-521-0600

LOCALIZATION, CLONING, AND IDENTIFICATION OF
THE *tdt* GENES FROM *Pseudomonas* sp. A19-6a AND
THEIR DISTRIBUTION IN PULP MILL MIXED LIQUOR
TREATMENT SYSTEMS

by

Christine A. Morgan

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Biology
Ottawa-Carleton Institute of Biology

Carleton University
Ottawa, Ontario, Canada

April 1999

Copyright © 1999, by Christine A. Morgan.

**National Library
of Canada**

**Acquisitions and
Bibliographic Services**

**395 Wellington Street
Ottawa ON K1A 0N4
Canada**

**Bibliothèque nationale
du Canada**

**Acquisitions et
services bibliographiques**

**395, rue Wellington
Ottawa ON K1A 0N4
Canada**

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-42801-X

Canada

The undersigned hereby recommend to the
Faculty of Graduate Studies and Research
acceptance of this thesis,

LOCALIZATION, CLONING, AND IDENTIFICATION OF THE *tdt* GENES FROM
Pseudomonas sp. A19-6a AND THEIR DISTRIBUTION IN PULP MILL
EFFLUENT TREATMENT SYSTEMS

submitted by
Christine A. Morgan, B. Sc.

In partial fulfilment of the requirements
For the degree of Doctor of Philosophy

Chair, Department of Biology
Dr. R. C. Wyndham

Thesis Supervisor
Dr. R. C. Wyndham

External Examiner
Dr. S. N. Liss

Carleton University, April 1999.

ABSTRACT

Resin acids are tricyclic diterpenic carboxylic acids that are toxic to aquatic life when released in high concentration in pulp mill mixed liquors. These naturally formed hydrocarbons are readily biodegradable by bacteria and fungi, although many of the mechanisms involved are still unknown. I report the enrichment isolation of 13 bacteria from the secondary treatment system at the Avenor bleached kraft pulp mill in Dryden, Ontario (Weyerhaeuser). These isolates are able to degrade various resin acid congeners commonly found in pulp mill mixed liquors. The collection was characterized biochemically, phenotypically, and genotypically for the purpose of selecting a candidate for resin acid catabolic gene probe development.

I report the localization, cloning, and sequencing of the *tdt* locus from *Pseudomonas* sp. A19-6a that comprises seven open reading frames. I demonstrate that two of the genes, *tdtL* and *tdtD*, that encode a coenzyme A ligase and a cytochrome P450 monooxygenase, respectively, are necessary for the degradation of abietic acid. Multiple sequence alignment confirms the identity of these two genes. The degree of amino acid similarity between TdtD and TerpC identifies it as a new member of the P450 superfamily. Although the requirement for the other open reading frames in resin acid metabolism is not known, their presence between the *tdtL* and *tdtD* genes suggests that they too may form a part of a *tdt* catabolic pathway. Identification of four of these genes was confirmed by multiple sequence alignment of the gene products with known

transcriptional regulators (*tdtR*), isomerase/decarboxylases (*tdtA*), dehydrogenases (*tdtB*), and thiolases (*tdtC*). Homology search of *tdtS* detected hypothetical proteins only.

Bacterial isolates collected from pulp mill mixed liquor secondary treatment systems located world wide were screened for the presence of the *tdt* genes. The results of this screening indicate that these genes are widely distributed geographically in genetic arrangements that are in some cases similar and in others different from A19-6a. The *tdtL* and *tdtD* gene probes may be suitable for use as tools to study bacterial populations able to degrade resin acids in pulp mill mixed liquor treatment systems.

ACKNOWLEDGEMENTS

How do I thank all of the people who made reaching this point possible. Mike Smith who recommended that Science was the only way to realize my dream. Ruth Lifeso whose encouragement helped me to realize this goal. Peter Kruuz who understood the sacrifices I made to embark upon this road. Suzanne Patterson who exemplified the point I wanted to reach. My partner Ian Nalder. He and Sue are life-long friends to cherish. Michelle Peel who befriended me when most I needed one. Dorothy McColl a peer, a colleague, and a friend. Miguel Providenti whose concern and friendship are endearing. Iain Lambert whose silent support has encouraged me to continue. Ed Bruggink and Joan Mallett, who are always available when I need a friend. To these individuals, and many others unmentioned, I give my heartfelt thanks for enabling this journey.

I give special thanks to Cam Wyndham for his stoicism. Without his enduring faith and unwavering belief in me I could not have achieved this end.

To my family, my appreciation for accepting my inaccessibility over the course of my academic career. To Mary for being supportive. To Jimmy for understanding. To Leslie for being my older brother. To Susan who is never far. Most of all to my father, Keith Morgan, for teaching me that giving your best is achievement enough.

My deepest gratitude goes to Mark Harper. Throughout the course of my studies he has stood both beside me and behind me. Across the peaks of my highs and through the valleys of my lows he has been steadfast in his support. I thank you.

This thesis is dedicated to my daughter, Lian, whose name whispers of the desecration we cause.

TABLE OF CONTENTS

ACCEPTANCE SHEET	II
ABSTRACT	III
ACKNOWLEDGEMENTS	V
TABLE OF CONTENTS	VII
ABBREVIATIONS	XII
LIST OF APPENDICES	XI
LIST OF TABLES	XIV
LIST OF FIGURES	XVII
CHAPTER 1	1
ENVIRONMENTAL SIGNIFICANCE OF TERPENOIDS	1
INTRODUCTION	2
General	3
BACKGROUND	4
Plant resin acid metabolism	4
Role of resin acids	8
Occurrence of resin acids in the environment	11
Resin acid toxicology	14
Government regulation of pulp mill wastewater toxicity	18
Treatment of pulp mill mixed liquors	19
EXPERIMENTAL DESIGN	20
Rationale	20
Hypothesis	22
Objectives	22
Methods	23

CHAPTER 2	24
ISOLATION AND CHARACTERIZATION OF RESIN ACID DEGRADING BACTERIA	24
INTRODUCTION	25
Resin acid chemistry	25
Biodegradation and biotransformation of resin acids	29
Isolation and characterization of degraders from pulp mill mixed liquors	31
Extraction and analysis of resin acids	33
Objective I	35
MATERIALS AND METHODS	36
Culture techniques and isolation of putative degraders	36
Biochemical characterization of resin acid degraders	37
Determination of resin acid degradation	38
Genotypic analysis of selected isolates	39
RESULTS	41
Isolation and characterization of bacteria	41
Resin acid utilization	43
Hierarchical cluster analysis of isolates	47
Genotypic analysis of isolates	47
DISCUSSION	50
Isolation and characterization of resin acid degraders	50
Resin acid utilization	52
Genotypic relatedness of isolates	54
CONCLUSION	55
 CHAPTER 3	 56
TRANSPOSON MUTAGENESIS AND SUBCLONING OF TRICYCLIC DITERPENE DEGRADATION GENES FROM <i>PSEUDOMONAS</i> SP. STRAIN A19-6A AND MUTANT A76-47	56
INTRODUCTION	57
Targeted mutagenesis	57
Transposon delivery to environmental isolates	58
Inverse polymerase chain reaction	60
Cloning strategies	60
Objective IIa	61
MATERIALS AND METHODS	62
Transposon mutagenesis	62
Total genomic DNA extraction	63
Southern blotting and hybridization	64

Inverse polymerase chain reaction	65
Direct plasmid cloning of target gene	66
Preparation and screening of lambda libraries	66
RESULTS	71
Transposon mutagenesis	71
Inverse PCR and direct plasmid cloning of target DNA	71
Mapping of transconjugant A76-47	73
Preparation, screening, and plasmid sub-cloning of cosmid libraries	75
DISCUSSION	77
Transposon mutagenesis	77
Direct plasmid cloning and mapping of A76-47	78
Genetic recombination	79
Genomic libraries	80
CONCLUSION	80
 CHAPTER 4	 81
 CHARACTERIZATION OF THE <i>TDT</i> LOCUS FROM <i>PSEUDOMONAS</i> SP. STRAIN A19-6A	 81
INTRODUCTION	82
Catabolic gene organization	82
An operon encoding dehydroabietic acid degradation	85
Objective IIb	85
MATERIALS AND METHODS	87
Sequencing	87
Preparation of probes	89
Knock-out experiments	89
RESULTS	92
Sequencing of cloned DNA	92
Open reading frames	96
Development of resin acid catabolic gene probes	111
Knock-outs	111
DISCUSSION	116
Sequence analysis of mutant A76-47	116
Sequencing of the <i>tdt</i> locus	117
Genes involved in resin acid degradation	118
Characterization of the <i>tdt</i> genes	122
Relationship between <i>tdt</i> and <i>dit</i> loci	128
Development of <i>tdt</i> probes	128
Gene knock-outs	129
CONCLUSION	130

CHAPTER 5 **131**

DISTRIBUTION OF *TDT* GENES IN RESIN ACID DEGRADING BACTERIA ISOLATED FROM PULP MILL MIXED LIQUOR SECONDARY TREATMENT SYSTEMS **131**

INTRODUCTION	132
Pulping processes	132
Wastewater treatment	134
Detoxification of pulp mill mixed liquors	135
Fate of resin acids in receiving waters	137
Objective III	137
MATERIALS AND METHODS	139
Determining probe specificity	139
16S rDNA sequencing	139
Screening of environmental samples	140
RESULTS	142
Specificity of <i>tdt</i> probes	142
16S rDNA sequence analysis	145
Screening of environmental samples	145
Phenotypic analysis of probe positive strains	147
Probe analysis of environmental isolates	147
DISCUSSION	152
Specificity of <i>tdt</i> gene probes	152
Phenotypic characterization of probe positive strains	153
Distribution of the <i>tdt</i> genotype	154
CONCLUSION	156

CHAPTER 6 **157**

DISCUSSION, CONCLUSIONS, AND FUTURE WORK **157**

GENERAL	158
Objective I	158
Objective II	159
Objective III	159
DISCUSSION	160
Collection of resin acid degrading bacteria	160
Gene localization by transposon mutagenesis	162
Resin acid biodegradation genes	163
Resin acid catabolic gene probes	166
CONCLUSIONS	167

FUTURE WORK	168
REFERENCES	170
APPENDIX A	192
APPENDIX B	195
APPENDIX C	215

ABBREVIATIONS

ABA	abietic acid
Ap	ampicillin
<i>aph</i>	aminoglucoside phosphotransferase
AQA	quinat in minimal media A
<i>bla</i>	β -lactamase
bp	base pair
CoA	coenzyme A
CPS	casitone, peptone, starch
CTAB	hexadecyltrimethyl ammonium bromide
CY	casitone, yeast extract
DHA	dehydroabietic acid
<i>dit</i>	diterpene
dNTP	dinucleotide triphosphate
ELISA	enzyme-linked immunosorbant assay
ERIC	enterobacterial repetitive intergenic consensus
FPP	farnesyl pyrophosphate
GC	gas chromatography
GGPP	geranyl geranyl pyrophosphate
GPP	geranyl pyrophosphate
HPLC	high pressure liquid chromatography
IPA	isopimaric acid
IPP	isopentenyl pyrophosphate
IS	insertion sequence
Kb	kilobase pair
Km	kanamycin
kPa	kilopascal
L/R	left/right
<i>lacZ</i>	β -galactosidase

LPA	levopimaric acid
MCS	multiple cloning site
ml	millilitre
mM	millimolar
MS	mass spectrometry
NAA	neoabietic acid
NPT	neomycin phosphotransferase
OD	optical density
ORF	open reading frame
<i>ori</i>	origin of replication
<i>oriT</i>	origin of transfer
PA	pimaric acid
PCR	polymerase chain reaction
PEG	polyethylene glycol
psi	pounds per square inch
RA	resin acid
RPM	revolutions per minute
SDS	sodium dodecyl sulphate
SLS	sodium N-lauroylsarcosine
SM	sodium, magnesium
SSC	Saline sodium citrate
STE	sodium, tris, EDTA
TCA	tricarboxylic acid
<i>tdt</i>	tricyclic diterpene
TE	tris, EDTA
TYE	tryptone, yeast extract
UV	ultraviolet
X-gal	5-bromo 4-chloro 3-indolyl- β -galactoside
YFG	your favorite gene

LIST OF TABLES

TABLE 1-1. Acute and sublethal toxicity of dehydroabiatic acid (DHA) to <i>Salmo gairdneri</i> Richardson in fresh water. Summary of exposure concentration and duration and accumulated tissue residues observed, adapted from Oikari et al, 1983.	16
TABLE 2-1. Selected chemical characteristics of resin acid congeners used in this study. Sources as indicated.	28
TABLE 2-3. Summary of physiological characteristics of resin acid degrading bacteria isolated from Avenor bleached kraft pulp mill, Dryden, Ontario (Weyerhaeuser). For a complete breakdown of carbon sources used by each isolate see Table A1, Appendix A.	42
TABLE 2-4. Summary of fatty acid methyl ester analysis of resin acid degraders A19-6a and D11-13. The the probability of the chromatogram matching the type strain indicated is included.	43
TABLE 3-1. Summary of transposon mutagenesis results for resin acid degraders A19-6a and D11-9. All values are averages unless otherwise stated. RA stands for resin acid.	72
TABLE 4-1. List of primers used to construct sequence of tdt operon. All position numbers refer to the plus strand of the sequence (see Figure B1, Appendix B). Template DNA was YFG23 for CM primers and 25b/f for 25b/f primers.	88
TABLE 4-2. Construction of pUC128::probe sub-clones from YFG23. Position numbers refer to plus strand of tdt locus sequence, see Figure B1, Appendix B.	90
TABLE 4-3. Summary of restriction enzymes used to construct suicide vectors for gene knock-outs. Position numbers refer to the plus strand of the tdt locus (see Figure B1, Appendix B). Insert size refers to target DNA only and does not include that which may have been contributed by using a pUC128 MCS site for cloning.	90
TABLE 4-4. Features summary of tdt locus and open reading frames. The complete sequence including features and amino acid translations is given in Figure B1 Appendix B.	98
TABLE 4-5. Percent similarity between putative tdt gene products and aligned sequences shown in Figures 4-7 to 4-13 inclusive.	99

TABLE 4-6. Resin acid degradation levels (percent) for <i>Pseudomonas</i> sp. A19-6a transconjugants A76-47, A103-37, and A104-38 and gene knock-out mutants P450-R1 and P450-R3. Degradation for the wild type A19-6a strain (positive control) was 100%. All cultures were grown on abietic acid except as noted. Level of degradation is based on the amount of resin acid in solution in comparison with an uninoculated negative control.	114
TABLE 5-1. Source list of pulpmill samples provided by R. Fulthorpe, University of Toronto, and screened with tdt gene probes.	141
TABLE 5-2. Summary of tdt gene probe positive Sac II fragments illustrated in Figure 5-2. All band sizes are in kilobases. Signal intensity is indicated relative to the A19-6a wild type control; (+) weak positive, (++) moderate positive, (+++) strong positive.	144
TABLE 5-3. Summary of band size and intensity for probe positive strains illustrated in Figure 5-5. All band sizes for these Sac II digests are in kilobases. Signal intensity is relative to the A19-6a wild type control; (+) weak positive, (++) moderate positive (+++ strong positive.	151
TABLE A1. Summary of phenotypic characteristics of resin acid degrading bacteria isolated from the Avenor bleached kraft pulp mill in Dryden, Ontario (Weyerhaeuser). Characteristics were used for hierarchical cluster analysis and to calculate Jaccard Correlation Coefficient (Sneath, 1973) illustrated in Figure 2-5. Symbols are + positive reaction, - negative reaction, and / weak or borderline reaction. Note that latter were scored as negative during analysis.	193
TABLE A2. Jaccard correlation coefficients (Sneath, 1973) for Avenor bleached kraft pulp mill isolates. Values based on phenotypic data presented in Table A1. Correlation calculated from $JC=a/(a+b+c)$ where a = number of traits in both isolates, b = number of traits in first isolate, and c = number of traits in second isolate. Note that weak or borderline reactions are counted as negative.	194
TABLE B1. Summary of expected and observed fragments for mapping of tdt gene probes in <i>Pseudomonas</i> sp. A19-6a.	213
TABLE B2. List of restriction enzymes and corresponding cut sites in tdt locus cloned from <i>Pseudomonas</i> sp. A19-6a. Recognition sites for both palindromic and non-palindromic cutters are included. Position number is based on plus strand as shown in Figure B1. Blunt end cutters are identified (BE).	214

TABLE C1. List of strains screened by dot blot hybridization for determining specificity of tdt gene probes. Percent abietic acid degraded by each strain as determined by HPLC analysis is included. Hybridization results are positive (+), negative (-), or weak (/).	216
TABLE C2. Environmental isolate collection screened for distribution of the tdt genotype. HPLC peak area and percent resin acid degraded in batch culture are indicated for each strain. Probe positive isolates are highlighted for ease of reference.	217
TABLE C3. Summary of substrates used by environmental isolates that were probe positive for the tdt genes. These data were used to generate the dendrogram illustrated in figure 5-4 and the Jaccard Correlation Coefficients presented in Table C4.	218
TABLE C4. Jaccard Correlation Coefficients for probe positive strains isolated from pulp mill mixed liquor treatment systems. Values are based on phenotypic characteristics presented in Table C3.	219

LIST OF FIGURES

- FIGURE 1-1. Interface between plant primary and secondary metabolic pathways. Core structures generated in phenylpropanoid, polyketide, and isoprenoid biosynthesis are shown. Adapted from Kramer and Kozlowski, 1979 and Ellis, 1997.6
- FIGURE 1-2. Isoprenoid biosynthesis. Dimethylallyl pyrophosphate and isopentenyl pyrophosphate are the starter group and repeating subunit, respectively, used in the synthesis of all terpenoid compounds. The abietane skeleton is also shown as an example of tricyclic diterpene structure. Adapted from Loomis and Croteau, 1980; Croteau and Johnson, 1985; and Hanson, 1972.7
- FIGURE 1-3. Overview of plant terpenoid ecology. Note that arrows merely denote direction and do not account for the fact that some components of the diagram may be more ecologically significant than others. Partly adapted from Langenheim, 1994..... 12
- FIGURE 1-4. General scheme for the elimination of toxins by fish via mixed function oxygenase systems. Adapted from Stegman and Hahn, 1994. 17
- FIGURE 2-1. Structures of resin acid congeners used in this study. The abietane group includes abietic acid (ABA), neoabietic acid (NAA), dehydroabietic acid (DHA), and levopimaric acid (LPA). Pimaric acid (PA) and isopimaric acid (IPA) belong to the pimarane group.....27
- FIGURE 2-2. Points of oxidative attack on dehydroabietic acid by bacteria and fungi.....31
- FIGURE 2-3. Resin acid utilization by abietic and dehydroabietic acid enriched bacteria isolated from pulp mill mixed liquor. Abbreviations are: ABA, Abietic acid; DHA, Dehydroabietic acid; NAA, Neoabietic acid; PA, Pimaric acid; IPA, Isopimaric acid; LPA, Levopimaric acid. Degradation is expressed as the amount of resin acid remaining in solution following eight days of incubation in the presence of the different isolates, relative to control (uninoculated) flasks. Pimaric acid and isopimaric acid belong to the pimarane-group of resin acids. The rest belong to the abietane group.45
- FIGURE 2-4. Growth versus ABA utilization by four ABA enriched isolates found in bleached kraft pulp mill mixed liquor. Symbols are x = growth measured as A_{600} and + = concentration of ABA in mg/L. All tests conducted in triplicate.46

- FIGURE 2-5. Dendrogram of similarities between resin acid degrading bacteria isolated from pulp mill mixed liquor. Cluster analysis using average linkage between groups was based on phenotypic characteristics listed in Table A1. The relative scale lists Jaccard Correlation Coefficients (Sneath, 1973) from Table A2 based on the same physiological characters. Groups of isolates discussed in the text are distinguished at the Jaccard similarity coefficient of 0.23 and include phenotype cluster 1 (isolates A19-3 down to A19-2), and phenotype cluster 2 (from D11-13 down to D11-3b)..... 48
- FIGURE 2-6. Agarose gel (1.5 %) of ERIC-PCR products of 10 resin acid degrading bacteria isolated from pulp mill mixed liquor. The gel was loaded with 10 μ l of product and run for approximately five hours at 105 V (constant voltage). The marker in outside lanes is a 100 basepair ladder. 49
- FIGURE 3-1. Composite map of A76-47 derived from hybridization with eight pBW30::Tn ϕ hoA'-1 probes used to identify the orientation of insert. The location of known A19-6a sequence (Chapter 4) is also shown. Positioning of the restriction sites in A76-47 and pBW30 are in relation to the same sites in Tn ϕ hoA'-1. Abbreviations as follows: pBW30, suicide vector; Tn ϕ hoA'-1, Tn5 derived transposable element; RP4 *oriT*, RP4 origin of transfer region; R6K *ori*, R6K origin of vegetative replication; *bla*, β -lactamase gene conferring ampicillin resistance; *lacZ*, β -galactosidase gene enabling colorimetric selection; *aph*, Tn5 aminoglycoside phosphotransferase gene conferring kanamycin resistance; and *tnp*, Tn5 transposase gene enabling transposition. 74
- FIGURE 3-2. Map of wild type A19-6a corresponding to the region surrounding the mutation site in A76-47. The arrowhead indicates the site of insertional inactivation by pBW30::Tn ϕ hoA'-1 in A76-47. The location of the probe cut from YFG23 used to construct the map is indicated. Regions of DNA represented by the plasmid sub-clones YFG23 and 25b/f and the phage clone 25b are included. 76
- FIGURE 4-1. Catabolic gene organization in *Pseudomonas putida* strain U (A) and *Acinetobacter calcoaceticus* (B). A generalized overview of each pathway is included. Adapted from Olivera *et al.*, 1998 (A) and Gerischer *et al.*, 1998 (B). 83
- FIGURE 4-2. Diterpene catabolic pathway (*dit* operon) of *Pseudomonas abietaniphila* strain BKME-9. Selected steps of the pathway elucidated by GCMS metabolite and gene knock-out analyses are also shown. Adapted from Mohn *et al.*, 1998 and Martin and Mohn, 1999..... 86

- FIGURE 4-3. YFG23/-40 sequence showing the MCS from pUC128 (underlined) and λ FIX II™ (italics). The *Xba* I site is shown in bold and the asterisk represents the stop codon of the *tdtD* gene. The beginning of the minus strand of the target DNA next to the filled-in (TC) bases (highlighted) is also indicated. Feature label is shown below the sequence..... 93
- FIGURE 4-4. Summary of sequences spanning: A. the λ library cloning site; B. the pBW30::TnphoA'-1 junction; and C. the point of recombination in A76-47. The *Xba* I site is shown in A. in bold and bases filled in during cloning are highlighted. Underline denotes pUC128 and italics λ cloning vectors. The boxed region in C. shows the RP4 *oriT* with the *nic* site as indicated. Start positions for *tnp*, *tral*, and *tdtL* DNA are also shown. Feature label is given below the sequence. 94
- FIGURE 4-5. Sequence of the mutated region of A19-6a (see below) showing similarity to the sequence of RP4 *oriT*. Two regions of similarity are indicated by shading. Identical residues between the two sequences (*) and the insertion site of the suicide vector (▼) are indicated below the sequence.95
- FIGURE 4-6. Sequence map of *tdt* locus isolated from *Pseudomonas* sp. A19-6a. BLASTX search (Altschul *et al*, 1990) identified the genes as a CoA ligase (*tdtL*), a putative transcriptional regulator (*tdtR*), an unknown (*tdtS*), an isomerase (*tdtA*), a dehydrogenase (*tdtB*), a thiolase (*tdtC*), and a cytochrome P450 monooxygenase (*tdtD*). Cloning sites (plus strand numbering) used in library construction are indicated. Sequenced regions and template used are also shown..... 97
- FIGURE 4-7. Multiple sequence alignment of TdtL, an acyl-CoA ligase, from *Pseudomonas* sp. A19-6a and several sequences identified by BLASTX search (see Table 2 for sequence identification and references). Boxed region is acyl-adenylate binding site (Babbitt *et al*, 1992). Site of insertional mutagenesis by TnphoA'-1 is indicated by an arrowhead. Residues the same (*) or similar (.) are indicated below the alignment. Sequences are aligned in order of decreasing similarity..... 101
- FIGURE 4-8. Multiple sequence alignment of *tdtR* gene product *luxR* and *mtrR* regulatory proteins. Residues the same (*) or similar (.) are indicated below the alignment. The helix-turn-helix region of *mtrR* is also indicated. See Table 4-5 for references and percent similarity between the sequences. Sequences are aligned in order of decreasing similarity. 102
- FIGURE 4-9. Multiple sequence alignment of *tdtS* gene product with hypothetical proteins from *Sphingomonas aromaticivorans* (*orf946*) and *Synechocystis* sp. (*slcslrf*). Residues the same (*) or similar (.) are indicated

- below the alignment. See Table 4-5 for references and percent similarity between the sequences. Sequences are aligned in order of decreasing similarity..... 103
- FIGURE 4-10. Multiple sequence alignment of *tdtA* gene product with five decarboxylase/isomerase proteins. Residues the same (*) or similar (.) are indicated below the alignment. See Table 4-5 for references and percent similarity between the sequences. Sequences are aligned in order of decreasing similarity. 105
- FIGURE 4-11. Multiple sequence alignment of *tdtB* gene product with six different known CoA dehydrogenases. Residues the same (*) or similar (.) are indicated below the alignment. See Table 4-5 for references and percent similarity between the sequences. Sequences are aligned in order of decreasing similarity. 106
- FIGURE 4-12. Multiple sequence alignment of *tdtC* gene product with four thiolases. Residues the same (*) or similar (.) are indicated below the alignment. Conserved active centre cysteine residues are boxed. See Table 4-5 for references and percent similarity between the sequences. Sequences are aligned in order of decreasing similarity. 107
- FIGURE 4-13. Multiple sequence alignment of P450_{tdtD} from *Pseudomonas* sp. A19-6a and several sequences identified by BLASTX search (see Table 2 for sequence identification and references). The boxed region is the heme-binding site (Nelson *et al.*, 1993). Absolutely conserved residues are indicated by an arrow above the alignment and those that are the same (*) or similar (.) are indicated below the alignment. Sequences are aligned in order of decreasing similarity. 109
- FIGURE 4-14. Protein sequence alignment of CoA ligases coded for by *tdtL* from *Pseudomonas* sp. A19-6a and *dit orf1* from *Pseudomonas abietaniphila* BKME-9. The latter sequence is currently unpublished and was provided by the authors (Martin and Mohn, 1999). The proposed acyl-adenylate binding site is highlighted (Babbitt *et al.*, 1992). Identical residues (*) and similar residues (.) are indicated below the alignment. 111
- FIGURE 4-15. Map of probes, based on *tdt* locus, developed for the study of resin acid gene distribution in environmental isolates. The eight restriction site maps shown were used to confirm that cloned probes corresponded with gene arrangement in A19-6a. A complete list of expected and observed results is presented in Table B1, Appendix B. 113

- FIGURE 4-16. Composite map of *tdt* gene inserts and the pRR1 suicide vector used to determine which genes were involved in abietic acid degradation. MCS3 was generated by cloning a linker into the *Not* I site of the original MCS. The arrow on the insert map indicates the point of insertional inactivation by pBW30::Tn*phoA*'-1. 113
- FIGURE 4-17. Proposed pathway for degradation of abietic acid in *Pseudomonas* sp. A19-6a. 128
- FIGURE 5-1. Generalized flow diagram of pulping processes used in the production of paper products. Adapted from McKean, 1980. 134
- FIGURE 5-2. Autoradiographs of selected strains used to determine specificity of the *tdt* gene probes. For a complete list of strains screened, see Table C1, Appendix C. Fragment sizes and intensity relative to the A19-6a wild type control are listed in Table 5-2. Fragment sizes indicated by arrowheads are in Kb. 143
- FIGURE 5-3. Phylogenetic tree of 16S rDNA sequences from six strains exhibiting various degrees of similarity to the *tdt* gene probes. Isolates with the same fragmentation pattern as A19-6a are indicated with an asterisk. Number of asterisks denotes probe intensity. Other strains have dissimilar band patterns as shown in Figure 5-2. 146
- FIGURE 5-4. Dendrogram of phenotypic similarities between *tdt*⁺ strains isolated from pulp mill mixed liquor from locations as indicated. Scale is based on the Jaccard correlation coefficient between the positive control strain A19-6a and those indicated in bold (see Table C4 Appendix C)..... 149
- FIGURE 5-5. Autoradiographs of *tdt* probe positive strains isolated from pulp mill mixed liquor treatment systems from various geographic locations. Genomic *Sac* II digests were used to generate the observed fragments. Fragment sizes can be found in Table 5-3. Fragment sizes indicated by arrowheads are in Kb. Gels are loaded in the same order as the dendrogram illustrated in Figure 5-4..... 150
- FIGURE B1. Complete sequence of *Pseudomonas* sp. A19-6a *tdt* locus. The sequence is presented in double strand format with the amino acid sequence (outline font) next to the coding strand. The site of insertional inactivation by pBW30::Tn*phoA*'-1 in A76-47 is indicated by an arrowhead. Nucleotide numbers refer to the plus strand and the base nearest right. Amino acid numbers follow a similar pattern except where translated regions overlap in which case the nearest right and first amino acids are both identified. Stop codons (★) are indicated by asterisks. **Ribosomal binding**

sites are shown in bold, *AT-rich inverted repeats* in italics, AT-rich regions are boxed, and hairpin structures are underlined. Primers are highlighted and appropriately identified with the direction of extension. Cloning sites for library construction are also shown..... 196

CHAPTER 1

Environmental significance of terpenoids

INTRODUCTION

While most of the recalcitrant organic compounds that occur within the environment are considered to be of xenobiotic origin (Bartha, 1990) there are also many examples of biosynthetically produced molecules that are classed as pollutants under certain conditions. Resin acids are among these and pose little ecological threat under natural release conditions, but have long been known to be toxic to ecosystems where their presence is not a part of the natural carbon cycle (Ebling, 1931; Hagman, 1936). Research into resin acid toxicity, the factors affecting it, and mechanisms for detoxification has only begun to emerge in the past 30 years. There was a two decade hiatus in research on microbial remediation of resin acids, after the preliminary work in the 1960s. The reason for this is due, at least in part, to the focus on xenobiotics and their presence in the environment, at the expense of natural compounds that appear at unusually high concentrations in particular environments where they should not (Alloway and Ayres, 1993).

At the start of this project in 1994, research into resin acids had been revived in the areas of chemical analysis, fish toxicity, and the ability of microorganisms to degrade these compounds. Studies in the latter case dealt primarily with isolations from soil and the biochemistry of pathways as determined by the chemical analysis of metabolites. Investigations into the molecular aspects of resin acid degradation by bacteria indigenous to pulp mill

mixed liquor treatment systems where these compounds become concentrated had not been undertaken. Herein lay the impetus for my project.

Technical advances in molecular biology have increased the reliability of studies designed to determine the effect of environmental variables on microbial population dynamics (Fulthorpe and Wyndham, 1989; Herrick *et al.*, 1993 and Erb and Wagner-Dobler, 1993). In fact, catabolic gene probes are now routinely used to monitor bacterial populations involved in biodegradation of a wide range of pollutants (Greer *et al.*, 1993). Thus, my project has been the development of DNA probes based on the catabolic genes involved in the degradation of resin acids by bacteria native to pulp mill mixed liquor secondary treatment systems, for the purpose of monitoring these populations *in situ*. This original research comprises both basic and applied aspects and the scope of this thesis is all inclusive of the research I have conducted on this project since its inception in May, 1994.

General

This thesis is divided into six chapters and three appendices and includes a review of the literature on topics related to the environmental significance of resin acids and findings from the research. Chapter 1 includes background on resin acid metabolism, toxicology, their occurrence in the environment, and the monitoring of their release by regulatory agencies. In addition, the project rationale and hypothesis and the experimental approach used in conducting this research are outlined. Chapter 2 covers the isolation and characterization of

resin acid degrading bacteria and relates this to similar results reported in the literature. A summary of the chemistry of resin acids and environmental factors that affect biodegradation and toxicity is included. The third and fourth chapters encompass the molecular aspects of developing probes based on resin acid catabolic genes. A comparison of findings from this research with known bacterial catabolic systems is made. Chapter 5 is a limited study into the distribution of resin acid degrading bacteria based on the probes developed for the *tdt* genes and describes methods for the treatment of resin acids thereby reducing the toxicity of pulp mill mixed liquors released into the environment. Chapter 6 summarizes the research and outlines areas for future research. The appendices are a compilation of the research data that support the findings reported herein.

BACKGROUND

Plant resin acid metabolism

Resin acid biosynthesis is a part of plant secondary metabolism that is related to but largely separated from the primary metabolic processes (Ellis, 1997). The interface between isoprenoid, polyketide, and phenylpropanoid biosynthesis and primary metabolism is illustrated in Figure 1-1. The shikimate shunt from 2-phosphoenol-pyruvate leads to the phenylpropanoid pathways which produce most plant phenolics (Ellis, 1997). Isoprenoids and polyketides are both derived from acetyl coenzyme A along pathways that produce cyclase

dependent structures or quinone type compounds, respectively. As can be seen, except for the point of interface, these core pathways function separately from primary metabolism. These pathways represent a drain on cellular resources that results in the production of an array of compounds that act in a variety of plant defence and communication mechanisms (Ellis, 1997). The product patterns shown are not exclusive to a single core pathway. Collectively these core structures lead to the synthesis of thousands of different secondary metabolites that are divided into six broad categories (Kramer and Kozlowski, 1979); tannins are involved in lignification, polyketides have anti-microbial activities, alkaloids are anti-grazing compounds, flavanoids act as both attractants and repellents, terpenoids are used in disease resistance, and steroids may be involved in membrane permeability (Kozlowski and Pallardy, 1997).

Resin acids are generated by the isoprenoid pathway through the head-to-tail-joining of isoprene (C_5H_8) subunits. Figure 1-2 summarizes the biosynthesis of isoprene-based compounds which form the largest group of plant organics that are both taxonomically (Croteau and Johnson, 1985) and ecologically specific (Langenheim, 1994). Synthesis begins with the condensation of two units of acetyl coenzyme-A by acetoacetyl-CoA thiolase in a reaction whose equilibrium favors the reverse. This reaction occurs in the cytoplasm and is mirrored in the mitochondria by a similar enzyme involved in the terminal step of fatty acid oxidation (Qureshi and Porter, 1981). Note that the thiolases that catalyse this reaction are widely distributed amongst both prokaryotes and eukaryotes. A second condensation occurs between

acetoacetyl-CoA and a third molecule of acetyl CoA by β -hydroxy- β -methylglutaryl-CoA synthase. Two molecules of NADPH are then used for the reduction of this intermediate to mevalonic acid which is the committed precursor in isoprenoid biosynthesis. Activation occurs by two successive phosphorylations followed by decarboxylation/dehydration to yield isopentenyl pyrophosphate (IPP) or dimethylallyl pyrophosphate which are the repeating subunit and starter group, respectively.

FIGURE 1-1. Interface between plant primary and secondary metabolic pathways. Core structures generated in phenylpropanoid, polyketide, and isoprenoid biosynthesis are shown. Adapted from Kramer and Kozlowski, 1979 and Ellis, 1997.

FIGURE 1-2. Isoprenoid biosynthesis. Dimethylallyl pyrophosphate and isopentenyl pyrophosphate are the starter group and repeating subunit, respectively, used in the synthesis of all terpenoid compounds. The abietane skeleton is also shown as an example of tricyclic diterpene structure. Adapted from Loomis and Croteau, 1980; Croteau and Johnson, 1985; and Hanson, 1972.

Terpenoid biosynthesis is energetically expensive in terms of substrates, cofactors, and pathway enzymes; few of which are shared despite reaction similarities, i.e. hydroxylations, reductions, and glucoside linkages (Gershenzon, 1994). There are also costs for storage of these compounds in specialized structures that enable the plant to direct defence mechanisms while reducing self-injury; many terpenoids are phytotoxic. Since many of these compounds are coupled with differentiation and they represent a reservoir of valuable materials plants may recycle them when they are no longer needed. Another mechanism for reducing net costs includes tight regulation of pathway enzymes which may only be produced in limited quantities as a result of induction. While enzymes in the primary pathways are substrate specific, those that catalyze cyclase reactions are multi-product so that complex mixtures of compounds can be generated from a single pathway. From an evolutionary standpoint the high cost of terpene biosynthesis may be considered counter-selective, however, the complex ecology that exists between plants and their environment demands that a repertoire of mechanisms be readily available to deal with the demands imposed by a broad range of interacting organisms (Ellis, 1997).

Role of resin acids

Terpenoids represent a bewildering array (>15,000 have been characterized) of plant metabolites that are all derived from acetyl-CoA (Langenheim, 1994). While most are lipophilic some oxygenated monoterpenes and terpene glucosides are relatively water soluble. Monoterpenes are constituents of oleoresin with varying volatilities that primarily serve as toxicants,

repellents, and inhibitors to insect and fungal infestations (Russell and Berryman, 1976). In the case of bark beetles these same compounds are chemoattractants that result in concentrated invasions through the production of pheromones by the primary colonizers using substrates sequestered from the host oleoresin (Gijzen *et al*, 1993). The lower terpenoids may also act as solvents for the less volatile diterpenes that seal wounds to form a physical barrier to subsequent attacks (Loomis and Croteau, 1980).

The diterpenes are a major group of the higher terpenoids ($\geq C_{20}$) that are classified into 20 structural types that may be specific to certain tree families, eg., the labdanes, kauranes, abietanes, and pimaranes of legumes and conifers (Langenheim, 1994). Resin acids fall into the latter two structural groups and represent part of a complex water-insoluble mixture of plant exudates commonly known as colophony (rosin) which is obtained by the distillation of oleoresin to produce turpentine (Hanson, 1972). While the exact composition is species specific, crude rosin is comprised of up to 90 percent resin and fatty acids in varying proportions (Lawrence, 1969) and up to 15 percent neutrals which are primarily alcohols (McSweeney *et al*, 1987).

Plant resin acids are important defence compounds that are formed in the epithelial cells of sapwood (Croteau and Johnson, 1985) and secreted into resin ducts where they are stored under pressure until wounding triggers their release (Kramer and Kozlowski, 1969). Heartwood is resin saturated material that may provide secondary flow when required and determine the success of pathological attack by borer insects and decay producing fungi. Following release of

oleoresin and the evaporation of the volatile fraction polymerization of conjugated double bonds by photooxidation leads to the formation of an impermeable barrier that is highly resistant to decay (Croteau and Johnson, 1985).

The industrial use of terpenoids is seen in the production of naval stores from by-products generated during the wood pulping process (Lawrence, 1969). The volatile constituents are released as crude turpentine which is later distilled into commercially useful products such as flavourings, resins, disinfectants, solvents, mining compounds, and rubberprocessing aids (McSweeney *et al*, 1987). Compounds with much higher boiling points are then steam fractionated to produce terpenes, pine oil, and tall oil rosin (Lawrence, 1969) that are used for the production of resins, soaps, emulsifiers, surfactants, tacifiers, adhesives, paper sizings, and printing inks (McSweeney *et al*, 1987).

In addition to the above, some terpenoids have been investigated for their therapeutic potential in antimicrobial, antiulcer, and cardiovascular treatments (Feliciano *et al*, 1993). Resins, saps, and even tree bark were often used by natives of Northern Europe and North America to treat wounds or various types of skin lesions. A study by Soderberg and colleagues (1990) showed that many of the abietanes inhibited the growth of Gram positive bacteria to varying degrees. Homeopathic remedies for colds often rely on pine essences as expectorants and the surfactant properties of resin acids makes some of their esters useful in the production of cosmetics (Feliciano *et al*, 1993). Diterpene inhibition of platelet aggregation suggests that it may be useful in the

development of drugs to treat thrombosis (Cheung *et al*, 1994). The antifeedant properties of resin acids also makes them good candidates for the development of natural insecticides (Xie *et al*, 1993).

Occurrence of resin acids in the environment

Figure 1-3 is a simplified overview of plant terpenoids and their relationship with the environment. As illustrated they have a complex ecology that is influenced by both abiotic and biotic factors in their synthesis and natural and anthropogenic factors in their release. Abiotic factors influencing synthesis include increases in light intensity which result in accumulations (Langenheim, 1994) and the abundance of limiting nutrients (K, N, P) that promote plant growth. This latter generally leads to reduced production of all secondary metabolites due to a reallocation of resources (Gershenzon and Croteau, 1991). Biotic pressures influencing synthesis include interactions that trigger defense responses and the need for dispersion vectors to promote reproduction and ensure species survival (Langenheim, 1994). Since both abiotic and biotic components can be seasonally and geographically variable it is not surprising that terpenes are used in plant taxonomy (Croteau and Johnson, 1985).

There are three principal routes by which terpenoid compounds can enter the environment following synthesis. Plants sequester these compounds as defensive mechanisms against potential infestations by pathogenic organisms or grazing by herbivores. The presence of pine gum on tree bark or the release of latex from legumes are examples of terpenes exuded by plants in response to

FIGURE 1-3. Overview of plant terpenoid ecology. Note that arrows merely denote direction and do not account for the fact that some components of the diagram may be more ecologically significant than others. Partly adapted from Langenheim, 1994.

wounding. Environmental emission occurs by leaching, volatilization, and their appearance as surface coatings. These coatings act as protection from high temperatures or evapotranspiration rates and may also increase plant fitness (Langenheim, 1994). Leaching is a result of litter decomposition and volatilization is important in the attraction of pollinators, dispersion vectors, and entomophages. Both of these release mechanisms are important in determining the dynamics of local ecosystems.

The third route is by anthropogenic release and accounts for the most significant occurrence of resin acids in the environment. The mechanical debarking of logs and the chemo/thermo/mechanical pulping of wood chips in the production of paper products results in the generation of extremely large volumes of mixed liquors that contain a wide range of toxic and neutral compounds in both dissolved and suspended forms (Allan *et al*, 1972). Resin acids are a part of the acidic fraction (Suntio *et al*, 1988) that to some extent is recovered for industrial and commercial applications (Lawrence, 1969 and McSweeney *et al*, 1987). The majority, however, is released to the environment in wastewater streams either directly or following secondary treatment (Allan *et al*, 1972). A recent study by Tian and colleagues (1998) showed that resin acids were also being released by storm runoff in areas where logs were being stored. In both cases, the hydrophobic nature of the resin acids released into receiving waters leads to their accumulation in sediments and their persistence in the environment (McLeay and Associates, 1986).

Resin acid toxicology

As plant defense compounds, terpenoids are of necessity toxic to the organisms against which they are directed. The volatile fraction of oleoresin is fatally toxic to bark beetles and limits the growth of their fungal symbionts (Gijzen *et al*, 1993). Those boring insects not overcome by monoterpene vapours may become engulfed by exudates at the wound site in a process called pitching out by the host tree. The success of this counterattack is highly dependant upon the amount and composition of oleoresin and the general health of the host at the time of infestation (Croteau and Johnson, 1985). Plant herbivores such as pine sawflies exhibit reduced growth rates and increased development times and mortality due to the toxic effects of resin acids during early larval stages (Langenheim, 1994). In later stages, however, they even sequester these compounds for use against predators and may actually metabolize them to some degree (Bjorkman and Gref, 1993). Terpenoid compounds have also been shown to act as antifeedants towards mammalian grazers that may associate their bitter taste to their being potentially poisonous (Langenheim, 1994). There is also evidence suggesting that terpenoids may interfere with steroid metabolism in cattle leading to reproductive dysfunction (Croteau and Johnson, 1985). Volatization and leaching of terpenes has also been shown to have an inhibitory effect on seed germination leading to complex plant-plant interactions that influence species abundance, diversity, and succession (Langenheim, 1994).

Aquatic organisms may also be exposed to plant secondary metabolites as a result of their release in pulp mill wastewater. The exact composition of

these mixed liquors is highly dependent upon feedstock, production and treatment processes, and the occurrence of process malfunctions. Nevertheless, all mixed liquors generally contain some combination of resin and fatty acids, phenols, and neutrals many of which may become chlorinated during bleaching (McLeay and Associates, 1987). Resin acids are a part of the acidic fraction of mixed liquors (Suntio *et al*, 1988) and have been shown to exert both acute and sub-lethal toxic effects (see McLeay and Associates, 1987 and Taylor *et al*, 1988 for comprehensive reviews). Table 1-1 is a summary of the bioaccumulation of dehydroabietic acid (DHA) for trout exposed to acute and sub-lethal concentrations. In this experiment (Oikari *et al*, 1983) the physiological effects observed included changes in blood chemistry and volume, heart, liver, and kidney tissues, and generalized and specific metabolic activities. Note that both the acute and sub-lethal dose levels are well within the discharge concentrations observed for both untreated ($<20\text{--}15,100\ \mu\text{g/l}$) and biotreated ($<1\text{--}5800\ \mu\text{g/l}$) mixed liquors (Taylor *et al*, 1988). In addition to the specific biomarkers of growth, development, reproduction, resistance and stress; behavioural, and survival effects are used as indicators of toxicity.

In general terms pulp mill mixed liquors and resin acids in particular have been examined using a variety of popular toxicity tests in an effort to determine the significance of their release to the environment. For example, results from a *Salmonella*/microsome assay of whole mixed liquor found that it contained mutagenic compounds that resulted in frameshift and base pair substitution mutations (Douglas *et al*, 1980). Nevertheless, this same test had previously failed to implicate resin acids as the source of the mutagenic activity; neoabietic

acid excepted (Nestmann *et al*, 1979). These results of course do not extrapolate well to whole organisms that may possess mechanisms for modifying or eliminating toxic substances.

TABLE 1-1. Acute and sublethal toxicity of dehydroabiatic acid (DHA) to *Salmo gairdneri* Richardson in fresh water. Summary of exposure concentration and duration and accumulated tissue residues observed, adapted from Oikari *et al*, 1983.

Exposure		
	ACUTE	SUB-LETHAL
[average DHA]	1.2 mg/L (ppm)	20 µg/L (ppb)
duration	4 days	30 days
Tissue residues [DHA]		
liver, kidney, heart	75-100 µg/g	1-2 µg/g
intestine, gill, skin	35-50 µg/g	3-4 µg/g
red/white muscle, gonads	<20 µg/g	<0.5-2 µg/g
brain, spleen	35-50 µg/g	1-2 µg/g
plasma	237 µg/ml	5 µg/ml

Recent studies have begun to focus on induction of mixed-function oxygenase activity as an indication of mixed liquor toxicity (Hodson, 1996). Figure 1-4 illustrates the general mechanism for eliminating toxins by cytochrome P450 monooxygenase based systems. These can be found in any tissue with the highest levels usually found in the liver. Detoxification is a two-step process involving the introduction of epoxide or hydroxyl groups by oxidation and conjugation of polar constituents to increase solubility and enable

excretion (Stegeman and Hahn, 1994). Several studies have shown that whole mill mixed liquors can induce mixed function oxygenase activity (Williams *et al*, 1996; Martel *et al*, 1996; Servos *et al*, 1992; Munkittrick *et al*, 1992; and Munkittrick *et al*, 1994), however, others have demonstrated that resin acids may not be the cause (Hewitt *et al*, 1996 and Ferguson *et al*, 1992.). Nevertheless, Johnson and colleagues (1995) have shown that the activity of 7-ethoxyresorufin-O-deethylase, the enzyme commonly used to detect mixed function oxygenase induction, did not correlate with the concentration of conjugated resin acids in the bile of trout exposed to mixed liquor from an integrated newsprint mill. They also found that the range of concentrations necessary for acute exposure and the minimum required to elicit a slight physiological response (170 $\mu\text{g/L}$ and 85 $\mu\text{g/L}$, respectively) was very narrow.

FIGURE 1-4. General scheme for the elimination of toxins by fish via mixed function oxygenase systems. Adapted from Stegman and Hahn, 1994.

In some recent studies researchers have begun to look at the neurotoxicity of resin acids to fish and have found that threshold levels are generally lower than those detected in brain tissue following exposure (Zheng and Nicholson, 1996). They have also found that these responses may not be linked with indicators normally equated with toxicity but may still cause subtle behavioural changes that could affect survival (Zheng and Nicholson, 1998).

While much of the research conducted in this area deals with commercially important fish species some studies have looked at the effect of total mixed liquor and resin acid exposure on other organisms. Here again toxicity results are highly variable depending upon experimental conditions used (Taylor *et al*, 1988). A recent study by Hickey and Martin (1995) showed that high resin acid mixed liquors had an impact on benthic invertebrates and that sub-lethal tests needed to be developed. For humans, contact dermatitis and asthma have been linked to exposure to oxidized or modified forms of resin acids (Gafvert, 1994). Resin acid degradation products such as retene and fichelite have also been shown to exert toxic effects (Burggraaf *et al*, 1996 and Liss *et al*, 1997).

Government regulation of pulp mill wastewater toxicity

In 1992 the Ontario Ministry of Environment and Energy (1994) cited most Ontario pulp and paper mills for one or more environmental discharge infractions relating to biochemical oxygen demand, total suspended solids, or

phosphorus/sulphide emissions (Ontario Ministry of Environment, 1992). On 19 May of that same year the Federal Government announced that more stringent environmental regulations governing pulp and paper mixed liquors were to be imposed by 1 December, 1992 (Madore, 1992). Parameters targeted by the new legislation included reductions in the release of total suspended solids, biochemical oxygen demand, organochlorines, acutely lethal mixed liquors and elimination of process additives that contribute to the formation of dioxins and furans. Resin acids were not specifically targeted even though the need for stricter controls on the release of these compounds had been identified (Murray, 1992).

In January 1993 the Ontario government announced its own proposed regulations that required mills to submit plans as to how they would achieve target dates and monitor discharges to ensure that limits were not exceeded (Ontario Ministry of Environment, 1993). Here again resin acids were included only as a part of the biochemical oxygen demand and total suspended solids parameters for which permissible levels were lower than stipulated by the Federal legislation.

Treatment of pulp mill mixed liquors

Even though release of resin acids is not specifically regulated they can be readily eliminated by effective treatment prior to discharge. Secondary treatment has been the most common method of resin acid removal and

systems used for the treatment of municipal wastes have been adapted to meet the needs of the pulp and paper industry (Taylor *et al.*, 1988). Mixed liquor detoxification usually occurs in phases that are designed to remove suspended solids, dissolved organics, and chemical precipitates if necessary. Approaches include the use of aerated lagoons, activated sludge basins, and anaerobic processes either singly or in some combination. Mixed liquors may also be pre-treated with lime to enhance biological treatability (Bennett *et al.*, 1988) or passed through a constructed wetland to remove residual contamination (Thut, 1989). While these mechanisms are generally effective at reducing resin acid toxicity, system failures due to a lack of knowledge of the factors affecting resin acid degradation rates (Zender *et al.*, 1994) and inadequate monitoring (Beak Consultants, 1987) has lead to the release of mixed liquors that do not meet toxicity standards (Lavallee *et al.*, 1993). This latter point is particularly important since resin acid levels and composition are wood species dependant so that increases in mixed liquor toxicity occur due to process malfunctions or to feedstock changes. These may have either gross or subtle influences on the microbial ecology of the system (O'Connor *et al.*, 1992).

EXPERIMENTAL DESIGN

Rationale

The release of plant secondary metabolites into the environment occurs by both natural and man-made routes. Natural routes include mechanisms designed to release these compounds in response to specific plant needs and

the uncontrolled release brought about by plant senescence. In both of these cases nature has devised ways of dealing with the onslaught of substances that under normal conditions may be considered deleterious. This, however, cannot be said for the anthropogenic release of resin acids as a result of the production of pulp and paper products.

By definition a pollutant is a compound (manmade or otherwise) that occurs in an environment and/or at a concentration that is harmful to one or more organisms within that environment (Alloway and Ayres, 1993). Based on this we can see that ecosystem complexities are specifically designed to deal with the controlled release of resin acids that occurs as part of normal plant growth and metabolism. We can also surmise that plant secondary succession is one of nature's ways of dealing with the wholesale release that may occur as a result of natural disasters. In the case of pulp mill mixed liquors resin acid release may occur continuously and in such volumes that natural mechanisms that may be present in receiving environments become quickly saturated and are unable to deal with and/or contain subsequent effects. Herein lies the rationale for this project.

Bacterial degradation of resin acids in existing treatment facilities is well known and subject only to the vagaries of pulp and paper production processes. The use of molecular biological techniques to monitor the effect of environmental variables on microbial ecology have become increasingly popular (Fulthorpe and Wyndham, 1989; Sayler and Layton, 1990; Herrick *et al*, 1993; Erb and Wagner-Dobler, 1993; among others). DNA or RNA probes based on catabolic genes or

their transcription products can be effective ways of monitoring degradation levels without the time consuming process of using culture techniques that may have the added drawback of sample bias. Furthermore, any experimental technique that is developed for use as a monitoring tool should have a sufficiently broad specificity so that a range of environmental samples can be analyzed while still ensuring that results correlate to the parameter in question.

Hypothesis

The hypothesis upon which this research is based is that resin acid degrading bacteria can be isolated from environments where these compounds become concentrated. These isolates can then be used to develop genetics based techniques suitable for studying the distribution of similar organisms in treatment systems and receiving waters from a range of geographical locations. Ideally these tools, when coupled with knowledge of local environmental variables that can influence degradation rates, will serve as efficient methods of monitoring *in situ* resin acid ecology to ensure that utilization is optimal and the toxicity of mixed liquors is reduced.

Objectives

The first objective of this research entailed the isolation and characterization of candidate bacteria from an existing pulp mill treatment facility. Those isolates that were most effective at utilizing resin acids as a sole source of carbon and energy were selected for the development of DNA probes based on

resin acid catabolic genes. This second objective comprises the bulk of the research undertaken for this project. The final objective was to confirm the usefulness of the probes in assessing the ecology of resin acid degradation through the distribution of specific genes in resin acid degraders isolated from a range of environments. Subsequent analyses that are necessary to validate the applicability of using the probes as monitoring tools are addressed under the section on future work.

Methods

The isolation of candidate resin acid degrading bacteria was based on enrichment sub-culturing of pulp mill mixed liquor mixed liquor inoculum grown in defined media containing only abietic or dehydroabietic acids as sources of carbon and energy. Randomly selected isolates were characterized both phenotypically and genotypically. Candidate degraders were selected for transposon mutagenesis in order to localize genes involved in resin acid catabolism. A range of molecular techniques were employed in the isolation of mutated genes from transconjugants that had lost the ability to degrade abietic acid in liquid culture. Cloned genes were identified by similarity search and gene knock-outs were used to confirm that they were necessary for resin acid degradation. Probes from these same genes were used to determine the distribution of this particular genotype in pulp mill mixed liquor samples from treatment systems world wide.

CHAPTER 2

Isolation and characterization of resin acid degrading bacteria

INTRODUCTION

Resin acid chemistry

In addition to exploring the origin and fate of resin acids in the environment it is important to understand their chemistry since this plays a significant role in determining their distribution. These compounds are described as tricyclic, diterpenic, monocarboxylic acids with characteristically low water solubilities (Kramer and Kozlowski, 1969). Figure 2-1 illustrates six of the eight congeners commonly found in pulp mill mixed liquors. The first four, abietic acid (ABA), neoabietic acid (NAA), dehydroabietic acid (DHA), and levopimaric acid (LPA), are of the abietane type which have conjugated double bonds and methyl groups on the C13 carbon. Pimaric acid (PA) and isopimaric acid (IPA) are pimaranes that carry a methyl and a vinyl group on the C13 carbon and have unconjugated double bonds that cannot become conjugated (Mutton, 1962).

Not all resin acids found in mixed liquors are initially formed in the production of oleoresin (Simonsen and Barton, 1961). Neoabietane, levopimarane, and pimarane are examples of primary resin acids that isomerize to other secondary congeners such as abietic and isopimaric acids. Dehydroabietic acid is a tertiary form that results from photolytic dehydrogenation of abietic acid to form an aromatic C-ring (Icail and Icail, 1994). Table 2-1 lists some of the chemical properties of the six congeners used in this study. All have high melting points in the range of 150°-212° C. With the

exception of dehydroabietic acid which has the formula $C_{20}H_{29}O_2$ and a molecular weight of 301.43, the general formula for resin acids is $C_{20}H_{30}O_2$ with a molecular weight of 302.44 (Suntio *et al*, 1988). Absorption maxima for these compounds is at the short end of the ultraviolet spectrum in the range 237.5 to 272.5 nm. The conjugated diene ($C=C-C=C$) is the chromophoric group that has a basic absorption wavelength of 253 nm when the double bonds are on the same ring and 214 nm when they occur on different rings (Pinder, 1960). Carbon substituents and double bonds are responsible for increasing these basal wavelengths. Abietic acid is one of the best known secondary acids that is formed by the isomerization of levopimaric or neoabietic acids under acidic conditions. It has four asymmetric centres leading to a total of 24 possible conformations. Retene is a common dehydrogenation product of most congeners that when oxidized gives an orange compound known as retenequinone ($C_{18}H_{16}O_2$).

As weak acids resin acids are subject to dissociation as determined by their ionization constant. For abietic and dehydroabietic acids pK_a is 6.4 and 5.7, respectively (Nyren and Bach, 1958). At low pH values these compounds are undissociated and primarily exist in a colloidal state. As pH increases dissociation occurs and they become more soluble up to approximately pH 7.5 where maximum dissolution is reached (Werker and Hall, 1997). From an environmental perspective it is this property of resin acids that is most significant. As illustrated in Table 2-2 acute toxicities are pH dependent within a very narrow range. As hydrophobic compounds the unionized species have a tendency to partition to similar environments which leads to increased sediment and

bioaccumulation levels at the near neutral pH of receiving waters (McLeay and Associates, 1979 and Taylor *et al*, 1988)

FIGURE 2-1. Structures of resin acid congeners used in this study. The abietane group includes abietic acid (ABA), neoabietic acid (NAA), dehydroabietic acid (DHA), and levopimaric acid (LPA). Pimaric acid (PA) and isopimaric acid (IPA) belong to the pimarane group.

Abietic acid (ABA)

Neoabietic acid (NAA)

Dehydroabietic acid (DHA)

Levopimaric acid (LPA)

Pimaric acid (PA)

Isopimaric acid (IPA)

TABLE 2-1. Selected chemical characteristics of resin acid congeners used in this study. Sources as indicated.

congener	type	melting point (C)*	UV absorbance	ionization constant**	standard solubility**
ABA	secondary	175	241***	6.4 +/-0.1	4.3 mg/L
NAA	primary	169	250***	na	na
DHA	tertiary	174	257.5	5.7 +/-0.1	6.6 mg/L
LPA	primary	150	272.5***	na	na
PA	primary	212	250	na	na
IPA	secondary	164	237.5	na	na

* Hanson, 1972

** Nyren and Black, 1958

*** Pinder, 1960

TABLE 2-2. Range of acute toxicity to fish of resin acids commonly found in pulp mill mixed liquors. Toxicity is pH dependent and lowest at high pH. Adapted from Taylor et al., 1988.

resin acid	acute toxicity 96h-LC ₅₀ (mg/L)	pH
abietic acid	0.2-2.38	6.4-8.0
neoabietic acid	0.6-1.7	7.5
dehydroabietic acid	0.77-2.10	6.3-9.0
levopimaric acid	0.6-1.0	6.4-6.5
pimaric acid	0.32-1.0	6.4-7.1
isopimaric acid	0.22-1.0	6.4-8.0

Biodegradation and biotransformation of resin acids

As naturally formed compounds it is not surprising that terpenoids are easily biodegraded. Button (1984) demonstrated that terpene based food chains could be found in aquatic environments where these compounds were abundant and that their presence in these ecosystems was probably due to rain run-off from plant surfaces (Button *et al*, 1981). Early work in the study of resin acid degradation dealt with the analysis of metabolites formed by bacteria and fungi isolated from a variety of terrestrial sources.

In a study of *Flavobacterium resinovorum* Biellman *et al*, (1968) found that contrary to expectations the aromatic ring was not the first point of attack. Rather, it was the A-ring that was hydroxylated at the C3 position (see Figure 2-2) followed by oxidation to a ketone and decarboxylation; the latter of which may or may not have been enzyme mediated. In subsequent research they discovered that the C7 carbon on the B-ring was also attacked and that both steps were necessary before dihydroxyl activation of the aromatic ring (Biellman *et al*, 1973a). At this point two alternative routes to complete mineralization became evident, one involved C-ring cleavage to yield 2-isopropyl malic acid and two single ring acid products, and the other kept the A and B rings intact as a bicyclic triketone until the C-ring was completely removed. In another study these same researchers found that strains of *Alcaligenes eutrophus* and *Pseudomonas* sp. that degraded dehydroabietic acid did so by oxidation at C7 only, prior to dihydroxylation and removal of the aromatic ring. The bicyclic ketone formed is then hydroxylated and oxidized to produce two diacids. Cross

and Myers (1968) in a study of abietic acid degradation found that metabolites included 5 α -hydroxyabietic acid, a ketone (7 β -hydroxy-13-isopropyl-8 ξ -podocarp-13-en-15-oic acid), and an epoxy- γ -lactone.

FIGURE 2-2. Points of oxidative attack on dehydroabietic acid by bacteria and fungi.

Resin acid degradation is not restricted to aerobic bacteria. Wilkins and Panadam (1987) in a New Zealand pulp mill study detected six metabolites generated by the anaerobic transformation of abietic acid. Their results suggested that there were two different pathways from which the more recalcitrant products retene and fichelite were formed. In the former case dehydrogenation of the C-ring leads to the enzymatic formation of dehydroabietic acid which is decarboxylated and demethylated at C10 followed by successive

dehydrogenation of the A and B rings to give dehydroabietin, 1,2,3,4-tetrahydroretene, and retene. The second pathway involved saturation of the substrate to produce abietan-18-oic acid which when decarboxylated yields fichtelite. These results supported earlier work by other researchers.

Fungi are also involved in resin acid biodegradation as illustrated by Ekman and Sjöholm (1979) who detected transformation products from the fungus *Fomes annosus* growing on dehydroabietic acid. Wang *et al.*, (1995) demonstrated that a mixed inoculum of wood-inhabiting fungi could reduce resin acid concentrations in wood chips by up to 96 percent for some congeners. James Kutney's lab in the Department of Chemistry, at the University of British Columbia has done extensive work on the biodegradation and biotransformation of dehydroabietic, abietic, and isopimaric acids by the zygomycete *Mortierella isabellina* (see Kutney *et al.*, 1981a, 1981b, 1982a, 1982b, 1983, 1985, and 1988). In fact, some research initiatives have begun to exploit the use of fungi as pretreatments for improving production processes and reducing mixed liquor toxicities (Farrell *et al.*, 1993; Brush *et al.*, 1994; and Duran *et al.*, 1994)

Isolation and characterization of degraders from pulp mill mixed liquors

More recent work has begun to focus on the use of genetics based techniques to monitor degradation levels in an effort to determine the environmental factors that influence them (Sayler and Layton, 1990). An example of such a study is one conducted by Greer and associates (1993) who used catabolic gene probes to isolate and monitor bacteria involved in petroleum

degradation of contaminated soils and subsequently mark a selected isolate with reporter genes in an effort to monitor its fate in the environment. In this particular case the *bphC*, *clcA*, *xylE*, *ndoB*, and *alkB* genes involved were already known. With resin acids, however, catabolic genes have only been recently targeted for isolation and characterization with the objective of developing similar tools to monitor pulp mill mixed liquors and wastewater receiving waters.

A necessary first step is of course the isolation of target resin acid degraders that may be used in the localization and isolation of resin acid catabolic genes. Unlike earlier work that dealt with the isolation of degraders from garden and forest soils, recent research has begun to focus on bacteria found in pulp mill mixed liquor treatment systems. William Mohn's lab at the University of British Columbia has done extensive work in this area and they have isolated by enrichment a total of 65 resin acid degrading bacteria from a bleached kraft pulp mill mixed liquor, a sequencing batch reactor, and forest soil (Mohn *et al*, 1998a). These researchers found that resin acid degradation levels and substrate preferences for both resin acids and alternative carbon sources were quite variable for these isolates. Nevertheless, they did manage to identify several strains that are proficient in the degradation of dehydroabietic acid (Mohn, 1995 and Bicho *et al*, 1995), some of its chlorinated congeners (Mohn and Stewart, 1997), and two isopimarane degraders (Wilson *et al*, 1996). Ribosomal DNA sequencing identified isolates belonging to several genera including *Ralstonia*, *Sphingomonas*, *Pseudomonas*, and *Zoogloea* and they have recently proposed that the latter and three of the *Pseudomonas* strains represent

newly isolated species (Mohn *et al*, 1998). Furthermore, the strain identified as *Pseudomonas abietaniphila* has been used to elucidate the genetics involved in the catabolism of dehydroabietic acid and probes based on isolated genes were used to monitor this organism in a consortium culture growing on bleached kraft mill mixed liquor (Mohn *et al*, 1998). In this work they also confirmed that C7 oxidation and dihydroxyl activation of the aromatic ring occurred prior to its cleavage.

Extraction and analysis of resin acids

Incumbent upon the successful isolation of bacteria involved in the catabolism of novel substrates is the ability to detect these compounds in aqueous media. In the case of resin acids this generally involves the extraction of the solution with an organic solvent to isolate and concentrate the analyte and separation of extractants by chromatography so that specific substrates and possible degradation products can be identified. Extraction methods include liquid extraction using ethyl acetate under acidic conditions (Kutney *et al*, 1985) or alkaline extraction using methyl-*tertiary*-butyl ether (Voss and Rapsomatiosis, 1985). Solid phase extraction using resins to fractionate samples by eluting absorbed analytes with different solvents is also used (Leach and Thakore, 1973). As with any experimental technique these methods have several limitations including sample foaming, isomerization of some resin acids thereby skewing results, and variable recoveries depending upon the sample extracted (Taylor *et al*, 1988). One of the principle drawbacks with the use of these techniques is the need to methylate the analyte prior to detection. This has lead

to interest in the use of supercritical fluid extraction which does not require derivatization of the analyte prior to quantification and analysis (Yamada *et al*, 1995). Derivatization using pentafluorobenzyl bromide vice diazomethane to enable analysts to distinguish target acids from their native esters has also been reported (Dethlefs *et al*, 1996).

Separation and detection of resin acids routinely includes gas chromatography (GC) and high pressure liquid chromatography (HPLC) coupled with mass spectrometry (MS) or UV absorption/flourescence spectroscopy, respectively (Taylor *et al*, 1988). GCMS has long been a popular method for analysing resin acids due to the ability to quantify the analyte as well as detect metabolites on the basis of changes in fragmentation patterns. In fact, this technique was used extensively to determine the metabolic pathways used by the aerobic and anaerobic bacteria already discussed (Beillman *et al*, 1968, 1973a, 1973b, and Wilkins and Panadam, 1987). HPLC methods are also popular for routine quantitative analysis (Richardson *et al*, 1992 and Volkman *et al*, 1993) but are unsuitable for the determination of metabolites since analyte identification is based solely on retention time. Gas/liquid chromatography of underivatized resin acids detected by flame ionization has also been reported (Gref, 1988) but does not appear to be widely used. In an effort to reduce the analysis time required by these conventional techniques researchers have begun to develop ELISA-based methods for the detection of resin acids in pulp mill mixed liquors (Li *et al*, 1994 and 1996). This method has the advantage of not requiring extensive sample preparation, however, it is unable to distinguish

between individual congeners (although pimarane and abietane types as a whole show little cross-reactivity) and does not give quantifiable results.

Objective I

The foregoing constitutes the theoretical basis upon which the first objective of my research is based. Bacteria capable of using resin acids as a sole source of carbon and energy have been isolated from a variety of environments including pulp mill mixed liquors and existing treatment systems. These organisms represent a variety of taxa exhibiting variable phenotypes.

MATERIALS AND METHODS

Culture techniques and isolation of putative degraders

Batch enrichment culture was used to isolate putative resin acid degraders. Minimal media A (Wyndham, 1986) containing 0.005% yeast extract as a source of growth factors was provided as a supplement to the abietic acid (Aldrich Chemical Co.) or dehydroabietic acid (Pfaltz and Bauer Inc.) carbon source. Stock solutions (300 mM) of carbon sources were dissolved in dimethylformamide (Analar Ltd.) and added prior to autoclaving. Solid selective media were prepared in the same manner as liquid media with the addition of 1.6% agar. A preliminary enrichment in 20 μ M ABA or DHA using 20 ml of mixed liquor and sediment, from the Avenor bleached kraft pulp mill in Dryden, Ontario (Weyerhaeuser), as inoculum was unsuccessful at obtaining suitable isolates after six weeks of successive 1 ml transfers even though concentrations were comparable to those found *in situ*.

A second enrichment using 470 μ M ABA and 330 μ M DHA was then performed with 1 ml transfers made every 24 hours. An aliquot from each culture prior to inoculation and at the time of transfer was retained for HPLC analysis to determine the extent of resin acid degradation. Isolates were obtained by serial dilution onto selective media of cultures A19 (ABA grown) and D11 (DHA grown). Fifty randomly selected colonies from each dilution were patched onto TYE (1% tryptone, 1% yeast extract, 0.5% NaCl) and then selected

for further analysis on the basis of differences in colony colour, size, shape, and/or margin. Following two successive purifications on selective media a total of 13 putative resin acid degrading bacteria were isolated; six from ABA enrichment and seven from DHA enrichment.

To ensure that healthy cultures were used for phenotypic analysis, biochemical testing of isolates was performed on cells transferred from selective media and grown on TYE. Source cultures were maintained on selective media and frequently renewed from stocks frozen in 20 percent glycerol at -70°C.

Biochemical characterization of resin acid degraders

Morphological and biochemical analyses were used to characterize each isolate. Colony morphology was determined by visual observation of growth on CPS (200 mg/L K₂HPO₄, 40 mg/L MgSO₄·H₂O, 4 drops saturated FeCl₃, 500 mg/L bactopectone, 500 mg/L bactocasitone, 1% soluble starch, 1% glycerol, 1.6% agar). King's B agar (King *et al.*, 1954) was used to determine whether colony colour was due to the presence of diffusable fluorescent pigments using *Pseudomonas putida* as a positive control.

Standard protocols, as described in Smibert and Kreig (1994), were used for Gram stain, motility, oxidase, catalase, and glucose fermentation tests. Cellular morphology and motility were observed using an Olympus Phase-contrast microscope. Nitrogen fixation was determined using minimal media A (with and without nitrogen) supplemented with succinate (Sigma). Temperature

growth range between 4°C and 41°C was determined on CPS for DHA isolates and 4 mM succinate in minimal media A for ABA isolates. The range of carbon sources utilized by each isolate was determined by Biolog fingerprinting as specified by the manufacturer (Hayward, Calif). Growth after eight days incubation on each of six resin acid congeners (1 mM in minimal media A) was also assessed.

Strains A19-6a and D11-13 were selected for fatty acid methyl ester analysis (K. Bernard and S. Tessier, Health Canada, Ottawa, Ont.) with the MIDI microbial identification system (MIDI, Newark, Del.) using gas chromatography (Hewlett Packard 5890) and flame ionization detection.

Determination of resin acid degradation

To determine the range of resin acids used by each isolate pure batch cultures were tested by HPLC (Millipore). Analysis was based on the direct measurement of the amount of resin acid in solution in the extracellular medium in comparison with a known standard and an uninoculated negative control. Sample preparation, including negative controls, consisted of two 10 minute spins at 12,000g in a microcentrifuge. Separation of residual resin acids in supernatants was achieved by a Novapak C₁₈ endcapped column 15 cm long, 3.9 mm inner diameter, and 4.0 µm bead diameter using a nonpolar solvent (80% methanol/20% distilled water (1% glacial acetic acid)). Flow rate was set to maintain an operating pressure of approximately 2000 psi (13,790 kPa). Detection was determined by UV absorbance at 242 nm for ABA, 257.5 nm for

DHA, 244 nm for neoabietic acid, 250 nm for pimaric acid, 237.5 for isopimaric acid, and 243 nm for levopimaric acid. The wavelengths were determined empirically from the standards using a full wavelength scan (Beckman DU-65 spectrophotometer) and selecting for maximum analyte absorption and minimal solvent interference. Standards were made by dissolving ultra pure congeners (>85%) from Helix Biotech Corporation (Vancouver, B.C.) in 100 % methanol (OmniSolv HPLC grade). For routine analysis of ABA cultures a standard made from the lab grade abietic acid (70 %) used for preparing cultures was preferred.

GCMS analysis of consortium cultures was used to identify metabolites formed during abietic acid degradation. Twenty-five milliliter aliquots taken from a one liter culture at six hour intervals were alkaline (pH9) extracted twice for 10 minutes using methyl-*tertiary*-butyl ether (Voss and Rapsomatiotis, 1985). The pooled organic phases were reduced to 1 ml and the analyte methylated with diazomethane. Samples were then analyzed using a Hewlett Packard 5890 gas chromatograph and 5971A mass spectrometer. Three microliters of sample was injected in the splitless mode and scanned for 40 minutes using helium as the carrier gas and an oven temperature of 80°C (initial) increasing to 280°C (16°C/min).

Genotypic analysis of selected isolates

ERIC-PCR (Versalovic *et al.*, 1991 and deBruijn, 1992) was used to determine genotypic similarity between 10 isolates. ERIC 1R (5'-ATG TAA GCT CCT GGG GAT TCA C-3') and ERIC2 (5'-AAG TAA GTG ACT GGG GTG AGC

G-3') primers based on enterobacterial repetitive intergenic consensus sequences were used to amplify genomic DNA. PCR reactions (25 μ l) were overlaid with 100 μ l of mineral oil (Sigma) and contained 20% Gitschier buffer (83 mM $(\text{NH}_4)_2\text{SO}_4$, 335 mM Tris-HCl (pH 8.8), 33.5 mM MgCl_2 , 33.5 μ M EDTA, 150 mM β -mercaptoethanol), 0.8% bovine serum albumen (20 mg/ml, Boehringer Mannheim), 10% dimethyl sulfoxide, 1.25 mM each dNTP, 62.5 pmol each primer, 0.8 units of Vent[®] DNA polymerase (New England Biolabs), and 50 ng of genomic DNA. Reactions were incubated overnight in a Perkin Elmer Cetus 480 DNA Thermal cycler using an initial denaturation cycle of seven minutes at 95°C, 35 amplification cycles (94°C, 1 min; 56°C, 1 min; 68°C, 8 min), and a final extension cycle at 68°C for 16 min followed by an indefinite soak at 4°C. Ten microliters of product were run on a 1.5% agarose gel 20 cm long for approximately five hours at 105V.

RESULTS

Isolation and characterization of bacteria

The first enrichment resulted in the isolation of 24 cultures selected at random from abietic acid and dehydroabietic acid plates. These isolates, however, could not be sub-cultured onto selective media after having been grown on TYE and could not be tested as pure batch cultures for their ability to use resin acids as sole carbon source. The second enrichment was successful at isolating 13 confirmed resin acid degraders. Six of these isolates were enriched on abietic acid (A19-X) and the remaining seven on dehydroabietic acid (D11-X). Table 2-3 summarizes the physiological characteristics exhibited by these isolates. As indicated, all are Gram negative rods that do not produce diffusible pigments. Some isolates were motile and some exhibited oxidase and/or catalase activities. No isolate was able to ferment glucose although many were able to oxidize this carbon source. The Biolog test plates supported these results for all but A19-6a which was positive rather than negative for glucose utilization. Carbon utilization coefficients suggest that some isolates preferentially use carbohydrates as a carbon source and others prefer more complex substrates such as organic acids, amino acids, and tricarboxylic acid (TCA) cycle intermediates. A comprehensive list of the carbon utilization profile for each isolate is presented in Appendix A, Table A1.

Fatty acid methyl ester analysis of isolates A19-6a and D11-13 identified them as *Comamonas testosteroni* (P=0.750) and *Alcaligenes xylosoxidans* (P=0.376), respectively. Table 2-4 is a summary of the esters detected and the percentage at which they occur in these isolates.

TABLE 2-3. Summary of physiological characteristics of resin acid degrading bacteria isolated from Avenor bleached kraft pulp mill, Dryden, Ontario (Weyerhaeuser). For a complete breakdown of carbon sources used by each isolate see Table A1, Appendix A.

ISOLATE	Gram stain	Shape	Length (mm)	Width (mm)	Motility	Oxidase	Catalase	King's B (diffusible pigments)	Glucose		Carbon sources ^a	
									Oxidation	Fermentation	Carbohydrates	Acids ^b
A19-1	-	rod	2-5	0.5	-	/	/	-	/	-	0.34	0.52
A19-2	-	rod	2-5	0.5	-	-	-	-	/	-	0.37	0.55
A19-3	-	rod	1-1.5	0.5	-	+	/	-	/	-	0.17	0.50
A19-5	-	rod	1	0.5	+	/	+	-	-	-	0.14	0.27
A19-6a	-	rod	2	0.5	+	+	+	-	-	-	0.17	0.50
A19-6b	-	rod	1	0.5	+	-	+	-	-	-	0.14	0.35
D11-3b	-	rod	1-1.5	0.5	+	/	/	-	+	-	0.83	0.33
D11-9	-	rod	1	0.5	+	/	+	-	+	-	0.14	0.45
D11-13	-	rod	1	0.5	+	-	/	-	+	-	0.54	0.07
D11-23a	-	rod	1	0.5	+	/	/	-	+	-	0.77	0.30
D11-23b	-	rod	1-1.5	0.5	-	+	+	-	-	-	0.14	0.35
D11-25	-	rod	1	0.5	+	-	+	-	+	-	0.57	0.14
D11-37	-	rod	1	0.5	+	+	+	-	+	-	0.14	0.38

NOTES: +, positive reaction; -, negative reaction; /, weak-positive reaction.

^a Values represent ratio of carbon sources used to the number of carbohydrates or acids available.

^b Organic acids, amino acids, and TCA cycle intermediates.

TABLE 2-4. Summary of fatty acid methyl ester analysis of resin acid degraders A19-6a and D11-13. The probability of the chromatogram matching the type strain indicated is included.

A19-6a		D11-13	
FATTY ACID	PERCENT	FATTY ACID	PERCENT
10:0 3OH	3.64	12:0	4.25
12:0	3.27	12:0 2OH	2.32
14:0	0.30	14:0	1.02
15:0	0.43	12:0 ALDE	6.39
14:0 2OH	0.29	16:1/14:0 3OH	
16:1w7c	36.93	16: w7c	33.11
16:0	27.70	16:0	41.34
17:0 CYCLO	1.12	17:0 CYCLO	1.44
17:0	0.69	18:2 w6,9c/18:0 ANTE	1.8
16:1 2OH	1.32	18:1 w7c/w9t/w12t	
16:0 2OH	2.69	18:1 w9c/w12t/w7c	6.48
18:1 w7c/w9t/w12t	21.82	18:1 w12t/w9t/w7c	
18:1 w12t/w9t/w7c		18:0	1.86
18:0	0.36		
19:0 CYCLO w8c	0.34		
TYPE STRAIN		TYPE STRAIN	
PROBABILITY		PROBABILITY	
<i>Comamonas testosteroni</i>	0.750	<i>Alcaligenes xylosoxidans</i>	0.376

Resin acid utilization

To determine each isolates' ability to use a range of resin acids as sole carbon source pure batch cultures were incubated for eight days in the presence of six resin acid congeners. Figure 2-3 presents the resin acid utilization profile

for each isolate as determined by the amount remaining after eight days incubation in comparison with an uninoculated negative control. In general, substrate specificity and degradative efficiency was quite variable between the isolates. Isolates such as A19-3, A19-6a, A19-6b, and D11-37 were efficient degraders able to use both the abietane and pimarane type resin acids tested. Other isolates were more substrate specific such as A19-1 and D11-23b that were optimal degraders for isopimaric acid (IPA) and less efficient for all the others.

To determine the relative contribution made by the ABA isolates to abietic acid degradation by the consortium, growth curves for four of the six degraders were constructed. The consortium culture reached stationary phase within 12 hours of inoculation which coincided with 100 percent depletion of the carbon source (Morgan unpublished thesis, 1993). This same growth/utilization profile was also observed for isolate A19-6a, see Figure 2-4. For isolates A19-2, A19-5, and A19-6b growth rate was similar, however, they were only able to use a portion of the available carbon source (54%, 38%, and 59%, respectively).

In an attempt to detect metabolites generated by the degraders, aliquots of the consortium culture were tested by GCMS. Even though degradation results mirrored the ones observed by HPLC no metabolites were detected in the organic extracts.

FIGURE 2-3. Resin acid utilization by abietic and dehydroabietic acid enriched bacteria isolated from pulp mill mixed liquor. Abbreviations are: ABA, Abietic acid; DHA, Dehydroabietic acid; NAA, Neoabietic acid; PA, Pimaric acid; IPA, Isopimaric acid; LPA, Levopimaric acid. Degradation is expressed as the amount of resin acid remaining in solution following eight days of incubation in the presence of the different isolates, relative to control (uninoculated) flasks. Pimaric acid and isopimaric acid belong to the pimarane-group of resin acids. The rest belong to the abietane group.

FIGURE 2-4. Growth versus ABA utilization by four ABA enriched isolates found in bleached kraft pulp mill mixed liquor. Symbols are x = growth measured as A_{600} and + = concentration of ABA in mg/L. All tests conducted in triplicate.

Hierarchical cluster analysis of isolates

To gain an understanding of the relatedness between the various isolates the phenotypic characteristics listed in Table A1, Appendix A were used to construct a dendrogram by hierarchical cluster analysis. As illustrated in Figure 2-5 two distinct clusters of resin acid degraders can be seen. The smaller of the two groups contains only DHA degraders that preferentially use carbohydrates as carbon sources. The larger group contains both ABA and DHA isolates that tend to utilize organic and amino acids rather than carbohydrate type substrates (see Table 2-3).

Genotypic analysis of isolates

ERIC PCR (de Bruijn, 1992) was used to determine whether phenotypic relatedness was a reflection of genotype. Figure 2-6 is a comparison of the amplification products from 10 of the 13 isolates; repeated attempts to obtain products for A19-2, D11-3b, and D11-25 were unsuccessful. As illustrated, there are obvious similarities between some isolate pairs such as A19-3 and A19-6a or A19-6b and D11-23b. Others, such as A19-5 and A19-6b, show distinctly different amplification patterns. In the former case genotypic analysis supports the phenotypic relatedness identified by cluster analysis. For A19-5 and A19-6b, however, these phenotypically similar isolates are very different genetically.

FIGURE 2-5. Dendrogram of similarities between resin acid degrading bacteria isolated from pulp mill mixed liquor. Cluster analysis using average linkage between groups was based on phenotypic characteristics listed in Table A1. The relative scale lists Jaccard Correlation Coefficients (Sneath, 1973) from Table A2 based on the same physiological characters. Groups of isolates discussed in the text are distinguished at the Jaccard similarity coefficient of 0.23 and include phenotype cluster 1 (isolates A19-3 down to A19-2), and phenotype cluster 2 (from D11-13 down to D11-3b).

Figure 2-6. Agarose gel (1.5 %) of ERIC-PCR products of 10 resin acid degrading bacteria isolated from pulp mill mixed liquor. The gel was loaded with 10 μ l of product and run for approximately five hours at 105 V (constant voltage). The marker in outside lanes is a 100 basepair ladder.

DISCUSSION

Isolation and characterization of resin acid degraders

The first objective of my research was the enrichment isolation of resin acid degrading bacteria from pulp mill mixed liquor. A preliminary attempt using concentrations below the solubility level for abietic and dehydroabietic acids was unsuccessful at obtaining isolates. The reason for this is considered to be due to carbon limitation since the defined media contained only a single source at a level that was reflective of environmental samples but too low to sustain growth. A second enrichment well above saturation level yielded 13 isolates that all had similar colony/cellular morphologies when grown on resin acid minimal media. Biochemical characteristics (Table 2-3) were more variable and carbon utilization profiles resulted in clustering of the isolates into two distinct groups, (Figure 2-5). The two groups are distinguished by the preference of isolates to use carbohydrate or organic and amino acid type substrates; a result similar to that found by other researchers (Mohn, 1995). Degraders that preferred carbohydrates as carbon sources were all enriched on dehydroabietic acid and represented a minority in the collection. Those that used more organic and amino acids as substrates included both abietic and dehydroabietic acid degrading isolates. This dichotomy may prove useful in understanding the biochemistry of resin acid degradation by doing a comparative study of the two clusters since the occurrence of carbohydrates in mixed liquors tends to be limited following treatment(Suntio *et al.*, 1988).

While carbon utilization can be useful in characterizing bacteria in comparative terms they are generally not very helpful in identifying environmental isolates. This is due to the fact that organisms isolated in a similar manner from the same source may be expected to utilize many of the same substrates. For this reason isolates A19-6a, an abietic acid degrader from the larger cluster, and D11-13 from the carbohydrate group were selected for fatty acid methyl ester analysis in an attempt to determine if there was a taxonomic difference between the two clusters. Table 2-4 lists the composition of fatty acids for these two isolates which were most similar to *Comamonas testosteroni* ($P=0.750$) and *Alcaligenes xylosoxidans* ($P=0.379$), respectively. In 1993 Fulthorpe *et al.*, did an in depth cluster analysis of isolates from the same mixed liquor source used in this study and also detected *C. testosteroni* but not *Alcaligenes* spp., specifically. Their *C. testosteroni* isolates were a part of the C cluster which had a similar carbon utilization profile to that of A19-6a. Bacteria in this cluster C group were isolated from the aerated lagoon and settling pond of the secondary treatment facility. Note that while the A19-6a fatty acid methyl ester comparison of $P=0.750$ is high enough to assign this isolate to the genus *Comamonas* it in fact contradicts the Biolog assignment of *Pseudomonas* sp. ($P=0.528$) that was later supported by 16S rDNA sequencing (see Chapter 5). Furthermore, two of the main cellular fatty acids (16:1 w7c and 16:0) and one of the minor ones (10:0 3OH) are the same as *Pseudomonas abietaniphilia* strain BKME-9 (Mohn, 1998). These isolates share other similarities such as cellular morphology and carbon source utilization.

Resin acid utilization

All tricyclic diterpenes have a similar backbone of three fused rings with structural differences occurring in the location and degree of saturation and the type of substituent found on the C13 carbon. In accordance with the results of similar studies (Bicho *et al.*, 1995 and Mohn, 1995), the members of this collection are affected by these differences as indicated by the range and efficiency of resin acid congeners used by each isolate (see Figure 2-3). It is unfortunate that no patterns of utilization could be discerned from the results since this might have provided some insight into the biochemistry of resin acid degradation. Rather, some isolates like A19-6b were able to degrade all congeners of both structural types with equal efficiency, whereas others such as D11-23b degraded all congeners poorly, isopimaric acid excepted. This strain along with other pimarane degraders (A19-1, A19-2, A19-3, D11-3b, D11-9, D11-13, D11-23a, and D11-37) may prove of particular interest in future studies since some researchers have reported difficulties in isolating bacteria able to grow on this resin acid (Bicho *et al.*, 1995). Nevertheless, other researchers have since been successful (Wilson *et al.*, 1996), and Kutney and colleagues (1981) have found *Mortierella isabellina* to be particularly useful at the removal of this resin acid from pulp mill mixed liquor samples.

While enrichment protocols could affect subsequent utilization profiles this is not considered likely for the abietic acid isolates, eg., A19-1, A19-2, and A19-5, since they were able to use other congeners more efficiently. This is

particularly noteworthy since the lab grade carbon sources used are only approximately 70 percent pure so that contamination may have been sufficient to aid in the isolation of these degraders that prefer other congeners. The same argument appears to hold true for the carbohydrate group of isolates that were all enriched on DHA since they all use pimaric and/or isopimaric acid more efficiently. Other researchers have suggested that culture losses of this resin acid may be due to adsorption to cell or vessel surfaces (Mohn, 1995) or isomerization (Wang *et al*, 1995) to other congeners. Isomerization is not considered to be a source of removal here since no new peaks were detected during analysis. Adsorption losses are also not considered to be a factor because degradation was 100 percent for two of the isolates, D11-23a and D11-25, and the detection protocol used measures dissolved resin acids only and reports values relative to the dissolved fraction in an uninoculated negative control. This latter point is of particular importance when dealing with insoluble carbon sources such as resin acids or petroleum based hydrocarbons. With these substrates growth on the dissolved fraction shifts the equilibrium defined by the solubility constant so that precipitated resin acid is continuously dissolved until all available substrate is consumed (Morgan, 1993). With our protocol we have been able to quantitatively show resin acid degradation in a manner simple enough to be used for routine culture analysis, however, the technique has limited application in the detection of resin acids in environmental samples. This is because much of the undissolved species responsible for mixed liquor toxicity may be lost during sample preparation.

Growth/carbon utilization curves for four of the abietic acid isolates were constructed to determine their contribution to degradation of this resin acid by the consortium (see Figure 2-4). Consortium refers to the mixed culture from which the abietic acid degraders were isolated. Isolate A19-6a achieved the highest optical density and degraded all of the ABA detectable by HPLC within 12 hours of inoculation. This result was mirrored by a consortium culture tested in the same manner and GCMS analysis showed that all ABA was depleted after 18 hours (Morgan, 1993). The fact that this isolate has a higher growth rate than the other degraders may be its selective advantage in mixed cultures since the results of the congener utilization experiment suggest that A19-6b is better overall resin acid degrader than is A19-6a (see Figure 2-3). In the case of A19-2 and A19-5 lower cell densities may reflect carbon utilization profiles since both prefer pimaric acid over the other congeners. This may also explain their survival in the enrichment culture. In any case, the co-existence of phenotypically diverse degraders in an enrichment culture implies that resin acid degraders in pulp mill mixed liquor may also be taxonomically diverse; a fact that is supported by the literature (Hemmingway and Greaves, 1973; Bruce and Clark, 1994; Mohn, 1995; Mohn and Stewart, 1995, and Mohn, 1998).

Genotypic relatedness of isolates

Since phenotypic analysis suggested that two distinct groups of resin acid degraders were isolated and that some, such as A19-3/A19-6a, were probably sibling species, ERIC PCR was used to determine the genotypic relatedness of

these bacteria. As can be seen by the gel presented in Figure 2-5 isolates with high Jaccard correlation coefficients (0.98) based on phenotypic characteristics give the same amplification patterns as in the case of A19-3 and A19-6a. In other comparisons patterns differ, eg., A19-5 and A19-6b, even though the correlation coefficient is higher than it is for A19-6b/D11-23b at 0.87 and 0.82, respectively. These results illustrate the limitations of using phenotypic data as a sole basis for determining phylogenetic relationships among bacteria isolated from the same environment.

CONCLUSION

The first objective of my research was to isolate and characterize resin acid degrading bacteria that form a part of the natural microflora found in pulp mill mixed liquor. The results presented in this chapter support the preliminary assumptions made by my hypothesis in that a collection of degraders could be isolated by selective enrichment of inoculum originating from an environment that had already been exposed to resin acids. The findings also suggest that resin acid degraders may be found in a variety of taxa that co-exist by exploiting selective advantages such as the ability to use different resin acid congeners. Other selective advantages of these species, *in situ*, may not be immediately obvious.

CHAPTER 3

Transposon mutagenesis and subcloning of tricyclic diterpene degradation genes from *Pseudomonas* sp. strain A19-6a and mutant A76-47

INTRODUCTION

Targeted mutagenesis

The advent of molecular biological techniques has enabled the development of genetics based analytical tools as a means of studying microbial ecology (Sayler and Layton, 1990). Incumbent upon the development of probes based on resin acid catabolic genes is their localization in the genome of confirmed degraders. Transposon mutagenesis is a gene manipulation technique that can be used to localize targeted genes by phenotypic selection, insertional inactivation, and physical mapping of mutated genes. Tn5 is particularly amenable to this purpose due its high frequency of transposition (up to 10^{-5} per cell), its randomness of insertion, the polarity of the mutation due to termination signals carried by the element, and the stability of transposition that reduces the occurrence of insertion site rearrangements and secondary transposition events (de Bruijn and Lupski, 1984). Constructed transposons based on Tn5 carry these same properties in addition to the features added to meet specific experimental needs.

Wilmes-Riesenberg and Wanner (1992) developed a series of Tn5 derived elements that could be used to study transcription, translation, and cell surface localization. They were based on the transposon Tn*phoA* in which most of the *phoA* gene that is inserted into the left IS50 element is replaced by *lacZ* to

enable colorimetric selection of transcriptional or translational mutants. This is possible because it is the IS50R element that is responsible for providing transposition functions. Mutations in IS50L only affect transposition frequency if they occur in the inverted repeat sequence (Rothstein *et al.*, 1980). Phenotypic selection of mutants is via kanamycin resistance (*aph*) which is carried between the two insertion sequences (Wilmes-Riesenberg and Wanner, 1992) and is transcribed from the NPT II promotor that occurs at the inside right end of the IS50L element (Rothstein *et al.*, 1980). Transcriptional mutants are generated with the operon elements that produce blue colonies when the mutated gene is being actively transcribed. For translational mutants a protein fusion construct is used that requires the mutation to be in frame for *lacZ* activity to occur. Since the restriction sites carried by the Tn*phoA*' elements are known, mapping of the insertion region in a mutant is achieved by Southern hybridization using probes derived from the donor transposon (Southern, 1975).

Transposon delivery to environmental isolates

A necessary first step to this type of mutagenesis is delivery of the transposon to the recipient in a manner that will assure transposition. This is achieved through the use of a suicide vector delivery system that cannot be maintained by the recipient so that transposition must occur for the selectable marker to be effective. In the case of *Escherichia coli* a wide range of these vectors are available for use in both transduction and conjugation experiments.

Unfortunately, phage vectors generally have narrow host ranges so are unsuitable for transduction of environmental isolates (Deho *et al.*, 1992). There are, however, plasmid delivery systems available whose host range is sufficiently broad to enable efficient transfer to pseudomonads while ensuring that the vector is not maintained. Vectors derived from plasmids belonging to the Incompatibility P group are examples of such systems (Thomas and Smith, 1987).

The *TnphoA'* elements are carried on a plasmid designated pBW30 that is designed to be stably maintained in a suitable *E. coli* donor that can transfer the plasmid to the recipient. Under normal circumstances, the recipient should not be able to maintain the plasmid. Features of the vector include the β -lactamase (*bla*) gene to ensure that the system is maintained as an autonomously replicating unit in the donor host, an R6K *ori* for vegetative replication, and the RP4 *oriT* region for conjugal transfer between donor and recipient (Wilmes-Riesenberg and Wanner, 1992). In terms of the R6K *ori* the absolute requirement for the Pi protein in vegetative replication ensures that organisms that do not carry the *pir* gene (encoding Pi) cannot maintain the plasmid (Scott, 1984). For suicide vectors the gene is provided *in trans* by a lysogenic phage (λ *pir*) carried on the host chromosome. As with other conjugative plasmids (Clark and Warren, 1979) only the RP4 *oriT* region is *cis*-acting and thus strains such as SM10 (Simon *et al.*, 1983) that carry chromosomally encoded transfer genes can mobilize the vector for transfer to a target recipient. Taken together the pBW30::*TnphoA'* elements constitute a broad-host range, suicide delivery,

transposon mutagenesis system (Wilmes-Riesenberg and Wanner, 1992) suitable for localization of the resin acid degradation genes in selected degraders.

Inverse polymerase chain reaction

The advent of *in vitro* DNA synthesis by the Polymerase Chain Reaction (PCR) was a revolutionary step in molecular biology leading to a host of experimental techniques used in diagnostics, taxonomy, mutagenesis, cloning, and sequencing (Mullis and Faloona, 1987). Inverse PCR is among these techniques (Ochman *et al.*, 1988) and is useful for the amplification of DNA flanking the site of transposon mutagenesis. In this procedure genomic DNA is digested and then religated under conditions that promote recircularization of the fragments (Collins and Weissman, 1984). Primers that read outwards from the ends of the transposon are then used to amplify the unknown DNA that flanks the mutation site in the original mutant and falls between the primers in the recircularized fragment. In the case of Tn5 based transposons only a single primer that is complementary to the outside ends of the IS50 inverted repeats is necessary (Rich and Willis, 1990).

Cloning strategies

Once a target gene has been localized by transposon mutagenesis it becomes necessary to isolate it and the intact gene from the mutant and wild type genomes, respectively, prior to sequencing and probe development. Direct

cloning of restricted genomic DNA into plasmids is possible but is limited by the size of fragment that can be efficiently cloned (Old and Primrose, 1989). Phage based cloning systems are more amenable to the cloning of large fragments, particularly since many commercially available systems are designed to enhance the packaging and propagation of recombinants. Purified phage DNA can then be subcloned into high-copy number plasmids (Keen *et al.*, 1988) that can be maintained in *E. coli* for easy manipulation and molecular analysis.

Objective IIa

The foregoing summarizes the basic theoretical aspects of the preliminary work for the second objective of my research. Since little is known about the genetics of resin acid degradation, basic research at the molecular level is necessary prior to probe development and characterization of the genes.

MATERIALS AND METHODS

Transposon mutagenesis

Isolates A19-6a and D11-9 were selected for transposon mutagenesis based on their effectiveness at degrading resin acids (see Chapter 2) and their sensitivity to kanamycin. Mutagenesis was achieved by bacterial conjugation between the recipient and the donor *Escherichia coli* K-12 strain BW16948 (Wilmes-Riesenberg and Wanner, 1992) that carries an R6K/RP4 based suicide vector pBW30 with the Tn5 derived Tn*phoA*'-1 transposon. The latter was engineered to provide a colorimetric screening mechanism for detecting transcriptional fusions.

Overnight cultures of donor and recipient cells grown in TYE (1% tryptone, 1% yeast extract, 0.5% NaCl) were used for conjugation experiments. Donor cell cultures were supplemented with 50 µg/ml kanamycin (TYE_{Km}). Donor cells (1 ml) resuspended in sterile saline (0.85% NaCl) were filter mated overnight on TYE with 20 µl of recipient cells to give a donor:recipient ratio of approximately 20:1. Filters were resuspended in 4.5 ml of saline and 100 µl of a 10 fold dilution was plated on 4mM quinic acid (Sigma) in minimal media A supplemented with 50 µg/ml kanamycin (AQA_{Km}). Transconjugants were collected after approximately 96 hours incubation at 32°C and replica plated onto 1mM ABA plus kanamycin (ABA_{Km}) and AQA_{Km}. Replica plates were incubated 21 days

and transconjugants that appeared to be ABA⁻ and AQA⁺ were retained for batch culture testing.

Putative ABA⁻ mutants were purified once on AQA_{Km} and grown 48 hours on TYE_{Km} prior to inoculation of batch cultures (1 mM ABA in minimal media A) and preparation of frozen stocks (20% glycerol). Liquid cultures were incubated eight days at 32°C with shaking prior to testing for resin acid degradation by HPLC as previously described (Chapter 2). Cultures that showed no or reduced resin acid degradation relative to positive and negative controls were retested to confirm the ABA⁻ phenotype.

Total genomic DNA extraction

Total genomic DNA was extracted from overnight cultures (10 ml) grown in TYE. Cells were resuspended in 2.82 ml TE (10mM Tris pH 8.0, 1mM EDTA), 75 µl 20% SDS, and 15 µl Proteinase K (20 mg/ml, Boehringer Mannheim) and incubated one hour at 37°C. Five hundred microliters of 5M NaCl and 400 µl of CTAB/NaCl (10 % hexadecyltrimethyl ammonium bromide (Sigma), 0.7M sodium chloride) were added with thorough mixing after each addition. Following a 15 minute incubation at 65°C proteins and cellular debris were removed by extraction with an equal volume of chloroform/isoamyl alcohol (24:1). The aqueous phase was collected and treated with RNase (20 µg/ml) for one hour at 37°C. RNase was removed by phenol/chloroform/isoamyl alcohol (25:24:1) extraction and residual phenol was removed by chloroform/isoamyl alcohol

extraction. DNA was precipitated with 0.6 volumes of isopropanol and gentle mixing, collected by centrifugation for 10 minutes at 12,000g, and washed with 70% ethanol. The pellet was air dried and resuspended in an appropriate volume of TE. One microliter aliquots were run on a 0.7% agarose gel and tested by UV absorbance to determine quality and yield.

Southern blotting and hybridization

New England Biolab restriction enzymes were used for all digestions unless stated otherwise. Twenty microlitre restriction digests (16 hours) containing 5 µg of genomic DNA were run overnight on a 0.7 percent agarose gel. DNA was depurinated by UV exposure (5 min) and vacuum transferred (Vacugene™ XL, Pharmacia Biotech) to positively charged nylon membrane (Boehringer Mannheim). DNA was denatured 20 minutes with 1.5M NaCl, 0.5M NaOH, neutralized 20 minutes with 1M Tris pH 7.5, 3M NaCl, and transferred for 1 hour with 10X SSC. The membrane was rinsed briefly in 6X SSC and allowed to air dry prior to UV cross-linking (5 min) of the DNA. Membranes were pre-hybridized in 10 ml of hybridization solution (5X SSC, 0.02% SDS, 0.2% SLS, 2% skim milk) for a minimum of 1 hour at 68°C. DIG labelled probe (1-7 µl) was diluted in 10 ml of hybridization solution, boiled 10 minutes, ice chilled, and transferred to the hybridization tubes for incubation overnight at 68°C. Following hybridization, membranes were washed twice at room temperature for 5 minutes with 2X SSC, 0.1% SDS and twice at 68°C for 15 minutes with 1X SSC, 0.1% SDS.

For probe detection, membranes were equilibrated 1 minute in Buffer 1 (100mM Tris pH7.5, 150mM NaCl) and blocked for a minimum of 30 minutes at room temperature with Buffer 2 (Buffer 1 + 2% skim milk). One microlitre of Anti-DIG Fab conjugate (Boehringer Mannheim) was diluted in 10 ml of Buffer 2 and allowed to bind to the labelled probe for 30 minutes at room temperature. Unbound Anti-DIG was removed by two 15 minute washes at room temperature with Buffer 1 + 0.3% Tween 20 (Sigma). The alkaline phosphatase conjugate was activated by a 2 minute wash in Buffer 3 (100mM Tris-HCl pH 9.5, 100 mM NaCl). Chemiluminescent (CSPD ready-to-use) or colorimetric (NBT and X-Phos) substrates were then added in accordance with the manufacturers recommendations. Kodak X-OMAT AR films were developed using the manufacturer's protocol for GBX developer and fixer. Wet membranes (chemiluminescent detection only) for re-probing were stripped twice for 15 minutes at 37°C with 0.2N NaOH, 0.1% SDS, rinsed twice for 15 minutes at room temperature in 2X SSC, and stored at 4°C in 20 ml 2X SSC until required for prehybridization.

Inverse polymerase chain reaction

Inverse PCR of transconjugants A103-37 and A104-38 was used to develop probes of the mutated genes for localization of intact genes in the wild type A19-6a strain. Genomic DNA (10 µg) was digested with *Stu*I and religated under dilute conditions to promote recircularization of restricted fragments. A single 20-mer primer (5'GGTTCCGTTTCAGGACGCTAC-3') reading outward from

the ends of the IS50 elements was used to amplify target DNA. The reaction conditions used were determined empirically for each transconjugant using an orthogonal array to optimize variables (Cobb and Clarkson, 1994). Amplification conditions were as specified in Rich and Willis (1990).

Direct plasmid cloning of target gene

Direct cloning of an 11.4 Kb *NheI* fragment from transconjugant A76-47 into the *SpeI* and *XbaI* sites of pUC128 was also used to isolate target DNA flanking the *TnphoA'*-1 insert. A *Bam*HI fragment from S1 (pUC128 *SpeI*::A76-47 *NheI*) was sub-cloned into pUC18 for sequencing of the target DNA upstream from the transposon.

Preparation and screening of lambda libraries

Phage libraries of Lambda FIX[®] *Xho* I Partial Fill-In (Stratagene) clones were constructed based on the manufacturer's recommendations. Restriction enzyme concentrations for partial digestion of genomic DNA from wild type A19-6a (*Bam* HI) and mutant A76-47 (*Bgl* II) were determined empirically by 5-fold serial dilution of the restriction enzyme in a reaction mixture containing 5 µg of DNA per reaction. Duplicate reactions were incubated for one and two hours at 37°C. Digests were run on a gel, transferred to nylon membrane, and probed to determine the optimal concentration of enzyme for partial digestion. For library construction partial digests, 900 µl volumes, containing 1X reaction buffer, 1X

BSA (if necessary), 90 μ l of DNA, and restriction enzyme diluted 600 and 3000 times were incubated two hours at 37°C. The reaction mixture was extracted once with 300 μ l phenol/chloroform/isoamyl alcohol and DNA was precipitated overnight at -20°C with one tenth volume 10M ammonium acetate and an equal volume of isopropanol. Following centrifugation at 12,000g for 10 minutes and a 70% ethanol wash the air dried pellet was dissolved overnight at 4°C in 100 μ l of sterile distilled water.

Partial fill-in of the overhanging single stranded DNA resulting from restriction digestion was necessary to prepare the insert for ligation to the vector arms. The DNA, 55 μ l in 30 μ l 10X buffer and 202 μ l dH₂O, was heated at 65°C for ten minutes and held on ice during the addition of nucleotides (5 μ l each of 10mM dGTP and dATP) and Klenow polymerase (5 U/ μ l, Stratagene). The reaction was incubated at room temperature for 30 minutes followed by a second round of DNA heating, nucleotide/enzyme addition, and room temperature incubation. The reaction was stopped by the addition of 33 μ l 10X STE (1M NaCl, 200mM Tris-HCl pH 7.5, 100mM EDTA) and 100 μ l 1X STE. Enzyme was extracted with 300 μ l phenol/chloroform/isoamyl alcohol and DNA was precipitated four hours at -20°C with one fifth volume ammonium acetate and two volumes of absolute ethanol. The precipitated DNA was centrifuged 30 minutes at 12,000g, washed with 70% ethanol, dried and resuspended in 50 μ l dH₂O.

The insert DNA (0.4 μ l) was ligated to 1 μ g of lambda arms in a 5 μ l reaction containing 1X ligation buffer, 1mM rATP, and 2 units T4 ligase (Stratagene); with incubation overnight at 4°C. One microliter of each ligation reaction was packaged with Gigapack® III (Stratagene) packaging extracts as recommended by the manufacturer and incubated 105 minutes at room temperature. Libraries were stored at 4°C in 500 μ l SM buffer (5.8% NaCl, 2% MgSO₄ 50 mM Tris pH 7.5, 0.01% gelatin) and 20 μ l chloroform prior to screening.

Libraries were screened by plaque hybridization using DIG labeled probes (Boehringer Mannheim). One microlitre of undiluted and 10-fold diluted phage was added to 200 μ l of *Escherichia coli* XL1-Blue MRA (P2) cells (Stratagene), prepared in accordance with the manufacturers recommendation, and incubated at 37°C for 15 minutes. The suspension was added to 3 ml of CY top agar (0.5% NaCl, 0.2% MgSO₄·7H₂O, 0.5% yeast extract, 1.0% casitone, pH 7.5, 0.7% agarose LE (Boehringer Mannheim)), plated on prewarmed CY agar (1.6%) plates, and incubated overnight at 37°C. Plaques were transferred to sterile nylon membranes for 1 minute and blotted briefly on Whatman paper. The DNA was denatured 5 minutes by incubating membranes on two sheets of Whatman paper saturated with 0.5M NaOH, 1.5M NaCl, 0.1% SDS, blotting briefly on dry Whatman, and neutralized by incubating in 1.0M Tris pH 7.5, 1.5M NaCl soaked Whatman for 15 minutes. The membranes were equilibrated with 2X SSC for 15 minutes and allowed to air dry prior to UV-crosslinking (5 minutes). Protein and cellular debris were removed by saturating the membranes with 2 mg/ml

Proteinase K (Boehringer Mannheim) in 2X SSC and incubating at 37°C for 1 hour followed by two surface washes using Whatman soaked with distilled water. Membrane hybridization and detection was performed as described above. Tertiary screening of randomly selected probe positive plaques was necessary to obtain purified phage suspensions.

Purified phage suspensions were amplified by adding 2.5×10^4 phage per 300 μ l of host cells to give confluent lysis of cells. Phage were collected by overlaying plates with 5 mls of SM buffer and incubating overnight at 4°C. Phage DNA was extracted for plasmid sub-cloning by using a high multiplicity of infection protocol. Host P2 cells were grown at 32°C overnight in 50 ml CY supplemented with 0.2% maltose. One millilitre was transferred to 500 ml of CY broth and incubated at 37°C with vigorous shaking to $OD_{600} = 0.541$. Cells were infected with phage (25:1) and the culture incubated overnight to ensure complete lysis. Ten millilitres of chloroform was added to the lysed culture followed by a 30 minute incubation at 37°C with moderate shaking. Debris was removed by centrifugation for 10 minutes at 5000g in a Sorval centrifuge, The supernate was equilibrated to room temperature for one hour and bacterial nucleic acids were removed by a one hour room temperature incubation with 1 μ g/ml RNase and 50 units of DNase. NaCl to a final volume of 1M was added by gentle swirling for 20 minutes, incubated on ice 1 hour, and spun in a chilled rotor for 10 minutes at 5000g. Fifty grams of PEG 8000 was stirred into the supernatant (volume of 50 ml) slowly for one hour followed by incubation on ice 1 hour and a 15 minute centrifugation. The well drained viral pellets were

resuspended in a final volume of 10 ml SM buffer, extracted once with 10 ml chloroform, and stored overnight at 4°C. Phage coat was removed by a one hour incubation at 37°C with 0.5% SDS and 50 µg/ml Proteinase K, phenol/chloroform/isoamyl alcohol and chloroform/isoamyl alcohol extraction (as described above), and overnight precipitation with 0.6 volume isopropanol. Spooled DNA was washed in 70% ethanol, spun 10 minutes at 12,000g, dried, and resuspended in TE.

Phage DNA from M5 (λ ::A76-47;*Bgl*II) and 25b (λ ::A19-6a;*Bam*HI) was sub-cloned into pUC128 using *Xba*I and *Sac*I, respectively, prior to sequencing, see Chapter 4.

RESULTS

Transposon mutagenesis

Transposon mutagenesis using *TnphoA'*-1 was approximately 10 percent more efficient for the dehydroabietic acid isolate D11-9 than it was for the abietic acid degrader (see Table 3-1). Nevertheless, because HPLC results were less reliable for DHA the transconjugants were not used for further study. Transfer frequency for *Pseudomonas* sp. A19-6a was 2.31×10^{-6} resulting in a total of 21 ABA⁻ transconjugants following duplicate batch culture screening. Three mutants, A76-47, A103-37, and A104-38 that were confirmed by Biolog to be offspring of the wild type A19-6a, were selected for further analysis based on the consistency of the HPLC results between the two trials. Mutants A76-47 and A104-38 formed white colonies on AQA_{Km} plus X-GAL, whereas strain A103-37 formed blue colonies on the same medium; a characteristic that is indicative of a transcriptional/translational fusion.

Inverse PCR and direct plasmid cloning of target DNA

The series of optimization experiments using an orthogonal array (Cobb and Clarkson, 1994) to determine the correct combination of reaction components was successful in generating products for A103-37 and A104-38. Even though the IS50 primers used also generated PCR products for the wild type strain A19-6a (negative control), it was assumed that size differences

between these and some of the transconjugant bands meant that different regions of the genome were being amplified and presumably represented the target resin acid catabolic genes. While efforts to develop gene probes from these transconjugants were ongoing, it was discovered that the mutants A103-37 and A104-38 were ABA⁺ when grown on solid media even though they had twice tested ABA⁻ in liquid culture. For this reason they were abandoned from further study and all subsequent work focused on the mutant strain A76-47.

TABLE 3-1. Summary of transposon mutagenesis results for resin acid degraders A19-6a and D11-9. All values are averages unless otherwise stated. RA stands for resin acid.

PARAMETER	A19-6a	D11-9
Total number of matings	13	11
Donor:recipient ratio	17:1	18:1
Number of transconjugants per mating:		
blue colonies	3053	15931
white colonies	1467	8089
total colonies	4519	23977
Transfer frequency	2.31×10^{-6}	2.32×10^{-5}
Total number of transconjugants screened by patching	12029	8570
Total number of patches unable to grow on RA plates	257	179
Total number of mutants testing RA- in batch culture	80	53
Total number of mutants retested in batch culture	67	46

Concurrent efforts to clone out the mutated gene from A76-47 into pUC128 lead to the isolation of two putative clones S1 (pUC128 *SpeI*::A76-47 *NheI*) and X1 (pUC128 *XbaI*::A76-47 *NheI*). Sequencing of the pUC18::S1 *Bam*HI sub-clone revealed that the region that was thought to be target DNA upstream from the transposon insertion site was in fact very similar in sequence to the DNA found in a wide range of cloning vectors. This result suggested that the entire pBW30::Tn*phoA*'-1 suicide plasmid had incorporated itself into *Pseudomonas* sp. A19-6a to give the A76-47 mutant.

Mapping of transconjugant A76-47

In an effort to get a picture of how the 16948 donor suicide plasmid was incorporated into transconjugant A76-47 a series of probes derived from the plasmid was used to map the insertion. Figure 3-1 illustrates the map of mutant A76-47 and the orientation of the pBW30::Tn*phoA*'-1 insert. Restriction sites shown were the ones used for mapping and in some cases may also represent the ones used to generate the probes. The *Bam*HI fragment that generated the pBW30 probe corresponds to the S1 fragment that was presumed to be target DNA. Note that the map was constructed from the size of probe positive bands and the location of restriction sites in both the vector portion of the insert and the A76-47 genome are relative to known sites in Tn*phoA*'-1.

(A19-6a::pBW30::Tn $phoA'$ -1)

Preparation, screening, and plasmid sub-cloning of cosmid libraries

Having mapped the vector/transposon insertion site in A76-47, it was decided that cosmid cloning would be the most efficient method of cloning out a large portion of target DNA. Since a restriction map of the transposon was available (Wilmes-Riesenberg and Wanner, 1992), it was known that a *Bgl* II partial digest would give a clonable insert that could be probed with *tnp* and presumably provide the maximum amount of target DNA. Following tertiary screening of the A76-47 cosmid library a purified phage designated M5 was digested with *Xba* I to release the cloned fragment to enable sub-cloning to a plasmid vector suitable for further molecular analysis and sequencing (see Chapter 4). This plasmid subclone was designated YFG23 and was used for the majority of genetic manipulations described in this thesis.

Since an intact copy of the mutated gene was also required, a wild type library from *Pseudomonas* sp. A19-6a needed to be constructed. To do this it was first necessary to map the region of the gene in A19-6a to determine which restriction enzyme would be most suitable for library construction. This map is shown in Figure 3-2 and was generated with the aid of the probe shown above the map that was cut from YFG23. This probe was selected on the basis of preliminary sequence data from YFG23 that indicated that it included target DNA near the point of insertional inactivation in A76-47. Digestion of A19-6a with *Bam* HI generated a 3.7 Kb fragment of which approximately 2552 basepairs corresponded to the region on the other side of the mutation point. This was

considered sufficient to include the complete coding sequence of most genes. In addition, an similarity search of the probe sequence indicated that it was similar to genes coding for proteins (CoA ligase) with approximatey 550 amino acids (1650 bp) of which 278 (834 bp) were included in the probe. Tertiary screening of the wild type library with the same probe lead to the isolation and purification of phage 25b carrying an insert of approximately 25 Kb. Viral DNA was digested with *Sac* I to release a 4.9 Kb fragment that included the intact gene for sub-cloning into the recombinant plasmid 25b/f later used to sequence the entire mutated gene (see Chapter 4).

FIGURE 3-2. Map of wild type A19-6a corresponding to the region surrounding the mutation site in A76-47. The arrowhead indicates the site of insertional inactivation by pBW30::Tn*phoA*'-1 in A76-47. The location of the probe cut from YFG23 used to construct the map is indicated. Regions of DNA represented by the plasmid sub-clones YFG23 and 25b/f and the phage clone 25b are included.

DISCUSSION

Transposon mutagenesis

Selection of strains A19-6a and D11-9 was based on their ability to degrade resin acids efficiently combined with their sensitivity to kanamycin which is the selectable marker for *TnphoA'*-1 transposition. Transposition frequency for the two strains was 10^{-6} and 10^{-5} , respectively (see Table 3-1), which is 10 and 100 fold lower than reported by Wilmes-Riesenberg and Wanner (1992) for this element. These differences are directly attributable to strain variations since their experiments were conducted with *E. coli* only and environmental isolates were used in this case.

While the *TnphoA'*-1 transposon was designed to enable colourimetric selection of target mutants (Wilmes-Riesenberg and Wanner, 1992) it was still considered necessary to screen both white and blue colonies due to the nature of the mutation being sought. Since catabolic genes were the target, transconjugants were selected on quinate minimal medium because it is counterselective for the *E. coli* donor and is a simple enough carbon source that it should not interfere with the detection of mutants that had lost the ability to use resin acids as a sole source of carbon and energy. Based on this and the properties of operon mutants, target transconjugants should be amongst the white colonies. Blue colonies would have been actively transcribing something that was necessary for growth on this medium. In *Acinetobacter calcoaceticus* the enzymes involved in quinate transformation are peripheral to the central β -

ketoadipate pathway that converts protocatechuate into acetyl CoA and succinate (van Kleef and Duine, 1988 and Harwood and Parales, 1996). The genes for quinate transformation are located between the β -ketoadipate pathway genes and the genes involved in degradation of *p*-hydroxybenzoate, forming a supraoperonic cluster, involved in the degradation of aromatic compounds (Elsemore and Ornston, 1994 and Harwood and Parales, 1996). Since transcriptional readthrough of supraoperonic genes has been demonstrated, and the resin acid catabolic genes could be part of a similar genetic arrangement, it was felt that possible target mutants could be lost if both blue and white colonies were not screened for insertional inactivation. The fact that transconjugant A103-37 is blue and unable to degrade abietic acid in liquid media supports this assumption. Unfortunately, both A103-37 and A104-38 were still able to grow on solid media containing abietic acid as sole carbon source and as such were no longer considered suitable for the development of resin acid catabolic gene probes.

Direct plasmid cloning and mapping of A76-47

Direct plasmid cloning of the mutated gene from transconjugant A76-47 resulted in the generation of two putative recombinant plasmids (S1 and X1). Hybridization results with the *aph* gene probe suggested these plasmids contained wild type A19-6a DNA that flanked the transposon insertion site in A76-47. Sequencing, however, identified the target region as having 95.8 percent identity with the pBAD-GFPuv cloning vector (Cramer *et al.*, 1996). This

suggested that in transconjugant A76-47 the mutation event that had occurred to knock-out the gene involved in resin acid catabolism had in fact been integration of some or all of the vector/transposon system rather than simple transposition of the *TnphoA'*-1 element. A series of probes derived from the complete pBW30::*TnphoA'*-1 construct was used to generate the map illustrated in Figure 3-1 that shows the orientation of the transposon and vector in A76-47. A similar integration event appears to have occurred in A104-38 since this transconjugant carries the *bla* gene for ampicillin resistance that is not found in the wild-type A19-6a strain (data not shown). Both A76-47 and A104-38 show phenotypic ampicillin resistance that is not shown by the parent A19-6a.

Genetic recombination

We can think of recombination events as being one of three basic types (Weaver and Hedrick, 1992). The first type comprises site-specific recombination that requires the presence of a particular target site on the recipient DNA; integration of temperate viruses into bacterial genomes is an example of this. The remaining two types are random in nature and may require some degree of sequence similarity as in the case of homologous recombination or virtually none at all as in the case of illegitimate recombination. The latter is exemplified by transposition events. The detection of vector DNA sequences in A76-47 suggests transposition was not the mode of mutation observed here. Since homologous recombination is the next logical alternative, we must assume that there is some similarity between a sequence found on the pBW30::*TnphoA'*-1 construct and the region of insertion in the A19-6a/A76-47 genome. Others

have also observed whole plasmid integration during transposon mutagenesis as illustrated by Hofer and colleagues (1996) in their mutation of the polychlorinated biphenyl/chlorobenzoate degrader *P. putida* P111 using a minitransposon carried on a Pi protein dependant mobilizable vector.

Genomic libraries

For mutant A76-47 library construction, *Bgl* II was the obvious choice of enzyme since most of the insert would have been target DNA based on the location of the transposon sites (Wilmes-Riesenberg and Wanner, 1992). Mapping of the wild type strain A19-6a with a probe generated from the mutated gene in A76-47 confirmed that *Bam* HI would generate a 10 Kb insert likely to contain an intact copy of the target resin acid degradation genes (see Figure 3-2).

CONCLUSION

Results of the preliminary work involved in resin acid catabolic gene probe development indicate that transposon mutagenesis is a useful technique for localizing target genes even though vector loss may not always occur. Nevertheless, legitimate transposition events tended to be the majority as evidenced by the higher proportion of blue to white colonies.

CHAPTER 4

Characterization of the *tdt* locus from *Pseudomonas* sp. strain A19-6a

INTRODUCTION

Catabolic gene organization

Given that the enzymes of complex metabolic pathways must function in a coordinated fashion with an ordered mass transfer of substrates and products within the confines of the bacterial cell, one wonders if physical proximity is reflected in the organization of pathway genes in the genome. If we look at *E. coli* we find that the genes involved in fatty acid metabolism are spread out at numerous locations around the genome map (Nunn, 1986), but that those involved in carnitine and homoprotocatechuate metabolism are not (Eichler *et al.*, 1994 and Roper *et al.*, 1993). In Pseudomonads the housekeeping genes of central metabolic pathways and anabolic functions tend to be linked together in one half of the genome, with the other half carrying more random clusters of catabolic genes (Krawiec and Riley, 1990). These clusters, however, may themselves show genetic organization as illustrated by the two examples shown in Figure 4-1.

FIGURE 4-1. Catabolic gene organization in *Pseudomonas putida* strain U (A) and *Acinetobacter calcoaceticus* (B). A generalized overview of each pathway is included. Adapted from Olivera *et al.*, 1998 (A) and Gerischer *et al.*, 1998 (B).

A. Phenylacetic acid catabolon and pathway of *Pseudomonas putida* U

B. *pca-qui-pob* supraoperonic cluster of *Acinetobacter calcoaceticus*

Figure 4-1A is the phenylacetyl-CoA catabolon of *Pseudomonas putida* U that is a core pathway involved in the degradation of related substrates via the common phenylacetyl-CoA intermediate (Olivera *et al.*, 1998). While the function and requirement for orfs 1 and 2 have not been identified, all other genes in this system have been studied. As indicated, there are two regulatory proteins *phaN* (a repressor) and *phaM* (function unknown) and three promotor sites suggesting that expression of the pathway is tightly controlled. All of the genes involved in similar steps are linked, eg., β -oxidation, ring hydroxylation, and transport and may actually represent three contiguous operons. The ligase encoded by the *phaE* gene is essential for activation of phenylacetic acid, although pathways generating the phenylacetyl-CoA intermediate can still feed into the pathway downstream from this step. The *phaL* gene encodes a ring-opening enzyme. It is noteworthy that these authors found a high degree of short repeat sequences in this stretch of DNA and speculate that, because it is a convergent point pathway, these sequences may contribute to selective evolution by mediating genetic translocations so as to expand the usefulness of an existing catabolic pathway (Olivera *et al.*, 1998).

Figure 4-1B illustrates the supraoperonic cluster of *pca-qui-pob* genes in *Acinetobacter calcoaceticus* (Elsemore and Ornston, 1994). In this example three separate operons that have been independently studied have been shown to be genetically linked (Gerischer *et al.*, 1998). Here again the core protocatechuate pathway is tightly regulated by *pcaU* which also induces the *qui* operon, one of the two peripheral pathways that feed into protocatechuate illustrated here.

An operon encoding dehydroabietic acid degradation

During the course of this thesis, researchers in the Mohn lab at the University of British Columbia have been characterizing dehydroabietic acid catabolic genes in *Pseudomonas abietaniphila* (Mohn *et al.*, 1998); a bacterium isolated from bleached kraft pulp mill mixed liquor (Bicho *et al.*, 1995). Illustrated in Figure 4-2, this region of the chromosome includes 14 open reading frames, of which 12 have been identified. Knock-out mutants have confirmed that most of these genes are necessary for the degradation of dehydroabietic acid (Mohn *et al.*, 1998 and Martin and Mohn, 1999). Note that these researchers have confirmed the requirement for a multi-component ring hydroxylating enzyme complex in the catabolism of dehydroabietic acid as was predicted by the intermediates generated by *Flavobacterium resinovorum* in the Biellman *et al.*, (1973) study. In addition, the operon includes other enzymes common to pathways involved in the catabolism of novel substrates, eg., isomerase/decarboxylases, dehydrogenases, and proteins that mediate substrate transport. A regulator is also present and its function in controlling this pathway is suggested by the fact that tetracycline inhibition of protein synthesis also inhibited dehydroabietic acid degradation (Bicho *et al.*, 1995).

Objective IIb

Experimental evidence suggests that catabolic pathways include a variety of enzymes and that the genes for these enzymes are often linked in the chromosome. The development of probes based on resin acid degradation genes, with the appropriate specificity to monitor levels in existing treatment

facilities and receiving waters, can take advantage of these contiguous genes. Therefore, isolation of some or all of an operon involved in tdt (tricyclic diterpene) metabolism was the objective of the following part of my research.

FIGURE 4-2. Diterpene catabolic pathway (*dit* operon) of *Pseudomonas abietaniphila* strain BKME-9. Selected steps of the pathway elucidated by GCMS metabolite and gene knock-out analyses are also shown. Adapted from Mohn *et al.*, 1998 and Martin and Mohn, 1999.

MATERIALS AND METHODS

Sequencing

Recombinant plasmid templates (YFG23 and 25b/f) were prepared by SIGMA column miniprep using the manufacturer's recommended protocol. All sequencing was performed by the University of Ottawa Biotechnology Research Institute using an Applied Biosystems automated sequencer and fluorescent ddNTPs. For the YFG23 template a chromosomal walk in both directions beginning with universal M13/pUC forward and reverse primers was performed until sequence overlap was reached. The opposite strand of the *tdtD* gene was also sequenced. In the case of 25b/f a walk to the end of the *tdtL* gene in the forward direction only was performed. Table 4-1 is a summary of the primers used and includes the primer position and the region of the *tdt* locus that it sequenced. All CM primers were generated from YFG23 template and 25b/f primers from the 25b/f sub-clone. The full sequence as presented in Table B1, Appendix B was compiled using a standard word processor and the electropherograms generated from each primer. Open reading frames, amino acid translations, restriction site analyses, and protein sequence alignments were generated by DNAMAN (Lynon Biosoft, 1994-1997). BLASTX similarity search (Altschul *et al.* 1998) was used to identify the genes.

TABLE 4-1. List of primers used to construct sequence of *tat* locus. All position numbers refer to the plus strand of the sequence (see Figure B1, Appendix B). Template DNA was YFG23 for CM primers and 25b/f for 25b/f primers.

Primer name	Sequence (5'→3')	mer	GC/AT	Strand	SEQUENCE		
					Primer Position (5' end)	Begins	Ends
YFG23-40	CAGCACTGACCCTTTTG	17	9/8	-	NOTE 1	9180	8925
YFG23-M13	CACACAGGAAACAGCTATGAC	21	10/11	+	NOTE 2	n/a	n/a
CM1	CCACTCAATCCCAGATCA	18	9/9	-	8978	8926	8476
CM3	GGTTCGTTCCAGGACGCTAC	20	12/8	+	NOTE 3	n/a	n/a
CM4	GTACACCTTGATAGGTGG	18	9/9	+	NOTE 4	553	793
CM5	CAGGTCTCTCCAACAGCA	18	10/8	-	8549	8475	8100
CM6.1	ATATCCGCCCTCTTGCTG	18	10/8	-	8154	8099	7642
CM7	ATCCGGGATGCGCTTGCC	18	12/6	+	700	794	1132
CM8	CGGCAGTGGCGTGATGGT	18	12/6	+	1084	1133	1586
CM9.1	GAAGTCGCTTTGAGTGC	18	10/8	+	1450	1587	1979
CM10	ATTGCTGTGTGCTGGGCC	18	11/7	-	7623	7641	7191
CM11	AATCCCTATCAATGTCAAG	19	7/12	+	1931	1980	2459
CM12	CTGGCGATCTTCAGGCGC	18	12/6	-	7258	7190	6745
CM14	GATGTCGTCGATACCCAG	18	10/8	-	6779	6744	6332
CM15	CGCTGATCGCTACCTCG	18	12/6	+	2348	2460	2952
CM16	CGCGGCTATGGCTGGCGA	18	13/5	-	6387	6331	5845
CM17	TTAAACTTATTCTCTGCC	18	6/12	+	2886	2953	3410
CM18	GATCGGCGTACGCAGCGC	18	13/5	-	5900	5844	5383
CM19	TCGCCCCGAAGGTCAACG	18	12/6	+	3343	3411	3999
CM20	CAGCTCCACCAGCTTCAT	18	10/8	-	5427	5382	4882
CM21	CGGATTCTCTACGGTTCC	18	10/8	+	3927	4000	4500
CM22	GGATTTCTCGCTCACCAC	18	10/8	-	4933	4881	4366
CM23	TATCCGCCCTTGACCGAG	18	11/7	+	4427	4501	4909
CM24	ATAGCGGATCGGATCATG	18	9/9	-	4329	4365	3862
CM25a	GACGTGGTCGAGGCGACC	18	13/5	+	4871	4910	5332
CM26	CCGATGATGTCGACCAGG	18	11/7	-	2919	3861	3355
CM27	GGAGTTCTTACCTGCGC	18	11/7	+	7442	7641	8206
CM29	ATCGGGTCATCCGCGAGG	18	12/6	+	8097	8207	8699
CM31	GCTGCTGGCGCGCATCCG	18	14/4	+	8551	8700	9180
25b/f-M13f	CACACAGGAAACAGCTATGAC	21	10/11	-	NOTE 5	3134	2667
25b/f-1	ACGGAGGTTCCAGGACGAAGG	20	12/8	-	2776	2666	2232
25b/f-2	ACAGAGAGCTATGGAAGCTG	20	10/10	-	2302	2231	1713
25b/f-3	TGAGCTATGAGTGGCATCGG	20	11/9	-	1767	1712	1172
25b/f-4	CTGCCTGATCTACACCTCCG	20	12/8	-	1222	1171	599
25b/f-5	CGTCGACGACAACGACCAGG	20	13/7	-	685	598	1

NOTES: 1. Universal pUC/M13 primer used to start chromosomal walk on insert YFG23 of - strand.
 2. Universal pUC/M13 primer reads into *trp* of *TnphoA*-1 insert in A76-47.
 3. Primer reads from outside end of IS50R into RP4 *traI* of pBW30 insert in A76-47.
 4. Primer reads out from RP4 *tra* region of pBW30 insert into A76-47 target DNA.
 5. Universal pUC/M13 primer used to start chromosomal walk on insert 25b/f.

Preparation of probes

Table 4-2 is a summary of restriction enzymes used to construct the pUC128::probe sub-clones from YFG23. Both insert and vector sites are specified and double digests, when necessary, were performed in accordance with New England Biolabs recommendations. Regions of the *tdt* locus represented by the inserts and probes are indicated by the plus strand position number of the restriction site used for cloning and/or probe recovery. All recombinant plasmids were transformed into *E. coli* JM109 by electroporation using a Bio-Rad Gene Pulser based on the manufacturer's recommended protocol. Transformants were screened by blue/white selection on TYE supplemented with 250 µg/ml ampicillin. Plasmids were recovered by standard alkaline lysis. Target inserts were confirmed by restriction digestion and gel electrophoresis prior to probe recovery. To ensure that the cloned probes corresponded to the target regions of the wild type genome a series of eight A19-6a restriction digests were probed and fragment sizes and common bands observed were compared to that which was predicted based on the DNA sequence.

Knock-out experiments

Homologous recombination using a suicide vector carrying a portion of each open reading frame was used to determine whether each gene was necessary for abietic acid degradation. Fragments from the YFG23 plasmid sub-clones were cloned into pRR1 (R. Rashid, unpublished) as indicated in Table 4-3. The pRR1::*tdt* suicide constructs were transformed by electroporation into *Escherichia coli* SM10 λ *pir* for mating, selection (Ap^r), and screening as

described in Chapter 2. Homologous recombination in target regions was confirmed by hybridization using *tdt* and *bla* gene probes.

TABLE 4-2. Construction of pUC128::probe sub-clones from YFG23. Position numbers refer to plus strand of *tdt* locus sequence, see Figure B1, Appendix B.

PROBE CLONE	GENE	INSERT RE(s)	POSITION	VECTOR RE(s)	PROBE RECOVERY	POSITION	SIZE (bp)
Ligase-1	<i>tdtL</i>	<i>Bam</i> HI/ <i>Bgl</i> II	828-1662	<i>Bam</i> HI	<i>Eco</i> RI*	828-1662	834
Reg-12	<i>tdtR</i>	<i>Pst</i> I/ <i>Nhe</i> I	2013-2966	<i>Pst</i> I/ <i>Xba</i> I	<i>Eco</i> RI*	2013-2966	953
Unk-K	<i>tdtS</i>	<i>Cla</i> I	3019-4054	<i>Cla</i> I	<i>Cla</i> I	3019-4054	1035
Iso-7	<i>tdtA</i>	<i>Kpn</i> I/ <i>Fsp</i> I	3698-4914	<i>Kpn</i> I/ <i>Eco</i> RV	<i>Pst</i> I	4015-4566	551
Dehyd-15	<i>tdtB</i>	<i>Kpn</i> I/ <i>Fsp</i> I	5025-5830	<i>Kpn</i> I/ <i>Eco</i> RV	<i>Nsi</i> I*/ <i>Pst</i> I*	5025-5830	805
Thio-11	<i>tdtC</i>	<i>Pst</i> I	4566-6973	<i>Pst</i> I	<i>Pvu</i> I	5899-6729	830
P450-8	<i>tdtD</i>	<i>Cla</i> I/ <i>Bam</i> HI	8094-8896	<i>Cla</i> I/ <i>Bam</i> HI	<i>Cla</i> I/ <i>Bam</i> HI	8094-8896	802

* RE site in pUC128 multiple cloning site.

TABLE 4-3. Summary of restriction enzymes used to construct suicide vectors for gene knock-outs. Position numbers refer to the plus strand of the *tdt* locus (see Figure B1, Appendix B). Insert size refers to target DNA only and does not include that which may have been contributed by using a pUC128 MCS site for cloning.

KNOCK-OUT CLONE	GENE	SOURCE	POSITION	INSERT RE	VECTOR RE	SIZE (bp)
Ligase-7	<i>tdtL</i>	YFG23	828-1662	<i>Bam</i> HI/ <i>Bgl</i> II	<i>Bgl</i> II	834
Reg-3	<i>tdtR</i>	U/R-6	2374-2966	<i>Xho</i> I/ <i>Nhe</i> I	<i>Sal</i> I/ <i>Xba</i> I	592
Unk-K5	<i>tdtS</i>	Unk-K	3382-3906	<i>Eag</i> I/ <i>Sal</i> I	<i>Eag</i> I/ <i>Sal</i> I	524
Iso-7/2	<i>tdtA</i>	Iso-7	4578-4914	<i>Eag</i> I*	<i>Not</i> I	336
D14	<i>tdtB</i>	Thio-11	5025-5523	<i>Fsp</i> I/ <i>Sma</i> I	<i>Sma</i> I	498
Thio-11/7	<i>tdtC</i>	Thio-11	6533-6973	<i>Sal</i> I*	<i>Sal</i> I	440
P450-6	<i>tdtD</i>	P450-8	8298-8896	<i>Not</i> I*	<i>Not</i> I	598

* RE site in pUC128 multiple cloning site.

Transformation with a linear piece of DNA carrying a selectable marker was also used to knock-out the *tdtD* gene. A 2058 bp *Nar* I fragment from YFG23 was digested with *Not* I/*Sal* I to generate two arms, 820 and 712 bp in size. These fragments comprised all of the 5' upstream region including the first 77 nucleotides of the target gene and the enzyme's core structure, respectively. The 526 bp *Not* I/*Sal* I internal fragment was replaced by the *bla* gene which had been previously cloned into pUC128 so that it could be recovered by a *Not* I/*Xho* I double digest. Competent cells were made by inoculating 500 ml of L-broth with 1 ml of an overnight culture and incubating them at 32°C with vigorous shaking to an OD₆₀₀=0.553. Cells were cooled on ice 30 minutes and harvested in a chilled rotor at 5000g for 15 minutes. Pellets were washed twice with an equal and half volume of 300 mM sucrose prior to resuspension in 2 ml of the same solution. Cells were again chilled on ice 30 minutes after which 100 µl aliquots were incubated on ice with 5 µl of the ligation product for five minutes. Following transformation (25µF, 200 OHM, 2.5KV, time constant 3.4-4.0) cells were resuspended in SOC (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose) and incubated one hour before plating on L-agar supplemented with 50 µg/ml kanamycin. Cells were also transformed with 5 µl of dilute pKT240 (Bagdasarian *et al*, 1983) as a positive transformation control.

RESULTS

Sequencing of cloned DNA

Automated fluorescence sequencing of YFG23 using the universal pUC/M13-40 primer sequenced through the pUC128 and λ FIX II™ cloning sites into the minus strand of *Pseudomonas* sp. A19-6a DNA. Figure 4-3 is the first 500 bases of this sequence and illustrates the *Xba* I cloning site (bold) used to subclone the M5 phage DNA into pUC128. Note that the filled in bases (TC) of the λ FIX II™ *Xho* I cloning site are seen but not the *Bgl* II (AG) fill-in bases of the target DNA due to the orientation of the insert relative to the sequencing primer.

Sequencing of the YFG23 template provided a better understanding of how the pBW30::Tn*phoA'*-1 insert had recombined with the A19-6a chromosome to produce the transconjugant A76-47. The pUC/M13 reverse primer generated plus strand sequence that was identical to Tn5 starting at base 15 of the inside end of the IS50R element (see Figure 4-4A, sequence YFG23M13). The same primer used in inverse PCR was used to generated sequence YFG23CM3 (Figure 4-4B) which reads the last six bases from the outside end of IS50R into the transfer region of pBW30 (see map of pBW30::Tn*phoA'*-1 in Figure 3-1). Note that the guanine (G) shown above the sequence was not actually read due to compression from the preceeding adenine (A). This sequence also reads the junction between vector and transposon that occurs at base 3761 of the RP4 *traI* gene (Zieglin *et al*, 1991). The sequence continues up to base 4431 of the RP4 *traJ* gene and is followed by the YFG23CM4 sequence (Figure 4-4C) that reads

through the entire RP4 *oriT* into target DNA. This sequence was generated by a primer spanning bases 4551 to 4568 of the RP4 *traJ* gene. Therefore, the point of insert recombination with the chromosome in A76-47 is 67 bases downstream from the *nic* site of the RP4 *oriT* (Figure 4-4C).

FIGURE 4-3. YFG23/-40 sequence showing the MCS from pUC128 (underlined) and λ FIX II™ (italics). The *Xba* I site is shown in bold and the asterisk represents the stop codon of the *tdtD* gene. The beginning of the minus strand of the target DNA next to the filled-in (TC) bases (highlighted) is also indicated. Feature label is shown below the sequence.

```

1   ACGCCAGTGCCAAGCTTGCATGCATCGTGTACCGGGCCCCCCTCGAGGT
51  CGACGGTATCGATAAGCTTGATATCGAATTCCTGCAGCCCGGGGATCCA
101 CTAGTCTAGAGCTCGCGGCCGCGAGCTCTAATACGACTCACTATAGGGC
      Xba I
151 GTCGACTCGATCTTCAACCTTCATTGTTGTTTTCCCCTTCGCTTAATGGG
      |→Pseudomonas sp. A19-6a target DNA
201 TTCTTGAAGTGCTTTTGATCATAGAGCACGCCAGAGAAAACTAAACACT
251 TTTTTTAAATATGTTCAACTTAATGGCATTCCCGAACCCCTGCCCCCTAGG
301 TCGCCAGGTACGCCGCAGCAGCGGCCCAACGGGACGGCATTACCGCCCC
351 TTAAGCGATACCACTCAATCCCAGATCACATGTACGTGATGAGGACCACG
      * ← tdtD
401 CAGCTGCGCACCGATGATCTCCGGAGCCGGCTTGTCGGGATCCANGCGGA
451 TGTTCGGCATCANGTCNAGGATGGCGTTGAGCGCGCAATTGATTTCCGTC

```

FIGURE 4-4. Summary of sequences spanning: A. the λ library cloning site; B. the pBW30::Tn*phoA*'-1 junction; and C. the point of recombination in A76-47. The *Xba* I site is shown in A. in bold and bases filled in during cloning are highlighted. Underline denotes pUC128 and italics λ cloning vectors. The boxed region in C. shows the RP4 *oriT* with the *nic* site as indicated. Start positions for *tnp*, *tral*, and *tdtL* DNA are also shown. Feature label is given below the sequence.

A. YFG23M13

```

1   AATTCGAGCTCCACCGCGGTGGCGGCCGCTCTAGAGCTCGCGGCCGCGGA
                                Xba I
51  GCTCAATTAACCCCTCACTAAAGGGAGTCGACTCGATCTTGATCCCCTGCG
                                |→ tnp
101 CCATCAGATCCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGCAGGGCT

```

B. YFG23CM3

```

1   GATCAGGCGCAGAGTCAGCGAGGGCGCGGTATGCCCGATACGGCTCATGGA
      |→ tral
51  TGGTGTTCGGGTCGGGTGAATCTTGTTGATGGCGATATGGATGTGCAGG

```

C. YFG23CM4

```

1   GTTACGCCGGCGGTAGCCGGCCAGCCTCGCAGAGCAGGATTCCCGTTGAG
51  CACCGCCAGGTGCGAATAAGGGACAGTGAAGAAGGAACACCCGCTCGCGG
101 GTGGGCCTACTTCACCTATCCTGCCCGGCTGACGCCGTTGGATACACCAA
                                ✂ nic site
151 GGAAAGTCTACACCGAACCCTTTGGCAAAATCCTGTATATCATGTTGCGCA
                                ↑→ tdtL
201 TCAGCTTGAGGGTATCCGCGGGACGGTTCCAGTAGCCCTGGAACATGATG

```

In an effort to determine if there was sequence similarity between RP4 and the mutated region of A76-47, the 99 bases comprising the RP4 *oriT* region were aligned with both strands of the cloned DNA. As illustrated in Figure 4-5 there were two regions of similarity where 46 of the bases were identical to the minus strand ending 52 bases upstream from the insertion site. Attempts to compare more of the RP4 *oriT* region with this or the opposite strand merely served to decrease percent similarity and/or move the homologous area(s) further from the insertion site.

FIGURE 4-5. Sequence of the mutated region of A19-6a (see below) showing similarity to the sequence of RP4 *oriT*. Two regions of similarity are indicated by shading. Identical residues between the two sequences (*) and the insertion site of the suicide vector (▼) are indicated below the sequence.

```

241  GCTGTCGCGGAACTCGATATAGCGCGGCAGGGCGTAGTAGGGCACCGAGTCGGTCGCCCCA
    CGACAGCGCCTTGAGCTATATCGCGCCGTCCCGCATCATCCCGTGGCTCAGCCAGCGGGT
                                     *      *
301  GTGGAACAGCTCCTCCGGCGTCAGTCGGGCCCCCTGGTGCAGCACGGCGGTCACCTTGAC
    CACCTTGTCGAGGAGGCCGAGTCAGCCCCGGGGGACCACGTCGTGCCGCCAGTGGAAGTG
    ***      *****      *****
361  GTCATCCTCGCCCTTGTCGGAGGGCACGGCATGCACCGCAACCTCGGCGAGGTCCGGATG
    CAGTAGGAGCGGGAACAGCCTCCCGTGCCGTACGTGGCGTTGGAGCCGCTCCAGGCCTAC
421  GACGGCGAAGGCCGCTTCCATCTCGAAGCTGGAGATGTTCTCGCCACGCCGGCGCAGGTA
    CTGCCGCTTCCGGCGAAGGTAGAGCTTCGACCTCTACAAGAGCGGTGCGGCCGCGTCCAT
                                     * * * *      *      ***** * * * * ***
481  GTCCTTCTTGCGGTCGACGAAATAGAAGAAGCCGTCATCATCGAACTTGCCGATGTGCGC
    CAGGAAGAACGCCAGCTGCTTTATCTTCTTCGGCAGTAGTAGCTTGAACGGCTACAGCGG
    *****      *      *
541  AGTGTGGAACCATGTTGCGCATCAGCTTGAGGGTATCCGCGGGACGGTTCCAGTAGCC
    TCACACCTTGGTGTACAACGCGTAGTCGAACTCCCATAGGCGCCCTGCCAAGGTCATCGG
    ▲ pBW30::TnphoA'-1

```

Figure 4-6 is a map of 9180 base pairs of compiled sequence referred to as the *tdt* locus; sequence overlap confirmed that the DNA was contiguous. As illustrated, there are seven open reading frames (ORFs) of which the first reads from the minus strand and the remaining six read from the plus strand. Table 4-4 summarizes some of the features of the locus and each of the ORFs. As indicated, the locus is GC rich (62%) and most has been sequenced on both strands. Two identical 6-base putative ribosomal binding (AGGAGA) sites precede the *tdtL* and *tdtD* genes. There are 5-base putative ribosomal binding sites found upstream from *tdtR*, *tdtS*, and *tdtB* (GAGAA, AGAGG, and GAGAG, respectively). The AT-rich inverted repeat is 51 bases upstream from the *tdtL* start codon. Other AT rich regions are found preceding *tdtR*, *tdtS*, and *tdtD*. The hairpin structures at positions 99 and 2847 follow *tdtL* and *tdtR*, respectively, and the other two fall within coding regions. The entire nucleotide sequence including the amino acid translation for coding regions is given in Figure B1, Appendix B.

Open reading frames

BLASTX similarity search (Altschul *et al*, 1990 and 1997) could identify five of these ORFs (*tdtL*, *tdtA*, *tdtB*, *tdtC*, and *tdtD*) as potentially coding for known enzymes. The *tdtR* gene is classed as a putative regulator and the identity of the ORF designated *tdtS* is unknown. Multiple sequence alignment of the putative gene products was used to determine how similar they are to proteins whose function is already known. Table 4-5 is a summary of the percent similarity between the *tdt* gene products and the sequences to which they are aligned (see Figures 4-7 to 4-13, inclusive). For ease of reference,

sequence sources have been included in Table 4-5 rather than in individual figure titles.

FIGURE 4-6. Sequence map of *tdt* locus isolated from *Pseudomonas* sp. A19-6a. BLASTX search (Altschul *et al*, 1990) identified the genes as a CoA ligase (*tdtL*), a putative transcriptional regulator (*tdtR*), an unknown (*tdtS*), an isomerase (*tdtA*), a dehydrogenase (*tdtB*), a thiolase (*tdtC*), and a cytochrome P450 monooxygenase (*tdtD*). Cloning sites (plus strand numbering) used in library construction are indicated. Sequenced regions and template used are also shown.

TABLE 4-4. Features summary of *tdt* locus and open reading frames. The complete sequence including features and amino acid translations is given in Figure B1 Appendix B.

tdt LOCUS								
	A	C	G	T	AT	GC	Total	
Composition	1670	2753	2918	1839	3509	5671	9180	
Percentage	18	30	32	20	38	62	100	
Molecular weight (kDa)	ssDNA: 2837.41		dsDNA: 5660.6					
FEATURES		SEQUENCE				POSITION STRAND		
putative ribosomal binding sites	AGGAGA				1772	-		
	AGGAGA				7689	+		
	GAGAA				1928	+		
	AGAGG				2998	+		
	GAGAG				4946	+		
AT-rich inverted repeats	5'-AAAATAAAAAT-3'				1813	-		
hairpin structures	GCCAGGCTggggggcctAGCCTGGC				99			
	TCAAAAATaaattgaacagATTTTGA				2847			
	CTCATGCGccattgaacgtacCGCATGAG				3412			
	AGCTCGACCatgaGGTCGAGCT				4446			
sequenced regions	554 to 5332, 7641 to 9119					+		
	1 to 3034, 3355 to 9180					-		
regions not sequenced	1 to 553, 5333 to 7640, 9120 to 9180					+		
	3035 to 3354					-		
OPEN READING FRAMES								
GENE	POSITION (+ STRAND)	dNTPs	FRAME	MW (ssDNA)	AT/GC (%)	PROTEIN	AMINO ACIDS	MW
tdtL	1760-129	1632	-1	502.83	38/62	CoA ligase	543	59945
tdtR	1942-2676	735	+1	226.81	41/59	regulator	244	27228
tdtS	3012-4088	1077	+3	332.39	39/61	unknown	358	40096
tdtA	4088-4942	855	+2	264.32	35/65	isomerase	284	30044
tdtB	4957-5862	906	+1	280.64	36/64	dehydrogenase	301	31711
tdtC	5859-7031	1173	+3	363.02	34/66	thiolase	390	40950
tdtD	7700-8974	1275	+2	393.75	37/63	monooxygenase	424	47054

TABLE 4-5. Percent similarity between putative *tdt* gene products and aligned sequences shown in Figures 4-7 to 4-13 inclusive.

	Organism	Gene	Product	Similarity (%)	Reference
<i>tdtL</i>	<i>Streptomyces griseus</i> subsp. <i>griseus</i>	<i>nonL</i>	NonL	34.4	Woo <i>et al</i> , 1998
	<i>Escherichia coli</i>	<i>fadD</i>	acyl-CoA synthetase	27.1	Black <i>et al</i> , 1992
	<i>Populus tremuloides</i>	<i>Pt4CL1</i>	4-coumarate:CoA ligase	25.0	Hu <i>et al</i> , 1998
	<i>Alcaligenes</i> sp. BR60	<i>ORF8</i>	aryl-CoA ligase	22.7	Di Gioia <i>et al</i> , 1998
	<i>Pseudomonas oleovorans</i>	<i>alkK</i>	acyl-CoA synthetase	22.5	van Beilen <i>et al</i> , 1992
	<i>Rhodopseudomonas palustris</i>	<i>badA</i>	benzoate-CoA ligase	21.5	Egland <i>et al</i> , 1995
<i>tdtR</i>	<i>Vibrio harveyi</i>	<i>luxR</i>	regulatory protein	19.4	Showalter <i>et al</i> , 1990
	<i>Neisseria gonorrhoeae</i>	<i>mtrR</i>	transcriptional repressor	17.8	Pan and Spratt, 1994
<i>tdtS</i>	<i>Sphingomonas aromaticivorans</i>	<i>orf946</i>	hypothetical protein	28.3	Romine <i>et al</i> , 1999
	<i>Synechocystis</i> sp.	<i>cyclsrf</i>	hypothetical protein	21.9	Kaneko <i>et al</i> , 1995
<i>tdtA</i>	<i>Rhodobacter capsulatus</i>	<i>rcap</i>	2-hydroxyhepta-2,4-diene-1,7-dioate isomerase/5-carboxymethyl-2-oxo-hex-3-ene-1,7-dioate decarboxylase	39.6	Vlcek <i>et al</i> , 1997
	<i>Bacillus subtilis</i>	<i>yisK</i>	putative decarboxylase/isomerase	33.3	Seror, 1997
	<i>Escherichia coli</i>	<i>hpaG</i>	5-oxo-1,2,5-tricarboxylic-3-penten acid decarboxylase/isomer	29.4	Prieto <i>et al</i> , 1996
	<i>Escherichia coli</i>	<i>hpcE</i>	2-hydroxyhepta-2,4-diene-1,7-dioate isomerase	28.9	Roper and Cooper, 1993
	<i>Rhodococcus erythropolis</i>	<i>bphC</i>	homologous to <i>hpcE</i>	28.2	Kosono <i>et al</i> , 1997
<i>tdtB</i>	<i>Clostridium acetobutylicum</i>	<i>adh1</i>	beta-hydroxybutyryl-CoA dehydrogenase	44.3	Youngleson <i>et al</i> , 1989
	<i>Clostridium acetobutylicum</i>	<i>bhbD</i>	3-hydroxybutyryl-CoA dehydrogenase	43.8	Boynton <i>et al</i> , 1996
	<i>Pseudomonas putida</i>	<i>phaC</i>	3-hydroxyacyl-CoA dehydrogenase	37.0	Olivera <i>et al</i> , 1998
	<i>Pseudomonas fragi</i>	<i>faoA</i>	3-hydroxyacyl-CoA dehydrogenase	33.2	Sato <i>et al</i> , 1992
	<i>Escherichia coli</i>	<i>fadB</i>	3-hydroxyacyl-CoA dehydrogenase	32.8	DiRusso, 1990
<i>tdtC</i>	<i>Acinetobacter</i> sp. ADP1	<i>catF</i>	beta-ketoadipyl CoA thiolase	45.5	Doten, 1987
	<i>Pseudomonas putida</i>	<i>pcaF</i>	beta-ketoadipyl CoA thiolase	44.8	Harwood <i>et al</i> , 1994
	<i>Escherichia coli</i>	<i>fadA</i>	3-ketoacyl-CoA thiolase	44.4	DiRusso, 1990
	<i>Pseudomonas</i> sp. 61-3	<i>phbA</i>	beta-ketothiolase	44.1	Matsusaki <i>et al</i> , 1998
<i>tdtD</i>	<i>Pseudomonas</i> sp.	<i>terpC</i>	P-450 _{terp}	25.4	Peterson <i>et al</i> , 1992
	<i>Streptomyces griseus</i>	<i>soyC</i>	P-450 _{soy}	25.2	Trower <i>et al</i> , 1992
	<i>Agrobacterium tumefaciens</i>	<i>pinF1</i>	P450 _{pinF1}	21.6	Kanemoto <i>et al</i> , 1989
		<i>pinF2</i>	P450 _{pinF2}	21.3	
	<i>Pseudomonas incognita</i>	<i>linC</i>	P-450 _{lin}	20.8	Ropp <i>et al</i> , 1993
	<i>Pseudomonas putida</i>	<i>camC</i>	P-450 _{cam}	20.8	Unger <i>et al</i> , 1986

Figure 4-7 is an alignment of the putative *tdtL* gene product with five bacterial (*nonL*, *fadD*, *ORF8*, *alkK*, and *badA*) and one plant (*pt4Cl1*) coenzyme A ligase sequences. With the exception of the boxed area most of the sequence similarity occurs downstream from the point of insertional inactivation indicated by the arrowhead. The amino acids enclosed in the box are the putative acyl-adenylate binding site (Babbitt *et al*, 1992) that is highly conserved amongst all CoA ligase sequences. Percent similarity between *tdtL* and the aligned sequences ranges from 34.4 percent for *nonL* to 21.5 percent for *badA*, see Table 4-5.

Multiple sequence alignment of *tdtR* (Figure 4-8) revealed that it was most similar to the *luxR* gene of *Vibrio harveyi*; although overall similarity was low at 19.5 percent. Comparison with *mtrR* shows 17.8 percent similarity. Note that most of the similarity between the three sequences is in and around the helix-turn-helix region of the *mtrR* gene.

Figure 4-9 is a multiple sequence alignment between the *tdtS* product and two hypothetical proteins. Some clustering of conserved (*) and similar (.) residues can be seen, however, no particular region stands out as being more highly conserved. The *tdtS* gene product is 28.3 percent homologous to *orf946* from *Sphingomonas aromaticivorans* and 21.9 percent homologous to *syctlrf* from *Synechocystis* sp. Attempts to align *tdtS* with proteins whose functions were known were unsuccessful.

FIGURE 4-7. Multiple sequence alignment of TdtL, an acyl-CoA ligase, from *Pseudomonas* sp. A19-6a and several sequences identified by BLASTX search (see Table 2 for sequence identification and references). Boxed region is acyl-adenylate binding site (Babbitt *et al*, 1992). Site of insertional mutagenesis by *TnphoA'*-1 is indicated by an arrowhead. Residues the same (*) or similar (.) are indicated below the alignment. Sequences are aligned in order of decreasing similarity.

tdtL -----MSGIGHPDTWSVGERDVTIAALDRAVARHPDR-ILLD-FSGNLYTYSEVDSLSTRMAHALAA-
nonL MIDDVLDPASAAVP---TAEQCVLGPLLRRAAAAPAAPYALMP--DGDLTWYARTLRETEETAAALQA-
fadD MKKVWLNRYPADVPTINPDYQSLVDMFEQSVARYADQPAFVN--MGEVMTFRKLEERSRAFAAYLQQG
Pt4CL1 -MNPQEFIFRSKLPDIYIPKNLPLHSYVLENLSKSSKP-CLINGANGDVYTYADVELTARRVASGLNK-
ORF8 -----DKLPAGAHAFDDLLAR-----
alkK -MLGQMMRNQLVIG---SLVEHAARYHGAREVVSVETSG-----EVTRSCWKEVELRARKLASALGK-
badA -----MNAAAVTPP-----PEKFNFAEHLQTNRVRPDKT-AFVD--DISSLSFAQLEAQTRQLAAALRA-

tdtL LGVKPGETVLTMLDNNIDAVVCWLAINKLRAVSVPIINTALKGEFLRHQIADTGTQLVICEADYLPRILPL
nonL LGVVPGEVLVLSWLPNGPDALRAWYGVNLAGAVLVPLNIAYRGAILRQVIADSGAEVLICRPSLAARLEDS
fadD LGLKKGDRVALMMPNLLQYPVALFGILRAGMIVNVNPLYTPRELEHQLNDSGASAIIVISNFAHTLEKV
Pt4CL1 IGIQQGDVIMLFLPSSPEFVLAFLGASHRAGAMITAAFPSTPAELAKHAKASRAKLLITQACYYEKVKDF
ORF8 -----
alkK MGLTPSDRCATIANNIRHLEVYYAVSGAGMVCHTINPRLFIEQITYVINHAEDKVVLDDDTFLPIIAEI
badA IGVKREERVLLMLDGTDPVFAFLGAIYAGIVPVAVNTLLTADDYAYMLEHSRAQAVLVSGALHPVLKAA

tdtL AEQLS-EVSQILYRGERAEPASCRIPPI-----AALDEFRGEDDTPFTS-KPQPS-----
nonL -----ARPAD-----PIDT-HGDEP-----
fadD LTKSDHEVQRVIVSRPALRWSRARSTS-----LSRRTSLEK---PAAT-QADDP-----
Pt4CL1 VDKT--AVQHVLTRMGDQLSTAKGTV-----VNEFVKYIKRLVPKYH-LPDAISFRSALHNGYR
ORF8 ARES--DVKVMCVDSPDQASLFRAHT-----QADENEVP---QVDI-SPDDV-----
alkK DDAVGAVRTVLLPGPEDAAQDVEALAGRLATFRFVETALRADRAEFAEPVPAPRPWDP-----
badA HGSLP-KVKAFLVMAHNNNSNASEQMPG-----LIAYEDLIGQG--DDNYIWPVDVDEN-----

tdtL -----DLACLIYTSGTTGPKSGKCMISY-NFMCNLARLQLRAGPANE---NDVTITPLPLFHMNA
nonL -----QTVIYTSGTTGPKSGKVSSYAHLYSSCTAAFHGMAG--P---EDRYLLQLPLFHAGG
fadD MQYVKPELVPEDLAFQYTGTTGTVAKGAMLTHRNMLANLEQVNATYGPLLEPG-KELVVLTALPLYHIFA
Pt4CL1 -----VALPYSSGTTGLPKGVMLTHKGLITSVAQQVDGDNPNLYFHSDEVILCVLPMFHIYA
ORF8 -----AFWLYSSGSTGPKGVVHSHKGLITSVAQQVDGDNPNLYFHSDEVILCVLPMFHIYA
alkK -----EASSLCYTSGTTGNPKGVVLYSHRSTVLHSMTTAMPDTLNLS--ARDTILPVVPMFHVNA
badA -----AFWLYSSGSTGPKGVVHSHKGLITSVAQQVDGDNPNLYFHSDEVILCVLPMFHIYA

tdtL LC-VSIIASILVGARAAILPRFSVSNFQEVER-----SGATIASILGGMGGLLAQAPDNDAMKRCHQI
nonL TI--GAYGMLVHGGSVTVVPAFTTGEFVKELIR-----TGTTLCTLLGVMATYLLKQPPLPQDTAHLRA
fadD LTINCLLFIELGGQNLITNPRDIPGLVKELAK-----YPTAITGVNTLFNALLNNKEFQQLDFSSHL
Pt4CL1 LN-SMMLCGLRVGASILIMPKFEIGSLGLIEK-----YKVSIAFPVPPVMMIAKSPDLDKHDLSSLAM
ORF8 LG-NALSFPLSVGATVVLMAERPTPQACFKRMT---EHRPTIFCGAPTGYGGMLASADLPAKEAVALRL
alkK WG--TPYSAAMVGAKLVLPGLDGDASLSKLIAS---EGVSIALGVPPVWQGLLAAQAGNGSKSQSLTR
badA LG-NALTFPMTVGATTLMLGERPTPDAVFKRWLGGVGGVKPTVIFYGAPTGYAGMLAAPNLPSRDQVALRL

tdtL HTARGNPYTEETKQIWRERFRTKLVGNGYGLTEACVITSLA-----AGEYAAPGS-----SGKRIP
nonL AYVIP--FTEG-ATEFSKRFG--VPVRALFNMTVSCPVLSAF-----DHHPGV-----PMHCGEFRP
fadD SAGGGMPVQQVVAERWVKLTG--QYLLEGYGLTECAPLVSVNP---YDIDYHSGS-----IGLFVP
Pt4CL1 IKSGGAPLGKELEDTVRAKFPQ-ARLGQGYGMTAAGPVLAMCLAFAKEPFDIKPGA-----CGTVVR
ORF8 CSSAGEALPQDIGERWTRHFG--VEIIDGIGSTEMLHIFLSN-----VPGEVHYGS-----TGHPVP
alkK VVVGGSACPASMIREFNDIYG--VEVIHAWGMTLSPPFGTANTP-LAHHVDSLPDEKLSLR-KSQGRFPY
badA ASSAGEALPAEIGQRFRHFG--LDIVDGGIGSTEMLHIFLSN-----LPDRVRYGT-----TGWPVP

tdtL DFDVRIVDDND---QELPAN--TPGEIVLRPLRPDIMFQGYWNRPADTLKLMRN-MWFHTGDIGKFDDDG
nonL GIAARVDDHD---REVADG--EAGELVLRADRPWSFLSGYLGRPAETAAVWRN-GWFHTGDTFRRAPDG
fadD STEAKLVDDDD---NEVPPG--QPGELCVK--GPQVML-GYWQRPDATDEIKN-GWLHTGDIAMVDEEG
Pt4CL1 NAEKIVDPETG--VSLPRN--QPGEICIR--GDQIMK-GYLNDEPATSRTIDKEGWLHTGDIGYIDDDG
ORF8 GYEVELRGEDGQALAPNADGSDIDGLYIK--GPSAAL-MYWNNRDKSRDTFQG-GWTKAGDKYQRNPNG
alkK GVELKIVNDEG---IRLPEDGRSKGNLMAR--GHWVIK-DYFHS--DPGSTLSD-GWFSTGDVATIDSDG
badA GYQIELRGDGG---GPVADG--EPGDLYIH--GPSSAT-MYWGNRAKSRDTFQG-GWTKSGDKYVRNDDG

tdtL FFYFVDRKKOYLRRRGENISSFEMEAAFAVHPDLAEVAVHAVPSDKGEDDVKVTAVLH-QGARLTPEELF
nonL GLVFVDRKKDAIRRRGENISSFEVEAQAHAHPGVLEAAAVAVPGDEGEDEVLLVVAADRPSPAPVPAALL
fadD FLRIVDRKKDMLVSGFNVPNEIEDVVMQHPGVGEVAAVGVPSGSSGEAVKIFVVKKDP--SLTEESLV
Pt4CL1 ELFIIVDRKLIKELIKYKGFQVAPTELEALLIAHPEISDAAVVGLKDEDAGEVPVAFVVKs-EKSQATEDEIK
ORF8 TYTYAGRNDMLKVSGIYVSPFEVEATLVQHPAVLEAAVIGHTDGEGLTKTKAYVVLK-PGQQAGAADLQ
alkK FMTICDRAKDIIKSGGEWISTVELESIAIAHPHIVDAAVIAARHEKWDERPLLIIVKS-PNSELTSGEVC
badA SYTYAGRDDMLKVSGIYVSPFEIEATLVQHPGVLEAAVGVVADEHGLTKPKAYVVPVPR-PGQTLSETELK

tdtL HWATDSVPYYALPRYIEFRDSLKPKNPQGRVLKYQLRDEGKTADTWLDLDTDIKVAKK----
nonL EFLRERLAHFMLPRYIRVLPPLPKTPTGKPTKHTLRAEGVAGTWDREAAGIRIRREKIV-
fadD TFCRRQLTGYKVPKLVEFRDELPSNVGKILRRELRLDEARGKVDNKA-----
Pt4CL1 QYISKQVIFYKRIKRVFFIEAIPKAPSGKILRKNLKEKLPGI-----
ORF8 AFVKERLAPYKYPRFIEFLDELPKTATGKIQRFRRLRERDQAG-----G
alkK NYFADKVARWQIPDAAI FVEELPRNGTGKILKNRLREKYGDILLRSSSSVCE-----
badA TFIKDRLAPYKYPRSTVFVAELPKTATGKIQRFKLREG-VLG-----

FIGURE 4-9. Multiple sequence alignment of *tdtS* gene product with hypothetical proteins from *Sphingomonas aromaticivorans* (*orf946*) and *Synechocystis* sp. (*slcs1rf*). Residues the same (*) or similar (.) are indicated below the alignment. See Table 4-5 for references and percent similarity between the sequences. Sequences are aligned in order of decreasing similarity.

TDTs	MLIDLHAHAPHDPDYDQHPYWGPAFESQPDGDIKLRVGDWILTLGAPERKKALREAHARG
orf946	MIIDAHHLVAPSSLYAHRSNLIVSGGQYGKNYRAKVSDGDLQESAEQN-----
sycslrf	MMKRIAVEEAFVTQEIADEWANVIATGAKNEPGFRKMGETILADSPGTRIHERLIDLGA
	*. . . *
TDTs	ETLNVAEYMAKWRDPNHRLAAMDAAGQSAQVLS-VPSHCYMYWAEAEFAVPFARKVNDVL
orf946	-----VKIMDAVGTDVQLLSRPFLTNGTARWEDIVDWTCDTNDMI
sycslrf	G-----RIAHMDATGIDIQVIS--ITSPGVQVFEPELATTLARQANDQL
	. *** * * . . *
TDTs	AGYCSAAPNRLMFWAHAPLNVP---HEAAKEIRRACTELGA KGLVAGGSNFGGMEFD---
orf946	ARTVKMFNPNRFRGVGALPQQAGRPVESLFDEIERVVDDLGFVGILLNPDPSSEGMNGSPPL
sycslrf	SEAIKNHPDRFAGLA AVAPQSP---QTAALIEIRAIQKLGCGVLIN-SHTGGEYLD---
	. * * . * * * * . *
TDTs	-SPELYPVWEALCDLPLIFVHGYNQSVTWGKQANTDRFETTAIVGMNYDETKCLWYLIN
orf946	GDPYWYPLYRRLCELDLPAHIHSG--QCCNGRESYDEHFIAEEGLAITS-----VYR
sycslrf	-DQKYWEIFAAAEANSAPIYLHPR-TPSPMISPFLDYGLYFAGWGFTVETATHGLRLIM
	. * . *
TDTs	GGVLDREFPNLKVIYITHGGGFAPYQLGRLAQT----NPNLDVYHNKKPFLDYLSNFYFDVE
orf946	ANVFDRFPDLKLMISHGGGPIPIYQIGRWRSHREMARAQGRIAADAPSFDEILRKFWFDTV
sycslrf	GGVFDHFPKLTIVLGHMGEGLPYWLQRLDNRYSLQVEIGAVPKMPRLPSEYFKDHFLITT
	. * * . * * . * * * . *
TDTs	LHE--VPMRQALVDIIGADRILYGSNFGGS----DAVRHDLTEGLRLSDEDLQKIRWKN-
orf946	LHN--PDSL DLLVKT VGYDRCCFGTERPGSGDGIDPLTRRHYDDLKPVIEGLTSLDDEAR
sycslrf	SGVCSPPALRHVLEVLGSDAVLFAADYPYES---VEEAVSFLDNAPISDQERRKIYQTNA
	. . . * * . . .
TDTs	ACELLHIDPAKLGKVG--A
orf946	ACIFEHNARKLFTRLDV--
sycslrf	EKLFR LGSTNGSNDTLFA-

BLASTX similarity search identified *tdtA* as a putative isomerase/decarboxylase gene with amino acid similarity ranging from 39.6 to 28.2 percent for three known and two hypothetical proteins, see Table 4-5. Figure 4-10 is a multiple sequence alignment of these gene products that shows that most of the similarity observed is in the second half of the protein.

The *tdtB* gene was found to be similar to two CoA dehydrogenases from *Clostridium acetobutylicum* (see Table 4-5). Multiple sequence alignment (Figure 4-11) with these and three other known dehydrogenases revealed that there were 52 conserved residues (*) and 39 similar (.) residues between the six sequences. Note that the amino acid sequences for PhaC, FaoA, and FadB are considerably longer than for TdtB, Adh1, and BhbD.

Figure 4-12 is a multiple sequence alignment between the *tdtC* gene product and four known thiolases. Homology with all sequences is high ranging from 45.5 to 44.1 percent (see Table 4-5). As indicated by the alignment many residues are identical between the sequences with one noteworthy region near the carboxy terminus that has 15 out of 17 conserved amino acids. Note the conserved cysteine residues (boxed) that are part of the proposed active centres of thiolase enzymes (Sato *et al.*, 1992).

FIGURE 4-10. Multiple sequence alignment of *tdtA* gene product with five decarboxylase/isomerase proteins. Residues the same (*) or similar (.) are indicated below the alignment. See Table 4-5 for references and percent similarity between the sequences. Sequences are aligned in order of decreasing similarity.

<i>tdtA</i>	-----MKLAR-----
<i>Rcap</i>	-----MKLVR-----Y--GE-----R--
<i>yisK</i>	-----MKFAT-----GELYN-----RMF
<i>hpaG</i>	MKGTIFAVALNHRSQLDAWQEAFFQSSPYKAPPKTAVWFIKPRNTVIGCGEPIPFPPQGEKVLSGATVALIVGKTATKVREE
<i>hpeC</i>	MKGTIFAVALNHRSQLDAWQEAFFQSSPIKAPPKTAVWFIKPRNTVIGCGEPIPFPPQGENLLSGATVALIVGKTATKVREE
<i>bphC</i>	-----MRVAN-----
.	
<i>tdtA</i>	-CAYFETG-----SFWAVVDPER-----DEVHAIAGE-----F
<i>Rcap</i>	-GAE-----KPGLLHVDGTIR-----DLSRVVRDIDGK-----AL
<i>yisK</i>	VGLIID-----DEKIMDLQKAEK-----KLFELETIPGS-----LIECI
<i>hpaG</i>	DAAEYIAGYALANDVSLPEESFYRPAIKAKCRDGFCEPIGETVALSNVDNLTIIYTEINGRPADHWNTADLQRNAAQLLSAL
<i>hpeC</i>	DAAEYIAGYALANDVSLPEESFYRPAIKAKCRDGFCEPIGETVALSNVDNLTIIYTEINGRPADHWNTSDLQRNAAQLLSAL
<i>bphC</i>	-QSG-----RLKLVGDGVG-----VDVEVASQGR-----FT
.	
<i>tdtA</i>	SS-----WAPAVARGAGISALRLSGS-----ALPLSKVRLLPIDP-----VNRVVVAGANYAKHLAE
<i>Rcap</i>	SD-----EGLSKLFWTDAGKLPPIP-----EGVR-LGAPVAG-----VGKLIAGLNYADHAE
<i>yisK</i>	AE-----GDKFVAHARQLAEWAKKPNDELGSFMYSLSEVKLHAPIPK-----PSKNIICIGKNYRDHAIE
<i>hpaG</i>	SEFATLNPGDAILLGTPOARVEIQPGD--RVRVLAEGFPFLENPVVDEREVTTRKSFPPTLPHPHGTLFALGLNYADHASE
<i>hpeC</i>	SEFATLNPGDAILLGTPOARVEIQPGD--RVRVLAEGFPFLENPVVDEREVTTRKSFPPTLPHPHGTLFALGLNYADHASE
<i>bphC</i>	SD-----PQAVFDRWDEFRLWAESID--FHGAEPIVESDLQAPVPR-----PGQIFAVGVNRYDHVEE
.	
<i>tdtA</i>	DFGLAS-PPQPVAFLKAYGALIGAHDPRIYPPLT-EELDHEVELVAVIG--SAEIDLEDPLACVLGYTVGNDSARDLQR
<i>Rcap</i>	AGLTV--PPEPVVFMKATSAICGPDDEAPRGS-EKLDWEVELAFVIGTRAKYVTLEDAMDHVAGYMICNDVSERSFQI
<i>yisK</i>	MGSEADIPEHPMVFTKSPVTVTGHGDIVKSHEEVTSQLDYEGLAVVIGKSGTRISKEDAYDHVFGYTI VNDITARDLQK
<i>hpaG</i>	LEFKP--PEEPLVFLKAPNTLTGDNQTSVRPNNI-EYMHYEAEELVVVIGKQARNVSEADAMDYVAGYTVCN DYAIRDYLE
<i>hpeC</i>	LEFKP--PEEPLVFLKAPNTLTGDNQTSVRPNNI-EYMHYEAEELVVVIGKQARNVSEADAMDYVAGYTVCN DYAIRDYLE
<i>bphC</i>	SGLSL--PEAPFVFTKFPGSITGPYNNIELPDGS---VDFEVELVAVIGRAARHVDASNGWDYVAGVTLGQDL SERQQL
.	
<i>tdtA</i>	SGPAGIGMDLFAAKSQDRRTGLGPWIVTKDEFVPGSP-KLRLTLKVNGEIRQDGSTAEMTWVVGQVRFVQQRSSFACGD
<i>Rcap</i>	E----RSGQWVKGKSHDTFAPLGPWLVTKDEIKD--PQALDLWLVNNGVRRQTGTATMVYGVAFVLVHYLSQFMTLEPGD
<i>yisK</i>	-----RHKQFFIGKSLDTTCPMGPVLVHKSSIQE--PERLKVETRVNGELRQSGSASDMI FSIPELIETLSKGMTLEAGD
<i>hpaG</i>	N----YYRPNLRVKS RDGLTPMLSTIVPKEAIPD--PHNLTLRTFVNGELRQQGTTADLIFSVPFLIAYLSEFMTLNPGD
<i>hpeC</i>	N----YYRPNLRVKS RDGLTPMLSTIVPKEAIPD--PHNLTLRTFVNGELRQQGTTADLIFSVPFLIAYLSEFMTLNPGD
<i>bphC</i>	SG--PAPQQFSLGKSYAGFAPIGPVLVTPDEFAD--PDDIGVSTRLSGDEMQRSTRHLIFPIPELVAYLSSILPLSPGD
.	
<i>tdtA</i>	VLFTGSPAGVGMGTG--LFLNPGDVVEATIEGIGTLRNVVSEKSRV-----
<i>Rcap</i>	IITGTGTPPGVGMGSRPPKFLHAGDVIELGITGLGSQRQVVVQA-----
<i>yisK</i>	I IATGTPSGVGKGFTPPKFLRSGDKIDITIDPIGTLNQGIG-----
<i>hpaG</i>	MIATGTPKGLSDVVP-----GDEVVVEVEGVGRLVNRIVSEETA-----K
<i>hpeC</i>	MIATGTPKGLSDVAM-----K-----
<i>bphC</i>	LIFTGTPEGIGWTREPKRLISTDDELVSTADVIGEMRHRFIPSRHHSATSSARSSTNV--
.. **.* *	

FIGURE 4-11. Multiple sequence alignment of *tdtB* gene product with six different known CoA dehydrogenases. Residues the same (*) or similar (.) are indicated below the alignment. See Table 4-5 for references and percent similarity between the sequences. Sequences are aligned in order of decreasing similarity.

cdtB -----
adh1 -----
bhbD -----
phaC -----
faoA MIYEGKAITVTALESGLVELKFDLKGESVNFNRLTLNELRQAVDAIKADASVKGVIVSSGKDVFI V GADITEFVENFKLPDAELIAGNLEANKI FSDFE
fadB MLYKGDITLYLDWLEDGIAELVFDAPG-SVNKLDATVASLGEAIGVLEQQSDLKGLLLRSNKAAFI V GADITEFLSLFLVPEEQLSQWLHFANSVFNRL

cdtB -----
adh1 -----
bhbD -----
phaC -----
faoA DLNVPTVAANGIALGGGLEMC LAADFVMDASAKIGLPEVKLG IY PGFGGT VRLPRLIGVDNAVEWIASGKENRAEDALKVSAVDVVTADKLGAALD
fadB DLFVPTTAAVNGYALGGGCECVLATDYRLATPDLRIGLPETKLGIMPFGGGSVRMPRLGADSALEI IAGKDVGDQALKIGLVDGVVKA EKLVEGAKA

cdtB -----
adh1 -----
bhbD -----
phaC -----
faoA LIKRAISGELDYKAKRQPKLEKLLKNAIEQMMAFETARGFVAGQAGPNYPAPVEAKTIQKAANFGRDKALEVEAAGFAKLAKTSASNCLIGLFLNDQEL
fadB VLRQAINGDLWKAKRQPKLEPLKLSKIEATMSFTIAGKMVAQTAGKHYPAPITAVKTI EAAARFGREEALNLENKSFVPLAHTNEARALVGI FLNDQYV

cdtB -----MSIDKEEKIVVVGAGLMTGIAHGFASSGYPTVLVDNTAESLARARDNIGKILSDGVALGKVSASAEASLARLVTSGDLSEAARGADWLVE
adh1 -----MEKIFVLGAGTMGAGIVQAFQKGYEVIVRDIKDEFVERGIAGINKGLTKQVAKGMTEEDKEAILSRISGTTDMKLAAD-CDLVVE
bhbD -----MKKVCVIGAGTMGSGIAQAFAAKGFEVLRDIKDEFVDRGLDFINKNLSKLVKKIGKIEEATKVEILTRISGTVDLNMAAD-CDLVIE
phaC -----MGALASTVQVAVIGAGAMGAGIAQVAAQAGHPVKLYDNRPAAAAQAVAGIERQLARLVEKGLRAVEREMISLRICPVDLTLEALAD-AGLVIE
faoA KKGAKVYDKIAKDVQAAVLGAGIMGGGIAYQSASKGTPILMKDINEHGIEQGLAEAAKLLVGRVDKGRMTPAKMAEVLNGIRPTLSYGD FGN-VDLVVE
fadB KKGAKKLT KDVETPKQAAVLGAGIMGGGIAYQSASWKGPVVMKDINDKSLTLG MTEAAKLLNKQLERGIKIDGLKLAGVISTIHPTLDYAGFDR-VDIVVE
. *

cdtB TVSEKLEVKRAVVAQAEPLAPSAL IATNTSALS VTEIAASLAAPERVIGMHFFNPVHKMKLVELVRGLATSDEVLRTRALCDAMGKTSIVVNESPLGT
adh1 AAIENMKIKKEI FAE LDGICKESTI LASNTSSLSITEVASATKRPKVIGMHFFNPAPVMKLV EIRGIATSQETFD AVKQLSVAIGKEPVEVAEAPGFV
bhbD AAVERMDIKKQIFADLDNICKPETILASNTSSLSITEVASATKTNDKVIGMHFFNPAPVMKLV EIRGIATSQETFD AVKETSIAIGKDPVEVAEAPGFV
phaC AIVENLQVKQALFSQLEALCTADCI IASNTSSLSITSLAAGLARPPQVVG MHFFNPAPIMALVEVVSGLATEPAVAACIYDTAQAWGKQPVHTRSTPGFI
faoA AVVENPKVKQAVLAEVENHVREDAL IASNTSTISISLLAKALKRPENFVGMHFFNPVHMMPLVEVIRGEKSSDLAVATTVA YAKKMGKNPIVVNDCPGFL
fadB AVVENPKVKAVLAETE QYVRQD TVLASNTSTIPISELANALERPENFCGMHFFNPVHRMPLVEIIRGEKSSDETIAKVVAWASKMGKTPIVVNDCPGFF
. *

cdtB TSRMSALLGNEAMYMLQEGTASAEIDTALRLGFNHPMGPLELGDLTGWDRLSVLRYLHQTGLG--EKFRPCFLIIKMVAAGRLGRKTHGHVYRYEDGKR
adh1 VNRI LIPMINEATFIYQEGIASVEDIDAA M KYGANHPMGPLALGD LIGLDVCLAIMDVL FNETGD-SKYRASSILRKYVRAGWLGRKSGKGFYDYS----
bhbD VNRI LIPMINEAVGILAEGIASVEDIDKAMKLGANHPMGPLELGD LIGLDICLAIMDVL YSETGD-SKYRPHTLLKKYVRAGWLGRKSGKGFYDYSK---
phaC VNRVARPFYAESLRLQEGAADCASLDALLRDAGGFRMGAFELTDLIGHDVNYAVTCSVDFAFYGD FRFQPSLVQKELVDAGRLGRKTGGFYSYAEGAE
faoA VNRVLFPYFGGF AKLV SAG-VDFVRIDKVM EK-FGWPMGPAYLLDVVGIDTGHHRDVM AEGFPDRMKDDRRSAIDALYEAKRLGQKNGKGFYAYEADKK
fadB VNRVLFPYFAGFSQLLRDG-ADFRKIDKVM EKQFGWPMGPAYLLDVVGIDTAHHAQAVMAAGFPQRMQKDYRDAIDALFDANRFGQKNGLGFWR YKEDSK
* * *

cdtB VPGSGLKASAL-----
adh1 -----
bhbD -----
phaC RPAPVELHSSTKAEACVIEGQLGVLQPLVERLRQNGIVVTVRAGSGVIVQGDATLALS DGRLASQRAREDGLRLNLVLLDLALDYSTASRIASWSGOTTE
faoA GKQKKLVDSSV-----LEVLKPIVYEQRD-----VTDEDI INWMMI PLCLETVR
fadB GKPKKEEDA AV-----EDILAEVSQPKRD-----FSEEEI IARMMI PMVNEVVR

cdtB -----
adh1 -----
bhbD -----
phaC SARDQAVALLQ RAGLKVTAVADLPGLVVLRTVAMLANEAADAVLQGVGSAADIDLAMRAGVNYPCGPLAWATNIGIAHTLRVLDNLQRSYGESRYRPSLL
faoA CLEDGIVETAAEADMGLVYGIGFP-----LFRGGALRYIDSIGVAEFVALADQYAE L GALYHPTAKLREMAKNGQSFFG
fadB CLEEGIIATPAEADMALVYGLGFP-----PFHGGAFRWLDTLGSARYLDMAQQYQHGLPLYEVEPEGLRNKARHNEPYPVPVEPARPVGDLKTA--

cdtB -----
adh1 -----K
bhbD -----
phaC LRRCEAKGGT LHD-
faoA -----
fadB -----

FIGURE 4-12. Multiple sequence alignment of *ttdC* gene product with four thiolases. Residues the same (*) or similar (.) are indicated below the alignment. Conserved active centre cysteine residues are boxed. See Table 4-5 for references and percent similarity between the sequences. Sequences are aligned in order of decreasing similarity.

<i>ttdC</i>	MKDIVIVDAIRTPIGR-FRGLAGVRADHLGALVIGKLLERVD-LSPAQVDDVIFGCVTQIGEQSANIAR
<i>catF</i>	MKHAYIVDAIRTPFGR-YAGGLAAVRADDLGAIPAAALIERNPSVNWAQVDDVIYGCANQAGEDNRNVGR
<i>pcaF</i>	MRDVFICDAIRTPIGR-FGGALAGVRADDLAAVPLKALIEPNPAVQWDQVDEVFFGCANQAGEDNRNVAR
<i>fadA</i>	MEQVVIVDAIRTPMGRSKGGAFRNVRAEDLSAHLMRITLLARNPALEAAALDDIYWGCVQQTLEQGFNIAR
<i>phbA</i>	MNEVVIVAATRTAIGS-FQGALSAIPATELGAIVIRRLLEQTG-IDAAQIDEVILGQVLTAGAG-QNPAR
	* * * * * . * * * . * . . . * . . . *
<i>ttdC</i>	TALLGAGWPVTIPGLTIDRKC [~] GSSEVAVHLAAGAIAAGSADIVVAGGAENMSRVPMGSNREIHG-----
<i>catF</i>	MSALLAGLPVEVPATTVNRLCGSSLDATAAARAIAKAGEAHLIIAGGVESMSRAPYVMGKSEGAFGRTQK
<i>pcaF</i>	MALLLAGLPESIPGVTLNRLCASGMDAIGTAFRAIASGEMELIAGGVESMSRAPFVMGKAESGYSRNMK
<i>fadA</i>	NAALLAEVPHSVPAVTVNRLCGSSMQALHDAARMIMTGDAQACLVGGVEHMGHVPMSHGVDVDFHP-----
<i>phbA</i>	QTAIKAGLPHTTPALTLNKVCGSGLKAVHLAVQAIRCGDAELVIAGGQENMSLAPYVLPKAR----TGLR
	. . . * * * * . * * * . * . . . * * * * *
<i>ttdC</i>	-----EAFGWMAAERYELTSQGEAAERMADKVLSRDELDDYAFASHSRAAAAADAGHFANEIVEVP
<i>catF</i>	IEDTTMGWRFINPKLKAMYGVDTMPQTAENVAEQFGIQREDQDQFAYTSQORTAAAQAKGYFAKEIIVPVT
<i>pcaF</i>	LEDTTIGWRFINPLMKSQYGVDSMPETADNVADDYQVSRADQDAFALRSQKAAAAQAAGFFAEIIVPVR
<i>fadA</i>	-----GLSRNVAKAAGMMGLTAEMLARMHGISREMQDAFAARSHARAWAATQSAEFKNEIIPITG
<i>phbA</i>	MGHAQLQDSMIQDGLWDAFNDDYHMGITAENLAQKYEISREAQDAFSAASQKAAAAIEGGRFQSEITPIL
	. * . * . * * . . * . . . * * * *
<i>ttdC</i>	VETLREHGLSEPVAPLRGDETIRRDTSREKLSTLKTSFRPETGRITAGNSSQISDGAAALLMSAETAGR
<i>catF</i>	IPQRKGEP-----VVIATDEHPRASTTLEGLAKLKGVPKPE-GSVTAGNASGINDGAAVLIASDEAVAQ
<i>pcaF</i>	IAHKKGE-----TIVERDEHLRPETTLEALTKLKPVNGPD-KTVTAGNASGVNDGAAALILASAEAVKK
<i>fadA</i>	GHDADGVL-----KQFNDEYVIRPETTVEALATLRPAFDPVNGMVTAGTSSALS DGAAAMLVMSESAHE
<i>phbA</i>	IPQRKGEP-----LVFDTDEQPRIDSTAQALAKLKPAFQKD-GSVTAGNSSTLNDGAAVLLASAAKQE
	* * * . . . * . . . * * * . . . * * * . . . *
<i>ttdC</i>	LGLRARARIRAF [~] TTVGSDPTLMLTGRSP-LLQVLEKAGLGIDDLFEVNEAFASVPLAWMCETG--VS
<i>catF</i>	YQLKARAKIIASTTVGIEPRIMGFAPAPAIKLLKQANLTL [~] DQMDVIELNEAF [~] AAQALACTRDLGLA-DD
<i>pcaF</i>	HGLTPRARVLGMASGGVAPRVMGIGPVPVRKLTERLGVAVSDFDVIELNEAFASQGLAVLRELGVA-DD
<i>fadA</i>	LGLKPRARVRSM [~] AVVGC [~] DP [~] SIMGYGVPVASKLALKKAGLSASDIGVFEMNEAF [~] AAQILPCIKDLGLIEQI
<i>phbA</i>	LGLPVLARIKAYASAGVDP [~] SIMGIGVPVATRLTLQKAGWSIEDLDLIEANEAF [~] AAQSLAVGKELG---WD
	* * . . . * * . * . . . * * * * . * . *
<i>ttdC</i>	HERLNVNGGAIALGHPLGASGARLMTTLLNELERRGARYGLQAT [~] CCAGGMGTATIIERLD-
<i>catF</i>	DARVNPNGGAIALGHPLGASGARLVTTALNQLEQSGGKYALCSMCI [~] GVGQGIALLIERV--
<i>pcaF</i>	APQVNPNGGAIALGHPLGMSGARLVLTALHQLKSGGRKGLATM [~] CVGVGQGLALAIER--V
<i>fadA</i>	DEKINLNGGAIALGHPLGCSGARISTTLLNLMERKDVQFGLATM [~] CI [~] GLGQGIATVFERV--
<i>phbA</i>	TSKVN [~] VNGGAIALGHPIGASGARILVSLVHELIRRDGKKGLATL [~] CI [~] GGGQGVGLAIER---
	. . * * * * * * * . * * * * . . * * * *

Figure 4-13 illustrates the alignment of the putative cytochrome P450 monooxygenase coded for by *tdtD*. The boxed region corresponds to the heme-binding region (Nelson *et al.*, 1993) that contains an absolutely conserved cysteine residue in the eighth position. The glutamine and arginine residues indicated by arrows, that form a salt bridge in the core structure, are also universally conserved amongst P450s (Graham-Lorence and Peterson, 1996).

Figure 4-14 is a sequence alignment between the gene products of *tdtL* and *dit orf1* from *Pseudomonas abietaniphila* strain BKME-9. This latter sequence is truncated and includes 434 amino acids of which 374 are identical (*) and 32 are similar (.) to the A19-6a CoA ligase. Overall similarity is 93.5 percent and includes the proposed acyl-adenylate binding site (highlighted) which is identical in the two enzymes (Babbitt *et al.*, 1992). At the nucleotide level similarity is 78 percent. The 37 nucleotides that are between *orf1* and *ditl* (Figure 4-2) are 54 percent homologous to bases 89 to 50 of the minus strand of the *tdt* locus (data not shown). Similarities between the other genes in the two loci are less than those presented in the alignments already shown.

FIGURE 4-13. Multiple sequence alignment of P450_{tdtD} from *Pseudomonas* sp. A19-6a and several sequences identified by BLASTX search (see Table 2 for sequence identification and references). The boxed region is the heme-binding site (Nelson *et al*, 1993). Absolutely conserved residues are indicated by an arrow above the alignment and those that are the same (*) or similar (.) are indicated below the alignment. Sequences are aligned in order of decreasing similarity.

tdtD MSGPAHSNLEQVFANVASNYRGADVDLHAVYREMREKS-PVLPENFMARLGVPISIAGLDPNRPFTTLFKY
 terpC MDARATIPHEIARTVILPQGYADDEVIYPAFKWLRLDEQ-PLAMAHIEGYDPMWIATKHADVMQIGKQPGL
 soyC MTDATTTPQTTDAFAFPSNRSCPYQLP-DGYAQLRDTPGPLHRVTLYDGRQAWVVTKHEAARKLLGDPRL
 pinF1 MIANSSTDVSVADQKFLNVAKSNDIDP-DAVPISRLDS-EGHSIFAWEPRKRPFLRRREDGIFLVLRADHI
 pinF2 -----MEERRVSISSITWRFPMLFAPVDDVTTTIDDLT--LDPYPIYRRMRVQNPVVHVASVRRFTLTKA
 linC -----MERPD LKNPDLYTQQVPHDIFARLRREE-PVYWN PESDGS GFVAWL RHKDI IEVSRQPLL
 camC -MTTETIQSNANLAPLP PHVPEHLVDFD MYNPSNLSAGVQEAWAVLQESNVPDLVWTRCNGGHWIATRG

tdtD DDVMA---VMRDATNFT-SGFIAEGLGSFFDGLILTAMDGEAHKNIRSLQLQPVFMPETVNRWKETKIDRV
 terpC FSNAEGSEILYDQNEAFMRSISGGCPHVID--SLTSM DPPHTAYRGLTLNWFPASI-RKLEENIRRI
 soyC SSNR-----TDDNFPATSPRFEAVRESPQ----AFIGLDPPEHGTRRRMTISEFTVKRI-KGMRPEVEEV
 pinF1 FLLGTDPRTRQIETELMLNRGVKAGAVDFDIDHSM LFSNGETHGKRSSGLSKAFSFRMV-EALRPEIAKI
 pinF2 FDTKM---VKDDPSRFS-SDDPSTPMKPAFQAHTLMRKDGTEHARERMAMARAFAPKAIADHWAPIYRDI
 linC ESSAYEN-GGHRIFNENEVGLTNAGEAAVGV--PFISLDPPVHTQYRKVIMPALSPARL-GDIEQRIRVR
 camC QLIRE---AYEDYRHFSSECPFIPREAGEAYDFIPTSM DPPEQRFALANQVVGMPVV-DKLENRIQEL

tdtD IREEYLRPMVASKRADIM-EFALYFPPIRVISLIGFPEDRPEEIEQYAAWALAILAGPQVDPEK--AAAA
 terpC AQASVQRLLD FDGECDFMTDCALYYP LHVMTALGVPEDDEP---LMLKLTQDFFG--VHEPDEQAVAAP
 soyC VHGFLDEMLAAGPTADLV SQFALPVPSMVICRL LGVPYADHE---FFQDASKRLVQ--STDAQ-----
 pinF1 TECLWDDLQ-KVDDFNFTEMYASQLPALTIASVLGLPSEDTP---FFTRLVYKVS--CLS PSWR--DEE
 pinF2 VNEYLDRLP-RGDTVDLFAEICGPVAARILAHILGICEASDV---EIIRWSQRLIDG-AGNFGWR--SEL
 linC AEALIERIP-LGEEVDLP LLSAPLPLLT LAELLGLDPDCWY---ELYNWTNAFVG--EDDPE-----F
 camC ACSLIESLR-PQGQCNFTEDYAEFPPIRIFMLLAGLPEEDIP---HLKYLTQDMTR---PDGS-----

tdtD RGAMEAAQALYDVVKVVAQRR-----AEGATGDDLICRLIRAIEYEGRSLLDDHEITTFVRSLLPAASET
 terpC RQSADEAARRFHETIATFYDYFNGFTVDRRSCPKDDVMSLLANSKLDGNYIDDKYINAYYVAIATAGHDT
 soyC --SALTARNDLAGYLDGLITQFQ-----TEPGAGLVGALVADQLANGEIDREELISTAMLLLIAGHET
 pinF1 FEEIEASAIELQDYVRSVIADSG-----RRMRDDFLSRYLKAVREAGTLSPIEIMQLMLIILAGSDT
 pinF2 FERSDEANAEMNCLFNLDLVKKHR-----SAPNPSAFATMLNAPD---PIPLSQIYANIKIAIGGVNE
 linC RKSPEDMAKVLGEFMGFCQELFE---SRRANPGPD IATLLANAEINGQPVALRDFIGNLTTLTVGGNET
 camC -MTFAEAKEALYDYLIP IIEQRR-----QKPGTDAISIVANGQVNGRPITSDEAKRMCGLLLVGGLDT

tdtD TTTRFGTLM TLLERPELLARIREDRSLVGKAIDEAVRY-EPVATFKVRQAAKDVEIRGVAIPKGAMVSC
 terpC TSSSSGGAIIGLSRNPEQLALAKSDPALIPRLVDEAVRWTA PVKSF-MRTALADTEVRGONIKRGDRIML
 soyC TASMTSLSVITLLDHPEQYAALRADRS LVPGAVEELLRYLA IADIAGGRVATADIEVEGHLIRAGEGVIV
 pinF1 TRTAMVMVTALALQNPALWSSLRGNQSYVAAAVEEGLRFEPVGSFP-RLALKDIDL DGYVL PKGSLAL
 pinF2 PRDALGTILYGLLTNPEQLEEVK-RQQCWGQAFEEGLRWVAPIQASS-RLVREDTEIRGFIVPKGDIVMT
 linC TRNSISHTIVTLSQQPDQWDILRQRPPELLKTATAEMVRHAS PVLHMR-RTAMEDTEIGGQAIAGDKVVL
 camC VVNFLSFSMEFLAKSPEHRQELIERPERIPAACEE LLRRFSLVADGR--ILTS DYE F HGVQLKKGQILL

tdtD IVTSANRDEDAFENADTFDIDR-RAKPSFGFGFGPHMCIGQFVAKTEI-NCALNAILDLMPNIRLD PDKP
 terpC SYPSANRDEEVFSNPDEFDITR-FPNRHLGFGWGAHMCLGQHLAKLEM-KIFFEELLPKLKSVELS--GP
 soyC VNSIANRDGT VYEDPDALDIHR-SARHHLAFGFGVHQCLGQNLARLEL-EVILNALMDRVPTLR LA--VP
 pinF1 SVMSGLRDEKHYEHPQLFDVGRQMRWHLGFGAGVHRCLGETLARIEL-QEGLRTLRRAPNLAVVG DWP
 pinF2 IQASANRDEDDVFEDGESFNVFR-PKSAHQSEFGSGPHHCPGAQISRQTVGAIMLPILFDRFPDMILP-HPE
 linC WYASGNRDES VFSDADRFDVTR-TGVQHVGFSGSQHVCGSR LAEMQL-RVVFEILSTRVKRFELC--SK
 camC PQMLSGLDERENACPMHVDFSR-QKVSHTTFGHGSHLCLGQHLARREI-IVTLKEWLTRIPDFS IAPGAQ

tdtD APEIIG---AQLRGPHHVHVIWD-----
 terpC PRLVAT---NFVGGPKNVPIRFTKA-----
 soyC VEQLVLRPGTTIQGVNELPVTW-----
 pinF1 RMMGHG--GIRRATDMVKLSFDL-----
 pinF2 LVQWRG---FGFRGPINLPVTLR-----
 linC SRRFRS---NFLNGLKNLNVVLVP-----K
 camC IQHKSG----IVSGVQALPLVWD PATTKAV-

FIGURE 4-14. Protein sequence alignment of CoA ligases coded for by *tdtL* from *Pseudomonas* sp. A19-6a and *dit orf1* from *Pseudomonas abietaniphila* BKME-9. The latter sequence is currently unpublished and was provided by the authors (Martin and Mohn, 1999). The proposed acyl-adenylate binding site is highlighted (Babbitt *et al.*, 1992). Identical residues (*) and similar residues (.) are indicated below the alignment.

```

tdt ligase      MSGIGHPDTWSVGERDTVIAALDRAVARHPDRILLDFSGNLYTTYSEVDSLSTRMAHALAA
dit orf1 ligase -----

tdt ligase      LGVKPGETVLTMLDNNIDAVVCWLAINKLRAVSVPIINTALKGEFLRHQIADTGTQLVICE
dit orf1 ligase -----EFLRHQIADTGTGLVICE
                      *****.*****

tdt ligase      ADYLPRILPLAEQLSEVSQILYR---GERAEPASCRIPIAALDEFRGEDDTPFTSKPQPS
dit orf1 ligase ADYLSRISAIADQLTDVHRVLYRNASGQQVTSPSAPFPIESLDAHRGSDDSAFERKPQPS
                      **** * .*.***. .*** *.. *.. **..** ** *..* *****

tdt ligase      DLACLIYTSGTTGPSKGCMISYNFMCNLARLQLRAGPANENDVTITPLPLFHMNALCVSI
dit orf1 ligase DLACLIYTSGTTGPSKGCMISYNFMCNLARLQLRAGPAEDDVTITPLPLFHMNALCVSI
                      ***** * *****

tdt ligase      IASILVGARAAILPRFSVSNFWQEVERSGATIASILGGMGGLLAQAPDNDAMKRCHGQIH
dit orf1 ligase IASILVGARAAILPRFSVSNFWPEVERSGATIASILGGMGGLLAQAPDNEAMLRCRGQIH
                      ***** *****..** **..****

tdt ligase      TARGNPYTEETKQIWRERFGTKLVGGNGYGLTEACVITSLAAGEYAAPGSSGKRIPDFDV
dit orf1 ligase TARGNPYTEETKQIWRERFGTRLVGGNGYGLTEACVVTSLAAGEYAAPGSSGKRIADFV
                      *****..*****..***** *****

tdt ligase      RIVDDNDQELPANTPGEIVLRPLRPDIMFQGYWNRPADTLKLMRNMWFTGDIKGFDDDG
dit orf1 ligase RIVDEQDNEVPGGTPGEIVVRPQRPDIMFQGYWRRPEDTQKLMRNMWFTGDVGKFDDEG
                      ****..*.*.*****..** ***** ** ** *****..*****.*

tdt ligase      FFYFVDRKKDYLRRRGENISSFEMEAAFAVHPDLAEVAVHAVPSDKGEDDVKVTAVLHQG
dit orf1 ligase FFYFVDRKKDYLRRRGENISSFEMEAAFAVHPALSEVAVHAVPSDKGEDDVKVTAVLHEN
                      ***** ** *.*****..

tdt ligase      ARLTPEELFWATDSVPYYALPRYIEFRDSLKPNPQGRVLKYQLRDEGKTADTWDLEDTD
dit orf1 ligase TELAPEALFWAADTVPYALPRYIEFRDSLKPNPQGRVLKYLLRDEGKTATTWDLDDTD
                      .*.** *****.*.***** ***** *****..***

tdt ligase      IKVAK-K
dit orf1 ligase IKVAKR-
                      *****

```

Development of resin acid catabolic gene probes

Figure 4-15 illustrates a series of probes based on the *tdt* genes developed for the study of resin acid gene distribution in environmental isolates. To ensure that probe positive results for the A19-6a positive control were as expected, based on the sequence, a series of restriction digests were probed to confirm common probe positive bands. This figure depicts the locations of the recognition sites for the eight restriction enzymes used in this experiment. The expected fragmentation patterns were supported experimentally for all probes except *tdtS*. In this case some unexpected *tdtS* fragments were found to be identical in size to those detected with the *tdtD* probe. For a complete list of expected and observed fragments and the location of known restriction sites see Appendix B, Tables B1 and B2, respectively.

Knock-outs

To determine whether these genes were necessary for abietic acid degradation, suicide vectors containing a portion of each gene were constructed to knock-out the functional genes in the wild type A19-6a strain. Figure 4-16 shows a map of the different fragment inserts and a map of the suicide vector pRR1 used to deliver them. The latter was derived by digestion of pUTCm (de Lorenzo and Timmis, 1994) by *Sa*I and religation of the 3.6 Kb fragment (R. Rashid, unpublished) that was later modified to include the MCS3 linker shown (M. Providenti, unpublished). Unfortunately, homologous recombination and gene knock-out was only successful with the *tdtD* construct. In fact, most of the putative transconjugants generated by the knock-out experiments remained

ABA⁺ and did not carry the *bla* gene even though phenotypically they were ampicillin resistant.

FIGURE 4-15. Map of probes, based on *tdt* locus, developed for the study of resin acid gene distribution in environmental isolates. The eight restriction site maps shown were used to confirm that cloned probes corresponded with gene arrangement in A19-6a. A complete list of expected and observed results is presented in Table B1, Appendix B.

FIGURE 4-16. Composite map of *tdt* gene inserts and the pRR1 suicide vector used to determine which genes were involved in abietic acid degradation. MCS3 was generated by cloning a linker into the *Not* I site of the original MCS. The arrow on the insert map indicates the point of insertional inactivation by pBW30::TnphoA'-1.

MCS: *Not* I *Sfi* I *Hin* DIII *Sal* I *Pst* I *Bgl* II

MCS3: *Eag* I *Nsi* I *Swa* I *Sma* I *Bam* HI *Xba* I *Sal* I *Pst* I *Sph* I *Hin* DIII
Not I *Sfi* I *Hin* DIII *Sal* I *Pst* I *Bgl* II

Table 4-6 is a summary of HPLC results for all confirmed transconjugants generated in this study. As previously described the *tdtL* gene was completely knocked out by insertional inactivation by pBW30::Tn*phoA*'-1. This fact was confirmed by both hybridization and sequence analysis. Mutants A103-37 and A104-38 were generated in the same manner, however, for these transconjugants the ABA⁻ phenotype is only exhibited in liquid culture. In addition, hybridization demonstrated that for the A103-37 transconjugant, mutation occurred as a result of transposition and not recombination as was the case for the other two (data not shown).

TABLE 4-6. Resin acid degradation levels (percent) for *Pseudomonas* sp. A19-6a transconjugants A76-47, A103-37, and A104-38 and gene knock-out mutants P450-R1 and P450-R3. Degradation for the wild type A19-6a strain (positive control) was 100%. All cultures were grown on abietic acid except as noted. Level of degradation is based on the amount of resin acid in solution in comparison with an uninoculated negative control.

GENE	MUTANT	CARBON SOURCE	% DEGRADED	TRIAL
<i>tdtL</i>	A76-47	abietic acid	0	I
<i>tdtL</i>	A76-47	abietic acid	17	II
unknown	A103-37	abietic acid	0	I
unknown	A103-37	abietic acid	0	II
unknown	A104-38	abietic acid	0	I
unknown	A104-38	abietic acid	0	II
<i>tdtD</i>	P450-R1	abietic acid	70	I
<i>tdtD</i>	P450-R1	abietic acid	71	II
<i>tdtD</i>	P450-R1	dehydroabietic acid	82	I
<i>tdtD</i>	P450-R3	abietic acid	76	I
<i>tdtD</i>	P450-R3	abietic acid	76	II

For the *tdtD* gene, 17 putative knock-out mutants were generated of which two, P450-R1 and P450-R3, were confirmed recombinants by hybridization with the *bla* gene probe. Unfortunately, efforts to map the point of recombination between vector and target region in both of these mutants were unsuccessful. In comparison with the wild type A19-6a strain, most restriction digests produced one or two extra *tdtD*⁺ bands for P450-R1 that were approximately 2.6 Kb and 4.2 Kb in size. These bands were also *aph*⁺ and similar in size to the uncut and linear suicide vector used to conduct the knock-out experiment. While this result suggested that recombination may have occurred near one of the ends of the target region, none of the restriction enzymes tested could corroborate this hypothesis. Note that the *tdt*⁺/*aph*⁺ bands were also similar in size to the pUC128::*tdtD* probe plasmid used to generate the knock-out insert. In the case of P450-R3 some digests produced bands that were slightly larger for the mutant than they were for the wild type, however, none of them were *aph*⁺ and as such were not useful in determining the point of recombination.

A second attempt was then made to physically as well as phenotypically demonstrate *tdtD* knock-out using transformation and a double-crossover event between the target region and a linear piece of DNA. This resulted in a single ABA⁺ mutant that had incorporated the *aph* marker into its genome somewhere outside of the target region.

DISCUSSION

Sequence analysis of mutant A76-47

Automated fluorescence sequencing of YFG23 using a pUC/M13 forward primer identified the phage and plasmid cloning sites used to isolate target DNA from the mutant A76-47 (see Figure 4-3). This sequence reads into DNA that was identified by similarity search as the C terminal end of a cytochrome P450 monooxygenase (see Figure 4-13).

Sequencing with the reverse primer provided irrefutable evidence that insertional inactivation of the mutated resin acid catabolic gene was due to integration of the transposon and vector. As illustrated in Figure 4-4A, *Bgl* II digestion of the mutant A76-47 cut at the inside end of the IS50R element ensuring that the maximum amount of target DNA was cloned into phage λ M5 and plasmid YFG23. The YFG23M13 sequence reads approximately 400 bases of the *tnp* sequence and is identical to the IS50R element of native Tn5 (Beck *et al.*, 1981). The YFG23CM3 (Figure 4-4B) sequence generated by the inverse PCR primer that reads out from the outside ends of the transposon (Rich and Willis, 1990) indicated that the transposon was located in the *tral* gene of the RP4 *oriT* region of pBW30 (Zieglin *et al.*, 1991). Conjugal transfer of the plasmid must have started at the *nic* site. The transposon/plasmid junction is 951 bases upstream from this point. If transposition of Tn*phoA*'-1 was the mechanism of mutation in A76-47, then the junction between transposon and target DNA should have been at the IS50L-O and IS50R-O ends. Instead, the junction point

was adjacent to *oriT* of the suicide plasmid (Figure 4-4C). Sequence YFG23CM4 identifies the insertion site as being 67 bases downstream from the *nic* site. Sequence comparison of the RP4 *oriT* and the mutated region shows 46 percent similarity between the two sequences in two areas ending 52 bases upstream from the insertion site (see Figure 4-5). While sequence similarity must have been sufficient to initiate recombination it cannot be considered enough to qualify as homologous. The mutation thus represents an illegitimate recombination event. Hofer and colleagues (1996) reported similar results and found that vector integration may have actually increased the stability of the mutation. It would be interesting to sequence the other junction between the insert and chromosome in A76-47 to determine if there was any deletion of DNA as a result of this mutagenic event.

Sequencing of the *tdt* locus

Compilation of the sequence data confirmed that a locus involved in the catabolism of tricyclic diterpene compounds had been isolated from *Pseudomonas* sp. A19-6a (see Figure 4-6). This series of *tdt* genes is found on a 9180 bp segment localized by the insertion of pBW30::Tn*phoA*'-1. Note that all genes are transcribed in the same direction except for *tdtL*. Table 4-4 lists general features of the *tdt* locus that is 62 percent GC rich and in agreement with the characteristic range for pseudomonads (Bergey, 1984). The presence of hairpin structures following *tdtL* and *tdtR* indicates that they may be independently transcribed from the other genes. The presence of a promoter like sequence (AT-rich inverted repeat) before *tdtL* also suggests that this gene is expressed. Expression may be controlled by *tdtR*, which is divergently

transcribed, as many operon regulators appear to be (DiMarco *et al.*, 1993; Olivera *et al.*, 1998; Gerischer *et al.*, 1998). The presence of putative ribosomal binding sites before *tdtL*, *tdtR*, *tdtS*, and *tdtD* indicates that these genes may be translated. The fact that there is sequence overlap between the stop and start codons of *tdtS/tdtA* and *tdtB/tdtC* suggests that these genes are co-transcribed (see Figure B1, Appendix B). Since the presence of these features can only allude to the metabolic importance of these genes more work is necessary to elucidate how this locus functions genetically.

Genes involved in resin acid degradation

Before discussing the identification of the *tdt* genes it is important to consider what one might expect to find in the characterization of resin acid catabolic genes. I discussed several examples of bacterial resin acid degradation pathways in Chapter 2. These were elucidated on the basis of GCMS data that identified intermediates formed but not the catalytic mechanisms used to generate them. Nevertheless, based on our knowledge of protein functions we can speculate on the types of enzymes involved in these reactions by looking at changes made between substrate and product. For example, Biellman *et al.*, (1973a; 1973b) reported that the first step towards the degradation of dehydroabietic acid by *F. resinovorum*, *A. eutrophus*, and a *Pseudomonas* sp. strain was oxidation of the C3 and/or C7 carbons of the substrate. Cytochrome P450 monooxygenases are well suited to this since their principle reaction chemistry involves the transfer of a single oxygen atom from molecular oxygen to a substrate (Guengerich, 1991). They are not, however,

restricted to this reaction mechanism and are known to catalyse reductions, dehalogenations, epoxidations, substrate rearrangements, dealkylations, and hydroxylations. In fact, the latter is exemplified by the biosynthesis of abietic acid in which two separate P450 enzymes are responsible for generating the carboxyl group through successive hydroxylations following diterpene cyclization (Funk and Croteau, 1994).

Following oxidation, activation of the C-ring is necessary prior to complete degradation of the substrate. In these bacterial systems activation takes the form of dihydroxylation of the C11 and C12 carbons (Biellman *et al.*, 1973a and 1973b). Flavin-monooxygenases such as *p*-hydroxybenzoate hydroxylase in *Acinetobacter calcoaceticus* are able to catalyze the single hydroxylation of *p*-hydroxybenzoate to protocatechuate (DiMarco *et al.*, 1993) prior to subsequent degradation to succinate and acetyl CoA (Gerischer *et al.*, 1998). Dihydroxylation requires the activity of multicomponent aromatic ring dioxygenases (van der Meer *et al.*, 1992) such as the *tod* system of *P. putida* F1 which comprises a reductase, a ferredoxin, and an iron-sulphur protein (Zylstra *et al.*, 1988). Oxidative activation of substrates is not the only form seen in bacterial systems. The β -oxidation of long-chain fatty acids in *E. coli* cannot occur without acyl-CoA synthetase activation (Black *et al.*, 1992).

Coenzyme A activation of substrates is well known in the anaerobic degradation of xenobiotics and is now being discovered in aerobic systems (Villemur, 1995). In the anaerobic degradation of benzoate by *Rhodopseudomonas palustris* a benzoate-CoA ligase catalyses the first step of

the pathway and defines the substrate affinity exhibited by this organism (Egland *et al.*, 1995). Aerobic degradation of benzoate in a denitrifying *Pseudomonas* sp. also occurs via an acyl-CoA intermediate (Altenschmidt *et al.*, 1993). In both of these systems the CoA ligase is inducible and catalyzes the addition of an acyl group to the carboxy moiety to generate a thioester (Villemur, 1995). There are several reasons why the acyl CoA activation of a substrate may be favored in the catabolism of novel substrates. In the case of *R. palustris* the benzoate CoA ligase generates a concentration gradient that promotes uptake of the substrate particularly when environmental levels are low (Egland *et al.*, 1995); a function not unlike that of *fadD* in *E. coli* (Nunn, 1986). One very good reason for this form of activation is the increased reactivity of thioesters in comparison to their oxygen ester counterparts (Zubay, 1993). In addition, the CoA group may act as a recognition site to enhance substrate binding to other enzymes in the catabolic pathway, as is observed for *E. coli* fatty acid synthetase. CoA thioesters may also aid in thermodynamically unfavorable reactions such as the carboxylation of phenol by denitrifiers in a manner analogous to the substrate concentration gradient mentioned above (Villemur, 1995).

As important as these early steps are, the need for substrate transport must not be overlooked. Passive diffusion of substrate is an unlikely means of substrate transport even for hydrophobic compounds. This is demonstrated by fatty acid metabolism in *E. coli* which has a saturable transport system that facilitates transfer of long- and medium-chain fatty acids across the outer membrane (Nunn, 1986). In *P. putida* the *pcaK* gene codes for a transport protein that plays a role in the chemotactic response of this organism to 4-

hydroxybenzoate (Harwood *et al.*, 1994). In addition, the phenylacetic acid pathway of this bacterium has two transport genes, a permease and an outer membrane channel protein (Oliver *et al.*, 1998). In *A. calcoaceticus* the *pcaK* and *quiX* genes encode a putative transporter and porin, respectively (Gerischer *et al.*, 1998).

Other enzymes common to these catabolic pathways include isomerases, dehydrogenases, and thiolases. Isomerases are responsible for substrate rearrangements. Examples include the conversion of maleylpyruvate to fumarylpyruvate during benzoate oxidation in *Pseudomonas* sp. strain KB740 (Niemetz *et al.*, 1995) or in the *cis/trans* isomerization of enoyl-CoA in *E. coli* fatty acid metabolism (Nunn, 1986 and Zubay, 1993). Like many isomerases, dehydrogenases and thiolases are often part of multi-enzyme complexes that are involved in substrate oxidations. For example, phenylacetic acid oxidation in *P. putida* occurs via a multi-enzyme complex that includes two hydratases, a dehydrogenase, and a ketothiolase (Olivera *et al.*, 1998). Similar units are found in carnitine and long-chain fatty acid metabolism in *E. coli* where the dehydrogenase (operating in reverse) mediates substrate reduction (Eichler *et al.*, 1994 and Nunn, 1986). The thiolase mediates the final step in *E. coli* fatty acid β -oxidation to form a truncated fatty acid and acetyl CoA (Nunn, 1986). In the β -ketoadipate pathway the addition of a second acyl CoA group to the substrate enable terminal hydrolysis to form acetyl CoA and succinate (Harwood *et al.*, 1994). Recall also that thiolases catalyse the first step in the formation of the isoprene subunit from which all resin acids are constructed (Qureshi and Porter, 1981).

Since these pathways have high demands in terms of cellular energy needs, cofactors, and reducing equivalents it is not surprising that transcription is tightly regulated in many of them. The *fadR* gene functions as a repressor in fatty acid oxidation and helps maintain a balance between synthesis and degradation in *E. coli* (DiRusso *et al.*, 1992). Induction of 4-hydroxybenzoate degradation in *A. calcoaceticus* is mediated by *pobR* activation of *pobA* and *pcaU* induction of the protocatechuate pathway (Gerischer *et al.*, 1998). In *P. putida* it is the *pcaR* gene that induces a minimum of four of the enzymes required by the β -ketoadipate pathway (Harwood *et al.*, 1994). Taken together it can be seen that the degradation of complex substrates occurs via equally complex catabolic pathways.

Characterization of the *tdt* genes

Of the seven open reading frames carried on this segment of DNA six were identified by BLASTX similarity search (see Table 4-5). The first gene of the locus, *tdtL*, codes for a putative acyl CoA ligase with 27.1 percent similarity to the *fadD* gene product of *E. coli*. Multiple sequence alignment (Figure 4-7) with this and several other CoA ligases identifies the presence of the putative acyl-adenylate binding region (Babbitt *et al.*, 1992) that is indicative of these genes. This suggests that thioesterification is an important step in abietic acid degradation since insertional inactivation of this gene resulted in the loss of the ABA⁺ phenotype in A76-47 (see Table 4-6). In fact, it may be the very first step in a manner similar to that seen for benzoate degradation in *Pseudomonas* sp.

KB640 (Altenschmidt *et al.*, 1993). This could account for A19-6a's ability to sequester substrate and grow at a higher rate than the other abietic acid degrading isolates in the consortium (see Figure 2-4). None of the other isolates had homologues to *tdtL* (see Table C1, Appendix C). If A19-6a ligated CoA to the substrate during or immediately following entry into the cell then a concentration gradient would be generated that could favour increased uptake and provide this isolate with a greater proportion of available substrate that would be translated into a higher growth rate. Similar steps in long-chain fatty acid uptake are proposed (Nunn, 1986). It is interesting to note the 25 percent sequence similarity between *tdtL* and *pt4CL1*, a CoA ligase involved in lignin biosynthesis in aspen (Hu *et al.*, 1998). One wonders if there is some kind of evolutionary relationship between these genes since the substrate recognized by the *Pt4CL1* ligase is similar in structure to monoterpenes.

Downstream from *tdtL* and reading in the opposite direction is a putative regulator designated *tdtR* (see Figure 4-6). A protein sequence alignment between *tdtR* and the *luxR* and *mtrR* gene products is shown in Figure 4-8. Homology to both of these sequences is low, 19.5 and 17.8 percent, respectively, and occurs primarily in and around the helix-turn-helix region characteristically found in DNA binding proteins (Pan and Spratt, 1994). The *luxR* gene product of *Vibrio harveyi* acts as a transcriptional activator (Swartzman *et al.*, 1992). In the *Vibrio luxR* regulatory system, which controls luminescence, another regulatory component may also be involved since the system could not be expressed in *E. coli* as had been done for *V. fischeri* luminescence (Showalter *et al.*, 1990). The *mtrR* gene product of *Neisseria*

gonorrhoeae controls antibiotic resistance by altering the permeability of the cell by repressing the synthesis of the envelope lipoproteins encoded by *mtrC* and *mtrD* (Pan and Spratt, 1994). Mutations in *mtrR* also affect susceptibility to detergents which is interesting since resin acids are commonly used as surfactants and soaps in manufacturing processes (Lawrence, 1969). It should be noted that repressors are sometimes shown to be homologous to transcriptional activators solely due to the presence of helix-turn-helix regions in both of these proteins (Neidle *et al.*, 1989). Therefore the nature of a novel regulator (whether an activator or repressor) cannot be determined from the sequence alone. The *tdtR* gene may be an activator or a repressor that acts to regulate *tdtL* in the presence of abietic acid. Repressor regulation of CoA ligase genes is found in several catabolic systems including benzoate metabolism in *Rhodopseudomonas palustris* (Kim and Harwood, 1991), the *pha* operon of *Pseudomonas putida* U (Olivera *et al.*, 1998), and *fadD* of *E. coli* (DiRusso, 1988) to name a few. Note that the range of inducers and substrates recognized by these genes is sometimes quite broad (Kim and Harwood, 1991 and Fernandez-Valverde *et al.*, 1993).

Attempts to identify the gene product of *tdtS* were unsuccessful. The similarity search gave low E-values only for the hypothetical proteins shown in Figure 4-9.

The *tdtA* gene putatively encodes a isomerase/decarboxylase type protein that is 39.6 percent homologous to a *Rhodobacter capsulatus* protein (see Table 4-5 and Figure 4-10). It is also 29.4 percent similar to the *hpaG* gene product of

E. coli that is a bifunctional enzyme responsible for the decarboxylation and isomerization of intermediates in homoprotocatechuate metabolism (Roper and Cooper, 1993 and Roper *et al.*, 1993). The similarity search of the putative *tdtB* gene product identified it as very similar to CoA-dependent dehydrogenases from several Gram-negative and Gram-positive bacteria (see Figure 4-11). In *E. coli* fatty acid metabolism the *fadB* gene product provides a substrate reduction function to the multienzyme oxidation complex. Mohn and colleagues (1998) also detected isomerase/decarboxylase and dehydrogenase genes in their resin acid catabolic operon, so finding similar sequences here is not entirely unexpected. Nevertheless, since neither Mohn nor I have demonstrated that these genes are necessary for resin acid degradation, the functions of *tdtA* and *tdtB* remain unknown.

The presence of a thiolase gene, *tdtC*, is an interesting result since these enzymes often catalyze the final step in degradative pathways (Nunn, 1986 and Harwood *et al.*, 1994). The conserved cysteine residues (boxed) in Figure 4-12 confirm the identification of *tdtC* as a thiolase gene since these are proposed to be part of the active centres (Sato *et al.*, 1992). While the literature suggests that finding regulatory, isomerase/decarboxylase, dehydrogenase, and thiolase genes together on a catabolic operon is not unusual, further molecular and biochemical analyses are necessary to demonstrate whether these genes are active in resin acid degradation. The fact that they are proximal between two genes whose resin acid catabolic activity has been demonstrated does not mean they too are involved, although genome organization in related *Pseudomonas* spp. strains suggests that this is a possibility.

The final gene product of the *tdt* locus for which resin acid catabolic activity has been demonstrated (Table 4-6) is a cytochrome P450 monooxygenase (see Figure 4-13). This gene is identified by the heme-binding region (boxed) and conserved glutamine, arginine, and cysteine residues (arrows) that are universal amongst these enzymes (Graham-Lorence and Peterson, 1996). Note that the gene knock-out mutant P450-R1 was still able to degrade dehydroabietic acid. A possible metabolic pathway as illustrated in Figure 4-17, involves desaturating the C-ring of ABA to produce the aromatic DHA prior to oxidation of the substrate. The degree of saturation is the only structural difference between the two congeners. Metabolism of ABA into a pathway for DHA metabolism, in a manner similar to quinate metabolism in *Acinetobacter calcoaceticus* (Elsemore and Ornston, 1994), would be an efficient way of handling these compounds with a minimum of genetic redundancy. The removal of 70 percent of the abietic acid available by this mutant and 76 percent by the P450-R3 mutant is considered to be due to cellular uptake, without further metabolism, as growth could not occur. Uptake may continue if the CoA ligase is in fact sequestering the substrate following entry into the cell, in an effort to maintain a favorable concentration gradient. Evidence for this could be found by testing the P450 mutants for the build-up of metabolites formed by the preceding steps of the pathway using highly purified abietic and dehydroabietic acids separately as sole carbon sources. This may easily identify whether *tdtD* is involved in desaturation and/or oxidation of resin acid substrates. Note that the amino acid similarity of TdtD to other P450s is well below the 40 percent deemed necessary for inclusion in an existing family (Nelson *et al.*, 1993). Thus *tdtD* is

considered a new member of the P450 superfamily of genes. Recall that these enzymes are also involved in abietic acid synthesis (Funk and Croteau, 1994) and in the elimination of resin acids in fish (see Figure 1-4).

FIGURE 4-17. Proposed pathway for degradation of abietic acid in *Pseudomonas* sp. A19-6a.

Taken together these genes form the *tdt* locus. The *tdtL* and *tdtD* genes are required for growth on abietic acid. Some or all of the other *tdt* genes may also be necessary for resin acid degradation by bacteria found in pulp mill mixed liquor treatment systems. Testing probes based on these genes with a variety of environmental isolates should help to determine if in fact this is the case.

Relationship between *tdt* and *dit* operons

At the time of writing, sequence data for the *dit* operon illustrated in Figure 4-2 was not available. Although this data has not as yet been published the authors, Vince Martin and Bill Mohn, were kind enough to provide me with the data for thesis analysis (Martin and Mohn, 1999). Figure 4-14 is an alignment between the *tdtL* and *orf1* gene products that are both identified as CoA ligases. With amino acid sequence similarity at 93.5 percent these proteins represent members of the same ligase group. This might be expected if both are involved in resin acid catabolism. Involvement in DHA degradation has not been determined for the *dit orf1* gene. Nevertheless, with only 78 percent similarity at the nucleotide sequence level, they can be considered distinctly different. Hybridization results suggest that the entire *tdt* locus is present in BKME-9 in the same genetic arrangement as in A19-6a (see Chapter 5). This similarity is unlikely to be due to the *dit* gene *orf1* similarity due to the low (78 %) sequence identity. For this reason and the fact that the regions downstream from the ends of the coding sequences bear little resemblance to one another, overlap between the *tdt* and *dit* loci at this point is considered unlikely.

Development of *tdt* probes

A region from each of the seven open reading frames was sub-cloned into pUC128 for making purified probes that could be used to detect each gene individually. To ensure that these probes corresponded to the sequence as it occurs in the wild type A19-6a strain, the restriction digests illustrated in Figure 4-15 were screened with each probe. Expected and observed results (Table B1,

Appendix B) confirmed the locus map illustrated in Figure 4-6. The presence of anomalous bands with the *tdtS* probe that mirrored *tdtD* fragment sizes cannot be explained by cross-hybridization because the 38 percent similarity between these two sequences is considered too low to cause interference under high stringency hybridization conditions. Since membrane stripping was considered complete (i.e. no other probes produced similar unexpected bands) the reason for the anomolous fragments is not clear.

Gene knock-outs

Unfortunately, gene knock-out was only successful for *tdtD*, demonstrating that the P450_{tdtD} enzyme is necessary for abietic acid degradation (see Table 4-6). While dehydrogenation of the substrate may be the function of this enzyme, the HPLC data presented cannot be considered proof of function since disappearance of substrate is insufficient to show this. GCMS analysis using ultra-pure congeners and the P450 mutants may shed more light on this aspect. The inability to map a confirmed recombination site may merely reflect the type of recombination events that seem to characterize this organism. These events do not conform to those predicted based on *E. coli*. This theory is further supported by the incorporation of the *tdtD-aph* construct somewhere other than where it was targeted.

CONCLUSION

The results presented in Chapters 3 and 4 further support the hypothesis that DNA probes based on resin acid catabolic genes can be constructed from bacteria isolated from contaminated environments. In addition, the genes that may be involved in degradation of these compounds show some similarity in sequence and organization to known catabolic operons from taxonomically related and unrelated bacteria. Some genes, such as *tdtL* and *tdtC*, have products that are similar to enzymes involved in the production of plant secondary metabolites. For *tdtD*, the gene product may represent a new member of the cytochrome P450 monooxygenase superfamily that includes enzymes involved in abietic acid carboxylation and mixed function oxygenase detoxification of resin acids in fish. Nevertheless, it is important to note that genetic manipulation of environmental isolates may not always produce the results predicted based on theory as evidenced by vector integration during transconjugation.

CHAPTER 5

Distribution of *tdt* genes in resin acid degrading bacteria isolated from pulp mill mixed liquor secondary treatment systems

INTRODUCTION

Pulping processes

Understanding the pulping process is a first step toward effective wastewater treatment. Regardless of species, wood is composed primarily of carbohydrates (up to 75 %) in the form of long chain, unbranched, cellulose fibers and short chain, branched, hemicellulose molecules (McKean, 1980). The next major component is lignin which accounts for up to 28 percent of the wood content and acts as a wood stabilizer between and around the polysaccharide chains in the fiber walls. The final group of constituents is composed of extractives accounting for only 5 percent maximum of the wood content. All components contribute to varying degrees to the biochemical oxygen demand, characteristic colour, and toxicity of mill mixed liquors. The purpose of the pulping process is the separation of the wood fibres from the remaining components and subsequent reaggregation to form end products as illustrated in Figure 5-1.

Debarking is the first step in the process and the extent to which the bark is removed is determined by the end product. Rough grades require little or no debarking. Fine paper production requires complete debarking in addition to removal of wood knots and extensive cleaning. These processes generate high volume wastewater streams, the toxicity of which, is determined by wood species (Swartz and MacDonald, 1962). After debarking the wood is chipped prior to digestion by either mechanical, semi-chemical, or chemical means. Mechanical

pulping gives high yields of low grade fibers that are used for non-permanent papers such as newsprint, wrappings, wallpaper, etc (McKean, 1980). The weak paper fibers have high residual lignin contents, although extractives are removed and account for much of the mixed liquor toxicity. Semi-chemical pulping uses mild chemicals to soften lignin and increase removal rates while reducing mechanical defiberization, to yield a stronger end product that is used as wood corrugates. Mixed liquor volumes and toxicity are considered moderate.

Chemical kraft pulping as illustrated in Figure 5-1 is by far the most common pulping process used. This method uses caustic and acid extractions at high temperature to remove approximately 80 percent of the lignins and most of the resin and fatty acids to produce pulp that is used for printing and hygiene products (Shimp and Owens, 1993). Most of the remaining lignin is dissolved in the bleaching stage which uses oxidizing agents to make the lignin more soluble. Subsequent alkaline extraction removes low-molecular weight extractives and neutralizes acids that can lead to product degradation (McKean, 1980). Recovery of process streams is important and contributes to the cost-effectiveness associated with this method. Turpentine is recovered as relief gases generated during cooking (Swartz and MacDonald, 1962) and the black liquor generated by process and washing streams is evaporated and burned to recover the inorganics used in the white cooking liquor (McKean, 1980). Condensates generated during chemical recovery may be used as pulp wash water prior to release and may be high in resin and fatty acids due to foaming. Tall oil is recovered by skimming these sodium soaps from the black liquor (Swartz and MacDonald, 1962).

FIGURE 5-1. Generalized flow diagram of pulping processes used in the production of paper products. Adapted from McKean, 1980.

Wastewater treatment

Treatment of wastewaters is a sequential process that targets the removal of principal components at each stage. Primary clarification is designed to remove suspended solids and has little effect on total toxicity but may reduce

biochemical oxygen demand by up to 10 percent (Bonsor *et al.*, 1988). Secondary treatment targets dissolved organics, colloidal materials, and fines that are the principal contributors to mixed liquor toxicity. Aerated stabilization basins are the most common method of detoxifying mixed liquors including the removal of up to 50 percent of the adsorbable organic halogens. Basins generally occur in series, with initial lagoons being aerated and having the highest removal rates. Seasonal variations in temperate regions are common and resin/fatty acid removal can be as high as 90 percent under normal operating conditions (McKean, 1980). Activated sludge systems are also common and usually have higher removal rates but more variable results due to inadequate retention times and a reduced ability to handle loading shocks.

Detoxification of pulp mill mixed liquors

Biological treatment by anaerobic or aerobic processes are important aspects of mixed liquor detoxification. Studies have begun to focus on the effectiveness of indigenous microflora and activated sludges in existing treatment facilities and the use of bioreactors as supplemental treatment systems. Stuthridge and colleagues (1991) did an in depth study of a bleached kraft mill in New Zealand that uses two lagoon treatment systems to process different mixed liquor streams. System A, which received general mill wastewaters, relied heavily on mechanical aeration and was effective at mixed liquor detoxification within the short 4.5 day retention period. System B was used to separately treat condensates and alkaline bleaching streams and was less effective at detoxifying mixed liquors even after 45 day retention times. In this system, natural aeration predominated in the long-term storage lagoon.

Most of the resin acid toxicity generated by the mill flows through this part of the system and subsequent research showed removal and/or transformation was associated with both aerobic and anaerobic phases of the system (Zender *et al.*, 1994). In this study the researchers found that aerobic conditions favoured resin acid removal. They also found that abietic acid degradation occurred by several routes involving hydrogenation, hydroxylation, or desaturation and decarboxylation; the latter of which may be the mechanism used by *Pseudomonas* sp. A19-6a. Liss and Allen (1992) did a microbiological study of similar treatment systems and found that wood species and temperature affected community dynamics that would likely impinge upon detoxification prior to release. These studies demonstrate the importance of understanding what environmental factors affect degradation levels so that treatment systems can be optimized for maximum resin acid detoxification prior to release (Liu *et al.*, 1996).

While most secondary treatment relies upon aerobic lagoons experimental research has begun to investigate the usefulness of anaerobic systems. Anaerobic treatment may be coupled to aerobic systems due to the inherent toxicity of pulp mill wastewaters to anaerobic microflora (Murray and Richardson, 1993). Following primary settling, biocatalytic removal of peroxide and acidogenesis prepares the mixed liquor for methanogenic removal of organic acids and aerobic post treatment which produces the activated sludge that feeds the peroxide removal tank (Welander, 1988). Acidogenic sludges are also recycled and the addition of metal ions removes insoluble organics and inorganic salts that inhibit methanogenesis. Nitrogen and phosphorus may also be added although they are not required to the same degree as in aerobic systems (Murray

and Richardson, 1993). While these systems are very effective at removing adsorbed organic halogens (Graves and Joyce, 1994) they are less efficient at resin acid removal without prior acclimation of the anaerobic community (McFarlane and Clark, 1988).

Fate of resin acids in receiving waters

Even though process changes and improvements to secondary treatment systems have reduced the amount of resin acids released to receiving waters, it is important to consider their fate after discharge. Several studies conducted near a kraft mill on Lake Superior found that dehydroabietic acid could be detected in sediments up to six kilometers from the discharge point (Fox, 1977), and that the half life for this resin acid was 21 years (Brownlee *et al.*, 1977). Partitioning into sediments is not unusual, owing to the hydrophobic nature of these compounds. It is for this reason that a better understanding of the anaerobic fate of resin acids is necessary. Tavendale and colleagues (1997a and 1997b) found that a 50 percent reduction in total resin acids occurred anaerobically within 264 days, and that more recalcitrant and toxic biotransformation products accounted for much of the dehydroabietic acid removal. This of course increases the risk of chronic sub-lethal exposure of aquatic species. This area of research has received little attention.

Objective III

The final objective of my research was to use the catabolic gene probes developed from *Pseudomonas* sp. strain A19-6a to study the distribution of

bacteria with similar genotypes found in aerated lagoon samples collected from a variety of secondary treatment systems world wide. This study provides a basis for future work, applying widely conserved gene probes to monitor the population dynamics of resin acid degrading bacteria *in situ*.

MATERIALS AND METHODS

Determining probe specificity

To test the specificity of the probes for detecting resin acid degraders and/or taxonomically related bacteria the 13 Avenor bleached kraft pulp mill isolates (see Chapter 2), two strains isolated from pulp mill mixed liquor by the University of British Columbia, six strains isolated from Mount Norquay in Banff National Park in Alberta, and 10 type strains from our collection were screened by dot-blot hybridization with each probe. Genomic DNA (1.0, 0.5, 0.1 μ g) was diluted in 100 μ l TE and vacuum transferred in a dot-blot manifold (Biorad Co.) to nylon membranes in accordance with the manufacturer's recommendations. Following transfer, DNA was denatured, neutralized, and equilibrated by incubating the wet membrane for 15 minutes each on Whatman soaked with the appropriate solution. After air drying and UV cross-linking, membranes were probed and detected as described in Chapter 3. A secondary screening of putative probe positive isolates was performed with *Sac* II digested DNA resolved by gel electrophoresis as described in Chapter 3, to eliminate any false positives that may have been detected using the dot-blot protocol.

16S rDNA sequencing

The first 500 basepairs (approximately) of the 16S rDNA sequence from isolates A19-6a, A19-6b, D11-9, 16LW, and NQ5 were determined by MIDI Labs, Newark, NJ, USA, using primers that anneal to *Escherichia coli* 16S rDNA

positions 005 and 531. The 16S rDNA sequence for isolate BKME-9 was kindly provided by Dr. W. Mohn, University of British Columbia.

Screening of environmental samples

To determine the distribution of the *tdt* genotype in other putative resin acid degraders, pulp mill samples from a variety of locales were screened by spot hybridization. Ten millilitre pulp mill samples were kindly provided by Dr. R. Fulthorpe, University of Toronto from the sources listed in Table 5-1. Isolates were obtained by serial dilution onto ABA or DHA plates and random selection of a variety of colonies following 14 days incubation at 30°C. Isolates were purified once on selective media and transferred to TYE/2 in minimal media A until growth was sufficient to prepare both batch cultures with ABA or DHA and glycerol stocks. Following 10 days incubation at 30°C ABA or DHA cultures were spun twice for 10 minutes at 12,000g and the supernatant tested by HPLC as described in Chapter 2.

For preliminary probing, genomic DNA from all isolates was extracted using a modified alkaline lysis protocol. Overnight cultures (2.5 ml TYE) were pelleted and resuspended in 600 µl 0.1M NaOH, 0.25% SDS, neutralized with 300 µl ice cold 3M sodium acetate pH 4.8, incubated on ice 10 minutes, and spun 15 minutes at 4°C, 12,000g to remove protein and cell debris. DNA was precipitated overnight at -20°C with 0.6 volume isopropanol, recovered by centrifugation, washed, dried, and dissolved in 10 µl of TE (overnight at 4°C). One microlitre aliquots from each extraction were spotted onto membrane and

the membranes were treated as indicated in Chapter 3 for plaque-lift hybridization.

TABLE 5-1. Source list of pulpmill samples provided by R. Fulthorpe, University of Toronto, and screened with *tdt* gene probes.

LOCATION	SAMPLE SITE	DATE REC'D	FEEDSTOCK	CODE
Aracruz, Brazil	Centre of aerated lagoons 1-6 inclusive	29 Nov. 1996	eucalyptus	A/D1-6a...
Comwall	Activated sludge system C2LL2	18 Dec. 1996	softwood	96C/R#A/D
Domtar, Windsor	Cells 4 and 15 (L4,L15) of aerated stabilization basin	4 Feb. 1997	softwood	L4/15-A/D#
Japan	Unknown location in biomembrane treatment system	26 Feb. 1997	unknown	7057J-A/D#
Lewiston, Idaho	Aerated stabilization basins I, II, III	14 Apr. 1997	softwood	7104-1,2,3 A/D#
Paparo, New Zealand	Aerated stabilization basins NZ Pond 2 4/8, NZ Pond 2 3/4, NZ Pond 4 3/4, NZ Pond 4 19/3	14 Apr. 1997	eucalyptus pine	7104-4,5,6,7 A/D#
Potlatch	Site 1 and Final of aerated stabilization basin	10 Jun. 1997	softwood	P1/F-A/D3
Paparo, New Zealand	Unknown site from aerated stabilization basin	10 Jun. 1997	eucalyptus pine	7161-NZ A/D#
Alpac, Alberta	Treated outflow from activated sludge system	10 Jun. 1997	aspen	AL-A/D#
Domtar, Windsor	Cell 9 of aerated stabilization basin	10 Jun. 1997	softwood	D9-A/D#
Repar	Out and Dock sites of aerated stabilization basin	29 Aug. 1997	softwood	RO/D-A/D#
Paparo, New Zealand	Unknown site from aerated stabilization basin	29 Aug. 1997	eucalyptus pine	NZO-A/D#
TEMBEC	unknown	29 Aug. 1997	unknown	TEM-A/D#

Genomic DNA from putative probe positive isolates was extracted using a scaled down CTAB/NaCl protocol (Chapter 3) and rescreened by dot-blot hybridization (see above). Isolate DNA samples that were confirmed probe positive were subsequently digested with *Sac* II and probed to determine if *tdt* gene arrangements were the same as that observed for A19-6a.

RESULTS

Specificity of *tdt* probes

To determine the specificity of the *tdt* probes for detecting genes involved in resin acid degradation from taxonomically similar or different bacteria a total of 31 strains were screened by dot-blot hybridization. From this preliminary screening 15 strains were retested by Southern hybridization to yield six confirmed probe positives. Figures 5-2 depicts the fragments generated by *Sac* II digests for each of the *tdt* probes. This restriction enzyme was selected because the same fragment is positive for all probes in the A19-6a wild type control. As indicated, BKME-9 and A19-6b show the same probe positive pattern as A19-6a; although band intensity is reduced in the A19-6b isolate (for the same DNA loading). For the other four strains band patterns differ in size and intensity from that shown for the wild type control (see Table 5-2). Note that in the preliminary screenings A19-2 was consistently negative and A19-5 was consistently positive for all *tdt* genes. For a summary of the dot-blot results including HPLC data for each strain, see Table C1, Appendix C.

FIGURE 5-2. Autoradiographs of selected strains used to determine specificity of the *tdt* gene probes. For a complete list of strains screened, see Table C1, Appendix C. Fragment sizes and intensity relative to the A19-6a wild type control are listed in Table 5-2. Fragment sizes indicated by arrowheads are in Kb.

TABLE 5-2. Summary of *tdt* gene probe positive *Sac* II fragments illustrated in Figure 5-2. All band sizes are in kilobases. Signal intensity is indicated relative to the A19-6a wild type control; (+) weak positive, (++) moderate positive, (+++) strong positive.

PROBE	LIGASE	REGULATOR	ISOMERASE	DEHY- DROGENASE	THIOLASE	MONO- OXYGENASE
STRAIN/GENE	<i>tdtL</i>	<i>tdtR</i>	<i>tdtA</i>	<i>tdtB</i>	<i>tdtC</i>	<i>tdtD</i>
A19-2	nil	nil	nil	nil	nil	4.2
	n/a	n/a	n/a	n/a	n/a	+
A19-5	nil	nil	nil	nil	nil	nil
	n/a	n/a	n/a	n/a	n/a	n/a
A19-6a	10.0	10.0	8.5	10.0	8.5	10.0
	+++	+++	+++	+++	+++	+++
A19-6b	10.0	10.0	8.5	10.0	8.5	10.0
	++	++	++	++	++	++
D11-9	2.8, 2.0	2.9	3.2	3.4	2.7	2.8, 2.5
	++	+++	+	+	+	+
D11-37	nil	nil	nil	nil	nil	nil
	n/a	n/a	n/a	n/a	n/a	n/a
16LW	2.4	nil	2.0	nil	nil	4.2
	+++	n/a	++	n/a	n/a	+++
BKME-9	10.0	10.0	8.5	10.0	8.5	10.0
	+++	+++	+++	+++	+++	+++
NQ-5	4.2	nil	2.0	nil	nil	4.2
	++	n/a	+	n/a	n/a	++

16S rDNA sequence analysis

In an effort to understand how different *tdt* fragmentation patterns related to the phylogeny of probe positive isolates, the first 500 bases of the 16S rDNA sequences from strains A19-6a, A19-6b, D11-9, 16LW, BKME-9, and NQ5 were compared. Figure 5-3 is a phylogenetic tree of the six isolates that clearly shows the separation between the β and γ subgroups of Proteobacteria. Note also that fragmentation patterns cross phylogenetic groups i.e. the wild type pattern (*) is found in both groups and different patterns for some of the γ subgroup isolates are seen. A combined percent similarity/distance matrix based on the six nucleotide sequences is included for ease of reference.

Screening of environmental samples

To determine the distribution of the *tdt* genotype amongst bacteria isolated from a variety of locales approximately 275 isolates from pulp mill mixed liquor samples collected worldwide were screened with each probe. A subset of 28 strains were rescreened to yield a group of 13 isolates from 6 of the 26 samples tested. Table C2, Appendix C, summarizes these results including the percent ABA degraded in batch culture. Note that virtually all of the isolates from the Aracruz mill in Brazil were poor degraders whereas those from the other sample sites were generally efficient at using abietic acid as sole carbon source.

FIGURE 5-3. Phylogenetic tree of 16S rDNA sequences from six strains exhibiting various degrees of similarity to the *tdt* gene probes. Isolates with the same fragmentation pattern as A19-6a are indicated with an asterisk. Number of asterisks denotes probe intensity. Other strains have dissimilar band patterns as shown in Figure 5-2.

		PERCENT HOMOLGY					
DISTANCE	STRAIN	A19-6a	A19-6b	BKME-9	D11-9	NQ5	16LW
	A19-6a		78.8	96.6	96.0	95.4	80.7
	A19-6b	0.212		78.4	77.4	79.3	88.9
	BKME-9	0.034	0.216		95.6	94.0	80.1
	D11-9	0.040	0.226	0.044		94.2	80.0
	NQ5	0.046	0.207	0.060	0.058		81.1
	16LW	0.193	0.111	0.199	0.200	0.189	

Phenotypic analysis of probe positive strains

All strains that were found to consistently probe positive to one or more of the *tdt* genes were screened for substrate utilization in Biolog test-plates. The profile generated was used for hierarchical cluster analysis to determine the phenotypic relatedness between isolates collected from geographically distant locales. Figure 5-4 is the dendrogram generated from the data listed in Table C3, Appendix C, based on average linkage between groups. The scale below the tree is based on the Jaccard correlation coefficients presented in Table C4 between A19-6a and isolates shown in bold. As found previously with the Avenor bleached kraft pulp mill isolates (Chapter 2), two distinct groups emerge based on the utilization of a broad range of carbohydrate substrates (strains 16LW and NQ5) or organic acids, amino acids, and TCA cycle intermediates (all others). The major group that preferentially uses amino and organic acid carbon sources is composed of isolates that originate from diverse locations (i.e. the grouping of A19-6a (Ontario), BKME-9 (British Columbia), and the Brazil isolates). Note also that both the carbohydrate-degrading and the amino- and organic acid-degrading groups have representatives that cross the phylogenetic boundary between the β and γ subgroups of Proteobacteria.

Probe analysis of environmental isolates

Figure 5-5 depicts the autoradiographs generated by probing the 19 strains positive for the *tdt* genes. For the six probes tested the first five strains (A2a to BKME-9) are at least 90 percent similar phenotypically (Figure 5-4 and Table C4) and are from geographically distant locations. These isolates give a

probe positive band identical to that observed for the A19-6a wild type strain. Note that the gel is loaded in the same order as the phenotypic tree (Figure 5-4) and that these isolates along with A19-6b and 7104-1 A1 are the only ones that are positive for all genes. The Idaho isolates group into three distinct phenotypic clusters which in turn generated three different *Sac* II fragment patterns. All other strains gave different probe positive patterns with varying intensities as indicated in Table 5-3. The P450, ligase, and dehydrogenase probes were positive for all strains and the thiolase probe was positive for 16 strains. The isomerase and regulator probes were positive for only 8 and 13 of the strains, respectively.

FIGURE 5-4. Dendrogram of phenotypic similarities between *tdt*⁺ strains isolated from pulp mill mixed liquor from locations as indicated. Scale is based on the Jaccard correlation coefficient between the positive control strain A19-6a and those indicated in bold (see Table C4 Appendix C).

FIGURE 5-5. Autoradiographs of *tdt* probe positive strains isolated from pulp mill mixed liquor treatment systems from various geographic locations. Genomic *Sac* II digests were used to generate the observed fragments. Fragment sizes can be found in Table 5-3. Fragment sizes indicated by arrowheads are in Kb. Gels are loaded in the same order as the dendrogram illustrated in Figure 5-4.

CYTOCHROME P450 MONOOXYGENASE
(*tdtD*)

ACYL COENZYME A LIGASE
(*tdtL*)

ACYL COENZYME A DEHYDROGENASE
(*TdtB*)

ACYL COENZYME A THIOLASE
(*tdtC*)

DECARBOXYLASE/ISOMERASE
(*tdtA*)

A2a
A2b
A1a
A1b
BKME-9
A19-6a
D11-9
A19-6b
7104-1 A3
7104-1 A4
7104-1 D4
7104-1 A1
7104-1 A2
7057 J A6
7104-2 A3
AL-A8
7104-2 D4
16LW
NQ5

PUTATIVE REGULATOR
(*tdtR*)

TABLE 5-3. Summary of band size and intensity for probe positive strains illustrated in Figure 5-5. All band sizes for these *Sac* II digests are in kilobases. Signal intensity is relative to the A19-6a wild type control; (+) weak positive, (++) moderate positive (+++) strong positive.

PROBE STRAIN/GENE	LIGASE <i>tdtL</i>	REGULATOR <i>tdtR</i>	ISOMERASE <i>tdtA</i>	DEHY- DROGENASE <i>tdtB</i>	THIOLASE <i>tdtC</i>	MONO- OXYGENASE <i>tdtD</i>
A2a	9.8 +++	9.2 +++	9.2 +++	9.0 +++	10.0 +++	9.8 +++
A2b	9.8 +++	9.2 +++	9.2 +++	9.0 +++	10.0 +++	9.8 +++
A1a	9.8 +++	9.2 +++	9.2 +++	9.0 +++	10.0 +++	9.8 +++
A1b	9.8 +++	9.2 +++	9.2 +++	9.0 +++	10.0 +++	9.8 +++
BKME-9	9.8 +++	9.2 +++	9.2 +++	9.0 +++	10.0 +++	9.8 +++
A19-6a	9.8 +++	9.2 +++	9.2 +++	9.0 +++	10.0 +++	9.8 +++
D11-9	3.4, 2.1 +/+++	3.3 +	nil n/a	3.1 +++	3.2 +++	3.2, 2.6 ++/++
A19-6b	9.8 ++	9.2 ++	9.2 ++	9.0 ++	10.0 ++	9.8 ++
7104-1 A3	4.5, 1.2 +/+++	1.6 ++	nil n/a	4.3, 2.5 ++,+	4.3, 2.5 ++,+	4.6, 1.6 ++,+
7104-1 A4	4.5, 1.2 +/+++	1.6 ++	nil n/a	4.3, 2.5 ++,+	4.3, 2.5 ++,+	4.6, 1.6 ++,+
7104-1 D4	4.6, 1.2 +/+++	1.6 ++	nil n/a	4.3, 2.5 ++,+	4.3, 2.5 ++,+	4.6, 1.6 ++,+
7104-1 A1	9.3, 4.6 +/+++	8.6 +	8.9 +	8.7 ++	9.0 ++	9.5, 3.0 ++,+
7104-1 A2	9.3, 4.6 +/+++	nil n/a	nil n/a	8.7 +	9.0 +	9.5, 3.0 ++,+
7057J A6	1.7 ++	7.1 +	7.1 +	7.1 ++	7.3 +	7.5, 3.0 ++,+
7104-2 A3	9.8 +	nil n/a	nil n/a	9.3 +	9.3 +	4.1 +
AL-A8	4.3 +	nil n/a	nil n/a	2.8 +	4.6 ++	2.0 +
7104-2 D4	5.1 ++	4.7 +	nil n/a	4.7 ++	nil n/a	5.0, 3.0 ++,+
16LW	2.4 +	nil n/a	nil n/a	5.4, 1.2 +,++	nil n/a	4.8, 2.2 +++,+
NQ-5	4.3 ++	nil n/a	nil n/a	2.4 +	nil n/a	4.8, 2.2 +++/+

DISCUSSION

Specificity of *tdt* gene probes

Specificity testing of the *tdt* gene probes demonstrated that they were not taxonomically specific in that all type *Pseudomonas* spp. strains tested were negative for hybridization (Table C1, Appendix C). *Pseudomonas abietaniphila* strain BKME-9 that is a known resin acid degrader (Bicho *et al.*, 1995 and Mohn *et al.*, 1998) and *Pseudomonas* sp. D11-9 that was enriched on dehydroabietic acid from the same Avenor bleached kraft pulp mill mixed liquor sample as A19-6a (see Chapter 2) are both positive for all the *tdt* genes. The probes were able to detect homologues in strains from the β subgroup of the Proteobacteria (see Figure 5-3), suggesting that they have the appropriate level of specificity to be useful as tools for monitoring *in situ* resin acid degradative gene distribution. In addition, the negative result with the *tdtD*-P450 probe with *Streptomyces griseus* DNA that is known to have a P450 gene, indicates that this probe is specific enough to detect only those genes involved in resin acid degradation under high stringency hybridization conditions. This is further supported by the ligase gene probe result for BKME-9. It is known that *tdtL* is at least 78 percent homologous at the nucleotide sequence level to *orf1* of the *dit* operon from *Pseudomonas abietaniphila* (Mohn *et al.*, 1998). The *tdtL* probe showed an identical fragment size in A19-6a and BKME-9, which, considering the intensity of the hybridization band, is unlikely to be due to the *dit orf1* gene in BKME-9. In fact, the 96.6 percent similarity in 16S rDNA sequences between A19-6a and BKME-9 indicates that they are very closely related but still distinct species. A two

percent 16S rDNA sequence divergence is generally accepted as a species divide (Bennasas *et al.*, 1996; Moore *et al.*, 1996). The difference in band intensity (Figure 5-2) between these strains and A19-6b suggests that horizontal transfer is not a factor in the distribution of the *tdt* genotype, even though the *Sac* II restriction sites appear to have been conserved amongst these strains. Note that not all probe positive strains (16LW and NQ-5) had homologues to all *tdt* genes. This suggests that even though the ligase and P450 monooxygenase are required for abietic acid degradation the other genes may not be necessary or their function may be carried out by different enzymes. In the case of D11-9, similarity was high for the *tdtL* and *tdtR* genes and much lower for the others, which may be linked to congener preference since this isolate is most efficient at using pimelic acid as a sole carbon source.

Phenotypic characterization of probe positive strains

Carbon utilization profiles for the collection of probe positive isolates from the various mixed liquor samples reflected a phenotypic distribution similar to that seen for the Avenor bleached kraft pulp mill isolates described in Chapter 2 and other collections reported in the literature (Mohn, 1995). Here again two groups are seen with the smaller of the two having a preference for carbohydrate type substrates (see Figure 5-4 and Table C3, Appendix C). The presence of β and γ Proteobacteria in both groups is not uncommon when substrate utilization is used to determine strain relatedness. Some clustering of isolates from the same source is seen as in the case of the Brazil samples and many of the ones from Idaho (7104-1 A3, A4, and D4). Some differences between the Idaho

isolates are also observed as in the case of the Basin II isolates (7104-2 A3 and A4) which are well outside the cluster of Basin I isolates (see Figure 5-4). Perhaps this is a reflection of environmental differences between the sample sites which may favour the growth of different strains. It is well known that these factors can affect the observed microflora (Liss and Allen, 1992) and that different parts of the same treatment system can have significantly different environmental conditions (Zender *et al.*, 1994). The fact that the majority of the isolates from the collection are not positive for the *tdt* genes but are still able to use abietic acid as sole carbon source suggests that there are alternative resin acid degradative pathways. This is not unexpected since it is seen in other systems. For example, *ortho* and *meta* aromatic ring cleavage pathways are extended by the addition of peripheral genes that are substrate specific (van der Meer *et al.*, 1992). If the *tdt* genes are peripheral to a central pathway, then they would not detect different genes that are performing similar functions in strains that have evolved alternative metabolic routes. This hypothesis makes sense since horizontal transfer does not seem to be a factor in the occurrence of this genotype.

Distribution of the *tdt* genotype

In terms of *tdt* gene distribution phenotypic grouping is reflected in the probe positive results illustrated in Figure 5-5. The Aracruz mill isolates from Brazil (A2a, A2b, A1a, A1b) cluster with BKME-9 and have the same *tdt* genotype as A19-6a, including the degree of similarity as reflected by band intensity (see Table 5-3). The fact that they are poor abietic acid degraders may reflect feedstock variations since the Aracruz mill uses eucalyptus, a member of

the Myrtaceae (Dr. M. Canny, personal communication), which is not noted for having significant concentrations of abietane type resin acids (Langenheim, 1994). Since congener specificity in probe positive isolates is known (eg., D11-9's preference for pimaric acid) the similarity to A19-6a may be due to strain relatedness. This is supported by the fact that the other isolates that are equally able to degrade abietic acid, but are phenotypically more distant, have *tdt* gene arrangements that bear little resemblance to the wild type strain, including which genes are present and their degree of similarity. Conversely, if the *tdt* genes are congener specific then the conserved gene arrangement in A19-6b may be a reflection of abietic acid enrichment, and the reduced signal intensity may then be due to taxonomic differences alone. If the latter is the case, then these probes may be too specific to monitor overall resin acid degradation levels.

For the Idaho isolates, phenotypic clustering is reflected in the probe positive results. The two Basin I clusters (7104-1 A3, A4, D4 and A1, A2, respectively) have different *tdt* gene arrangements but isolates within each cluster have the same arrangement. The 91 percent phenotypic similarity between 7104-1 A1 and A2 may be insufficient to treat these strains as siblings, since A19-6a and BKME-9 are distinct strains that are 90 percent similar at the phenotypic level. The Basin II isolates from Idaho and those from all the other samples have different probe positive results including which of the *tdt* genes are present, the location of *Sac* II restriction sites, and the degree of similarity between the probes and target DNA (see Figure 5-5 and Table 5-3). All of these isolates are well removed phenotypically from the other strains and have varying abilities to use abietic acid as sole carbon source.

CONCLUSION

Based on these results the *tdtL* and *tdtD* gene probes for the CoA ligase and cytochrome P450 monooxygenase, respectively, appear to be the best candidates for use as genetics based tools for monitoring resin acid degrading populations *in situ*. This, because these probes are able to detect homologues in the greatest number of strains tested. Homologues for the other genes were less common and, since their role in abietic acid degradation has not been determined, they may lack the specificity necessary for use as catabolic gene probes.

CHAPTER 6

Discussion, conclusions, and future work

GENERAL

My research hypothesis was that resin acid degrading bacteria could be isolated from environments where these compounds are found in abundance. These same microorganisms could be used to develop tools for determining the contribution of specific populations to resin acid biodegradation in existing pulp mill mixed liquor treatment systems and receiving waters. Overall, the results presented in this thesis have borne out this hypothesis and most of my original research objectives were met.

Objective I

The first objective was the isolation and characterization of resin acid degrading bacteria from an mixed liquor sample taken from an aerated lagoon secondary treatment system. Chapter 2 describes the results of an abietic acid and dehydroabietic acid enrichment protocol that was successful in isolating 13 bacteria able to degrade various resin acid congeners in addition to the ones used for enrichment isolation. Characterization of these isolates using standard biochemical tests, phenotypic analysis, and genotypic fingerprinting was useful in identifying siblings and determining which strains would be most suitable for use in meeting my second objective.

Objective II

The second objective was the localization and identification of resin acid catabolic genes that could be used to develop DNA probes. The results of this basic research are described in two chapters that include the molecular aspects of gene localization (Chapter 3) and the sequence analysis of targeted genes (Chapter 4). The end result was the cloning of a *tdt* locus comprising seven open reading frames from which probes were developed for use in meeting the final objective of my research. Six of these were identified by similarity search and two of these genes were shown to be necessary for abietic acid degradation.

Objective III

Since the ultimate application of my research is to use the *tdt* probes as monitoring tools to assess bacterial population dynamics during resin acid biodegradation *in situ*, it was necessary to first confirm their usefulness for such an application. This was my third objective, that of necessity became limited in scope to the study of the distribution of the cloned genes in bacteria isolated from secondary treatment facilities at pulp mills located world wide. These results are described in chapter 5 and suggest that the probes for the *tdtD* and *tdtL* genes encoding a cytochrome P450 monooxygenase and a coenzyme A ligase, respectively, are the most suitable for use as potential monitoring tools.

Having discussed the individual results and conclusions in each chapter it would be redundant to repeat them here. Nevertheless, there are aspects of the overall results that have emerged from compiling the data and reviewing the

literature that pique one's interest and generate questions for future research. It is to these aspects that I now turn my attention.

DISCUSSION

Collection of resin acid degrading bacteria

In the characterization of my Avenor bleached kraft pulp mill isolate collection I tested each strain's ability to use six of the most common resin acid congeners found in pulp mill mixed liquors. Other researchers have reported the isolation of resin acid degrading bacteria from mixed liquor treatment facilities that show a preference for abietanes (Mohn, 1995 and Bicho *et al.*, 1995) or pimaranes (Wilson *et al.*, 1996), generally with little utilization of congeners from the other structural type. In contrast, my isolates have a relatively broad congener range, see Figure 2-3, and many are equally able to use both abietane and pimarane types. The most significant difference between my enrichment and maintenance protocols and that used by others is the purity of resin acids provided as a substrate. The bacteria isolated by Mohn's group were enriched and maintained on high purity congeners ($\geq 90\%$) whereas mine were enriched on lab grade (70 % pure) resin acids that contained various unidentified congeners or other compounds. The high purity congeners were used only for the substrate utilization experiment (Figure 2-3). Oddly enough, most of the isolates seem to prefer pimaranes even though they were enriched on abietic or dehydroabietic acid. This suggests that low purity enrichment may give a slightly

better picture of the culturable bacteria involved in the biodegradation of complex mixtures of organics. It also explains why several different isolates were maintained in the enrichment consortia since congener niches were still available to increase species diversity.

Since the goal of isolation and characterization was to find the best candidate (A19-6a) for catabolic gene probe development the remaining collection lies dormant but may be useful in future studies. For instance, why does A19-6b that was difficult to culture separately from A19-6a grow more slowly yet appear to be a better degrader of resin acids as a whole? Why does it harbour all of the same genes in what appears to be the same genetic arrangement as A19-6a but not have the same growth characteristics as A19-6a? Does the reduced similarity suggested by band hybridization intensity (see Figures 5-2 and 5-5) indicate that A19-6b contains variants of the *tdt* genes that express different metabolic capabilities? What is the difference between the abietic acid isolates A19-1 and A19-2 that makes the former prefer isopimaric acid and the latter pimaric acid when phenotypic analysis suggested they are siblings? How did A19-5 survive in the enrichment when it is a poor resin acid degrader overall? Why do all of the dehydroabietic acid enriched isolates show a preference for pimarane type congeners? Obviously there is a wealth of questions to be answered in the further characterization of this collection, although the practicality of such research needs to be considered in the design of subsequent studies.

Gene localization by transposon mutagenesis

Transposon mutagenesis has long been a popular method of physically localizing target DNA (de Bruijn and Lupski, 1984) and in a manner of speaking was successful in finding genes involved in abietic acid degradation by *Pseudomonas* sp. A19-6a (Chapter 3). While in theory this form of illegitimate recombination is supposed to result in the loss of the suicide vector, this was not the case here; a result that is echoed in the literature (Hofer *et al.*, 1996). Why did the suicide vector incorporate with the transposon in a large fraction of the mutants? Was this simply an artifact of sample size? Does this type of illegitimate recombination improve the stability of the mutation over normal transposition events, as implied by Hofer's results, and if so why? Did it play a role in the 10-fold difference in mutation rate observed for A19-6a and D11-9? If so, is this an event peculiar to *Pseudomonads* which Hofer and colleagues were also studying? Since D11-9 is amenable to mutation, prefers pimaranes to abietanes, and has homologues to the *tdt* genes it might be useful to do some parallel catabolic gene probe development with this isolate.

With respect to the A19-6a transconjugants A103-37 and A104-38 why were they able to use abietic acid when grown on solid media when they could not in liquid culture? How was the S1 sub-clone from A76-47 actually generated and why is X1 approximately a kilobase smaller than expected? While most of these questions are purely rhetorical they serve to underline how little we know about the genetics of environmental isolates. They also illustrate how assumptions can obscure results by pointing us in false directions. For example, had I not assumed vector loss as a result of transposition I could have

prescreened my transconjugant collection for legitimate transposition mutants and possibly had more reproducible inverse PCR results. In addition, had I replicated my solid media results at the same time as the batch culture analysis of my transconjugant collection I could have eliminated non-target mutants sooner.

Resin acid biodegradation genes

Compilation of the sequence generated by the A19-6a wild type and A76-47 mutant sub-clones (Chapter 4) demonstrated that a single contiguous region of DNA 9180 basepairs in length had been isolated. Included on this segment of DNA are seven open reading frames that have been called *tdt* genes (for tricyclic diterpene metabolism). Two of them, *tdtD* and *tdtL*, are involved in the degradation of abietic acid by *Pseudomonas* sp. A19-6a. This is an interesting result since similar genes are involved in the biosynthesis of this resin acid and in the elimination of resin acids by fish. More specifically, *tdtD* codes for a cytochrome P450 monooxygenase that, based on HPLC analysis of knock-out mutants, appears to be involved in abietic acid degradation but not dehydroabietic acid degradation. The only structural difference between these resin acid congeners is the degree of saturation of the C-ring. Since abiotic dehydrogenation of abietic acid results in the formation of dehydroabietic acid, and P450 monooxygenases are known to catalyze this type of reaction (Geungerich, 1991), perhaps dehydrogenation of ABA to produce an aromatic is a function of TdtD. The *qui* genes of *A. calcoaceticus* illustrate that aromatization prior to degradation is a way to extend the substrate specificity of central pathways. If the *tdt* genes are peripheral to a central resin acid catabolic

pathway, then perhaps ABA is converted to DHA prior to degradation. This metabolic route has been suggested for abietic acid degradation (Zender *et al.*, 1994). In addition, Martin and Mohn (1999) have shown that abietic acid cannot be degraded by BKME-9 mutants that lack aromatic ring hydroxylation activity. In fact, because their isolate is so closely related taxonomically to A19-6a and cross hybridizes with all of the *tdt* probes, A19-6a may also harbour homologues to some or all of the *dit* genes. It would be very interesting indeed to probe A19-6a and BKME-9 to compare the presence and genetic arrangement of all the *tdt* and *dit* open reading frames, and to determine if these loci are genetically linked.

The fact that a monooxygenase was found to be involved in abietic acid degradation is not a surprising result since their presence has been implicated by oxygenated metabolites generated during resin acid degradation (see Liss *et al.*, 1997 for a review and Mohn *et al.*, 1998a). Since P450 monooxygenases represent a superfamily of enzymes involved in a wide range of metabolic activities including transformation or biodegradation of organics and the biosynthesis of vitamins, steroids (which are also terpenoids), and corticoids (Harayama *et al.*, 1992) it is understandable that they can be involved in resin acid biosynthesis, biodegradation, and biodegradation. The fact that they are found performing these functions in widely different taxa, i.e. plants, bacteria, and fish, respectively, is evidence to their ancient ancestry which purports a single common progenitor dating back 3.5 billion years (Nelson *et al.*, 1993 and Gotoh, 1993).

The *tdtL* gene that encodes a coenzyme A ligase that is essential to abietic acid degradation in A19-6a is also found in BKME-9. The results in Chapter 5 show an homologous DNA fragment and the sequence data of Martin and Mohn (1999) show that the *dit* locus *orf1* has some similarity at both the amino acid and nucleotide levels. While this gene is unlikely to represent a physical overlap between the *tdt* and *dit* loci perhaps it represents an homologous one. The *tdtL* and *dit orf1* genes may code for resin acid specific CoA ligases. Since BKME-9 contains all the *tdt* genes in what appears to be the same genetic arrangement as in A19-6a, perhaps my strain has homologues to the *dit* genes on the other side of *tdtL*. If it does and they are all arranged as in BKME-9 then how did this occur when both species are distinct and geographically isolated from one another?

Like the P450, the presence of a Co-A ligase in abietic acid degradation brings the concept of enzyme structural conservation full circle in that these enzymes also function in isoprenoid synthesis and are commonly involved in the production of plant secondary metabolites. In fact, this was the reason for including the aspen ligase (*Pt4CL1*) in the multiple sequence alignment. Aspen is the primary feedstock used by the Alpac Alberta mill from which AL-A8 was isolated; a *tdtL*⁺ strain. Similar structural conservation can be seen for the *tdtA* encoded thiolase, an enzyme that catalyzes the first step of the isoprenoid pathway, the last step of many catabolic pathways, and is ubiquitous in nature. These results illustrate how biological systems can be adapted to meet new evolutionary challenges thereby generating the bewildering array of diversity that arises from the sequential recombination of four simple molecules.

Resin acid catabolic gene probes

In curbing the scope of the thesis due to time constraints, catabolic gene probes were not validated *in situ*. The suitability of the *tdt* probes as tools for monitoring resin acid degrading bacteria is still in question. Nevertheless, some factors of interest did emerge from the limited isolate distribution study I did undertake. For instance, homologues of these genes, and in particular the P450 and the ligase genes, exist in culturable bacteria in mixed liquor treatment systems from as far afield as Japan, Brazil, Idaho, Alberta, and British Columbia. Recall that the probes originated from a northern Ontario bacterium. Some of these isolates appear to have a genetic arrangement similar to A19-6a whereas others have obvious differences in terms of restriction enzyme fragmentation patterns and/or band hybridization intensities. While it would serve little purpose to do any further analysis on these isolates, the results do suggest that homologues to the *tdt* genes may be found in pulp mill mixed liquor secondary treatment systems world wide. This may be an indication that, at the very least, the *tdtD* and *tdtL* probes may be useful for *in situ* monitoring of resin acid biodegrading populations. This is further supported by the presence of homologues in degraders that prefer other congeners; although a direct correlation between degradation and gene presence has yet to be determined.

The presence of probe positive isolates in environments where different structural types of resin acids may predominate begs the question of how broad is the substrate range of these enzymes? Most researchers find that structural

differences between the abietane and pimarane type resin acids appear to affect an isolate's ability to use a given congener. My results suggest that broad congener specificity is seen and that similar genes may be involved. Could the CoA ligase be providing the recognition handle necessary for enzymes in a pathway to identify similar substrates that may otherwise be overlooked? What of the fact that I use abietic acid as a preferred model substrate rather than dehydroabietic acid as most researchers do? The degree of saturation of other congeners is more similar to ABA than to DHA, so that the use of the saturated substrate ABA for enrichment may have had an influence on the substrate specificity of the resulting isolate collection.

CONCLUSIONS

Having drawn individual conclusions relevant to each objective in the appropriate chapter, I find that there is little more that I can add here. Cloning and characterization of the *tdt* genes and the study regarding their distribution in pulp mill mixed liquor treatment systems barely scratches the surface of their ecological significance. Even though the results of my research supported my hypothesis, they have also asked more questions than they have answered. This said is it not in all honesty the true purpose of scientific research? Good experimental design demands that the hypothesis upon which it is based be flexible enough to deal with the unexpected results that research often generates. Now is the time to rethink my original hypothesis to include the

results described herein so that further research into the genetics and ecology of resin acid biodegradation can begin.

FUTURE WORK

When I think of future work two principle areas of research that are of immediate interest emerge from my results. Many of the questions discussed above are of interest, but are less important than the need to increase our understanding of the genetics and ecology of the *tdt* genes.

In terms of the genetics of the *tdt* genes there is much work to be done to determine their role in resin acid degradation. Even though *tdtD* and *tdtL* appear to be involved in abietic acid degradation, this has only been inferred by mutation and biochemical analysis is still necessary to confirm their actual function. The knock-out mutants P450-R1 and P450-R3 could be useful in this regard. The need for the other open reading frames in resin acid catabolism that lie between *tdtD* and *tdtL* is also unknown. Here again mutation and/or biochemical analysis are necessary.

What of the relationship between the *tdt* and *dit* genes of A19-6a and BKME-9, respectively? I consider this investigation to be very important since it could significantly increase our understanding of the genetics involved in bacterial resin acid degradation.

Future work on the ecological significance of the *tdt* genes deals with the validation of these probes for use as monitoring tools and the actual application of them to *in situ* studies. Validation can be achieved by mixed culture analysis under variable environmental conditions to determine whether changes in degradation levels correlate with changes in hybridization data. This will provide an understanding of the population dynamics of bacteria involved in resin acid degradation, in addition to confirming a relationship between the *tdt* genes and mixed liquor detoxification. It will also be necessary to determine whether these probes can detect homologues in the non-culturable fraction of bacterial communities, since these microorganisms could contribute as much or more to resin acid biodegradation levels as the culturable populations studied here. Application analysis is also necessary to confirm their suitability for *in situ* monitoring. A two year study of the *tdt* genes in the Avenor bleached kraft pulp mill secondary treatment system would be ideal for this and could be conducted concurrently with the mixed culture analysis. In both cases both molecular and chemical analyses would be necessary.

REFERENCES

- Allan, L., E. K. Kaufman, and J. Underwood. 1972. Pollution in the pulp and paper industry. The Council on Economic Priorities, New York. pp. 504.
- Alloway, B. J. and D. C. Ayres. 1993. Chemical principles of environmental pollution. First Edition. Chapman and Hall, Glasgow. Pp. 291.
- Altenschmidt, U., B. Oswald, E. Steiner, H. Herrmann, and G. Fuchs. 1993. New aerobic benzoate oxidation pathway via benzoyl-coenzyme A and a 3-hydroxybenzoyl-coenzyme A in a denitrifying *Pseudomonas* sp. J. Bacteriol. 175(15): 4851-4858.
- Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids Res. 25(17): 3389-3402.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215: 403-410.
- Asperger, O. and H.-P. Kleber. 1991. Distribution and diversity of bacterial cytochromes P-450. In: Microbial and plant cytochromes: biochemical characteristics, genetic engineering, and practical implications. Frontiers in Biotransformation. Vol. 4. K. Ruckpaul and H. Rein (Eds.).
- Atlas, R.M. and A.K. Bej. 1994. Polymerase chain reaction. In: Methods for General and Molecular Bacteriology. (P. Gerhardt, R.G.E. Murray, W.A. Wood, and N.R. Krieg Eds.). American Society for Microbiology, Washington. pp. 418-435.
- Babbitt, P. C., G. L. Kenyon, B. M. Martin, H. Charest, and M. Sylvestre. 1992. Ancestry of the 4-chlorobenzoate dehalogenase: analysis of amino acid sequence identities among families of acyl-adenyl ligases, enoyl-CoA hydratases/isomerases, and acyl-CoA thioesterases. Biochem. 31: 5594-5604.
- Bagdasarian, M. M., E. Amann, R. Lurz, B. Ruckert, and M. Bagdasarian. 1983. Activity of the hybrid *trp-lac(tac)* promotor of *Escherichia coli* in *Pseudomonas putida*. Construction of broad-host-range, controlled-expression vectors. Gene. 26: 273-282.
- Bartha, R. 1990. Isolation of microorganisms that metabolize xenobiotic compounds. In: Isolation of Biotechnological Organisms from Nature. D. P. Labeda (Ed.). McGraw Hill Publ., Co., New York. pp. 283-307.

Bauer, E. 1992. Production, detection, and treatment of resin acids in the pulp and paper industry. Unpublished report. Faculty of Forestry, University of British Columbia. 68 pp.

Beak Consultants. 1987. Non-conventional contaminants in pulp and paper mixed liquors. Prepared for the Ontario Forest Industries Association. Pp106.

Beck, E., G. Ludwig, E. A. Auerswald, B. Reiss, and H. Schaller. 1982. Nucleotide sequence and exact location of the neomycin phosphotransferase gene from transposon Tn5. *Gene*. 19: 327-336.

Bennasas, A., R. Rosello-Mora, J. Lalucat, and E.R.B. Moore. 1996. 16S rRNA gene sequence analysis relative to genomovars of *Pseudomonas stutzeri* and proposal of *Pseudomonas balearica* sp. nov. *Int. J. Syst. Bacteriol.* 46: 200-206.

Bennett, D. H., D. Reeser, A. G. Campbell, and C. M. Fraser. 1988. Pre-treatment of CTMP mixed liquor by lime to reduce resin acids, sulfite, and toxicity. In: Proceedings of the 1988 TAPPI Environmental Conference. Pp. 199-207.

Bergey. 1984. *Bergey's Manual of Systematic Bacteriology*. Williams and Wilkens Baltimore/London.

Bicho, P. A., V. Martin, and J. N. Saddler. 1995. Growth, induction, and substrate specificity of dehydroabietic acid-degrading bacteria isolated from a kraft mill mixed liquor enrichment. *Appl. Env. Microbiol.* 61(9): 3245-3250.

Biellman, J. F., R. Wennig, P. Daste, and M. Raynaud. 1968. Microbial degradation of diterpenic acids. *Chem. Comm.* 1968, 168-169.

Biellmann, J.F., G. Branlant, M. Gero-Robert, and M. Poiret. 1973a. Dégredation bactérienne de l'acide déhydroabiétique par *Flavobacterium resinovorum*. *Tetrahedron* 29: 1227-1236.

Biellmann, J.F., G. Branlant, M. Gero-Robert, and M. Poiret. 1973b. Dégredation bactérienne de l'acide déhydroabiétique par un *Pseudomonas* et une *Alcaligenes*. *Tetrahedron* 29:1237-1241.

Bjorkman, C. and R. Gref. 1993. Survival of pine sawflies in cocoon stage in relation to resin acid content of larval food. *J. Chem. Ecol.* 19(12): 2881-2890.

Black, P. N., C. C. DiRusso, A. K. Metzger, and T. L. Heimert. 1992. Cloning, sequencing, and expression of *fadD* gene of *Escherichia coli* encoding acyl coenzyme A synthetase. *J. Biol. Chem.* 267(35): 25513-25520.

- Bonsor, N., N. McCubbin, and J. B. Sprague. 1988. Municipal Industrial Strategy for Abatement (MISA): Kraft mill mixed liquors in Ontario. Queen's Printer for Ontario, Toronto. Pp: 260.
- Brownlee, B., M. E. Fox, W. M. J. Strachan, and S. R. Joshi. 1977. Distribution of dehydroabietic acid in sediments adjacent to a kraft pulp and paper mill. J. Fish. Res. Board Can. 34: 838-843.
- Bruce, K.D., W.D. Hiorns, J.L. Hobman, A.M. Osborn, P. Strike, and D.A. Ritchie. 1992. Amplification of DNA from native populations of soil bacteria by using the polymerase chain reaction. Appl. Env. Microbiol. 58(10): 3413-3416.
- Bruce, M. E. and T. A. Ckark. 1994. *Klebsiella* and nitrogen fixation in pulp and paper mill mixed liquors and treatment systems. Appita. 47(3): 231-237.
- Brush, T.S., R.L. Farrell, and C. Ho. 1994. Biodegradation of wood extractives from southern yellow pine by *Ophiostoma piliferum*. TAPPI Journal. 77(1): 155-159.
- Burggraaf, S., A. G. Langdon, A. L. Wilkins, and D. S. Roper. 1996. Accumulation and depuration of resin acids and fichtelbergite by the freshwater mussel *Hyridella menziesi*. Env. Tox. Chem. 15(3): 369-375.
- Button, D.K. 1984. Evidence for a terpene-based food chain in the Gulf of Alaska. Appl. Env. Microbiol. 48(5): 1004-1011.
- Button, D.K., B.R. Robertson, and K.S. Craig. 1981. Dissolved hydrocarbons and related microflora in a fjordal seaport: sources, sinks, concentrations, and Kinetics. Appl. Env. Microbiol. 42(4): 708-719.
- Cheung, H. T. A., S.-L. Fu, and M. A. Smal. 1994. Inhibition of platelet aggregation by diterpene acids from *Pinus massonai* resin. Arzneimittel-Forschung/Drug Res. 44(1): 17-25.
- Clark, A. J. and G. J. Warren. 1979. Conjugal transmission of plasmids. Ann. Rev. Genet. 13: 99-125.
- Cobb, B. D. and J. M. Clarkson. 1994. A simple procedure for optimising the polymerase chain reaction (PCR) using modified Taguchi methods. Nucleic Acids Res. 18: 3801-3805.
- Collins, F. S. and S. M. Weissman. 1984. Directional cloning of DNA fragments at a large distance from an initial probe: a circularization method. Proc. Natl. Acad. Sci. 81: 6812-6816.

- Crameri, A., E. A. Whitehorn, E. Tate, and W. P. C. Stemmer. 1996. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat. Biotechnol.* 14: 315-319.
- Cross, B.E. and P.L. Myers. 1968. The bacterial transformation of abietic acid. *Biochem. J.* 108:303-310.
- Croteau, R. and M.A. Johnson. 1985. Biosynthesis of terpenoid wood extractives. In: *Biosynthesis and Biodegradation of Wood Components*. (T. Higuchi Ed.). Academic Press, Inc., Orlando. pp. 379-439.
- de Bruijn, F. J. and J. R. Lupski. 1984. The use of transposon Tn5 mutagenesis in the rapid generation of correlated physical and genetic maps of DNA segments cloned into multicopy plasmids – a review. *Gene*. 27: 131-149.
- de Bruijn, F.J. 1992. Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergenic consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. *Appl. Env. Microbiol.* 58(7): 2180-2187.
- de Lorenzo, V. and K. N. Timmis. 1994. Analysis and construction of stable phenotypes in Gram-negative bacteria with Tn5- and Tn10-derived minitransposons. *Methods in Enz.* 235: 386-405.
- Deho, G., G. Bertoni, and A. Polissi. 1992. Bacteriophage P4-derived shuttle vectors for cloning and transposon mutagenesis in *Pseudomonas putida* and other Gram-negative bacteria. In: *Pseudomonas: Molecular Biology and Biotechnology*. E. Galli, S. Silver, and B. Witholt. (Eds.). American Society for Microbiology, Washington. 443 pages.
- Dethlefs, F., K. O. Gerhardt, and H.-J. Stan. 1996. Gas chromatography/mass spectrometry of 13 resin acids as their PFB esters. *J. Mass Spec.* 31: 1163-1168.
- Di Gioia, D., M. Peel, F. Fava, and R. C. Wyndham. 1998. Structures of homologous composite transposons carrying *cbaABC* genes from Europe to North America. *Appl. Env. Microbiol.* 64(5): 1940-1946.
- DiMarco, A. A., B. A. Averhoff, E. E. Kim, and L. N. Ornston. Evolutionary divergence of *pobA*, the structural gene encoding *p*-hydroxybenzoate hydroxylase in an *Acinetobacter calcoaceticus* strain well-suited for genetic analysis. *Gene*. 125:25-33.

DiMarco, A. A., B. Averhoff, and L. N. Ornston. Identification of the transcriptional activator *pobR* and characterization of its role in the expression of *pobA*, the structural gene for *p*-hydroxybenzoate hydroxylase *Acinetobacter calcoaceticus*. *J. Bact.* 175(14): 4499-4506.

DiRusso, C. C. 1988. Nucleotide sequence of the *fadR* gene, a multifunctional regulator of fatty acid metabolism in *Escherichia coli*. *Nucleic Acids Res.* 16(16): 7995-8009.

Douglas, G. R., E. R. Nestmann, J. L. Betts, J. C. Mueller, E. G.-H. Lee, H. F. Stich, R. H. C. San, R. J. P. Brouzes, A. L. Chmelauskas, H. D. Paavila, and C. C. Walden. 1980. Mutagenic activity in pulp mill mixed liquors. In: *Water chlorination-environmental impact and health effects*. Vol. III. R. L. Jolley, W. A. Brungs, R. B. Cummings, and V. A. Jacobs (Eds.). Ann Arbor Science, Ann Arbor, Michigan. Pp. 865-880.

Duran, N., E. Esposito, L. H. Innocentini-Mei, and V. P. Canhos. 1994. A new alternative for kraft E1 mixed liquor treatment. A combination of photochemical and biological methods. *Biodegradation* 5:13-19.

Ebling, G. 1931. Results of chemical and bioassay tests on pulp mill mixed liquors. *Vom Wasser.* 5: 192-200.

Egland, P. G., D. A. Pelletier, M. Dispensa, J. Gibson, and C. S. Harwood. 1997. A cluster of bacterial genes for anaerobic benzene ring biodegradation. *Proc. Natl. Acad. Sci.* 94: 6484-6489.

Egland, P. G., J. Gibson, and C. S. Harwood. 1995. Benzoate-coenzyme A ligase, encoded by *badA*, is one of three ligases able to catalyze benzoyl-coenzyme A formation during anaerobic growth of *Phodopseudomonas palustris* on benzoate. *J. Bacteriol.* 177(22): 6545-6551.

Eichler, K., F. Bourgis, A. Buchet, H.-P. Kleber, and M.-A. Mandrand-Bertholet. 1994. Molecular characterization of the *cai* operon necessary for carnitine metabolism in *Escherichia coli*. *Mol. Microbiol.* 13(5): 775-786.

Ekman, R. and R. Sjöholm. 1979. Hydroxylation of dehydroabiatic acid by *Fomes annosus*. *Acta Chemica Scandinavica* 33:76-78

Ellis, B. E. 1997. Metabolism of defence and communication. In: *Plant Metabolism*. D. T. Dennis, D. H. Turpin, D. D. Lefebvre, and D. B. Layzell (Eds.) Addison, Wesley Longman Ltd., Edinburgh Gate, Harlow, Essex. pp. 148-160.

- Elsmore, D.A. and L.N. Ornston. 1994. The *pca-pob* supraoperonic cluster of *Acinetobacter calcoaceticus* contains *quiA*, the structural gene for quinate-shikimate dehydrogenase. J. Bac. 176(24): 7659-7666.
- Erb, R.W. and I. Wagner-Dobler. 1993. Detection of polychlorinated biphenyl degradation genes in polluted sediments by direct DNA extraction and polymerase chain reaction. Appl. Env. Microbiol. 59(12): 4065-4073.
- Farrell, R.L., R.A. Blanchette, T.S. Brush, Y. Hadar, S. Iverson, K. Krisa, P.A. Wendler, and W. Zimmerman. 1993. Cartapip™: a biopulping product for control of pitch and resin acid problems in pulp mills. J. Biotech. 30: 115-122.
- Feliciano, A. S., M. Gordaliza, M.A. Salinero, and J.M. Miguel del Corral. 1993. Abietane acids: sources, biological activities, and therapeutic uses. Planta Med. 59:485-490.
- Ferguson, M. L., M. R. Servos, K. R. Munkittrick, and J. Parrott. 1992. Inability of resin acid exposure to elevate EROD activity in rainbow trout (*Oncorhynchus mykiss*). Water Poll. Res. J. Canada. 27(3): 561-573.
- Fernandez-Valverde, M., A. Reglero, H. Martinez-Blanco, and J. M. Luengo. 1993. Purification of *Pseudomonas putida* acyl coenzyme A ligase active with a range of aliphatic and aromatic substrates. Appl. Env. Microbiol. 59(4): 1149-1154.
- Fox, M. E. 1977. Persistence of dissolved organic compounds in kraft pulp and paper mill mixed liquor plumes. J. Fish. Res. Board Can. 34:798-804.
- Fulda, M., E. Heinz, and F. P. Wolter. 1994. The *fadD* gene of *Escherichia coli* K12 is located close to the *md* at 39.6 min of chromosomal map and is a new member of the AMP-binding protein family. Mol. Gen. Genet. 242: 241-249.
- Fulthorpe, R. R., S. N. Liss, and D. G. Allen. 1993. Characterization of bacteria isolated from a bleached kraft pulp mill wastewater treatment system. Can. J. Microbiol. 39: 13-24.
- Fulthorpe, R.R. and D.G. Allen. 1994. Evaluation of Biolog MT plates for aromatic and chloroaromatic substrate utilization tests. Can. J. Microbiol. 40: 1067-1071.
- Fulthorpe, R.R. and R.C. Wyndham. 1989. Survival and activity of a 3-chlorobenzoate-catabolic genotype in a natural system. Appl. Env. Microbiol. 55: 1584-1590.

Funk, C. and R. Croteau. 1994. Diterpenoid resin acid biosynthesis in conifers: characterization of two cytochrome P450-dependent monooxygenases and an aldehyde dehydrogenase involved in abietic acid biosynthesis. *Arch. Biochem. Biophysics*. 308(1): 258-266.

Gäfvert E. 1994. Allergenic components in modified and unmodified rosin: chemical characterization and studies of allergenic activity. Scandinavian University Press, Oslo. 36 pp.

Galas, D.J. and M. Chandler. 1989. Bacterial insertion sequences. In: *Mobile DNA*. (D.E. Berg and M.M. Howe Eds.). American society of Microbiologists, Washington. pp. 109-162.

Gerischer, U., A. Segura, and L. N. Ornston. 1998. PcaU, a transcriptional activator of genes for protocatechuate utilization in *Acinetobacter*. *J. Bact.* 180(6): 1512-1524.

Gershenzon, J. 1994. Metabolic costs of terpenoid accumulation in higher plants. *J. Chem. Ecol.* 20(6): 1281-1328.

Gershenzon, J. and R. Croteau. 1991. Terpenoids. In: *Herbivores, Their Interactions with Secondary Metabolites, Vol 1, The Chemical Participants*. G. A. Rosenthal and M. R. Berenbaum (Eds.). Academic Press, New York. pp. 165-219.

Gijzen, M., E. Lewinsohn, T.J. Savage, and R.B. Croteau. 1993. Conifer monoterpenes: Biochemistry and bark beetle chemical ecology. In: *Bioactive Volatile Compounds from Plants*. American Chemical Society. pp. 8-22.

Gotoh, O. 1993. Evolution and differentiation of P-450 genes. In: *Cytochrome P450. Second Edition*. T. Omura, Y. Ishmura, and Y. Fujii-Kuriyama (Eds.).

Government of Canada. 1990. Canada's Green Plan: for a healthy environment. Department of Environment. Supply and Services Canada, Canada. 174 pp.

Graham-Lorence, S. and J. A. Peterson. 1996. Structural alignments of P450s and extrapolations to the unknown. *Methods Enz.* 272: 315-326.

Graves, J. W. and T. W. Joyce. 1994. A critical review of the ability of biological treatment systems to remove chlorinated organics discharged by the paper industry. *Water SA.* 20(2): 155-160.

Greer, C., L. Masson, Y. Comeau, R. Brousseau, and R. Samson. 1993. Application of molecular biology techniques for isolating and monitoring pollutant-degrading bacteria. *Water Poll. Res. J. Canada*. 28(2): 275-287.

Gref, R. 1988. Gas chromatographic analysis of underivatized resin acids. *J. Chromat.* 448: 428-432.

Gross, G. G. 1985. Biosynthesis and metabolism of phenolic acids and monolignols. In: *Biosynthesis and biodegradation of wood components*. T. Higuchi (Ed). Academic Press Inc., Orlando. pp. 229-271.

Guengerich, F. P. 1991. Reactions and significance of cytochrome P-450 enzymes. *J. Biol. Chem.* 266(16): 10019-10022.

Hagman, N. 1936. Resin acids in fish mortality. *Paperi Puu*. 18: 32-41.

Hahlbrock, K. and H. Grisebach. 1979. Enzymatic controls in the biosynthesis of lignin and flavonoids. *Ann. Rev. Plant Physiol.* 30: 105-130.

Hanson, J. R. 1972. Introduction and nomenclature. In: *Chemistry of Terpenes and Terpenoids*. A. A. Newman (Ed.). Academic Press Inc. (London) Ltd.

Hanson, J. R. 1972. The Di- and Sesterterpenes. Part 1: The diterpenes. In: *Chemistry of terpenes and terpenoids*. A. A. Newman (Ed.). Academic Press Inc. (London) Ltd. pp: 155-199.

Harayama, S., M. Kok, and E. L. Neidle. 1992. Functional and evolutionary relationships among diverse oxygenases. *Ann. Rev. Microbiol.* 46: 565-601.

Harwood, C. S. and R. E. Parales. 1996. The β -ketoadipate pathway and the biology of self-identity. *Ann. Rev. Microbiol.* 50: 553-590.

Harwood, C. S., N. N. Nichols, M.-K. Kim, J. L. Ditty, and R. E. Parales. 1994. Identification of the *pcaRKF* gene cluster from *Pseudomonas putida*: involvement in chemotaxis, biodegradation, and transport of hydroxybenzoate. *J. Bact.* 176(21): 6479-6488.

Hemingway, R. W. and H. Greaves. 1973. Biodegradation of resin acid sodium salts. *TAPPI J.* 56(12): 189-192.

Herrick, J.B., E.L. Madsen, C.A. Batt, and W.C. Ghiorse. 1993. Polymerase chain reaction amplification of naphthalene-catabolic and 16S rRNA gene sequences from indigenous sediment bacteria. *Appl. Env. Microbiol.* 59(3): 687-694.

- Hewitt, M. L., J. H. Carey, D. G. Dixon, and K. R. Munkittrick. 1996. Examination of bleached kraft mill mixed liquor fractions for potential inducers of mixed function oxygenase activity in rainbow trout. In: Environmental Fate and Effects of Pulp and Paper Mill Mixed liquors. M. R. Servos, K. R. Munkittrick, J. H. Carey, and G J. Van Der Kraak (Eds.). pp. 79-94.
- Hickey, C. W. and M. L. Martin. 1995. Relative sensitivity of five benthic invertebrate species to reference toxicants and resin-acid contaminated sediments. *Env. Tox. Chem.* 14(8): 1401-1409.
- Hodson, P. V. 1996. Mixed function oxidase induction by pulp mill mixed liquors: advances since 1991. In: Environmental Fate and Effects of Pulp and Paper Mill Mixed liquors. M. R. Servos, K. R. Munkittrick, J. H. Carey, and G J. Van Der Kraak (Eds.). pp. 349-358.
- Hofer, B., R. Blasco, M. Megharaj, M. Seeger, D. McKay, R.-M. Wittich, D. H. Pieper, and K. N. Timmis. 1996. Pseudomonads designed for bioremediation and circumvention of undesirable routes. In: Molecular Biology of Pseudomonads. T. Nakazawa, K. Furukawa, D. Haas, and S. Silver (Eds.). American Society for Microbiology, Washington.
- Hu, W.-J., A. Kawaoka, C.-J. Tsai, J. Lung, K. Osakabe, H. Ebinuma, and V. L. Chiang. 1998. Compartmentalized expression of two structurally and functionally distinct 4-coumarate:CoA ligase genes in aspen (*Populus tremuloides*). *Proc. Nat. Acad. Sci.* 95: 5407-5412.
- Ian, F., A. L. Wilkins, and T. R. Healy. 1998. Accumulation of resin acids in sediments adjacent to a log handling area, Tauranga Harbour, New Zealand.
- Icil, S. and H. Icil. 1994. Photolysis of abietic acid to dehydroabietic acid and hydrogen. *Spec. Ltrs.* 27(8): 1041-1048.
- Inouye, M., Y. Tokada, N. Muto, T. Beppu, and S. Horinouchi. 1994. Characterization and expression of a P-450-like mycinamicin biosynthesis gene using a novel *Micromonospora-Escherichia coli* shuttle cosmid vector. *Mol. Gen. Genet.* 245: 456-464.
- Johnson, K., K. Mattsson, J. Tana, T. R. Stuthridge, J. Hemming, and K. J. Lehtinen. 1995. Uptake and elimination of resin acids and physiological responses in rainbow trout exposed to total mill mixed liquor from an integrated newsprint mill. *Env. Tox. Chem.* 14(): 1561-1568.
- Kaneda, T. 1991. Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance. *Microbiol. Rev.* 55: 288-302.

- Kanemoto, R. H., A. T. Powell, D. E. Akiyoshi, D. A. Regier, R. A. Kerstetter, E. W. Nester, M. C. Hawes, and M. P. Gordon. 1989. Nucleotide sequence and analysis of the plant-inducible locus *pinF* from *Agrobacterium tumefaciens*. J. Bacteriol. 171(5): 2506-2512.
- Keen, N. T., S. Tamaki, D. Kobayashi, and D. Trollinger. 1988. Improved broad-host-range plasmids for DNA cloning in Gram-negative bacteria. Gene. 70: 191-197.
- Keen, N. T., S. Tamaki, D. Kobayashi, and D. Trollinger. 1988. Improved broad-host-range plasmids for DNA cloning in Gram-negative bacteria. Gene. 70: 191-197.
- Kim, M.-K. and C. S. Harwood. 1991. Regulation of benzoate-CoA ligase in *Rhodopseudomonas palustris*. FEMS Microbiol. Lett. 83: 199-204.
- Kozlowski, T. T. and S. G. Pallardy. 1997. Physiology of Woody Plants. Second Edition. Academic Press, Inc. San Diego. 411 pp.
- Kramer, P.J. and T.T. Kozlowski. 1979. Physiology of Woody Plants. Academic Press, New York. 811 pp.
- Krawiec, S. and M. Riley. 1990. Organization of the bacterial genome. Microbiol. Rev. 54(4): 502-539.
- Kutney, J.P., E. Dimitriadis, G.M. Hewitt, M. Singh, and B. R. Worth. 1983. Studies related to biological detoxification of kraft pulp mill mixed liquor. III The biodegradation of abietic acid with *Mortierella*. Helv. Chim. Acta 65:1343-1358.
- Kutney, J.P., E. Dimitriadis, G.M. Hewitt, P. H. Salisbury, M. Singh, J. A. Servizi, D. W. Martens, and R. W. Gordon. 1983. Studies related to biological detoxification of kraft pulp mill mixed liquor. The biodegradation of 12,14-dichlorodehydroabietic acid with *Mortierella isabellina*. Helv. Chim. Acta 66:921-928.
- Kutney, J.P., E. Dimitriadis, G.M. Hewitt, P. H. Salisbury, M. Singh, J. A. Servizi, D. W. Martens, and R. W. Gordon. 1983. Studies related to biological detoxification of kraft pulp mill mixed liquor. The biodegradation of 14-chlorodehydroabietic acid with *Mortierella isabellina*. Helv. Chim. Acta 65:1343-1358.
- Kutney, J.P., J.-D. Berset, G.M. Hewitt, and M. Singh. 1988. Biotransformation of dehydroabietic, abietic, and isopimaric acids by *Mortierella isabellina* immobilized in polyurethane foam. Appl. Environ. Microbiol. 54:1015-1022

Kutney, J.P., J.-D. Berset, G.M. Hewitt, and M. Singh. 1988. Biotransformation of dehydroabietic, abietic, and isopimaric acids by *Mortierella isabellina* immobilized in polyurethane foam. *Appl. Environ. Microbiol.* 54:1015-1022

Kutney, J.P., L. S. L. Choi, G.M. Hewitt, P. H. Salisbury, and M. Singh. 1985. Biotransformation of dehydroabietic acid with resting cell suspensions and calcium alginate-immobilized cells of *Mortierella isabellina*. *Appl. Environ. Microbiol.* 49:96-100

Kutney, J.P., M. Singh, E. Dimitriadis, G.M. Hewitt, P. H. Salisbury, B. R. Worth, J. A. Servizi, D. W. Martens, and R. W. Gordon. 1981. Studies related to biological detoxification of kraft pulp mill mixed liquor. II The biotransformation of isopimaric acid with *Mortierella isabellina*. *Helv. Chim. Acta* 59:3350-3355.

Kutney, J.P., M. Singh, G. Hewitt, P. H. Salisbury, B. R. Worth, J. A. Servizi, D. W. Martens, and R. W. Gordon. 1981. Studies related to biological detoxification of kraft pulp mill mixed liquor. II The biodegradation of dehydroabietic acid with *Mortierella isabellina*. *Helv. Chim. Acta* 59:2234-2341.

Kutney, J.P., G.M. Hewitt, P. H. Salisbury, M. Singh, J. A. Servizi, D. W. Martens, and R. W. Gordon. 1983. Studies related to biological detoxification of kraft pulp mill mixed liquor. The biotransformation of 12-chlorodehydroabietic acid with *Mortierella isabellina*. *Helv. Chim. Acta* 66:2191-2197.

Langenheim, J.H. 1994. Higher plant terpenoids: A phytocentric overview of their ecological roles. *J. Chem. Ecology*. 20(6): 1223-1279.

Lavallee, H.-C., L. Rouisse, and R. Paradis. 1993. A comprehensive study of the Quebec pulp and paper mill mixed liquors as an instrument to define a strategy for abatement. *Pulp and Paper Can.* 94(11): T393-T399.

Lawrence, R.V. 1969. Naval stores products from southern pines. *Forest Products Journal*. 19(9): 87-92.

Leach, J.M. and A.N. Thakore. 1973. Identification of the constituents of kraft pulping mixed liquor that are toxic to juvenile coho salmon (*Oncorhynchus kisutch*). *J. Fish. Res. Board Can.* 30(4): 479-484.

Li, K., M. Chester, J. P. Kutney, J. N. Saddler, and C. Breuil. 1994. Production of polyclonal antibodies for the detection of dehydroabietic acid in pulp mill mixed liquors. *Anal. Ltrs.* 27(9): 1671-1688.

Li, K., T. Chen, C. Breuil, and J. N. Saddler. 1996. A comparison of gas chromatographic and immunochemical methods for quantifying resin acids. In: *Environmental Fate and Effects of Pulp and Paper Mill Mixed liquors*. M. R.

Servos, K. R. Munkittrick, J. H. Carey, and G. J. Van Der Kraak (Eds.). pp. 119-127.

Liss, S. N. and D. G. Allen. 1992. Microbiological study of a bleached kraft pulp mill aerated lagoon. *J. Pulp and Paper Sci.* 18(6): J216-J220.

Liss, S. N., P. A. Bicho, and J. N. Saddler. 1997. Microbiology and biodegradation of resin acids in pulp mill mixed liquors: a minireview. *Can. J. Microbiol.* 75: 599-611.

Liu, H.-W., S.-N. Lo, and H.-C. Lavallee. 1996. Mechanisms of removing resin and fatty acids in CTMP mixed liquor during aerobic biological treatment. *TAPPI Journal.* 79(5): 145-154.

Loomis, W.D. and R. Croteau. 1980. Biochemistry of terpenoids. In: *The Biochemistry of Plants: A Comprehensive Treatise. Volume 4.* (P.K. Stumpf and E.E. Conn Eds.). Academic Press, Inc., New York. pp. 363-418.

Lynon-Biosoft. 1994-1997. Sequence analysis software: DNAMAN. Vaudreuil, Quebec.

Madore, O. 1992. The Pulp and Paper Industry: The impact of the new Federal Environmental Regulations. BP-299E. Research Branch, Library of Parliament, Ottawa. 17 pp.

Martel, P. H., T. G. Kovacs, and R. H. Voss. 1996. Mixed liquors from Canadian pulp and paper mills: a recent investigation of their potential to induce mixed function oxygenase activity in fish. In: *Environmental Fate and Effects of Pulp and Paper Mill Mixed liquors.* M. R. Servos, K. R. Munkittrick, J. H. Carey, and G. J. Van Der Kraak (Eds.). pp. 401-412.

Martin, V. J. J. and W. W. Mohn. 1999. A novel aromatic ring hydroxylating dioxygenase from the diterpenoid-degrading bacterium *Pseudomonas abietaniphila* BKME-9. *J. Bacteriol.* In press.

McFarlane, P. N. and T. A. Clark. 1988. Metabolism of resin acids in anaerobic systems. *Wat. Sci. Tech.* 20(1): 273-276.

McKean, W. T. 1980. Pulp and paper industry. In: *Introduction to Environmental Toxicology.* F. E. Guthrie and J. J. Perry (Eds.). Elsevier, New York. pp: 210-226.

McLeay and Associates. 1986. Aquatic toxicity of pulp and paper mixed liquors: a review. Environment Canada Report EPS 4/PF/1. Supply and Services Canada. 191 pp.

McLeay, D.J., C.C. Walden, and J.R. Munro. 1979. Influence of dilution water on the toxicity of kraft pulp and paper mill mixed liquor, including mechanisms of effect. *Water Res.* 13:151-158.

McSweeney, E.E., H.G. Arlt, Jr., and J. Russell. 1987. *Tall Oil and Its Uses - II*. Pulp Chemicals Association, Inc., New York.

Mohn, W. W. 1995. Bacteria obtained from a sequencing batch reactor that are capable of growth on dehydroabietic acid. *Appl. Env. Microbiol.* 61(6): 2145-2150.

Mohn, W. W. and G. R. Stewart. 1997. Bacterial metabolism of chlorinated dehydroabietic acids occurring in pulp and paper mill mixed liquors. *Appl. Env. Microbiol.* 63(8): 3014-3020.

Mohn, W. W., A. E. Wilson, P. Bicho, and E. R. B. Moore. 1998a. Physiological and phylogenetic diversity of bacteria growing on resin acids. *Syst. Appl. Microbiol.* In press.

Mohn, W. W., V. Martin, and Z. Yu. 1998b. Distribution and abundance in pulp mill mixed liquor treatment systems of a resin acid degrading bacterium and a catabolic gene encoding resin acid biodegradation. Eighth International Symposium on Microbial Ecology (ISME-8), Halifax, Canada, 9-14 August, 1998. Abst.

Mohn, W.W. 1995. Bacteria obtained from a sequencing batch reactor that are capable of growth on dehydroabietic acid. *Appl. Environ. Microbiol.* 61: 2145-2150.

Moore, E.R.B., M. Mau, A. Arnscheidt, E. C. Bottger, R. A. Hutson, M. D. Collins, Y. van de Peer, R. de Wachter, and K. N. Timmis. 1996. The determination and comparison of the 16S rRNA gene sequences of the genus *Pseudomonas* (sensu stricto) and estimation of the natural intrageneric relationships. *Syst. Appl. Microbiol.* 19: 478-492.

Morgan, C. A. and R. C. Wyndham. 1996. Isolation and characterization of resin acid degrading bacteria found in mixed liquor from a bleached kraft pulp mill. *Can. J. Microbiol.* 42: 423-430.

Morgan, C.A. 1993. Isolation and identification of resin acid (dehydroabietic and abietic acids) degrading bacteria from Canadian Pacific Forest Products, Dryden, Ontario pulp mill mixed liquor. Unpublished thesis, Carleton University, Ottawa. 39 pp. + appendices.

Mueller, J.C., J.M. Leach, and C.C. Walden. 1977. Detoxification of bleached kraft mill mixed liquors - a manageable problem. TAPPI Journal. 66(9): 135-137.

Mullis, K. B. and F. A. Faloona. 1987. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Meth. Enz. 155: 335-350.

Munkittrick, K. R., G. J. Van Der Kraak, M. E. McMaster, and C. B. Portt. 1992. Reproductive dysfunction and MFO activity in three species of fish exposed to bleached kraft mill mixed liquor at Jackfish Bay, Lake Superior. Water Poll. Res. J. Canada. 27(3): 439-446.

Munkittrick, K. R., G. J. Van Der Kraak, M. E. McMaster, C. B. Portt, M. R. Van Den Heuvel, and M. R. Servos. 1994. Survey of receiving-water environmental impacts associated with discharges from pulp mills. 2. Gonad size, liver size, hepatic EROD activity, and plasma sex steroid levels in white sucker. Env. Tox. Chem. 13(7): 1089-1101.

Murray, W. 1992. Pulp and Paper: The reduction of toxic mixed liquors. BP-292E. Research Branch, Library of Parliament, Ottawa. 23 pp.

Murray, W. D. and M. Richardson. 1993. Development of biological and process technologies for the reduction and degradation of pulp mill wastes that pose a threat to human health. Crit. Rev. Env. Sci. Tech. 23(2): 157-194.

Mutton, D. B. 1962. Wood resins. In: Wood Extractives and Their Significance to the pulp and paper Industries. W. E. Hillis (Ed.). Academic Press, New York.

Neidle, E. L., C. Hartnett, and L. N. Ornston. 1989. Characterization of *Acinetobacter calcoaceticus catM*, a repressor gene homologous in sequence to transcriptional activator genes. J. Bacteriol. 171(10): 5410-5421.

Nelson, D. R., T. Kamataki, D. J. Waxman, F. P. Guengerich, R. W. Estabrook, R. Feyereisen, F. J. Gonzalez, M. J. Coon, I. C. Gunsalus, O. Gotoh, K. Okuda, and D. W. Nebert. 1993. The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. DNA and Cell Biol. 12(1): 1-51.

Nestmann, E. R., E. G.-H. Lee, J. C. Mueller, and G. R. Douglas. 1979. Mutagenicity of resin acids identified in pulp and paper mill mixed liquors using *Salmonella*/mammalian-microsome assay. Env. Mutagenesis. 1: 361-369.

Nietmetz, R., U. Altenschmidt, S. Brucker, and G. Fuchs. 1995. Benzoyl-coenzyme-A 3-monooxygenase, a flavin dependent hydroxylase: purification,

some properties and its role in aerobic benzoate oxidation via gentisate in a denitrifying bacterium. *Eur. J. Biochem.* 227: 161-168.

Nunn, W. D. 1986. A molecular view of fatty acid catabolism in *Escherichia coli*. *Microbiol. Rev.* 50(2): 179-192.

Nyren, V. and E. Back. 1958. The ionization constant, solubility product, and solubility of abietic and dehydroabietic acid. *Acta Chem. Scan.* 12: 1516-1520.

O'Connor, B. I., T. G. Kovacs, and R. H. Voss. 1992. The effect of wood species composition on the toxicity of simulated mechanical pulping mixed liquors. *Env. Tox. Chem.* 11: 1259-1270.

Ochman, H., J. W. Ajioka, D. Garza and D. L. Hartl. 1989. Inverse polymerase chain reaction. In: *PCR Technology*. H. A. Erlich (Ed.). Stockton Press, New York. pp. 105-111.

Oikari, A., B.-J. Lonn, M. Castren, T. Nakari, B. Snickars-Nikinmaa, H. Bister, and E. Virtanen. 1983. Toxicological effects of dehydroabietic acid (DHAA) on the trout, *Salmo gairdneri* Richardson, in fresh water. *Water Res.* 17:81-89.

Old, R. W. and S. B. Primrose. 1989. Principles of gene manipulation: an introduction to genetic engineering. *Studies in Microbiology*. Vol. 2. N. G. Carr, J. L. Ingraham, and S. C. Rittenberg (Eds.). 438 pages.

Olivera, E. R., B. Minambres, B. Garcia, C. Muniz, M. A. Moreno, A. Ferrandez, E. Diaz, J. L. Garcia, and J. M. Luengo. 1998. Molecular characterization of the phenylacetic acid catabolic pathway in *Pseudomonas putida* U: the phenylacetyl-CoA catabolon. *Proc. Natl. Acad. Sci.* 95: 6419-6424.

Ontario Ministry of Environment. 1992. Report on the quality assurance and quality control data for the MISA pulp and paper sector. Water Resources Branch, Ontario Ministry of the Environment, Toronto.

Ontario Ministry of Environment and Energy. 1994. The 1992 non-compliance report. April 1994. Oqueen's Printer for Ontario, Toronto.

Ontario Ministry of Environment. 1993. Clean water regulation: pulp and paper sector. Ontario Ministry of the Environment, Toronto. 8 pp.

Pan, W. and B. G. Spratt. 1994. Regulation of the permeability of the gonococcal cell envelope by the *mtr* system. *Mol. Microbiol.* 11(4): 769-775.

Peterson, J. A., J.-Y. Lu, J. Geisselsoder, S. Graham-Lorence, C. Carmona, F. Witney, and M. C. Lorence. 1992. Cytochrome P450_{terp}: isolation and

purification of the protein and cloning and sequencing of its operon. J. Biol. Chem. 267(20): 14193-14203.

Pinder, A. R. 1960. The chemistry of the terpenes. Chapman Hall, London. Pp. 223.

Providenti, M. A., H. Lee, and J. T. Trevors. 1993. Selected factors limiting the microbial degradation of recalcitrant compounds. J. Ind. Microbiol. 12: 379-395.

Qureshi, N. and J. W. Porter. 1981. Conversion of aceto-coenzyme-A to isopentyl pyrophosphate. In: Biosynthesis of Isoprenoid Compounds. Volume 1. (J.W. Porter and S.L. Spurgeon, Eds). John Wiley and Sons, New York. pp. 47-94.

Rich, J.J. and D.K. Willis. 1990. A single oligonucleotide can be used to rapidly isolate DNA sequences flanking a transposon Tn5 insertion by the polymerase chain reaction. Nuc. Acids Res. 18(22): 6673-6676.

Richardson, D. E., J. B. Bremner, and B. V. O'Grady. 1992. Quantitative analysis of total resin acids by high-performance liquid chromatography of their coumarin ester derivatives. J. Chromat. 595: 155-162.

Rodriguez, A. M., C. Olano, C. Mendez, C. R. Hutchinson, and J. A. Salas. 1995. A cytochrome P450-like gene possibly involved in oleandomycin biosynthesis by *Streptomyces antibioticus*. FEMS Microbiol. Letters. 127: 117-120.

Ronald, S., M. A. Farinha, B. J. Allan, and A. M. Kropinski. 1992. Cloning and physical mapping of transcriptional regulatory (Sigma) factors from *Pseudomonas aeruginosa*. In: *Pseudomonas: molecular biology and biotechnology*. E. Galli, S. Silver, and B. Witholt (Eds.). American Society for Microbiology, Washington, D. C. pp. 249-257.

Roper, D. I. And R. A. Cooper. 1993. Purification, nucleotide sequence and some properties of a bifunctional isomerase/decarboxylase from the homoprotocatechuate degradative pathway of *Escherichia coli* C. Eur. J. Biochem. 217: 575-580.

Roper, D. I., T. Fawcett, and R. A. Cooper. 1993. The *Escherichia coli* C homoprotocatechuate degradative operon: *hpc* gene order, direction of transcription and control of expression. Mol. Gen. Genet. 237: 241-250.

Ropp, J. D., I. C. Gunsalus, and S. G. Sligar. 1993. Cloning and expression of a member of a new cytochrome P-450 family: cytochrome P-450lin (CYP111) from *Pseudomonas incognita*. J. Bacteriol. 175(18): 6028-6037.

Rothstein, S. J., R. A. Jorgenson, J. C.-P. Yin, Z. Young-Di, R. C. Johnson, and W. S. Reznikoff. 1980. Genetic organization of Tn5. Cold Spring Harbor Symp. Quant. Biol. 45: 99-105.

Russell, C. E. and A. A. Berryman. 1976. Host resistance to the fir engraver beetle. 1. Monoterpene composition of *Abies grandis* pitch blisters and fungus-infected wounds. Can. J. Bot. 54: 14-18.

Sato, S., M. Hayashi, S. Imamura, Y. Ozeki, and A. Kawaguchi. 1992. Primary structures of the genes, *faoA* and *faoB*, from *Pseudomonas fragi* B-0771 which encode the two subunits of the HDT multi-enzyme complex involved in fatty acid β -oxidation. J. Biochem. 111: 8-15.

Sayler, G. S. and A. C. Layton. 1990. Environmental application of nucleic acid hybridization. Ann. Rev. Microbiol. 44: 1663-1673.

Scott, J. R. 1984. Regulation of plasmid replication. Microbiol. Rev. 48: 1-23.

Servos, M. R., J. H. Carey, M. L. Ferguson, G. Van Der Kraak, H. Ferguson, J. Parrott, K. Gorman, and R. Cowling. Impact of a modernized bleached kraft mill on white sucker populations in the Spanish River, Ontario. Water Poll. Res. J. Canada. 27(3): 423-437.

Shimp, R. J. and J. W. Owens. 1993. Pulp and paper technologies and improvements in environmental emissions to aquatic environments. Tox. Env. Chem. 40: 213-233.

Showalter, R. E., M. O. Martin, and M. R. Silverman. 1990. Cloning and nucleotide sequence of *luxR*, a regulatory gene controlling bioluminescence in *Vibrio harveyi*. 172(6): 2946-2954.

Simon, R., U. Priefer, and A. Puhler. 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram negative bacteria. Bio/technology. November 1983: 784-791.

Simonsen, Sir J. and D. H. R. Barton. 1961. The terpenes. Vol. III. Cambridge University Press, London. Pp. 575.

Smibert, R. M. and N. R. Krieg. 1994. Phenotypic characterization. In: Methods for general and molecular bacteriology. P. Gerhardt, R. G. E. Murray, W. A. Wood, and N. R. Krieg (Eds.). American Society for Microbiology Press, Washington.

- Sneath, P.H. 1973. Numerical Taxonomy. W.H. Freeman and Company, New York.
- Soderberg, T. A., R. Gref, S. Holm, T. Elmros, and G. Hallmans. 1990. Antibacterial activity of rosin and resin acids *in vitro*. Scand. J. Reconstr. Hand Surg. 24: 199-205.
- Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503-517.
- Spurgeon, S.L. and J.W. Porter. 1981. Introduction. In: Biosynthesis of Isoprenoid Compounds. Volume 1. (J.W. Porter and S.L. Spurgeon, Eds). John Wiley and Sons, New York. pp. 1-40.
- Stegeman, J. J. and M. E. Hahn. 1994. Biochemistry and molecular biology of monooxygenases: current perspectives on forms, functions, and regulation of cytochrome P450 in aquatic species. In: Aquatic Toxicology: Molecular, Biochemical, and Cellular Perspectives. D. C. Malins and G. K. Ostrander (Eds.). pp. 87-206.
- Stuthridge, T.R., D.N. Campin, A.G. Langdon, K.L. Mackie, P.N. McFarlane, and A.L. Wilkins. 1991. Treatability of bleached kraft pulp and paper mill wastewaters in a New Zealand aerated lagoon treatment system. Wat. Sci. Tech. 24(3/4): 309-317.
- Sunito, L.R., W.Y. Shiu, and D. Mackay. 1988. A review of the nature and properties of chemicals present in pulp mill mixed liquors. Chemosphere. 17(7):1249-1290.
- Swartz, J. N. and R. C. MacDonald. 1962. Alkaline pulping. In: Pulp and Paper Science Technology. Vol 1. C. E. Libby (Ed.). McGraw-Hill Book Company, New York. pp: 160-239.
- Tavendale, M. H., P. N. McFarlane, K. L. Mackie, A. L. Wilkins, and A. G. Langdon. 1997b. The fate of resin acids-2: the fate of resin acids and resin acid derived neutral compounds in anaerobic sediments. Chemosphere. 35(10): 2153-2166.
- Tavendale, M. H., P. N. McFarlane, K. L. Mackie, A. L. Wilkins, and A. G. Langdon. 1997a. The fate of resin acids-1: the biotransformation and degradation of deuterium labelled dehydroabietic acid in anaerobic sediments. Chemosphere. 35(10): 2137-2151.

Taylor, B.R., K.L. Yeager, S.G. Abernathy, and G.F. Westlake. 1988. Scientific criteria document for the development of provincial water quality objectives and guidelines: Resin acids. Queen's Printer for Ontario. 50 pp. + appendices.

Thomas, C. M. and C. A. Smith. 1987. Incompatibility group P plasmids: genetics, evolution, and use in genetic manipulation. *Ann. Rev. Microbiol.* 41: 77-101.

Thut, R. N. 1989. Utilization of artificial marshes for treatment of pulp mill mixed liquors. In: *Constructed Wetlands for Wastewater Treatment: Municipal, Industrial, and Agricultural*. D. A. Hammer (Ed.). Lewis Publishers Inc., Chelsea. Pp. 239-244.

Tian, F., A. L. Wilkins, and T. R. Healy. 1998. Accumulation of resin acids in sediments adjacent to a log handling area, Tauranga Harbour, New Zealand. *Bull. Environ. Contam. Toxicol.* 60: 441-447.

Trower, M. K., R. Lenstra, C. Omer, S. E. Bucholz, and F. S. Sariaslani. 1992. Cloning, nucleotide sequence determination and expression of the genes encoding cytochrome P450_{soy} (soyC) and ferredoxin_{soy} (soyB) from *Streptomyces griseus*. *Mol. Microbiol.* 6(15): 2125-2134.

Turgay, K., M. Krause, and M. A. Marahiel. 1992. Four homologous domains in the primary structure of GrsB are related to domains in a superfamily of adenylate-forming enzymes. *Mol. Microbiol.* 6(4): 529-546.

Unger, B. P., I. C. Gunsalus, and S. G. Sligar. 1986. Nucleotide sequence of the *Pseudomonas putida* cytochrome P450_{cam} gene and its expression in *Escherichia coli*. *J. Biol. Chem.* 261(3): 1158-1163.

van Beilen, J. B., G. Eggink, H. Enequist, R. Bos, and B. Witholt. 1992. DNA sequence determination and functional characterization of the OCT-plasmid-encoded *alkJKL* genes of *Pseudomonas oleovorans*. *Mol. Microbiol.* 6(21): 3121-3136.

van der Meer, J. R., W. M. de Vos, S. Haratama, and A. J. B. Zehnder. 1992. Molecular mechanisms of genetic adaptation to xenobiotic compounds. *Microbiol. Rev.* 56(4): 677-694.

Van Kleef, M. A. G. and J. A. Duine. 1988. Bacterial NAD(P)-independent quinate dehydrogenase is a quinoprotein. *Arch. Microbiol.* 150: 32-36.

Versalovic, J., T. Koeuth, and J. R. Lupski. 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res.* 19: 6823-6831.

- Villemur, R. 1995. Coenzyme A ligases involved in anaerobic biodegradation of aromatic compounds. *Can. J. Microbiol.* 41: 855-861.
- Volkman, J. K., D. G. Holdsworth, and D. E. Richardson. 1993. Determination of resin acids gas chromatography and high-performance liquid chromatography in paper mill mixed liquor, river waters and sediments from Derwent Estuary, Tasmania. *J. Chromat.* 643: 209-219.
- Voss, R. H. and A. Rapsomatiotis. 1985. An improved solvent-extraction based procedure for the gas chromatographic analysis of resin acids in pulp mill mixed liquors. *J. Chromatogr.* 346: 205-214.
- Walden, C. C. and T. E. Howard. 1977. Toxicity of pulp and paper mill mixed liquors - a review of regulations and research. *Tappi J.* 60(1): 122-125.
- Walden, C.C., E.G.H. Lee, and J.C. Muller. 1978. Biological characteristics of pulp mill mixed liquors. Part I. Environmental Protection Service, Environment Canada. CPAR Report 678-1. 24 pp.
- Walia, S., A. Khan, and N. Rosenthal. 1990. Construction and applications of DNA probes for detection of polychlorinated biphenyl-degrading genotypes in toxic organic contaminated soil environments. *Appl. Env. Microbiol.* 56: 254-259.
- Wang, Z., T. Chen, Y. Gao, C. Breuil, and Y. Hiratsuka. 1995. Biological degradation of resin acids in wood chips by wood-inhabiting fungi. *Appl. Env. Microbiol.* 61(1): 222-225.
- Welander, T. 1988. An anaerobic process for the treatment of CTMP mixed liquor. *Wat. Sci. Tech.* 20(1): 143-147.
- Werker, A. G. and E. R. Hall. 1997. The influence of pH on resin acid solubility related to biodegradation kinetics of resin acid in pulp mill mixed liquor. In: TAPPI Proceedings of the 1997 Environmental Conference and Exhibit. Pp. 19-26.
- Wilkins, A. L. and S. Panadam. 1987. Extractable organic substances from the discharges of a New Zealand pulp and paper mill. *Appita* 40(3): 208-212.
- Williams, T. G., J. H. Carey, B. K. Burnison, D. G. Dixon, and H.-B. Lee. 1996. Rainbow trout (*Oncorhynchus mykiss*) mixed function oxygenase responses caused by unbleached and bleached pulp mill mixed liquors: a laboratory-based study. In: Environmental Fate and Effects of Pulp and Paper Mill Mixed liquors. M. R. Servos, K. R. Munkittrick, J. H. Carey, and G. J. Van Der Kraak (Eds.). pp. 379-388.

- Wilmes-Riesenberg, M. R. and B. L. Wanner. 1992. *TnphoA* and *TnphoA'* elements for making and switching fusions for study of transcription, translation, and cell surface localization. *J. Bacteriol.* 174(14): 4558-4575.
- Wilson, A. J., E. R. B. Moore, and W. W. Mohn. 1996. Isolation and characterization of isopimaric acid-degrading bacteria from a sequencing batch reactor. *Appl. Env. Microbiol.* 62(9): 3146-3151.
- Woo, A. J., W. R. Strohl, and N. D. Priestly. 1998. Nonactin biosynthesis: the product of *orfX* catalyses the formation of the furan ring of nonactic acid. Unpublished.
- Wyndham, R.C. 1986. Evolved aniline catabolism in *Acinetobacter calcoaceticus* during continuous culture of river water. *Appl. Environ. Microbiol.* 51(4): 781-789.
- Wyndham, R.C., C. Nakatsu, M. Peel, A. Cashore, J. Ng, and F. Szilagy. 1994. Distribution of the catabolic transposon Tn5271 in a groundwater bioremediation system. *Appl. Env. Microbiol.* 60(1): 86-93.
- Xie, Y., M. B. Isman, Y. Feng, and A. Wong. 1993. Diterpene resin acids: major active principles in tall oil against variegated cutworm, *Peridroma saucia* (Lepidoptera: Noctuidae). *J. Chem. Ecol.* 19(6): 1075-1084.
- Yamada, A., Y. Ezaki, K. Matsuo, T. Yarita, and A. Nomura. 1995. Supercritical fluid chromatography of free resin acids on an ODS-silica gel column. *J. Chromat.* 709: 345-349.
- Zender, J.A., T.R. Stuthridge, A.G. Langdon, A.L. Wilkins, K.L. Mackie, and P.N. McFarlane. 1994. Removal and transformation of resin acids during secondary treatment at a New Zealand bleached kraft pulp and paper mill. *Wat. Sci. Tech.* 29(5/6): 105-121.
- Zheng, J. and R. A. Nicholson. 1996. Influence of two naturally occurring abietane monocarboxylic acids (resin acids) and a chlorinated derivative on release of the inhibitory neurotransmitter γ -aminobutyric acid from trout brain synapses. *Bull. Environ. Contam.* 56: 114-120.
- Zheng, J. and R. A. Nicholson. 1998. Action of resin acids in nerve ending fractions isolated from fish central nervous system. *Env. Tox. Chem.* 17(9): 1852-1859.

Ziegelen, G., W. Pansgrau, B. Strack, D. Balzer, M. Kroger, V. Kruff, and E. Lanka. 1991. Nucleotide sequence and organization of genes flanking the transfer origin of promiscuous plasmid RP4. *DNA Seq.* 1 (5): 303-327.

Zubay, G. 1993. *Biochemistry*. Third Edition. Wm. C. Brown Communications Inc., Dubuque.

Zylstra, G. J., W. R. McCombie, D. T. Gibson, and B. A. Finette. 1988. Toluene degradation by *Pseudomonas putida* F1: genetic organization of the *tod* operon. *Appl. Env. Microbiol.* 54(6): 1498-1503.

This appendix presents raw data in support of the results presented in Chapter 2. Table A1 is a complete list of carbon sources used by resin acid degrading bacteria isolated from pulp mill mixed liquor. Selected phenotypic characteristics are included. Table A2 summarizes the Jaccard correlation coefficients for the 13 isolates based on the characteristics delineated in Table A1.

TABLE A1. Summary of phenotypic characteristics of resin acid degrading bacteria isolated from Avenor bleached kraft pulp mill in Dryden, Ontario (Weyerhaeuser). Characteristics were used for hierarchical cluster analysis and to calculate Jaccard Correlation Coefficient (Sneath, 1973) illustrated in Figure 2-5. Symbols are + positive reaction, - negative reaction, and / weak or borderline reaction. Note that latter were scored as negative during analysis.

ISOLATE	A19-1	A10-2	A19-3	A19-5	A19-6a	A19-6b	D113b	D11-9	D11-13	D11-23a	S11-23b	D11-25	D11-37
Gram stain	-	-	-	-	-	-	-	-	-	-	-	-	-
Shape	rod	rod	rod	rod	rod	rod	rod	rod	rod	rod	rod	rod	rod
Motility	-	-	-	+	-	+	+	+	+	+	-	+	+
Oxidase	/	-	+	/	/	-	/	/	-	/	+	-	+
Catalase	/	-	/	+	/	+	/	+	/	/	+	+	+
Oxidation	/	/	/	-	-	-	+	+	+	+	-	+	+
Fermentation	-	-	-	-	-	-	-	-	-	-	-	-	-
King's B	-	-	-	-	-	-	-	-	-	-	-	-	-
α-cyclodextrin	-	-	-	-	-	-	-	-	-	-	-	-	-
dextrin	-	-	-	-	-	-	+	-	+	+	-	+	-
glycogen	-	-	-	-	-	-	-	-	+	+	-	+	-
tween 40	+	+	+	+	+	+	+	+	-	+	+	-	+
tween 80	+	+	+	+	+	+	+	+	-	+	+	-	+
N-acetyl-D-galactosamine	-	-	-	-	-	-	+	-	+	+	-	+	-
N-acetyl-D-glucosamine	+	+	-	-	-	-	+	-	+	+	-	+	-
adonitol	-	-	-	-	-	-	+	-	-	+	-	-	-
L-arabinose	+	+	-	-	-	-	+	-	-	+	-	-	/
D-arabitol	+	+	-	-	-	-	+	-	-	+	-	-	-
cellobiose	-	-	-	-	-	-	+	-	+	+	-	+	-
i-erythritol	-	-	-	-	-	-	-	-	-	-	-	-	-
D-fructose	-	+	+	-	+	-	+	-	+	+	-	+	-
L-fucose	+	+	-	-	-	-	+	-	-	+	-	-	-
D-galactose	+	+	-	-	-	-	+	-	+	+	-	+	-
gentiobiose	-	-	-	-	-	-	+	-	+	+	-	+	-
α-D-glucose	+	+	+	-	+	-	+	+	+	+	-	+	+
m-inositol	-	-	-	-	-	-	+	-	-	+	-	-	-
α-lactose	-	-	-	-	-	-	+	-	+	-	-	+	-
α-D-lactose lactulose	-	-	-	-	-	-	+	-	+	-	-	+	-
maltose	-	-	-	-	-	-	+	-	+	+	-	+	-
D-mannitol	+	+	-	+	-	+	+	-	-	+	+	-	-
D-mannose	+	+	-	-	-	-	+	-	+	+	-	+	-
D-melibiose	-	-	-	-	-	-	+	-	+	+	-	+	-
β-methyl,D- glucoside	-	-	-	-	-	-	+	-	+	+	-	+	-
psicose	-	-	-	-	-	-	-	-	-	-	-	+	-
D-raffinose	-	-	-	-	-	-	+	-	-	+	-	-	-
L-rhamnose	-	-	-	-	-	-	+	-	-	+	-	-	-
D-sorbitol	-	-	-	-	-	-	+	-	-	+	-	-	-
sucrose	-	-	-	-	-	-	+	-	+	+	-	+	-
D-trehalose	-	-	-	-	-	-	+	-	+	+	-	+	-
turanose	-	-	-	-	-	-	+	-	+	+	-	+	-
xylitol	-	-	-	-	-	-	-	-	-	-	-	-	-
methyl pyruvate	+	+	+	+	+	+	-	+	+	-	+	+	+
mono-methyl succinate	+	+	+	+	+	+	+	+	-	-	+	-	+
acetic acid	+	+	+	+	+	+	+	+	-	+	+	-	+
cis-aconitic acid	+	+	+	+	+	+	-	+	+	-	+	+	+
citric acid	+	+	+	-	+	-	-	+	-	-	-	-	+
formic acid	+	+	-	-	-	+	-	-	-	-	+	-	-
D-galactonic acid lactone	+	+	-	-	-	-	-	-	-	-	-	-	-
D-galacturonic acid	+	+	-	-	-	-	+	-	-	-	-	-	-
D-gluconic acid	+	+	+	+	+	+	+	+	-	+	+	-	+
D-glucosaminic acid	-	-	-	-	-	-	-	-	-	-	-	-	-

D-glucuronic acid	-	-	-	-	-	-	-	-	-	-	-	-
α-hydroxy butyric acid	+	+	-	-	-	-	-	-	-	-	-	-
β-hydroxy butyric acid	+	+	+	+	+	+	-	+	+	-	+	+
γ-hydroxy butyric acid	-	-	+	-	+	-	+	-	-	+	-	-
p-hydroxy phenylacetic acid	+	+	-	-	-	+	-	+	-	-	+	-
itaconic acid	+	+	+	-	+	-	-	-	-	-	-	-
α-keto butyric acid	+	+	-	-	-	+	-	-	-	-	+	-
α-keto glutaric acid	+	+	+	-	+	-	-	+	-	+	-	-
α-keto valeric acid	-	-	-	-	-	-	-	-	-	-	-	-
D,L-lactic acid	+	+	+	+	+	+	+	+	-	-	+	+
malonic acid	+	+	+	-	+	-	-	-	-	-	-	-
propionic acid	+	+	+	+	+	+	+	+	-	+	+	-
quinic acid	-	-	+	+	+	+	+	+	-	+	+	-
D-saccharic acid	+	+	+	-	+	-	-	-	-	-	-	-
sebacic acid	-	-	-	+	-	+	-	-	-	-	+	-
succinic acid	+	+	+	+	+	+	+	+	+	+	+	+
bromo succinic acid	+	+	+	+	+	+	+	+	+	+	+	+
succinamic acid	+	+	-	-	-	+	-	+	-	-	+	-
glucuronamide	-	-	-	-	-	-	-	-	-	-	-	-
alaninamide	-	-	-	-	-	-	-	-	-	-	-	-
D-alanine	+	+	+	-	+	-	-	+	-	-	-	-
L-alanine	+	+	+	-	+	-	-	+	-	-	-	-
L-alanyl-glycine	-	+	-	-	-	-	-	-	-	-	-	-
L-asparagine	+	+	+	+	+	+	+	+	-	+	+	-
L-aspartic acid	+	+	+	+	+	+	+	+	-	+	+	-
L-glutamic acid	+	+	+	+	+	+	+	+	-	+	+	+
glycyl-L-aspartic acid	-	-	-	-	-	-	-	-	-	-	-	-
glycyl-L-glutamic acid	-	-	-	-	-	-	-	-	-	-	-	-
L-histidine	-	-	+	-	+	-	+	+	-	+	-	-
hydroxy L-proline	-	-	+	-	+	-	+	+	-	+	-	-
L-leucine	-	+	-	-	-	-	-	-	-	-	+	-
L-ornithine	-	-	-	-	-	-	-	+	-	+	-	-
L-phenyl alanine	-	-	-	-	-	-	-	-	-	-	-	-
L-proline	+	+	+	+	+	+	+	-	+	+	+	+
L-pyro glutamic acid	+	+	+	+	+	+	-	-	+	+	+	+
D-serine	-	-	-	-	-	-	-	+	-	-	-	-
L-serine	-	-	+	-	+	-	+	-	-	+	-	-
L-threonine	-	-	-	-	-	-	-	-	-	-	-	-
D,L-carnitine	-	-	+	-	+	-	-	-	-	-	-	+
γ-amino butyric acid	+	+	+	+	+	+	-	+	-	-	+	-
urocanic acid	-	-	+	-	+	+	-	-	-	-	-	+
inosine	-	-	+	-	+	-	-	+	-	-	-	+
uridine	-	-	-	-	-	-	-	-	-	-	-	-
thymidine	-	-	-	-	-	-	-	-	-	-	-	-
phenyl ethylamine	-	-	-	-	-	-	-	-	-	-	-	+
putrescene	-	-	+	-	+	-	-	+	-	-	-	+
2-amino ethanol	+	+	-	-	-	-	-	-	-	-	-	+
2,3-butanediol	-	-	-	-	-	-	-	-	-	-	-	-
glycerol	+	+	-	-	-	-	+	+	-	+	-	-
D,L-α-glycerol phosphate	+	+	-	-	-	-	-	-	-	-	-	-
glucose-1-phosphate	-	-	-	-	-	-	+	-	-	-	-	-
glucose-6-phosphate	-	-	-	-	-	-	+	-	-	-	-	-
abietic acid	+	+	+	+	+	+	+	+	-	-	-	+
dehydroabietic acid	-	-	+	-	+	+	+	+	+	+	+	+

[illegible]

This appendix presents raw data in support of the results presented in Chapter 4. Figure B1 is the complete 9180 bp *tdt* locus sequence in double strand format. It includes the amino acid sequence for transcribed regions which is shown next to the coding strand. Other features as listed in Table 4-4 are indicated as described in the figure caption. Table B1 summarizes the expected and observed fragments used to confirm the cloned probes in *Pseudomonas* sp. A19-6a, Figure 4-15 refers. A complete list of restriction sites for the entire *tdt* locus is presented in Table B2. Position numbers refer to plus strand numbering as illustrated in Figure B1.

FIGURE B1. Complete sequence of *Pseudomonas* sp. A19-6a *tdt* locus. The sequence is presented in double strand format with the amino acid sequence (outline font) next to the coding strand. The site of insertional inactivation by pBW30::Tn ϕ oA'-1 in A76-47 is indicated by an arrowhead. Nucleotide numbers refer to the plus strand and the base nearest right. Amino acid numbers follow a similar pattern except where translated regions overlap in which case the nearest right and first amino acids are both identified. Stop codons (★) are indicated by asterisks. **Ribosomal binding sites** are shown in bold, *AT-rich inverted repeats* in italics, AT-rich regions are boxed, and hairpin structures are underlined. Primers are highlighted and appropriately identified with the direction of extension. Cloning sites for library construction are also shown.

1 TGCCTGGCTACAACTCCCGCCGGCGGACGTTCCGGTTCGCGAATACTGACAGGGTGACCG
 ACGGACCGATGTTGAGGGCGGCCGCTGCAAGGCCAAGCGCTTATGACTGTCCCACTGGC

 61 GCGCGGCTGGCGCCGCGCCTTAGCTCCGTGGGTGGCCTGCCAGGCTGGGGGGCTAGCCTG
 CGCGCCGACCGCGGCGCGGAATCGAGGCACCCACCGGACGGTCCGACCCCCGATCGGAC

 121 GCCTGCGGGTCATTTCTTCGCGACTTTGATGTCGGTGTCTTCCAGGTCCCAGGTGTCGGC
 CGGACGCCCAGTAAAGAAGCGCTGAACTACAGCCACAGAAGGTCCAGGGTCCACAGCCG
 543 * K K A V K I D T D E L D W T D A

 181 CGTCTTGCCCTCGTCGCGCAACTGATATTTACGACGCGTCCCTGCGGGTTCTTCGGCAG
 GCAGAACGGGAGCAGCGCTTGACTATAAAGTCGTGCGCAGGGACGCCCAAGAAGCCGTC
 527 T K G E D R L Q Y K L V R G Q P N K P L

 241 GCTGTCGCGGAACCTCGATATAGCGCGCAGGGCGTAGTAGGGCACCGAGTCGGTCGCCCA
 CGACAGCGCCTTGAGCTATATCGCGCCGTCCCGCATCATCCCGTGGCTCAGCCAGCGGGT
 507 S D R F E I Y R P L A Y Y P V S D T A W

 301 GTGGAACAGTCTCTCCGGCGTCAGTCGGGCCCCCTGGTGCAGCACGGCGGTACCTTGAC
 CACCTTGTCGAGGAGGCCGCGAGTCAGCCCGGGGGACCACGTCGTGCCGCCAGTGGAAGT
 487 H F L E E P T L R A G Q H L V A T V K V

 361 GTCATCCTCGCCCTTGTCGGAGGGCACGGCATGCACCGCAACCTCGGCGAGGTCCGGATG
 CAGTAGGAGCGGGAACAGCCTCCCGTGCCGTACGTGGCGTTGGAGCCGCTCCAGGCCTAC
 467 D D E G K D S P V A H V A V E A L D P H

 421 GACGGCGAAGGCCGCTTCCATCTCGAAGCTGGAGATGTTCTCGCCACGCCGGCGCAGGTA
 CTGCCGCTTCCGGCGAAGGTAGAGCTTCGACCTCTACAAGAGCGGTGCGGCCGCGTCCAT
 447 V A F A A E M E F S S I N E G R R R L Y

 481 GTCCTTCTTGCGGTGCGACGAAATAGAAGAAGCCGTCATCATCGAACTTGCCGATGTCGCC
 CAGGAAGAACGCCAGCTGCTTTATCTTCTTCGGCAGTAGTAGCTTGAACGGCTACAGCGG
 427 D K K R D V F Y F F G D D D F K G I D G

▼ pBW30::Tnp^{hoA'} -1

```

541  AGTGTGGAACACATGTTGCGCATCAGCTTGAGGGTATCCGCGGGACGGTTCCAGTAGCC
      TCACACCTTGGTGTACAACGCGTAGTCGAACTCCCATAGGCGCCCTGCCAAGGTCATCGG
407    T H F W M N R M L K L T D A P R N W Y G

601  CTGGAACATGATGTCCGGTCGCAGTGGACGCAGCACGATTTACCGGGGGTATTGGCTGG
      GACCTTGTACTIONACAGGCCAGCGTCACCTGCGTCGTGCTAAAGTGGCCCCCATAACCGACC
387    Q F M I D P R L P R L V I E G P T N A P

661  TAGTTCCTGGTCGTTGTGCTCGACGATGCGCACGTGCGAAATCCGGGATGCGCTTGCCGGA
      ←-----25b/f-5                                CM7-----→
367  ATCAAGGACCAGCAACAGCAGCTGCTACGCGTGCAGCTTTAGGCCCTACGCGAACGGCCT
      L E Q D N D D V I R V D F D P I R K G S

721  AGAACCGGGAGCCGCATACTCGCCCGCCGCCAGGGAGGTGATCACACAGGCTTCGGTCAG
      TCTTGGCCCTCGGCGTATGAGCGGGCGGCGGTCCCTCCACTAGTGTGTCCGAAGCCAGTC
347    S G P A A Y E G A A L S T I V C A E T L

781  GCCGTAGCCGTTGCCACCGACCAGTTTGGTGCCGAAGCGCTCGCGCCAGATCTGTTTGGT
      CGGCATCGGCAACGGTGGCTGGTCAAACCACGGCTTCGCGAGCGCGGTCTAGACAAACCA
327    G Y G N G G V L K T G F R E R W I Q K T

841  TTCTTCGGTGTAGGGGTTGCCGCGTGCGGTATGGATCTGCCCATGGCAGCGTTTCATCGC
      AAGAAGCCACATCCCCAACGGCGCACGCCATACCTAGACGGGTACCGTCGCAAAGTAGCG
307    E E T Y P N G R A T H I Q G H C R K M A

901  GTCGTTGTCCGGCGCCTGGGCCAGCAGGCCGCCCATGCCCCCAGGATCGAGGCGATGGT
      CAGCAACAGGCCGCGGACCCGGTTCGTCCGGCGGGTACGGGGGGTCCCTAGCTCCGCTACCA
287    D N D P A Q A L L G G M G G L I S A I T

961  CGCTCCGGAACGCTCGACTTCCTGCCAGAAGTTGGAGACCGAGAAGCGCGGCAGGATGGC
      GCGAGGCCTTGCGAGCTGAAGGACGGTCTTCAACCTCTGGCTCTTCGCGCCGTCCTACCG
267    A G S R E V E Q W F N S V S F R P L I A

```

1021 GGCCCGAGCACCAACCAGGATGCTGGCGATGATGCTGACGCACAGGGCGTTCATGTGGAA
 CCGGGCTCGTGGTTGGTCCTACGACCGCTACTACGACTGCGTGTCCTCGCAAGTACACCTT
 247 A R A G V L I S A I I S V C L A N M H F

1081 CAGCGGCAGTGGCGTGATGGTTACGTCGTTTTTCGTTGGCCGGGCGGCGCGTAGTTGCAG
 CM8-----→
 GTCGCCGTCAACGCACTACCAATGCAGCAAAAGCAACCGGCCCGGCGCGCATCAACGTC
 227 L P L P T I T V D N E N A P G A R L Q L

1141 GCGAGCCAGGTTGCACATGAAGTTGTAGCTGATCATGCAGCCCTTCGACGGGCGGTTGGT
 CGCTCGGTCCAACGTGTACTTCAACATCGACTAGTACGTCGGGAAGCTGCCCCGGCCACCA
 207 R A L N C M F N Y S I M C G K S P G T T

1201 GCCGGAGGTGTAGATCAGGCAGGCAAGGTCGGAGGGTTGTGGCTTGCTGGTAAAAGGCGT
 ←-----24b/f-4
 CGGCCTCCACATCTAGTCCGTCCGTTCCAGCCTCCCAACACCGAACGACCATTTTCCGCA
 187 G S T Y I L C A L D S P Q P K S T F P T

1261 GTCGTCCTCGCCACGGAACCTCGTCCAGGGCGGCGATCGGAATCCTGCAACTGGCTGGCTC
 CAGCAGGAGCGGTGCCTTGAGCAGGTCCCGCCGCTAGCCTTAGGACGTTGACCGACCGAG
 167 D D E G R F E D L A A I P I R C S A P E

1321 GGCCCGTTCAACACGATAGAGGATTTGGCTGACCTCGCTGAGCTGCTCGGCTAGCGGGAG
 CCGGGCAAGTGGTGCTATCTCCTAAACCGACTGGAGCGACTCGACGAGCCGATCGCCCTC
 147 A R E G R Y L I Q S V E S L Q E A L P L

1381 GATGCGCGGTAAGTAGTCGGCCTCGCAGATGACCAGCTGGGTTCCCGTGTCGGCGATCTG
 CTACGCGCCATTTCATCAGCCGGAGCGTCTACTGGTCGACCCAAGGGCACAGCCGCTAGAC
 127 I R P L Y D A E C I V L Q T G T D A I Q

1441 GTGGCGCAGGAACTCGCCTTTGAGTGCGGTGTTGATCGGTACGCTGACCGCGCGGAGCTT
 CM9.1-----→
 CACCGGTCCTTGAGCGGAAACTCACGCCACAAGTCCATGCGACTGGCGCGCCTCGAA
 107 H R L F E G K L A T N I P V S V A R L K

1501 GTTGATGGCCAGCCAGCAGACGACCGCGTCTATATTGTTGTCCAGCATGGTCAGCACGGT
 CAACTACCGGTCGGTCGTCTGCTGGCGCAGATATAACAACAGGTCGTACCAGTCGTGCCA
 87 N I A L W C V V A D I N N D L M T L V T

1561 CTCGCCCCGCTTCACGCCCCAAGGCGGCCAGCGCATGAGCCATGCGTGTGATAGCGAGTC
 GAGCGGGCCGAAGTGC GG GTTCGCGGTACTCGGTACGCACAGCTATCGCTCAG
 67 E G P K V G L A A L A H A M R T S L S D

1621 AACTTCCGAATATGTGTATAGATTTCCACTGAAATCCAGCAGGATCCTGTCTGGATGCCT
 TTGAAGGCTTATACACATATCTAAAGGTGACTTTAGGTCGTCCTAGGACAGACCTACGGA
 47 V E S Y T Y L N G S F D L L I R D P H R

1681 TGCAACCGCTCGGTGAGTGCTGCGATGACGGTGTCTCTCTCGCCTACGGACCATGTGTC
 ACGTTGGCGAGCCAGCTCACGACGCTACTGCCACAGAGAGCGGATGCCTGGTACACAG
 27 A V A R D L A A I V T D R E G V S W T D

1741 CGGATGACCGATGCCACTCATAGCTCATCTCCTTTGGAGGCCTTCCCCGCTGTGGGATGG
 ←-----25b/f-3
 GCCTACTGGCTACGGTGAGTATCGAGTAGAGGAAACCTCCGGAAGGGGCGACACCCTACC
 7 P H G I G S M
 ←tdtL

1801 CAAATTTTATTTTGGTGGGAGAATGCCATTCTAAATTCAAATTTGTCACTTGTCTCCAGC
 GTTTAAATAAAAACCACCCTCTTACGGTAAGATTTAAGTTTAAACAGTGAACAGAGGTGCG

1861 TAAGATTTGAACAAATCCATGTGGTACTGCCTATATCCTGTGCTAATGTGACGAAAATTT
 ATTCTAAACTTGTTTAGGTACACCATGACGGATATAGGACACGATTACACTGCTTTTAAA

1 tdtR→
 M S R N A T K S S K T E P
 1921 ACCAACTGAGAAATCCCTATCAATGTCAAGGAATGCTACGAAGTCCAGCAAAACAGAGCCT
 CM1-----→
 TGGTTGACTCTTAGGGATAGTTACAGTTCCTTACGATGCTTCAGGTCGTTTTGTCTCGGA

14 I A P A R R K A A S A E K P S T A V K Q
 1981 ATCGCTCCTGCCAGGAGGAAAGCCGCCTCTGCAGAGAAGCCTTCCACTGCGGTAAACAG
 TAGCGAGGACGGTCCTCCTTTTCGGCGGAGACGTCTCTTCGGAAGGTGACGCCAATTTGTC

34 A V Q K N G R A S S V A G V K V K P N R
 2041 GCTGTTCAGAAAAATGGCCGCGCGTCATCGGTGGCCGGAGTCAAGGTAAAGCCCAACCGG
 CGACAAGTCTTTTTACCGGCGCGCAGTAGCCACCGGCCTCAGTTCCATTTCGGGTTGGCC

54 A V Q K N G R A S S V A G V K V K P N R
 2101 CGCTCCGAGGAGACCATTGGCAATATTCTCGCCGCCACCGAAGAGGTGGTGTCTGCGCTCT
 GCGAGGCTCCTCTGGTAACCGTTATAAGAGCGGCGGTGGCTTCTCCACCACGACGCGAGA

 74 G A D R I S I L D V C Q V A G V S R G T
 2161 GGGGCGGACAGGATCTCGATCCTGGATGTCTGTCTCAGGTGGCGGGGGTGTGCGGGGGCACC
 CCCC GCCTGTCTAGAGCTAGGACCTACAGACAGTCCACCGCCCCACAGCGCCCCGTGG

 94 F Y R Y F S S Q E D L L D T F S R H K R
 2221 TTCTATCGTTACTTCTCATCCCAGGAAGACCTTCTGGACACCTTCTCCCGGCACAAGCGG
 AAGATAGCAATGAAGAGTAGGGTCCTTCTGGAAGACCTGTGGAAGAGGGCCGTGTTGCC

 114 D S F H S S L S Q A L E D I A D P D A R
 2281 GACAGCTTCCATAGCTCTCTGTCCCAGGCCCTGGAGGATATAGCCGATCCTGACGCACGT
 ←-----25b/f-2
 CTGTCTGAAGGTATCGAGAGACAGGGTCCGGGACCTCCTATATCGGCTAGGACTGCGTGCA

 134 L N A L I A Y L D N Y L E Q S R A R R L
 2341 CTGAATGCGCTGATCGCCTACCTCGACAACCTATCTCGAGCAGAGTCGCGCACGTCTGTCTG
 CM15-----→
 GACTTACGCGACTAGCGGATGGAGCTGTTGATAGAGCTCGTCTCAGCGCGTGCAGCAGAC

 154 L V V A P E Y A L G F F R R I F H D S I
 2401 CTGGTGGTGGCGCCTGAGTATGCGCTGGGCTTTTCCGGCGCATTTTCCATGACTCCATC
 GACCACCACCGCGGACTCATACGCGACCCGAAAAAGGCCGCGTAAAAGGTACTGAGGTAG

 174 V R F Q D L L G I V F D E W D A R L G V
 2461 GTGCGCTTTTCAGGATCTGCTAGGCATTGTCTTCGACGAGTGGGACGCTCGGCTCGGAGTC
 CACGCGAAAGTCCTAGACGATCCGTAACAGAAGCTGCTCACCCTGCGAGCCGAGCCTCAG

 194 K L D R E L I C E M V I R Y V L S E I L
 2521 AAGCTCGATCGTGAACCTGATTGCGAAATGGTCATCCGCTATGTGCTGAGCGAGATCCTT
 TTCGAGCTAGCACTTGACTAAACGCTTTACCAGTAGGCGATACACGACTCGCTCTAGGAA

 214 V P E G S G R R L L P V R V R K L V A A
 2581 GTGCCGGAGGGTAGCGGTGCGCGTCTTCTTCTGTCCGTGTGAGAAAGCTCGTTGCTGCT
 CACGGCCTCCCATCGCCAGCGGCAGAAGAAGGACAGGCACACTCTTTTCGAGCAACGACGA

 234 L T I G S T S R S R R *
 2641 CTGACCATCGGTTCCACCTCCCGCTCTCGCCGCTGATCCGCCCCGCTTCCCGGCTGCCCT
 GACTGGTAGCCAAGGTGGAGGGCGAGAGCGGCGACTAGGCGGGCGCAAGGGCCGACGGGA

2701 CCGAGTGAGGGTGTCCGCTCTCCTTTCTTCTCGCGTCACACCCCGTCCCCCCTTTCCTT
 GGCTCACTCCCACAGGCGAGAGGAAAGAAGGAGCGCAGTGTGGGGCAGGGGGGAAAGGAA ←----

2761 CGTCCTGAACCTCCGTCCGGCTGGTTGCTCCGCCGAGGCCTGCGCGTGTCTGCTGCAATG
 -----25b/f-1
 GCAGGACTTGGAGGCAGGCCGACCAACGAGGCGGCTCCGGACGCGCACAGACGACGTTAC

2821 CTCTGGAATGGGCTTTTCCGTCTGTTTCAAAAATAAATTGAACAGATTTTGTATTATGA
 GAGACCTTACCGAAAAGGCAGACAAAGTTTTTATTTAACTTGTCTAAAACTAAATACT

2881 TCATTTTAACTTATTCTCTGCCTTGAAGCAGGCTGTCTGTCACCCCTTGTGTGCGTGTA
 -----cm17-----→
 AGTAAAATTTGAATAAGAGACGGAACTTCGTCCGACAGCACGTGGGGAACACACGCACAT

2941 CGGC TGGAAATA CGCGCTGCTAGCGCGAC ATTACACAACCACAAAAAATGT CACCTCAGA
 GCCG ACCTTTAT GCGCGACGATCGCGCTG TAATGTGTTGGTGT TTTTACAGTGGAGTCT

tdtS⇒
 1 M L I D L H A H A P H P D Y Y D Q
 3001 GGTACCGGACCATGCTCATCGATCTTCACGCTCATGCGCCTCATCCCGACTACTACGACC
 CCATGGCCTGGTACGAGTAGCTAGAAAGTCCGAGTACGCGGAGTAGGGCTGATGATGCTGG

18 H P Y W G P A F E S Q P D G D I K L R V
 3061 AGCATCCCTATTGGGGGCCTGCTTTTCGAGTCCCAGCCCGATGGCGATATCAAGCTGCGTG
 TCGTAGGGATAACCCCGGACGAAAGCTCAGGGTCGGGCTACCGCTATAGTTCGACGCAC

38 G D W I L T L G A P E R K K A L R E A H
 3121 TAGGCGATTGGATCCTGACCCTGGGGGCACCGGAGCGTAAGAAGGCGCTGCGCGAGGCGC
 ↑ *Bam*HI
 ATCCGCTAACCTAGGACTGGGACCCCGTGGCCTCGCATTCTTCCGCGACGCGCTCCGCG

66 A R G E T L N V A E Y M A K W R D P N H
 3181 ATGCCCCGTGGCGAAACGCTGAATGTGCGCGAGTACATGGCCAAGTGGCGTGACCCGAACC
 TACGGGCACCGCTTTGCGACTTACAGCGGCTCATGTACCGGTTACCGCACTGGGCTTGG

86 R L A A M D A A G Q S A Q V L S V P S H
 3241 ACCGCCTCGCGGCAATGGATGCTGCTGGCCAGAGCGCTCAGGTGCTGTGCGTGCCGAGCC
 TGGCGGAGCGCCGTTACCTACGACGACCGGTCTCGCGAGTCCACGACAGCCACGGCTCGG

106 C Y M Y W A E A E F A V P F A R K V N D
3301 ATTGCTACATGTACTGGGCCGAGGCCGAGTTCGCTGTGCCCTTCGCCCCGAAGGTCAACG
CM19-----
TAACGATGTACATGACCCGGCTCCGGCTCAAGCGACACGGGAAGCGGGCGTTCCAGTTGC

126 V L A G Y C S A A P N R L M F W A H A P
3361 ACGTGCTGGCCGATATTGCTCGGCCGCGCCCAATCGCCTGATGTTCTGGGCTCATGCGC
→
TGCACGACCGGCCTATAACGAGCCGGCGCGGGTTAGCGGACTACAAGACCCGAGTACGCG

146 L N V P H E A A K E I R R A C T E L G A
3421 CATTGAACGTACCGCATGAGGCGGCCAAGGAAATCCGCCGCGCCTGTACCGAGCTGGGGG
GTAACCTTGCATGGCGTACTCCGCCGGTTCCTTTAGGCGGCGCGGACATGGCTCGACCCCC

166 K G L V A G G S N F G G M E F D S P E L
3481 CCAAGGGGCTGGTGGCCGGCGGTTCCAACCTTCGGTGGCATGGAGTTCGATTCCCCGAGC
GGTTCCCCGACACCGGCCGCAAGGTTGAAGCCACCGTACCTCAAGCTAAGGGGGCTCG

186 Y P V W E A L C D L D L P I F V H G Y N
3541 TGTATCCGGTCTGGGAGGCGCTGTGCGATCTCGACCTGCCGATCTTCGTGCACGGCTATA
ACATAGGCCAGACCCTCCGCGACACGCTAGAGCTGGACGGCTAGAAGCACGTGCCGATAT

206 Q S V T W G K Q A N T D R F E T T A I V
3601 ACCAGTCCGTGACCTGGGGCAAGCAAGCCAATACCGACCGCTTCGAAACCACCGCCATCG
TGGTCAGGCACTGGACCCCGTTCGTTTCGGTTATGGCTGGCGAAGCTTTGGTGGCGGTAGC

226 G M N Y D E T K C L W Y L I N G G V L D
3661 TCGGCATGAACTACGATGAAACCAAGTGCCTCTGGTACCTGATCAATGGCGGTGTACTGG
AGCCGTACTTGATGCTACTTTGGTTCACGGAGACCATGGACTAGTTACCGCCACATGACC

246 R F P N L K V Y I T H G G G F A P Y Q L
3721 ACCGCTTCCCGAACCTGAAGGTCTACATACCCATGGCGGCGGCTTCGCTCCCTACCAGT
TGGCGAAGGGCTTGGACTTCCAGATGTAGTGGGTACCGCCGCCGAAGCGAGGGATGGTCA

266 G R L A Q T N P N L D V Y H N K K P F L
3781 TGGGACGCCTGGCGCAGACCAACCCGAACCTCGACGTCTATCACAACAAGAAGCCTTTCC
ACCCTGCGGACCGCGTCTGGTTGGGCTTGGAGCTGCAGATAGTGTGTTCTTCGGAAAGG

286 D Y L S N F Y F D V E L H E V P M R Q A
3841 TGGACTATCTGTCCAACCTTCTACTTCGACGTGGAGTTGCATGAGGTGCCCATGCGCCAGG
ACCTGATAGACAGGTTGAAGATGAAGCTGCACCTCAACGTACTCCACGGGTACGCGGTCC

306 L V D I I G A D R I L Y G S N F G G S D
 3901 CCCTGGTGCACATCATCGGTGCGGACCGGATTCTCTACGGTTCCAATTTTCGGTGGCAGCG
 ←-----CM26-----CM21-----→
 GGGACCAGCTGTAGTAGCCACGCCTGGCCTAAGAGATGCCAAGGTTAAAGCCACCGTCGC

326 A V R H D L T E G L R L S D E D L Q K I
 3961 ATGCCGTGCGCCATGACCTGACCGAAGGGCTGCGCCTGTCGGACGAGGATCTGCAGAAGA
 TACGGCACGCGGTACTGGACTGGCTTCCCGACGCGGACAGCCTGCTCCTAGACGTCTTCT

346 R W K N A C E L L H I D P A K L G K V G
 4021 TCCGCTGGAAGAATGCTTGCGAGCTGCTGCACATCGATCCGGCGAAATTGGGCAAGGTAG
 AGGCGACCTTCTTACGAACGCTCGACGACGTGTAGCTAGGCCGCTTTAACCCGTTCCATC

tdtA→
 366/1 A * M K L A R C A Y F E T G S F W A V V
 4081 GTGCCTGATGAACTGGCACGCTGCGCTTATTTTCGAGACTGGCTCGTTCTGGGCGGTGGT
 CACGGACTACTTTGACCGTGCGACGCGAATAAAGCTCTGACCGAGCAAGACCCGCCACCA

19 D P E R D E V H A I A G E F S S W A P A
 4141 CGATCCTGAGCGGGACGAGGTGCATGCCATCGCCGGCGAGTTTTCTCCTGGGCGCCTGC
 GCTAGGACTCGCCCTGCTCCACGTACGGTAGCGGCCGCTCAAAAGGAGGACCCGCGGACG

39 V A R G A G I S A L R L S G S A L P L S
 4201 CGTGGCCCGTGGGGCCGAATCAGTGCGCTACGGCTTTCCGGTTCGGCGCTGCCGTTGTC
 GCACCGGGCACCCCGGCCTTAGTCACGCGATGCCGAAAGGCCAAGCCGCGACGGCAACAG

59 K V R L L P P I D P V N R V V V A G A N
 4261 CAAGGTGCGCTTGCTTCCACCGATCGATCCGGTGAACCGGGTGGTGGTGGCAGGGGCCAA
 GTTCCACGCGAACGAAGGTGGCTAGCTAGGCCACTTGGCCCACCACCACCGTCCCCGGTT

79 Y A K H L A E D F G L A S P P Q P V A F
 4321 CTACGCCAAGCACCTGGCCGAGGACTTCGGCCTTGCCAGCCCGCCACAGCCAGTAGCCTT
 GATGCGGTTTCGTGGACCGGCTCCTGAAGCCGGAACGGTTCGGGCGGTGTCGGTCATCGGAA

99 L K A Y G A L I G A H D P I R Y P P L T
 4381 CCTGAAGGCCTACGGCGCTCTGATCGGCGCCCATGATCCGATCCGCTATCCGCCCTTGAC
 ←-----CM24-----CM23-----→
 GGACTTCCGGATGCCGCGAGACTAGCCGCGGGTACTAGGCTAGGCGATAGGCGGGAAGT

119 E E L D H E V E L V A V I G S A E I D L
 4441 CGAGGAGCTCGACCATGAGGTCGAGCTGGTGGCGGTAATCGGCAGTGCCGAGATCGATCT
 -----→
 GCTCCTCGAGCTGGTACTCCAGCTCGACCACCGCCATTAGCCGTCACGGCTCTAGCTAGA

139 E D P L A C V L G Y T V G N D V S A R D
 4501 GGAAGACCCGCTGGCCTGCGTGCTGGGATACACGGTCGGCAACGACGTCAGTGCGCGCGA
 CCTTCTGGGCGACCGGACGCACGACCCTATGTGCCAGCCGTTGCTGCAGTCACGCGCGCT

159 L Q R S G P A G I G M D L F A A K S Q D
 4561 CCTGCAGCGTAGCGGCCCGGCGGTATCGGCATGGACCTGTTGCGGGCCAAGAGCCAGGA
 GGACGTCGCATCGCCGGGCGGCCATAGCCGTACCTGGACAAGCGCCGGTTCTCGGTCCT

179 R T T G L G P W I V T K D E F P V G S P
 4621 CCGCACCCTGGCCTGGGGCCCTGGATCGTCACCAAGGACGAGTTCCCCGTCGGTTCGCC
 GGCGTGGTGACCGGACCCCGGGACCTAGCAGTGGTTCTGCTCAAGGGGCAGCCAAGCGG

199 K L R L T L K V N G E I R Q D G S T A E
 4681 CAAGCTGCGCCTGACCCTCAAGGTGAATGGCGAGATCCGCCAGGACGGCTCGACCGCCGA
 GTTCGACGCGGACTGGGAGTTCCACTTACCGCTCTAGGCGGTCTGCCGAGCTGGCGGCT

219 M T W D V G Q L V R F V Q Q R S S F A C
 4741 AATGACCTGGGATGTGCGTCAACTGGTCCGCTTCGTCCAGCAACGTTCCAGCTTCGCGCTG
 TTACTGGACCCTACAGCCAGTTGACCAGGCGAAGCAGGTCGTTGCAAGGTGCAAGCGGAC

239 G D V L F T G S P A G V G M G T G L F L
 4801 TGGCGACGTGTTGTTACCGGCTCACC GGCTGGCGTCGGCATGGGGACGGGACTGTTCTCT
 ACCGCTGCACAACAAGTGGCCGAGTGGCCGACCGCAGCCGTACCCCTGCCCTGACAAGGA

259 N P G D V V E A T I E G I G T L R N V V
 4861 CAACCCGGGGGACGTGGTCGAGGCGACCATCGAAGGAATAGGCACGCTGCGCAACGTGGT
 GTTGGGCCCCCTGCACCAGCTCCGCTGGTAGCTTCCTTATCCGTGCGACGCGTTGCACCA

284/1 S E K S R V * *tdtB*→ M S I D K E E K
 4921 GAGCGAGAAATCCCGCGTCTAAATCGAGAGCGAAACATGAGCATTGATAAAGAAGAAAAA
 -----CM22
 CTCGCTCCTTTAGGGCGCAGATTTAGCTCTCGCTTTGTA CTGTA ACTATTTCTTCTTTT

9 I V V V G A G L M G T G I A H G F A S S
 4981 ATCGTAGTGTTAGGTGCTGGCCTGATGGGGACCGGCATTGCGCACGGATTGCCAGTTCC
 TAGCATCACCATCCACGACCGGACTACCCCTGGCCGTAACGCGTGCCTAAGCGGTCAAGG

29 G Y P T V L V D T N A E S L A R A R D N
 5041 GGCTATCCCACCGTGCTGGTGGACACCAATGCGGAGTCGCTGGCCCGTGCCAGGGACAAT
 CCGATAGGGTGGCACGACCACCTGTGGTTACGCCTCAGCGACCGGGCACGGTCCCTGTTA

49 I G K I L S D G V A L G K V S A E S A E
 5101 ATCGGCAAGATCCTCAGCGACGGTGTGGCGCTCGGTAAGGTTTCCGCCGAGTCGGCAGAG
 TAGCCGTTCTAGGAGTCGCTGCCACACCGCGAGCCATTCCAAAGGCGGCTCAGCCGTCTC

 69 A S L A R L V T S G D L S E A A R G A D
 5161 GCGTCCCTGGCGCGCCTGGTCACCACTGGCGACCTGAGCGAGGCGGCGCGGGGCGCTGAC
 CGCAGGGACCGCGCGGACCAGTGGTCACCGCTGGACTCGCTCCGCCGCGCCCCGCGACTG

 89 W L V E T V S E K L E V K R A V V A Q A
 5221 TGGCTGGTGGAAACGGTCTCCGAGAAGCTGGAAGTGAAGAGGGCGGTGGTCGCCCAGGCC
 ACCGACCACCTTTGCCAGAGGCTCTTCGACCTTCACTTCTCCCGCCACCAGCGGGTCCGG

 109 E P L L A P S A I I A T N T S A L S V T
 5281 GAGCCCCTGCTGGCGCCCAGCGCGATCATTGCCACCAACACCTCGGCGCTCAGCGTGACC
 CTCGGGGACGACCGCGGGTCGCGCTAGTAACGGTGGTTGTGGAGCCGCGAGTCGCACTGG

 129 E I A A S L A A P E R V I G M H F F N P
 5341 GAGATCGCCGCATCGCTGGCCGCTCCCGAGCGGGTCATCGGCATGCACTTCTTCAACCCG
 CTCTAGCGGCGTAGCGACCGGCGAGGGCTCGCCCAGTAGCCGTACGTGAAGAAGTTGGGC

 149 V H K M K L V E L V R G L A T S D E V L
 5401 GTGCACAAGATGAAGCTGGTGGAGCTGGTGCCTGGCTTGGCGACCAGCGACGAAGTGCTG
 ←-----CM20
 CACGTGTTCTACTTCGACCACCTCGACCACGCACCGAACCGCTGGTCGCTGCTTCACGAC

 169 V R T R A L C D A M G K T S I V V N E S
 5461 GTGCGAACGCGGGCGCTGTGCGACGCCATGGGCAAGACCTCGATTGTGGTTAACGAGTCC
 CACGCTTGCGCCCGCGACACGCTGCGGTACCCGTTCTGGAGCTAACACCAATTGCTCAGG

 189 P G L T T S R M S A L L G N E A M Y M L
 5521 CCCGGGTTGACCACCAGTCGCATGTCCGCGCTGCTCGGCAACGAGGCCATGTACATGCTG
 GGGCCCAACTGGTGGTCAGCGTACAGGCGCGACGAGCCGTTGCTCCGGTACATGTACGAC

 209 Q E G T A S A E D I D T A L R L G F N H
 5581 CAAGAGGTACGGCCAGCGCCGAGGACATCGATAACGCCCTGCGTCTGGGTTTCAACCAT
 GTTCTCCCATGCCGGTCGCGGCTCCTGTAGCTATGGCGGGACGCAGACCCAAAGTTGGTA

 229 P M G P L E L G D L T G W D T R L S V L
 5641 CCCATGGGGCCCTTGGAGCTGGGTGACCTCACCGGCTGGGATAACCGCTGTGCGGTGCTG
 GGGTACCCCGGGAACCTCGACCCACTGGAGTGGCCGACCCTATGGGCGAACAGCCACGAC

249 R Y L H Q T L G E K F R P C P L I I K M
 5701 CGCTACCTGCATCAGACCCTGGGCGAGAAATCCGTCCTGCCCCTGATCATCAAGATG
 GCGATGGACGTAGTCTGGGACCCGCTCTTTAAGGCAGGGACGGGCGACTAGTAGTTCTAC

269 V A A G R L G R K T G H G V Y R Y E D G
 5761 GTTGCCGCTGGTCGCCTGGGACGCAAGACCGGACATGGCGTCTATCGCTACGAAGACGGC
 CAACGGCGACCAGCGGACCCTGCGTTCTGGCCTGTACCGCAGATAGCGATGCTTCTGCCG

1 K R V P G S G L K A S A L/M K D I V I V D
 5821 AAGCGGGTACCGGGTTCCGGGCTCAAGGCGAGTGCCTATGAAAGACATCGTGATCGTCG
 TTCGCCCATGGCCCAAGGCCGAGTTCCGCTCACGCGATACTTTCTGTAGCACTAGCAGC

9 A L R T P I G R F R G S L A G V R A D H
 5881 ATGCGCTGCGTACGCCGATCGGCCGTTCCGTGGCAGCCTGGCTGGTGTGCGTGCCGACC
 ←-----Cm18
 TACGCGACGCATGCGGCTAGCCGGCCAAGGCACCGTCGGACCGACCACACGCACGGCTGG

29 L G A L V I G K L L E R V D L S P A Q V
 5941 ACCTGGGCGCGCTGGTGATCGGCAAGCTGCTGGAGCGCGTTGATCTGTGCCCCGCGCAGG
 TGGACCCGCGCGACCACTAGCCGTTGACGACCTCGCGCAACTAGACAGCGGGCGCGTCC

49 D D V I F G C V T Q I G E Q S A N I A R
 6001 TCGATGACGTGATTTTCGGCTGTGTCACCCAGATCGGCGAGCAGTCGGCCAATATCGCGC
 AGCTACTGCACTAAAAGCCGACACAGTGGGTCTAGCCGCTCGTCAGCCGGTTATAGCGCG

69 T A L L G A G W P V T I P G L T I D R K
 6061 GCACCGCACTGCTTGGCGCTGGCTGGCCGGTAACCATTTCCCGGCTGACCATCGACCGCA
 CGTGGCGTGACGAACCGCGACCGACCGGCCATTGGTAAGGGCCGGA CTGGTAGCTGGCGT

89 C G S G E V A V H L A A G A I A A G S A
 6121 AATGCGGCTCTGGCGAGGTGGCGGTGCATCTCGCCGCCGGCGCTATCGCGGCCGGCTCCG
 TTACGCCGAGACCGCTCCACCGCCACGTAGAGCGGCGGCCGCGATAGCGCCGGCCGAGGC

109 D I V V A G G A E N M S R V P M G S N R
 6181 CCGACATCGTGGTCGCCGGTGGGGCGGAGAACATGAGCAGGGTGCCCATGGGCAGCAACC
 GGCTGTAGCACCGCGGCCACCCCGCCTCTTGTA CTGCTCCACGGGTACCCGTCGTTGG

129 E I H G E A F G W M A A E R Y E L T S Q
 6241 GCGAGATCCATGGCGAAGCCTTCGGCTGGATGGCGGCCGAGCGCTACGAGCTGACCAGCC
 CGCTCTAGGTACCGCTTCGGAAGCCGACCTACCGCCGGCTCGCGATGCTCGACTGGTCGG

149 G E A A E R M A D K V S L S R D E L D D
 6301 AGGGCGAGGCCGCCGAGCGTATGGCTGACAAGTGGTCGCTCAGCCGTGACGAGCTGGACG
 TCCCGCTCCGGCGGCTCGCATACCGACTGTTCCACACGGAGTCGGCACTGCTCGACCTGC

169 Y A F A S H S R A A A A A D A G H F A N
 6361 ACTATGCCTTCGCCAGCCATAGCCGCGCGGCGGCTGCGGCGGATGCCGGGCACTTCGCCA
 ←-----CM16
 TGATACGGAAGCGGTTCGGTATCGGCGCGCCGCCGACGCCGCCTACGGCCCGTGAAGCGGT

189 E I V E V P V E T L R E H G L S E P V A
 6421 ATGAAATCGTCGAGGTGCCGGTCGAAACCCTGCGTGAGCACGGCCTGAGCGAACCTGTGG
 TACTTTAGCAGCTCCACGGCCAGCTTTGGGACGCACTCGTGCCGGACTCGCTTGGACACC

209 P L R G D E T I R R D T S R E K L S T L
 6481 CGCCACTGCGTGGCGACGAACTATTCGCCGCGATACTTCACGGGAAAACTGTCGACTC
 GCGGTGACGCACCGCTGCTTTGATAAGCGGCGCTATGAAGTGCCCTTTTTGACAGCTGAG

229 K T S F R P E T G R I T A G N S S Q I S
 6541 TGAAGACCAGCTTTGCCCCGAAACCGGGCGCATCACCGCTGGTAACTCCTCACAGATTT
 ACTTCTGGTCGAAAGCGGGCCTTTGGCCCGCTAGTGGCGACCATTGAGGAGTGTCTAAA

249 D G A A A L L L M S A E T A G R L G L R
 6601 CCGACGGCGCGGCGGCGTTGCTGCTGATGTGCGCGGAAACGGCGGGGCGCCTGGGGCTCA
 GGCTGCCGCGCCGCCGCAACGACGACTACAGCCGCCTTTGCCGCCCCGCGGACCCCGAGT

269 A R A R I R A F T T V G S D P T L M L T
 6661 GGGCGCGCGCGCGGATTTCGTGCCTTCACCACCGTAGGCTCCGATCCGACGCTGATGCTCA
 CCCGCGCGCGGCCTAAGCACGGAAGTGGTGGCATCCGAGGCTAGGCTGCGACTACGAGT

289 G R S P L L Q V L E K A G L G I D D I D
 6721 CCGGCCGATCGCCGCTACTACAAAGTGTGGAGAAGGCCGCTGGGTATCGACGACATCG
 ←-----CM14
 GGCCGGCTAGCGGCGATGATGTTACGACCTCTTCCGGCCGGACCCATAGCTGCTGTAGC

309 L F E V N E A F A S V P L A W M C E T G
 6781 ACCTGTTTCGAGGTCAACGAGGCCTTTGCTTCTGTACCGCTGGCCTGGATGTGCGAAACTG
 TGGACAAGCTCCAGTTGCTCCGGAACGAAGACATGGCGACCGGACCTACACGCTTTGAC

329 V S H E R L N V N G G A I A L G H P L G
 6841 GCGTGTGCGACGAGCGGCTCAACGTCAATGGCGGGGCTATCGCCCTGGGGCACCCACTCG
 CGCACAGCGTGCTCGCCGAGTTGCAGTTACCGCCCCGATAGCGGGACCCCGTGGGTGAGC

349 A S G A R L M T T L L N E L E R R G A R
6901 GTGCCAGCGGCGCGCGGTTGATGACCACGCTGCTTAATGAGTTGGAGCGGCGCGGCGCCC
CACGGTCGCCGCGCGCCAACACTACTGGTGCACGAATTACTCAACCTCGCCGCGCCGCGGG
369 Y G L Q A I C C A G G M G T A T I I E R
6961 GCTATGGACTGCAGGCAATCTGCTGTGCTGGCGGTATGGGTACGGCGACCATCATCGAGC
CGATACCTGACGTCCGTTAGACGACACGACCGCCATACCCATGCCGCTGGTAGTAGCTCG
389 L D *
7021 GCCTCGATTGAGTATGTGCGGGATTCCCATGGTGCCGCTGGAGGAAATGCCGGCTGGACT
CGGAGCTAACTCATACAGCGCCTAAGGGTACCACGGCGACCTCCTTTACGGCCGACCTGA
7081 GCGCGATGCCATTTCTCGGGGGCGGGCAAGCGGCATGCTCAGCTCCAGCAGGCCCGTCCA
CGCGCTACGGTAAAGAGCCCCGCGCTTCGCCGTACGAGTCGAGGTCGTCCGGGCAGGT
7141 GGTCTGGGCGCACCGCCCGGCTCTGGCGCAGGCCTGGCTGGAGCTAATGAGTGGTTTCTA
CCAGACCCGCGTGGCGGGCCGAGACCGGTCCGGACCGACCTCGATTACTACCAAAGAT
7201 TCGGGATAGCCTGCTTGATGAGCGGCTAACGCGAACTGGTGCGCCTGAAGATCGCCAGCA
AGCCCTATCGGACGAACTACTCGCCGATTGCGCTTGACCACGCGGACTTCTAGCGGTCGT
7261 TCACCACCTGCCAGGCCTGCCAACTGGCGCGTAAGTCCGACACGGTCAGCGAAGAGGATA
AGTGGTGGACGGTCCGGACGGTTGACCGCGCATTACAGGCTGTGCCAGTCGCTTCTCCTAT
7321 TCGCCTGCCTGGCGACGGACAGCGATCGCTTCGCGCCGGCCGAGCAGGCGGCCCTGCGTT
AGCGGACGGACCGCTGCCTGTGCTAGCGAAGCGCGGCCGGCTCGTCCGCCGGGACGCAA
7381 TCGCAGAACTGTTGCGCGCTGACTACATGGCCATCGATGACGAGCACTATGCCCGCCTGT
AGCGTCTTGACAAGCGGCGACTGATGTACCGGTAGCTACTGCTCGTGATACGGGCGGACA
7441 CGGAGTTCTTCACCTGCGCGCAGGTGGTCGAACTCAATATGTACTGCGCCCTGATGCTGG
CM27-----→
GCCTCAAGAAGTGGACGCGCGTCCACCAGCTTGAGTTATACATGACGCGGGGACTACGACC

7501 CCGGTGGCCGCATGACTTATGTCCAGCAGGCCTATTGATCTTGTTGCTTGTAGTCACTCG
 GGCCACCGGCGTACTGAATACAGGTCGTCCGGATAACTAGAACACGAACATCAGTGAGC

7561 ATCTCACGGGATGTCTTGAGTAGTTAGTC**AAAAATGTAAAAAA**CGTCCAATAGTGGTCGA
 TAGAGTGCCCTACAGAACTCATCAATCAG**TTTTACATTTTTTT**GCAGGTTATCACCAGCT

7621 TGGATGCCTAAGATGGCCGGTAATACGCGGCCTTCCAGGCCACGTATTCCAACAATGCAC
 ACCTACGGATTCTACCGGCCATTATGCGCCGGAAGGTCCGGTGCATAAGGTTGTTACGTG

1 tdtD⇒ M S G P A H S N L E Q V F A
 7681 AATAAGAA**AGGAG**ACGCCTATGTCCGGCCAGCACACAGCAATCTGGAGCAGGTCTTCGC
 T TATTCTTTCCTCTGCGGATACAGGCCGGGTCGTGTGTCGTTAGACCTCGTCCAGAAGCG
 ←-----**cm10**

15 N V A S N Y R G A D V D L H A V Y R E M
 7741 CAACGTCGCGAGCAATTACCGTGGTGCCGACGTCGACCTGCACGCGGTGTATCGCGAGAT
 GTTGACGCGCTCGTTAATGGCACCACGGCTGCAGCTGGACGTGCGCCACATAGCGCTCTA

35 R E K S P V L P E N F M A R L G V P S I
 7801 GCGCGAAAATCCCCGTCCTCCCGGAGAACTTCATGGCACGCCTGGGCGTACCATCGAT
 CGCGCTTTTTAGGGGGCAGGAGGGCCTCTTGAAGTACCGTGCAGGACCCGCATGGTAGCTA

55 A G L D P N R P T F T L F K Y D D V M A
 7861 TGCCGGGCTTGATCCGAACCGACCTACTTTCACGTTGTTCAAGTACGACGACGTCATGGC
 ACGGCCCCGAAGTGGCTGGCTGGATGAAAGTGCAACAAGTTCATGCTGCTGCAGTACCG

75 V M R D A T N F T S G F I A E G L G S F
 7921 GGTGATGCGCGATGCGACCAACTTCACCAGCGGCTTCATAGCCGAAGGGCTTGGCTCGTT
 CCACTACGCGCTACGCTGGTTGAAGTGGTCGCCGAAGTATCGGCTTCCCGAACCGAGCAA

95 F D G L I L T A M D G E A H K N I R S L
 7981 CTTGATGGCCTGATCCTTACCGCGATGGACGGCGAGGCGCACAGAAGAACATCCGCAGCCT
 GAAGCTACCGGACTAGGAATGGCGCTACCTGCCGCTCCGCGTGTCTTGTAGGCGTCGGA

115 L Q P V F M P E T V N R W K E T K I D R
 8041 GCTGCAGCCGGTGTTTCATGCCCCGAGACGGTCAACCGCTGGAAGGAAACCAAGATCGATCG
 CGACGTCGGCCACAAGTACGGGCTCTGCCAGTTGGCGACCTTCCTTTGGTTCTAGCTAGC
cm29-

135 V I R E E Y L R P M V A S K R A D I M E
8101 GGT CAT CCG C GAG GAG TAC CTG CGG CCC ATG GTG GCC AGC AAG AGG GCG GAT ATA ATG GA
-----> <-----CM6.1
CCAGTAGGCGCTCCTCATGGACGCCGGGTACCACCGGTCGTTCTCCCGCCTATATTACCT

155 F A L Y F P I R V I Y S L I G F P E D R
8161 GTTCGCCCTGTACTTCCCTATTCTGTGTGATCTACTCGCTGATCGGCTTCCCCGAGGACCG
CAAGCGGGACATGAAGGGATAAGCACACTAGATGAGCGACTAGCCGAAGGGGCTCCTGGC

175 P E E I E Q Y A A W A L A I L A G P Q V
8221 CCCCAGGAAATCGAGCAGTACGCCGCCTGGGCGCTGGCCATCCTGGCTGGACCGCAGGT
GGGGCTCCTTTAGCTCGTCATGCGGCGGACCCGCGACCGGTAGGACCGACCTGGCGTCCA

195 D P E K A A A A R G A A M E A A Q A L Y
8281 CGACCCCCGAGAAGGCTGCGGCCGCGCGCGCGCGCGCGATGGAGGCCGCGCAGGCGCTATA
GCTGGGGCTCTTCCGACGCCGGCGCGCGCGCGCGCTACCTCCGGCGCGTCCGCGATAT

215 D V V K V V V A Q R R A E G A T G D D L
8341 CGACGTGGTCAAGGTTGTGGTGGCGCAGCGTCGCGCCGAGGGTGCCACAGGCGACGACCT
GCTGCACCAGTTCCAACACCACCGCTCGCAGCGCGGCTCCACGGTGTCCGCTGCTGGA

235 I C R L I R A E Y E G R S L D D H E I T
8401 AATCTGCCGGCTGATCAGGGCCGAGTACGAAGGCCGTTGCTGGACGATCACGAGATCAC
TTAGACGGCCGACTAGTCCCGGCTCATGCTTCCGGCAAGCGACCTGCTAGTGCTCTAGTG

255 T F V R S L L P A A S E T T T R T F G T
8461 CACCTTCGTCCGTTTCGCTGCTGCCGGCGGCCAGCGAGACGACTACGCGTACCTTCGGCAC
GTGGAAGCAGGCAAGCGACGACGGCCGCCGGTTCGCTCTGCTGATGCGCATGGAAGCCGTG

275 L M T L L L E R P E L L A R I R E D R S
8521 GTTGATGACACTGCTGTTGGAGAGACCTGAGCTGCTGGCGCGCATCCGTGAGGATCGTAG
<-----CM5 CM31----->
CAACTACTGTGACGACAACCTCTCTGGACTCGACGACCGCGCGTAGGCACTCCTAGCATC

295 L V G K A I D E A V R Y E P V A T F K V
8581 TCTGGTTGGAAAGGCCATCGACGAGGCGGTGCGCTATGAGCCTGTAGCCACCTTTAAGGT
AGACCAACCTTTCCGGTAGCTGCTCCGCCACGCGATACTCGGACATCGGTGGAAATTCCA

315 R Q A A K D V E I R G V A I P K G A M V
8641 ACGCCAGGCGGCCAAGGACGTGGAAATTCGCGGCGTGGCCATTCCTCAAGGGGGCTATGGT
TGCGGTCGCGCGGTTCTCTGCACCTTTAAGCGCCGACCGGTAAGGGTTCCCCCGATACCA

335 S C I V T S A N R D E D A F E N A D T F
8701 CTCGTGCATCGTCACCTCGGCCAATCGCGACGAGGACGCCTTCGAGAACGCCGACACCTT

GAGCACGTAGCAGTGGAGCCGGTTAGCGCTGCTCCTGCGGAAGCTCTTGCGGCTGTGGAA

355 D I D R R A K P S F G F G F G P H M C I
8761 CGACATCGACCGCCGCGCCAAGCCTTCGTTTCGGCTTCGGTTTCGGTCCGCACATGTGCAT

GCTGTAGCTGGCGGCGCGGTTTCGGAAGCAAGCCGAAGCCAAAGCCAGGCGTGACACGTA

375 G Q F V A K T E I N C A L N A I L D L M
8821 CGGCCAGTTCGTGCGCAAGACGGAAATCAATTGCGCGCTCAACGCCATCCTCGACCTGAT

GCCGGTCAAGCAGCGGTTCTGCCTTTAGTTAACGCGCGAGTTGCGGTAGGAGCTGGACTA

395 P N I R L D P D K P A P E I I G A Q L R
8881 GCCGAACATCCGCCTGGATCCCGACAAGCCGGCTCCGGAGATCATCGGTGCGCAGCTGCG

CGGCTTGTAGGCGGACCTAGGGCTGTTTCGGCCGAGGCCTCTAGTAGCCACGCGTCGACGC

415 G P H H V H V I W D *
8941 TGGTCCTCATCACGTACATGTGATCTGGGATTGAGTGGTATCGCTTAAGGGGCGGTAATG

←-----ca
ACCAGGAGTAGTGATGTACACTAGACCCTAACTCACCATAGCGAATTCCTCCGCCATTAC

9001 CCGTCCCGTTGGCGCCGCTGCTGCGGCGTACCTGGCGACCTAGGGGGCAGGGTTCGGGAA
GGCAGGGCAACCGCGGCGACGACGCCGCATGGACCGCTGGATCCCCCGTCCCAAGCCCTT

9061 TGCCATTAAGTTGAACATATTTAAAAAAGTGTTTAGTTTTCTCTGGCGTGCTCTATGA

ACGGTAATTCAACTTGTATAAATTTTTTTCACAAATCAAAAAGAGACCGCACGAGATACT

9121 TCAAAAGCACTTCAAGAACCCATTAAGCGAAGGGGAAAACAACAATGAAGGTTGAAGATC
AGTTTTTCGTGAAGTTCTTGGGTAATTCGCTTCCCTTTTGTTGTTACTTCCAACCTTCTAG

↑ BgIII

TABLE B1. Summary of expected and observed fragments for mapping of *tdt* gene probes in *Pseudomonas* sp. A19-6a.

ENZYME		<i>Ava</i> I	<i>Bcl</i> I	<i>Eco</i> RV	<i>Nar</i> I	<i>Pst</i> I	<i>Pvu</i> I	<i>Sac</i> I	<i>Sal</i> I
PROBE	# SITES	10	7	2	10	5	7	1	6
<i>tdtL</i>	expected	>1.8	1.7	>2.8	1.5; >0.6	>1.7	1.2; >0.9	>4.1	3.2
	observed	3.9; 1.7	2.0; 0.9	3.8	2.5; 1.8	3.1	1.5; 1.4	16.1; 5.0	3.1
<i>tdtR/L</i>	expected	>1.8	1.7	>2.8	1.5	>1.7	1.2	>4.1	3.2
	observed	1.7	2.0	3.9	2.6; 1.9	3.1	1.5	5.0	3.1
<i>tdtR</i>	expected	1.2; >1.8	1.7	>2.8	1.5; >0.6	>1.7	1.2; >0.9	>4.1	3.2
	observed	2.6; 2.5; 1.6; 1.4	2.9; 1.9; 1.1	3.5; 2.9	4.3; 3.1; 2.5; 2.2; 2.0; 1.8	2.0	1.9; 1.5; 1.1	4.2	3.0
<i>tdtS</i>	expected	1.3; 1.2	2.0; 0.8	4.2; 2.8	1.8	0.5; 2.0	1.7	>4.1	2.6; 3.2
	observed	2.5; 1.3	2.9; 1.2	3.0	2.8; 2.6; 2.3	1.4; 2.2	1.9; 1.1	7.6	1.4; 0.9
<i>tdtA</i>	expected	1.3	2.0	4.2	0.9; 1.8	2.4; 0.5	1.6; 1.7	>4.7; >4.1	2.6
	observed	1.7	2.3	4.5	4.3; 3.1; 2.3; 2.1; 1.5; 1.3	2.5; 0.9	1.9; 1.8	17.1; 14.3; 7.8; 5.1	2.6
<i>tdtB</i>	expected	1.6	2.7; 2.0	4.2	1.2; 0.9	2.4	1.6	>4.7	2.6
	observed	1.8	2.9; 2.3	4.5	4.4; 3.2; 1.6	2.4	1.8	7.8	2.6
<i>tdtC</i>	expected	1.6	2.7	4.2	2.0; 1.2	2.4	0.6; 0.8	>4.7	1.2; 2.6
	observed	1.8	2.8	4.5	4.4; 1.6	2.4	1.2	7.1	2.5; 1.4
<i>tdtD</i>	expected	>0.9	0.7; 2.7	>1.9	2.0	>1.1	>1.1; 0.8	>4.7	>0.9; 0.5
	observed	2.5; 1.3	3.7; 3.0; 1.1	4.4; 3.0	2.8; 2.6; 2.3	1.4; 2.2	1.1; 1.9	13.6; 7.6	1.4; 0.9

TABLE B2. List of restriction enzymes and corresponding cut sites in *tdt* locus cloned from *Pseudomonas* sp. A19-6a. Recognition sites for both palindromic and non-palindromic cutters are included. Position number is based on plus strand as shown in Figure B1. Blunt end cutters are identified (BE).

Restriction enzyme	Number of cut sites	Recognition site	POSITION (plus strand numbering)
<i>AatI</i> (BE)	7	AGG/CCT	1780 2798 4388 6801 7172 7275 7530
<i>AatII</i>	5	GACGT/C	362 3817 4548 7773 7914
<i>Acc65I</i> = <i>Asp718I</i>	3	G/GTACC	3001 3694 5826
<i>AccI</i>	7	GT/MKAC	494 680 3742 3907 6534 7773 8280
<i>AccIII</i> = <i>BspMII</i>	4	T/CCGGA	413 964 1739 8914
<i>AclI</i> = <i>Psp1406I</i>	1	AA/CGTT	4783
<i>AflII</i>	1	C/TTAAG	8984
<i>AflIII</i>	7	A/CRYGT	215 552 3307 4806 8504 8811 8956
<i>AhaIII</i> = <i>DraI</i> (BE)	2	TTT/AAA	2887 9082
<i>AhdI</i> = <i>Eam1 105I</i>	6	GACNNN/NGTC	1735 3848 5534 5797 6332 8276
<i>Alw44I</i>	3	G/TGCAC	2919 3588 5401
<i>AlwI</i>	29	GGATCNNNN/	881 953 1657 1670 2172 2179 2480 2320 2568 2669 3125 3138 4013 4015 4050 4136 4280 4409 4414 4652 4708 5103 6239 6696 7865 7987 8580 8891 8904
<i>ApaI</i>	3	GGGCC/C	331 4641 5651
<i>ApalI</i>	3	G/TGCAC	2919 3588 5401
<i>ApoI</i>	6	R/AATTY	1802 1833 1839 1915 5728 8664
<i>AscI</i>	1	GG/CGCGCC	5170
<i>Asp718I</i> = <i>Acc65I</i>	3	G/GTACC	3001 3694 5826
<i>AsuII</i>	1	TT/CGAA	3643
<i>AvaI</i>	11	C/YCGRG	1023 2374 3534 4864 5365 5521 7095 8060 8210 8222 8285
<i>AvrII</i>	1	C/CTAGG	9039
<i>Ball</i> = <i>MscI</i> (BE)	7	TGG/CCA	1508 3219 3268 7410 8135 8258 8678
<i>BamHI</i>	3	G/GATCC	1662 3130 8896
<i>BanI</i>	28	G/GYRCC	69 281 808 911 1198 2215 2409 3001 3146 3290 3694 3884 4080 4192 4406 5292 5826 6221 6434 6479 6646 6889 6900 6954 7051 7763 8381 9011
<i>BanII</i>	8	GRGCT/C	331 3413 4449 4641 5285 5651 5843 6658
<i>BbeI</i>	10	GGCGC/C	73 915 2413 4196 4410 5296 6483 6650 6958 9015
<i>BbsI</i>	8	GAAGACNN/	150 2252 2481 2596 4509 5819 6549 7726
<i>BbuI</i> = <i>SphI</i>	5	GCATG/C	393 3183 4166 5385 7117
<i>BbvI</i>	42	GCAGCNNNNNNNN/	351 642 898 1189 1349 1687 2138 2622 2680 2799 2943 3100

			3154 3248 3967 3976 4030 4088 4236 4576 4671 4893 5537 5564 5684 5872 5926 5953 6244 6380 6607 6916 8028 8046 8056 8281 8377 8463 8538 8922 8944 9004
<i>BciVI</i>	4	GTATCCNNNNN/	585 3552 4521 5674
<i>BclI</i>	7	T/GATCA	759 1170 2878 3700 5747 8412 9118
<i>BglI</i>	14	GCCNNNN/NGGC	101 469 1124 2097 4211 4725 5152 6201 7329 7365 7656 8249 8264 8330
<i>BglII</i>	1	A/GATCT	828
<i>BipI</i>	5	GC/TNAGC	1358 2566 5329 6339 7118
<i>BmrI</i>	2	ACTGGGNNNNN/	3323 292
<i>BpmI</i>	9	CTGGAGNNNNN NNNNNNNNNNN/	470 1839 2332 5991 6768 7079 7108 7199 7745
<i>Bpu10I</i>	8	CC/TNAGC	79 3277 4146 5113 5194 6465 6657 8547
<i>Bpu1102I</i>	5	GC/TNAGC	1358 2566 5329 6339 7118
<i>BsaAI</i> (BE)	2	YAC/GTR	7663 8953
<i>BsaBI</i> (BE)	2	GATNN/NNATC	1298 2176
<i>BsaHI</i>	23	GR/CGYC	70 318 359 912 2410 3785 3814 4193 4407 4545 4833 5161 5293 5483 5798 6480 6647 6955 7694 7770 7911 8736 9012
<i>BsaI</i>	6	GGTCTCN/	989 1564 2104 5241 8536 8704
<i>BsaJI</i>	74	C/CNNGG	86 168 332 402 579 599 644 750 881 915 941 1271 1284 1578 2105 2240 2303 2309 2793 3139 3140 3185 3319 3445 3481 3613 3752 3901 4188 4200 4207 4260 4338 4440 4633 4640 4653 4746 4864 4865 5090 5165 5273 5321 5486 5520 5521 5600 5642 5651 5717 5718 5775 5909 5942 6226 6248 6299 6650 6761 6883 6884 7047 7759 7843 8127 8211 8223 8247 8376 8652 8685 8715 9039
<i>BsaWI</i>	13	W/CCGGW	31 413 614 964 1739 3004 3149 3545 3925 4238 4288 5788 8914
<i>BseRI</i>	5	GAGGAGNNNNNNNNN/	301 2122 4457 6578 8126
<i>BsgI</i>	3	GTGCAGNNNNN NNNNNNNNNNN/	358 4032 7763

<i>BsiEI</i>	28	CGRY/CG	180 294 494 619 1296 1524 1694 2529 2598 3385 3638 4284 4536 4581 4734 5899 5903 6116 6172 6277 6442 6725 6729 7346 7360 8098 8301 8770
<i>BsiHKA I</i>	8	GWGCW/C	1030 2923 3592 4449 5405 6460 7426 9114
<i>Bsil</i>	3	C/TCGTG	8701 6849 8450
<i>BsiWI=Sun I</i>	1	C/GTACG	5889
<i>BsII</i>	58	CCNNNNN/NNGG	22 106 175 408 489 585 750 915 991 1187 1188 1272 1309 1791 1792 2267 2310 2583 2707 2776 3072 3073 3140 3248 3475 3613 3902 4303 4339 4603 4864 5090 5091 5199 5625 5643 5691 5836 6265 6305 6389 6489 6564 6788 7095 7144 7162 7334 7439 7496 7703 7783 8080 8142 8224 8290 8440 8789
<i>BsmAI</i>	12	GTCTCN/	989 1564 1719 1857 2104 4109 5241 7685 8057 8704 8489 8536
<i>BsmBI</i>	3	CGTCTCN/	7685 8057 8489
<i>BsmFI</i>	20	GGGACNNN NNNNNNN/	151 204 597 2286 2293 2515 2731 3074 3796 4166 4658 4863 4883 5022 5107 5148 5502 5720 5792 8988
<i>BsmI</i>	5	GAATGCN/	1827 1956 2349 4037 9064
<i>BsoBI</i>	11	C/YCGRG	1023 2374 3534 4864 5365 5521 7095 8060 8210 8222 8285
<i>Bsp120I</i>	3	G/GGCC	327 4637 5647
<i>Bsp1407I</i>	1	T/GTACA	5570
<i>BspDI=Cla I</i>	8	AT/CGAT	3019 4054 4284 4494 5609 7414 7856 8094
<i>BspEI=Bsp MII</i>	4	T/CCGGA	413 964 1739 8914
<i>BspLU11I</i>	4	A/CATGT	552 3307 8811 8956
<i>BspMI</i>	12	ACCTGCNNNN/	465 3583 4569 5714 5987 7275 7451 7461 7785 8127 7720 8266
<i>BspMII=Bs pEI</i>	4	T/CCGGA	413 964 1739 8914
<i>BsrBI (BE)</i>	9	GAG/CGG	1688 2663 2717 4150 5362 5370 6854 6947 7222
<i>BsrDI</i>	5	GCAATGNN/	2822 3259 3299 5015 5306
<i>BsrFI</i>	29	R/CCGGY	19 57 468 1123 1192 2096 3495 4172 4580 4817 4825 5011 5671 5902 6086 6156 6171 6195 6437 6720 6756 7069 7355 7500 7636 8047 8406 8482 8908

<i>BsrGI</i>	1	T/GTACA	5570
<i>BsrI</i>	22	ACTGGN/	298 539 592 801 1314 3318 3602 3721 3776 4098 4123 4370 4633 4767 5033 5183 5224 5534 6842 7240 7288 8824
<i>BssHII</i>	10	G/CGCGC	4553 5170 5947 6057 6663 6910 7456 8303 8558 8853
<i>BssSI</i>	3	C/TCGTG	8701 6849 8450
<i>BssSI</i>	3	C/TCGTG	6849 8450 8701
<i>BstAPI</i>	3	GCANNNN/NTGC	390 3258 3884
<i>BstBI</i> = <i>Fsp II</i>	1	TT/CGAA	3643
<i>BstD102I</i> (BE)	9	GAG/CGG	4150 5370 6854 6947 7222 1688 2663 2717 5362
<i>BstEII</i>	5	G/GTNACC	53 349 5178 5662 6089
<i>BstF5I</i> = <i>BseGI</i>	27	GG/ATG	367 417 706 1015 1039 1381 1674 1743 1796 2185 2241 2557 3046 3067 3258 4751 5642 6269 6402 6827 7570 7623 8033 8108 8264 8567 8870
<i>BstXI</i>	7	CCANNNNN/NTGG	1042 3099 3488 4333 4633 5055 8659
<i>BstYI</i>	13	R/GATCY	828 873 1662 2171 2472 2573 3130 4007 4018 4713 5108 6244 8896
<i>ClaI</i> = <i>BspDI</i>	8	AT/CGAT	3019 4054 4284 4494 5609 7414 7856 8094
<i>Csp45I</i>	1	TT/CGAA	3643
<i>CspI</i>	2	CG/GWCCG	3923 8804
<i>DraI</i> = <i>AhaIII</i> (BE)	2	TTT/AAA	2887 9082
<i>DraIII</i>	3	CACNNN/GTG	6489 6900 7786
<i>DrdI</i>	5	GACNNNN /NNGTC	148 1525 4457 6189 7699
<i>EaeI</i>	36	Y/GGCCR	177 1116 1506 1584 2055 2072 3217 3266 3367 3382 3442 3493 4335 4578 4605 5357 5591 5900 6046 6084 6169 6274 6722 7357 7408 7498 7505 7634 8133 8256 8298 8487 8649 8676 8718 8821
<i>EagI</i> = <i>XmaIII</i>	9	C/GGCCG	177 3382 4578 5900 6169 6274 6722 7357 8298
<i>Eam1105I</i>	6	GACNNN/NNGTC	1735 3848 5534 5797 6332 8276
<i>EarI</i>	3	CTCTTCN/	2135 5251 7306
<i>Ecl136II</i> (BE)	1	GAG/CTC	4447
<i>Eco47III</i> (BE)	3	AGC/GCT	818 3275 6282
<i>Eco52I</i>	9	C/GGCCG	177 3382 4578 5900 6169 6274 6722 7357 8298
<i>Eco56I</i>	13	G/CCGGC	19 468 1123 3495

			4172 6156 6171 6756 7069 7355 8406 8482 8908
<i>Eco57I</i>	4	CTGAAGNNNNN NNNNNNNNNNN/	3756 4403 6561 7266
<i>EcoICRI</i>	1	GAG/CTC	4447
<i>EcoNI</i>	4	CCTNN/NNNAGG	406 4337 5197 6786
<i>EcoRV</i> (BE)	2	GAT/ATC	3107 7319
<i>EheI</i> = <i>SfoI</i> (BE)	10	GGC/GCC	71 913 2411 4194 4408 5294 6481 6648 6956 9013
<i>EspI</i>	5	GC/TNAGC	1358 2566 5329 6339 7118
<i>FauI</i>	29	CCCGCNNNN/	24 118 218 574 751 1367 1794 2668 2689 3354 2193 2204 2270 4368 4515 4940 4143 5201 5363 5462 5692 5750 5816 5999 6635 6864 6966 7095 7440
<i>FseI</i>	1	GGCCGG/CC	6760
<i>FspI</i> = <i>MstI</i> (BE)	5	TGC/GCA	560 689 4910 5021 8931
<i>HgaI</i>	22	GACGCNNNNN/	206 307 636 888 1066 1515 2051 2341 2512 2723 3793 4822 4924 5150 5491 5611 5787 5789 6716 7702 8357 8744
<i>HpaI</i> (BE)	1	GTT/AAC	5511
<i>HphI</i>	27	GGTGANNNNNNN/	65 343 633 769 1320 2983 3740 4303 4642 4714 4807 4815 4930 5172 5661 5674 5966 6017 6566 6677 6710 7253 7442 7933 7936 8449 8704
<i>KasI</i>	10	G/GCGCC	69 911 2409 4192 4406 5292 6479 6646 6954 9011
<i>KpnI</i>	3	GGTAC/C	3005 3698 5830
<i>MboII</i>	26	GAAGANN NNNNNN/	127 150 223 517 731 834 2152 2257 2481 2596 2718 3015 3575 4028 4040 4514 4984 5268 5382 5824 6554 7259 7323 7439 7726 7972
<i>MfeI</i>	1	C/AATTG	8848
<i>Mlu113I</i>	1	CC/GCGG	580
<i>MluI</i>	2	A/CGCGT	215 8504
<i>MscI</i> = <i>BalI</i> (BE)	7	TGG/CCA	1508 3219 3268 7410 8135 8258 8678
<i>MslI</i> (BE)	8	CAYNN/NNRTG	3031 3184 5380 5572 5755 7919 8522 8955
<i>MspAI</i>	14	G/GWCC	165 411 1729 3007 3719 3923 4594 4618 4765 5009 8215 8270 8804 8942
<i>MstI</i> = <i>FspI</i> (BE)	5	TGC/GCA	560 689 4910 5021

			8931
<i>NaeI</i> (BE)	13	GCC/GGC	21 470 1125 3497 4174 6158 6173 6758 7071 7357 8408 8484 8910
<i>NarI</i> = <i>Mly113I</i>	10	GG/CGCC	70 912 2410 4193 4407 5293 6480 6647 6955 9012
<i>NcoI</i>	8	C/CATGG	881 3752 5486 5642 6226 6248 7047 8127
<i>NgoMI</i>	13	G/CCGGC	19 468 1123 3495 4172 6156 6171 6756 7069 7355 8406 8482 8908
<i>NheI</i>	3	G/CTAGC	112 1370 2959
<i>NlaIV</i> (BE)	55	GGN/NCC	71 167 283 329 330 548 590 730 810 913 1200 1422 1664 2217 2411 2652 3003 3076 3132 3148 3292 3479 3503 3696 3886 3941 4082 4194 4213 4315 4408 4638 4639 5010 5294 5648 5649 5828 5835 5907 6176 6223 6436 6481 6648 6698 6891 6902 6956 7053 7765 8383 8898 8913 9013
<i>NotI</i>	1	GC/GGCCGC	8298
<i>NruI</i> (BE)	5	TCG/CGA	39 140 7748 7794 8727
<i>PaeR7I</i> = <i>XhoI</i>	1	C/TCGAG	2374
<i>PfiFI</i>	6	GACN/NNGTC	1525 1711 4874 7302 7910 8345
<i>PfiMI</i>	5	CCANNNN/NTGG	991 2310 3902 5643 6489
<i>PleI</i>	11	GAGTCNNNN/	295 1624 2086 2390 2446 2524 3095 5082 5157 5523 6530
<i>PpuMI</i>	1	RG/GWCCY	165
<i>PshAI</i> (BE)	2	GACNN/NNGTC	2336 7518
<i>Psp1406I</i> = <i>AclI1</i>		AA/CGTT	4783
<i>PstI</i>	5	CTGCA/G	2013 4015 4566 6973 8046
<i>PvuI</i> = <i>XorI</i> I	7	CGAT/CG	1296 2529 4284 5899 6729 7346 8098
<i>PvuII</i> (BE)	2	CAG/CTG	1416 8935
<i>RsrI</i>	2	CG/GWCCG	3923 8804
<i>SacI</i> = <i>SstI</i>	1	GAGCT/C	4449
<i>SacII</i> = <i>SstII</i>	1	CCGC/GG	582
<i>Sall</i>	6	G/TCGAC	493 679 3906 6533 7772 8279
<i>SciI</i> (BE)	1	CTC/GAG	2376
<i>SfaNI</i>	29	GCATCNNNNN/	570 675 696 1029 1041 1371 1664 1740 3071 3248 3950 5359 5718 5870 6155 6392 6580 6703 7075 7267 7483 7613 7788 7914 7921 8571 8715 8826

			8868
<i>Sfci</i>	6	C/TRYAG	2009 4011 4562 6969 8042 8622
<i>Sfil</i>	2	GGCCNNNN/NGGCC	4211 7656
<i>Sgfl</i>	1	GCGAT/CGC	7346
<i>SgrAI</i>	5	CR/CCGGYG	19 468 4172 6156 6195
<i>SmaI</i> (BE)	2	CCC/GGG	4866 5523
<i>SmlI</i>	6	C/TYRAG	568 2374 4697 5842 7575 8984
<i>SphI</i> = <i>BbuI</i>	5	GCATG/C	393 3183 4166 5385 7117
<i>SplI</i> = <i>SunI</i> = <i>BsiWI</i>	1	C/GTACG	5889
<i>Spol</i> (BE)	5	TCG/CGA	39 140 7748 7794 8727
<i>Ssp1I</i>	1	TT/CGAA	3643
<i>Sspl</i>	1	AAT/ATT	2124
<i>SstI</i> = <i>SacI</i>	1	GAGCT/C	4449
<i>SstII</i> = <i>Sac II</i>	1	CCGC/GG	582
<i>StuI</i> (BE)	7	AGG/CCT	1780 2798 4388 6801 7172 7275 7530
<i>StyI</i>	17	C/CWWGG	881 1578 3445 3481 3752 4260 4653 5486 5642 5651 6226 6248 7047 8127 8652 8685 9039
<i>SunI</i> = <i>SplI</i> = <i>BsiWI</i>	1	C/GTACG	5889
<i>TatI</i>	15	W/GATCW	759 828 1170 1212 2878 3700 5747 5981 7536 8187 8412 8454 8919 8961 9118
<i>TspRI</i>	13	CAGTGNN/	305 546 628 1093 1653 2031 4228 4489 4555 4633 5190 6073 6490
<i>Tth111I</i>	6	GACN/NNGTC	1525 1711 4874 7302 7910 8345
<i>XcmI</i>	7	CCANNNNN /NNNTGG	445 2142 3488 5055 5649 7054 8692
<i>XhoI</i> = <i>PaeR7I</i>	1	C/TCGAG	2374
<i>XmaI</i>	2	C/CCGGG	4864 5521
<i>XmaIII</i> = <i>EagI</i>	9	C/GGCCG	177 3382 4578 5900 6169 6274 6722 7357 8298
<i>XmnI</i> (BE)	3	GAANN/NTTC	2249 3834 6502
<i>XorI</i> I= <i>PvuI</i>	7	CGAT/CG	1296 2529 4284 5899 6729 7346 8098

APPENDIX C

This appendix provides data in support of the results presented in Chapter 5. Table C1 shows the results of the dot blot screening of 31 strains tested for *tdt* gene probe specificity and includes HPLC data regarding percent abietic acid degraded in batch culture. Table C2 is a summary of the environmental isolates screened for distribution of the *tdt* genotype. HPLC data is included and probe positive isolates are highlighted. Table C3 lists the substrates used by each of the probe positive isolates and Table C4 is the Jaccard Correlation Coefficients generated from this phenotypic data; both of which form the basis for the dendrogram illustrated in Figure 5-4.

TABLE C1. List of strains screened by dot blot hybridization for determining specificity of *tdt* gene probes. Percent abietic acid degraded by each strain as determined by HPLC analysis is included. Hybridization results are positive (+), negative (-), or weak (/).

Strain	% ABA degraded	PROBE SPECIFICITY						
		tdtL	tdtR	tdtS	tdtA	tdtB	tdtC	tdtD
A19-1	26	-	-	-	-	-	-	-
A19-2	39	-	-	-	-	-	-	-
A19-3	99	+	+	+	+	+	+	+
A19-5	2	+	+	+	-	+	+	+
A19-6a	97	+	+	+	+	+	+	+
A19-6b	100	+	+	+	+	+	+	+
D11-3b	56	-	-	-	-	-	-	+
D11-9	52	+	+	+	+	+	+	+
D11-13	42	-	-	+	-	-	-	+
D11-23a	44	-	-	-	-	-	-	-
D11-23b	34	-	-	-	-	-	-	-
D11-25	6	-	-	-	-	-	-	+
D11-37	37	-	-	-	-	-	-	+
16LW	81	+	/	+	+	-	/	+
BKME-9	84	+	+	+	+	+	+	+
<i>Streptomyces griseus</i>	39	-	-	-	-	-	-	-
<i>Pseudomonas aeruginosa</i>	75	-	-	-	-	-	-	-
<i>Pseudomonas fluorescens</i>	77	-	-	-	-	-	-	-
<i>Pseudomonas putida</i>	63	-	-	-	-	-	-	-
<i>Pseudomonas cepacia</i>	69	-	-	-	-	-	-	-
<i>Alcaligenes calcoaceticus</i>	74	-	-	-	-	-	-	-
<i>Moraxella sp.</i>	79	+	+	-	-	-	-	+
<i>Vibrio natriegens</i>	51	-	-	-	-	-	-	-
NQ1	70	-	-	-	-	-	-	-
NQ2	65	-	-	-	-	-	-	-
NQ3	73	-	-	-	-	-	-	-
NQ4	55	-	-	-	-	-	-	-
NQ5	58	+	/	+	+	-	-	+
NQ6	78	-	-	-	-	-	-	-
<i>Comomonas testosteroni</i>	68	-	-	-	-	-	-	-
<i>Escherichia coli</i> 16948	not tested	-	-	-	-	-	-	+

TABLE C2. Environmental isolate collection screened for distribution of the *tdt* genotype. HPLC peak area and percent resin acid degraded in batch culture are indicated for each strain. Probe positive isolates are highlighted for ease of reference.

LOCATION	SAMPLES	CODE	ISOLATE	Abietic acid		ISOLATE	Dehydroabietic acid	
				PEAK AREA	% DEGRADED		PEAK AREA	% DEGRADED
Aracruz, Brazil	Lagoon 1 centre	L1C	A1a	502046	31	D1a	69775	27
			A1b	835571	0	D1b	77374	20
			A1c	860525	0	D1c	129224	0
			A1d	554049	24	D1d	87730	9
			A1e	612689	16	D1e	101254	0
			A1f	625245	14	D1f	78582	18
			A1g	603775	17	D1g	69225	28
			A2a	658792	10	D2a	78203	19
	Lagoon 2 centre	L2C	A2b	631881	14	D2b	97197	0
			A2c	446786	44	D2c	74381	23
			A2d	672925	16	D2d	69380	28
			A2e	918462	0	D2e	98714	0
			A2f	640001	20	D2f	71611	26
			A2g	403294	49	D2g	1220	99
			A2h	630850	21	D2h	80555	16
			A3a	568708	29	D3a	148079	0
	Lagoon 3 centre	L3C	A3b	581448	27	D3b	10807	89
			A3c	613479	23	D3c	81943	15
			A3d	647944	19	D3d	94264	2
			A3e	536436	25	D3e	94511	2
			A3f	634005	12	D3f	71516	26
			A4a	978038	0	D4a	69153	28
	Lagoon 4 centre	L4C	A4b	703093	2	D4b	60451	48
			A4c	560400	22	D4c	56058	52
			A4d	665601	7	D4d	53115	55
			A4e	751236	0	D4e	73232	37
			A4f	709150	1	D4f	83226	29
			A4g	700091	3			
	Lagoon 5 centre	L5C	A5a	529645	26	D5a	92432	21
			A5b	535883	10	D5b	65575	44
			A5c	846444	0	D5c	66445	43
			A5d	874425	0	D5d	86544	26
			A5e	824530	0	D5e	91816	22
			A5f	905169	0	D5f	84984	27
	Lagoon 6 centre	L6C	A6a	540046	9	D6a	62657	46
			A6b	641566	0	D6b	86318	26
			A6c	687079	0	D6c	99182	15
			A6d	330805	45	D6d	89493	24
						D6e	79697	32
						D6f	128291	0
Cornwall	C2LL2	C2LL2	A1	180476	56	D1	15577	87
			A2	139438	66	D2	37516	68
			A3	215705	48	D3	35956	69
			A4	157597	62	D4	61153	47
			A6	167337	68	D5	25397	76
			A7	306263	41	D6	52495	55
			A8	220566	58	D7	21052	82
						D8	77712	33
	RAS	RAS	A1	188112	64	D2	31809	73
			A2	219903	58	D3	33258	71
			A3	182792	65	D4	41661	64

Domtar, Windsor	L4	L4	A4	173391	67	D5	29411	75
			A5	181098	65	D6	30704	73
			A6	161900	69	D7	21147	82
			A7	176021	66	D8	36579	68
			A8	209509	60	D10	32694	72
			A1	331700	58	D1	54563	47
			A3	215549	73	D2	32001	69
			A5	224096	71	D3	111216	0
			A8	225415	71	D5	42106	59
			A9	532552	32	D6	44751	56
			A11	204266	74	D7	97487	4
			A12	571647	27	D8	33830	67
	L15	L15	A2	330268	58	D1	59028	42
			A3	200446	74	D2	75371	26
			A5	142286	82	D3	55124	46
			A7	185582	76	D4	48479	52
			A8	155891	80	D5	52570	48
						D6	40760	60
						D7	43320	58
Japan	unknown	7057J	A1	184647	76	D1	51564	49
			A2	99779	87	D2	88206	14
			A3	120010	85	D3	30178	70
			A5	201160	74	D4	90160	12
			A6	59812	92	D5	12206	88
			A7	58452	93	D6	16622	84
			A9	250775	68	D7	69367	32
			A10	290027	63	D8	60591	41
			A11	292493	63			
			A12	158334	80			
Lewiston, Idaho	Basin I	7104-1	A1	74052	86	D1	49085	55
			A2	63875	88	D3	50342	54
			A3	102540	81	D4	9310	91
			A4	146718	73	D5	20109	81
			A5	112307	79			
	Basin II	7104-2	A1	9784	98	D1	41631	62
			A2	453101	16	D2	39915	63
			A3	136073	75	D4	49381	54
			A4	146410	73			
			A5	103783	81			
	Basin III	7104-3	A1	101341	81	D1	91741	15
			A2	168630	69	D2	113739	0
						D3	118774	0
						D4	113516	0
Paparoa, New Zealand	NZ Pond 2 4/8	7104-4	A2	113354	79	D1	97956	10
			A3	213793	61	D2	97676	10
			A4	138135	75	D3	104261	4
						D4	19519	82
	NZ Pond 4 3/4	7104-6	A1	66117	88	D2	40929	62
			A2	193130	64	D3	49999	54
			A3	181923	66			
			A4	178804	67			
	NZ Pond 4 19/3	7104-7	A1	494367	9	D1	52045	52
			A2	154367	72			

Potlatch	Site 1	P1	A3	423297	22			
			A4	84465	84			
			A7	444752	18			
			A2	157400	81	D2	37819	55
			A3	157109	81	D3	17399	79
			A5	362505	57	D4	76052	10
			A6	207111	75			
			A7	248029	70			
			A8	281149	66			
			A1	209822	75			
Papro, New Zealand	Final	PF	A2	308055	63			
			A3	379036	55			
			A4	437111	48			
			A7	239121	71			
			A8	244346	71			
			A2	454537	32	D5	22995	73
			A3	251705	63	D6	31285	60
			A4	227720	66			
			A5	312264	54			
			A6	383422	43			
Alpac, Alberta	unknown	AL	A7	192128	71			
			A8	336789	50			
			A1	277503	67	D1	1502	98
			A2	202323	76	D2	37749	55
			A3	103688	85	D3	21418	75
			A4	230083	66	D4	24146	71
			A5	271162	60	D5	31406	63
			A6	172936	74	D6	31127	63
			A7	140467	79	D8	37118	56
			A8	76876	89			
Domtar, Windsor	#9	D9	A1	213778	74	D1	53913	36
			A2	202761	76	D2	15037	82
			A3	496973	40	D3	71260	15
			A4	245077	71	D4	46745	44
			A5	294746	65	D5	23516	70
			A6	245685	71	D6	58962	30
			A7	251255	70	D7	30257	61
			A8	210209	75	D8	35338	55
			A1	424275	29	D1	28676	66
						D2	53325	31
REPAR	Out	RO				D3	125293	0
						D2	64093	24
			A1	537696	10	D4	109521	0
			A2	227797	62	D3	104277	0
			A3	433369	27	D5	23516	70
			A4	297697	50	D1	42490	45
						D1	2020	98
			A1	489358	18	D2	1297	98
			A2	528689	11	D3	37357	56
			A3	295561	50	D4	74683	11
Papro, New Zealand	Out	NZO	A3	308523	48	D1	44632	43
			A4	374872	37	D2	72330	7
						D3	61094	22
TEMBEC	unknown	TEM						

TABLE C3. Summary of substrates used by environmental isolates that were probe positive for the *tdt* genes. These data were used to generate the dendrogram illustrated in figure 5-4 and the Jaccard Correlation Coefficients presented in Table C4.

Carbon source	A19-6a	A19-6b	D11-9	BKME9	16LW	NQ5	A1a	A1b	A2a	A2b	7057J A6	AL-A8	7104-1 A1	7104-1 A2	7104-1 A3	7101-1 A4	7104-1 D4	7104-2 A3	7104-2 D4
nil	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
a-cyclodextrin	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
dextrin	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
glycogen	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
tween 40	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
tween 80	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
N-acetyl-D-galactosamine	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
N-acetyl-D-glucosamine	-	-	-	-	+	+	-	-	-	-	-	+	-	-	-	-	-	-	-
adonitol	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+	-	-	+
L-arabinose	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-
D-arabitol	-	-	-	-	+	+	-	-	-	-	+	+	-	-	-	+	-	+	+
cellobiose	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
l-erythritol	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D-fructose	+	+	-	+	+	+	-	-	-	-	+	+	-	-	-	-	-	-	+
L-fucose	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D-galactose	-	-	-	-	+	+	-	-	-	-	+	+	-	-	-	-	-	-	-
gentiobiose	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
a-D-glucose	+	-	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	+
m-inositol	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
a-lactose	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
lactulose	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
maltose	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
D-mannitol	-	+	-	-	+	+	-	-	-	-	+	+	+	+	-	+	-	+	+
D-mannose	-	-	-	-	+	+	-	-	-	-	+	+	+	+	-	-	-	+	-
D-melibiose	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
B-methyl glucoside	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
psicose	+	-	-	+	+	-	+	-	-	-	+	-	-	-	-	-	-	-	+
D-raffinose	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L-rhamnose	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D-sorbitol	-	-	-	-	+	-	-	-	-	-	-	+	-	-	-	-	-	-	+
sucrose	-	-	-	-	+	+	-	-	-	-	-	+	-	-	-	-	-	-	-
D-trehalose	-	-	-	-	+	+	-	-	-	-	-	+	-	-	-	-	-	-	-
turanose	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
xylitol	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+	-	-	+
methyl pyruvate	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
mono-methyl succinate	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+
acetic acid	+	+	+	+	+	+	+	+	-	+	+	-	+	-	-	-	-	+	-
cis-aconitic acid	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
citric acid	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	-
formic acid	-	+	-	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-	+
D-galactonic acid lactone	-	-	-	-	+	+	-	-	-	-	+	+	-	-	-	-	-	-	-
D-galacturonic acid	-	-	-	-	+	+	-	-	-	-	+	+	+	+	+	+	+	-	-
D-gluconic acid	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	+
D-glucosaminic acid	-	-	-	-	-	+	-	-	-	-	+	+	-	-	-	-	-	-	-
D-glucuronic acid	-	-	-	-	-	+	-	-	-	-	+	-	+	+	+	+	+	+	-
a-hydroxy butyric acid	-	+	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-
B-hydroxy butyric acid	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Y-hydroxy butyric acid	+	+	-	+	+	-	+	+	+	+	-	-	+	-	-	-	-	-	+
p-hydroxy phenylacetid acid	-	+	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+
itaconic acid	+	+	-	+	+	-	+	+	+	+	-	-	-	-	-	-	-	-	-
a-keto butyric acid	-	+	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
a-keto glutaric acid	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-

α-keto valeric acid	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-
D,L-lactic acid	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
malonic acid	+	-	-	+	+	+	+	-	+	+	+	+	+	-	-	+	-	+	-
propionic acid	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
quinic acid	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
D-saccharic acid	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
sebacic acid	+	+	-	-	+	-	-	-	-	-	-	-	+	+	-	-	-	-	-
succinic acid	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
bromo succinic acid	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+
succinamic acid	-	+	+	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-
glucuronamide	-	-	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-
alaninamide	-	-	-	-	+	+	-	-	-	-	-	+	-	-	-	-	-	+	-
D-alanine	+	-	+	+	+	+	+	+	+	-	+	+	-	-	+	+	+	+	-
L-alanine	+	-	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-
L-alanyl-glycine	-	-	-	-	+	+	-	-	-	-	+	+	-	-	-	-	-	-	-
L-asparagine	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
L-aspartic acid	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
L-glutamic acid	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
glycyl-L-aspartic acid	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
glycyl-L-glutamic acid	-	-	-	-	+	+	-	-	-	-	+	+	-	-	-	-	-	-	-
L-histidine	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-
hydroxy L-proline	+	+	+	+	+	+	+	+	+	+	+	-	+	+	-	-	-	-	-
L-leucine	+	+	-	-	+	+	-	-	-	-	+	+	-	-	-	-	+	+	-
L-ornithine	-	-	+	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-
L-phenyl alanine	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	+	-	-
L-proline	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
L-pyro glutamic acid	+	+	-	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-
D-serine	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L-serine	+	-	-	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-
L-threonine	-	-	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-
D,L-carnitine	+	-	-	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-
Y-amino butyric acid	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
urocanic acid	+	+	-	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-
inosine	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	-
uridine	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-
thymidine	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
phenyl ethylamine	-	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-
putrescine	+	-	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	-	-
2-amino ethanol	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	-	-
2,3-butanediol	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
glycerol	-	-	+	-	+	+	-	-	-	-	+	+	-	-	-	-	-	-	-
D,L-a-glycerol phosphate	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
glucose-1-phosphate	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
glucose-6-phosphate	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-

[illegible]